

Handbook of Comparative and Historical Indo-European Linguistics

HSK 41.2

Handbücher zur Sprach- und Kommunikations- wissenschaft

Handbooks of Linguistics
and Communication Science

Manuels de linguistique et
des sciences de communication

Mitbegründet von Gerold Ungeheuer
Mitherausgegeben (1985–2001) von Hugo Steger

Herausgegeben von / Edited by / Edités par
Herbert Ernst Wiegand

Band 41.2

De Gruyter Mouton

Handbook of Comparative and Historical Indo-European Linguistics

Edited by

Jared Klein

Brian Joseph

Matthias Fritz

In cooperation with Mark Wenthe

De Gruyter Mouton

ISBN 978-3-11-052161-0
e-ISBN (PDF) 978-3-11-052387-4
e-ISBN (EPUB) 978-3-11-052175-7
ISSN 1861-5090

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2017 Walter de Gruyter GmbH, Berlin/Boston

Typesetting: Meta Systems Publishing & Printservices GmbH, Wustermark

Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen

Cover design: Martin Zech, Bremen

⊗ Printed on acid-free paper

Printed in Germany

www.degruyter.com

Contents

Volume 2

VIII. Italic

46. The documentation of Italic	733
47. The phonology of Italic	743
48. The morphology of Italic	751
49. The syntax of Italic	804
50. The lexicon of Italic	828
51. The dialectology of Italic	835
52. The evolution of Italic	858

IX. Germanic

53. The documentation of Germanic	875
54. The phonology of Germanic	888
55. The morphology of Germanic	913
56. The syntax of Germanic	954
57. The lexicon of Germanic	974
58. The dialectology of Germanic	986
59. The evolution of Germanic	1002

X. Armenian

60. The documentation of Armenian	1028
61. The phonology of Classical Armenian	1037
62. The morphology of Armenian	1080
63. The syntax of Classical Armenian	1097
64. The lexicon of Armenian	1115
65. The dialectology of Armenian	1132
66. The evolution of Armenian	1146

XI. Celtic

67. The documentation of Celtic	1168
68. The phonology of Celtic	1188
69. The morphology of Celtic	1203
70. The syntax of Celtic	1218
71. The lexicon of Celtic	1250
72. The dialectology of Celtic	1264
73. The evolution of Celtic	1274

XII. Tocharian

74. The documentation of Tocharian	1298
75. The phonology of Tocharian	1304
76. The morphology of Tocharian	1335
77. The syntax of Tocharian	1352
78. The lexicon of Tocharian	1365
79. The dialectology of Tocharian	1389

Volume 1

Preface	v
-------------------	---

I. General and methodological issues

1. Comparison and relationship of languages	1
2. Language contact and Indo-European linguistics	7
3. Methods in reconstruction	15
4. The sources for Indo-European reconstruction	20
5. The writing systems of Indo-European	26
6. Indo-European dialectology	62
7. The culture of the speakers of Proto-Indo-European	75
8. The homeland of the speakers of Proto-Indo-European	85

II. The application of the comparative method in selected language groups other than Indo-European

9. The comparative method in Semitic linguistics	93
10. The comparative method in Uralic linguistics	98
11. The comparative method in Caucasian linguistics	105
12. The comparative method in African linguistics	114
13. The comparative method in Austronesian linguistics	121
14. The comparative method in Australian linguistics	129

III. Historical perspectives on Indo-European linguistics

15. Intuition, exploration, and assertion of the Indo-European language relationship	138
16. Indo-European linguistics in the 19 th and 20 th centuries: beginnings, establishment, remodeling, refinement, and extension(s)	171
17. Encyclopedic works on Indo-European linguistics	210
18. The impact of Hittite and Tocharian: Rethinking Indo-European in the 20 th century and beyond	220

IV. Anatolian

19. The documentation of Anatolian	239
20. The phonology of Anatolian	249
21. The morphology of Anatolian	256
22. The syntax of Anatolian: The simple sentence	274
23. The lexicon of Anatolian	291
24. The dialectology of Anatolian	298

V. Indic

25. The documentation of Indic	309
26. The phonology of Indic	325
27. The morphology of Indic (old Indo-Aryan)	344
28. The syntax of Indic	377
29. The lexicon of Indic	409
30. The dialectology of Indic	417
31. The evolution of Indic	447

VI. Iranian

32. The documentation of Iranian	471
33. The phonology of Iranian	481
34. The morphology of Iranian	503
35. The syntax of Iranian	549
36. The lexicon of Iranian	566
37. The dialectology of Iranian	599
38. The evolution of Iranian	608

VII. Greek

39. The documentation of Greek	625
40. The phonology of Greek	638
41. The morphology of Greek	654
42. The syntax of Greek	682
43. The lexicon of Greek	695
44. The dialectology of Greek	710
45. The evolution of Greek	717

Volume 3

XIII. Slavic

80. The documentation of Slavic
81. The phonology of Slavic
82. The morphology of Slavic
83. The syntax of Slavic
84. The lexicon of Slavic
85. The dialectology of Slavic
86. The evolution of Slavic

XIV. Baltic

87. The documentation of Baltic
88. The phonology of Baltic
89. The morphology of Baltic
90. The syntax of Baltic
91. The lexicon of Baltic
92. The dialectology of Baltic
93. The evolution of Baltic

XV. Albanian

94. The documentation of Albanian
95. The phonology of Albanian
96. The morphology of Albanian
97. The syntax of Albanian
98. The lexicon of Albanian
99. The dialectology of Albanian
100. The evolution of Albanian

XVI. Languages of fragmentary attestation

101. Phrygian
102. Venetic
103. Messapic
104. Thracian
105. Sicilian
106. Lusitanian
107. Macedonian
108. Illyrian
109. Pelasgian

XVII. Indo-Iranian

- 110. The phonology of Indo-Iranian
- 111. The morphology of Indo-Iranian
- 112. The syntax of Indo-Iranian
- 113. The lexicon of Indo-Iranian

XVIII. Balto-Slavic

- 114. Balto-Slavic
- 115. The phonology of Balto-Slavic
- 116. The morphology of Balto-Slavic
- 117. The syntax of Balto-Slavic
- 118. The lexicon of Balto-Slavic

XIX. Wider configurations and contacts

- 119. The shared features of Italic and Celtic
- 120. Graeco-Anatolian contacts in the Mycenaean Period

XX. Proto-Indo-European

- 121. The phonology of Proto-Indo-European
- 122. The morphology of Proto-Indo-European
- 123. The syntax of Proto-Indo-European
- 124. The lexicon of Proto-Indo-European

XXI. Beyond Proto-Indo-European

- 125. More remote relationships of Proto-Indo-European

VIII. Italic

46. The documentation of Italic

Before defining the documentation of “Italic”, it is appropriate to establish how this term is employed, since its different uses still divide the methodological tendencies of comparative linguistics developed in the 20th century. In the genealogical perspective of the Indo-European languages introduced by A. Schleicher, the term “Italic”, as a correspondent of German “italisch”, denotes the unitary branch of Indo-European located in Italy, from which two sub-groups derive: Latin-Faliscan, on the one hand, and Sabellic, on the other. Moreover, the German terminology distinguished conservative and archaic features, called “uritalisch” with further distinction between “fruh-uritalisch” and “spat-uritalisch”, from common features determined by secondary and more recent developments, called “gemeinitalisch”. Also “kernitalisch”, referring to both Latin and Venetic, has been adopted (De Bernardo Stempel 2000). French scholarship used the term “italique commun” as corresponding more closely to “uritalisch” than to “gemeinitalisch” (Meillet 1928), while “italique” has been applied more extensively in a non-strictly genealogical sense, to include also other IE languages of Italy, such as Venetic, Sicel, and Messapic (Lejeune 1945, 1950; Poccetti 2001). In agreement with the findings of linguistic geography and language contact, Italian scholarship developed the thesis of Walde (1917), who challenged the prehistoric unity of the Latin and Sabellic groups, and therefore undermined their genealogical relationships in the sense of Schleicher’s *Stammbaumtheorie*. For example, Devoto (1929) considered the isoglosses between the two groups as resulting from common innovations which were understood by Pisani (1954) to have resulted from a “linguistic league” (*Sprachbund*) or, in other words, from a koiné (Triantafillis 2005). Both approaches were criticized by Rix (1994). Scholars adopting this view use the term “Italic” as a denotation of the Sabellic languages, in the sense of “non-Latin” or “anti-Latin”. This value has its correlate in history, where the Italic people represented by Oscan speaking Samnites are usually counter-posed to Rome embodying the Latin people. The term “Italic” has been used also in English, at various times and in different perspectives, in the above-mentioned meanings.

The term “Italic” is employed here to refer to the group of languages that represent the intermediate unity of Proto-Indo-European, according to Schleicher’s genealogical tree, while the term “Sabellic” is applied to the branch other than that of Latin and Faliscan. The languages belonging to the “Italic” group have a qualitatively and quantitatively various documentation. If one adopts a classification of languages according to quantitative, qualitative, and functional parameters of texts, a fundamental distinction may be made between Latin on the one hand, and Faliscan and the Sabellic languages on the other. In agreement with this classification, Latin is considered a language with “a large corpus” (*Grosscorpussprache*), similarly to Ancient Greek and Old Indic for example, while Faliscan and Sabellic are considered fragmentarily recorded languages (*Trümmersprachen*), like Gaulish, Celtiberian, Lydian, or Phrygian (Mayrhofer 1980; Untermann 1980, 1989). The condition of Latin as a “large-corpus language” is determined by the presence of a conspicuous literary tradition and of a considerable epigraphic corpus, encompassing an extended time span, as well as by a grammatical tradition

that informs us about the meta-linguistic competences of the speakers. Accordingly, Latin differs from “small-corpus languages” (*Kleincorpussprachen*) such as Gothic or Old English, whose documentation is limited to few types of texts and to a more constrained diachronic period.

Instead, languages with a fragmentary record are attested only by epigraphic means and usually by a very restricted number of text types. This condition offers a limited and partial image of the language, and impinges upon a synchronic and diachronic description of the grammar and of the lexicon, as well as upon the identification of sociolinguistic varieties and of the strategies employed in the pragmatics of communication. This classification also includes “residual” languages (*Restsprachen*), whose function and use can be comprehended in the perspective of sociolinguistic registers or technical and specialistic languages, located in conditions of diglossia.

Thus, a language can share more than one classification. For example, Gothic and Old Church Slavic are at the same time small-corpus languages and “residual” languages (*Restsprachen*), since they are represented only by a restricted class of texts. Similarly, *Trümmersprachen* can share the condition of *Restsprachen* if an epigraphic corpus coincides with a specialistic language, as in the case of Old Persian, which reflects to a large extent the chancery language (*Kanzleisprache*) of the Achaemenid dynasty.

The fragmentarily attested languages of the Italic branch (Faliscan and Sabellic languages) represent types of *Restsprachen*, since their documentation consists of fixed and repetitive formulas, characteristic of certain types of epigraphic texts (burial inscriptions, votive inscriptions, ritual texts). These formulas, pertaining to the specialistic language of the persons entrusted with the task of writing epigraphic texts, probably did not reflect the ordinary language, but were certainly a powerful vehicle of standardization and diffusion of the language. This can be clearly observed in Latin, where the high quantity of epigraphy, produced in every part of the Roman Empire until its latest epoch, hints at different types of linguistic variation: such variation is almost nonexistent on the syntactic and lexical level, minimal on the morphological level, and quite extensive on the phonetic level (Herman 1990). A similar phenomenon can be seen in the Sabellic languages, in particular in Oscan, which in the final phase of its documentation undergoes a process of standardization in the language of public inscriptions.

The documentation of Faliscan and of the Sabellic languages, amounting in all to approximately a thousand texts ranging from three to six centuries of the first millennium BCE, is similar to the documentation of other fragmentarily recorded (IE and non-IE) languages in the Mediterranean area. In contrast, the millennium-long documentation of Latin extends far beyond the history of Rome, coinciding with the linguistic and cultural history of Western Europe. For this reason, Latin occupies a rather remarkable position even in the IE domain, revealed as well by the evolutionary process that brought about its break-up into the Romance languages. This condition makes absolutely exceptional the documentation of Latin in its whole history for both IE and Romance linguistics.

Thus, Latin can be classified as a “large-corpus language” from the viewpoint of its global history, the high degree of its grammatical description, and the size of its lexicon. On the other hand, it can be classified as a “residual language” in regard to the role it has been exerting on certain registers (for example, even today in ecclesiastic and judiciary domains). Moreover, Latin can also be considered a fragmentarily attested language at least in two respects: 1) the non-documented varieties of the language, mainly related to the spoken levels, which form a more or less continuous flow of “submerged Latin”;

2) the type of Latin attested in the period preceding the appearance of the literary documentation, ranging from the 7th to the 3rd century BCE).

Pre-literary Latin, characterized by a scanty and qualitatively meagre documentation (for epigraphy, cf. Hartmann 2005), appears phonetically and morphologically very different from the literary language, so that it was almost unintelligible to the Roman speakers of the 2nd century BCE. The texts of pre-literary Latin still present considerable problems for interpretation and for grammatical description and give the impression of representing a language different from that codified by the literary and administrative standard, which is labelled *urbanitas* by its users.

Moreover, some documentary aspects of Latin can be regarded as items of a “small-corpus language”, such as Merovingian Latin (7th–8th centuries CE), its latest variety represented by a homogeneous corpus, different from the preceding texts (real diplomas, private and public acts, and scholastic tradition), also of interest in the perspective of the Romance languages (Calboli 1984).

This peculiar documentation of Latin makes it extremely difficult to define the borderline between the end of its documentation and the earliest documentation of the Romance languages. The beginning of the evolution toward the Romance languages is already apparent in Republican Latin and in particular categories of inscriptions or texts on papyrus that are less subjected to the standard language. Various identified regional differences and dialectal phases of Latin and their sociolinguistic involvements (Adams 2007, 2013) underlie the development toward Romance languages: the earliest is at the dawn of its documentation, reflecting varieties of several communities of archaic Latium (*Latinisch*); the second is the diffusion of the Roman model (*Lateinisch*) in ancient Italy; the third is the diffusion in the rest of the empire through different steps (*Roman[ic]us*).

The appearance of the documentation of Latin, as well as of other languages of ancient Italy, is by and large coincident with the spread of alphabetic writing in Italy. Traces of syllabic writing from the Aegean discovered in Italy on objects going back to the 2nd millennium BCE did not impinge on the indigenous communities. The diffusion of the alphabetic system in the first part of the 1st millennium BCE was favoured on the one hand by the imprint of Greek colonization, which – unlike the Phoenician colonization – tended to establish permanent settlements within a certain environment, and on the other hand by the different attitudes of the indigenous communities, where writing emerges as a hallmark of high social status. This is the reason why: a) the alphabetic writing of the Phoenicians, albeit attested in Italy earlier than Greek writing (9th–8th centuries), did not have any effect among the indigenous cultures; b) Greek writing, albeit conveyed by immigrants, merchants and craftsmen, was taken over by high social classes and members of the local aristocracies; c) introducing an alphabetic writing system was a means as well as an aspect of social transformations.

A social use of writing so different from the culture that exported the alphabetic system also determined – almost simultaneously with the adoption of the alphabet – new typologies of texts related to the function and nature of the material support of the texts themselves. For example, the earliest texts in Latium are written on luxury items. Here the writing represents an extra value to indicate the signature of the craftsman or alternatively signals gifts and exchanges among persons of the same social level. This is true, for example, of the inscription on the golden fibula of Praeneste, the earliest Latin document, whose authenticity has been wrongly contested (Franchi de Bellis 2007, 2011).

The diffusion of the Greek alphabet in Italy was not uniform in space and time. The first records of Greek writing are not located in areas directly influenced by Greek colonization (Magna Graecia and Sicily), but rather in peripheral regions with respect to the expansion of Greek civilization: signs of the Greek alphabet are attested in the first decades of the 8th century BCE on ceramic material discovered in Latium and in Bologna almost contemporaneously with the earliest manifestations of alphabetic writing in Greece, or at least preceding the earliest documents of the Western colonies by half a century.

In the same way, the earliest documents of the Italic group do not belong to the languages more directly exposed to contact with Greek colonization, namely the Sabellic languages in Southern Italy and in Sicily, but to languages located in the central area of the Tyrrhenian coast: Latin and Etruscan, two languages totally different from each other from a genealogical and a typological point of view, but in close contact around the course of the River Tiber.

The course of the Tiber was also the vehicle of reception and diffusion of Greek alphabets in the archaic age, and of their adjustment to the writings of the main linguistic groups of the Italic Peninsula, namely the Italic branch (Latin, Faliscan, and Sabellic languages) and Etruscan, a non-IE language. Therefore the earliest evidence for all these languages is located around the basin of this river. The leading role of the Euboic alphabet is intertwined with other Greek alphabetic types (e.g., the so-called “red” alphabets, such as the alphabet of Corinth and of the Doric colonies). The archaic alphabet used by the Sabellic language (Sabine and South-Picene) points to an origin and a development which are independent from the Latin and Etruscan alphabets, as shown by the shape of signs, the use of “dead letters” to indicate the different sounds of both palatal and velar vowels, and the introduction of a special sign for /f/.

However, the formation of each “national” alphabet has been also determined by inter-alphabetic contacts, as shown by the Etruscan origin of the sign <c> with the value of a voiceless (and not voiced!) stop in the Latin alphabet. In the earliest documentation no substantial alphabetic distinction emerges among the languages in contact in the Tiber basin. The stabilization and the creation of alphabets for each type of language can be outlined only from the 6th century BCE with the appearance of a clearer alphabetic distinction between Etruscan and Latin on the one hand, and between Faliscan and the Sabellic languages on the other. The sign of the fricative /f/, which is absent in the model alphabet, represents the most significant trace of such a distinction.

The River Tiber was an important linguistic and cultural borderline between Latin- and Etruscan-speaking areas (in its low course), and Etruscan- and Sabellic-speaking areas (in its middle course), until the early Imperial age. But simultaneously the Tiber formed also an area of strong contacts and convergences of these languages, owing to social mobility. Archaic documentations (7th–5th centuries BCE) of all these languages reflect this particular condition. Actually, some early Etruscan documents are found in a Latin area (e.g., the earliest texts discovered in Rome), or conversely Sabellic texts come from Etruscan districts. More stable settlements of Italic speaking people are attested in the Etruscan area. This is the case of the Faliscan language and of the speech of Capena, which is close to Sabine, in a territory strongly influenced by Faliscan culture. That is why Etruscan evidence may contribute to clarify phonetic or morphological facts of the Italic languages.

Pre-literary Latin (prior to 300 BCE) is represented by about 20 epigraphic remains of variously fragmentary sizes (some of them consist of only a few letters). The longest and more important texts (known partly from the late 19th century, e.g. the *Lapis Niger* or the *Duenos* vase inscriptions, and partly from some thirty or so years ago, e.g. the *Lapis Satricanus*) still give rise to heavy interpretive problems. This makes the hermeneutics of pre-literary Latin not so different from that of fragmentarily attested languages. An even more serious problem (owing to the complications of the manuscript tradition) is raised by the scanty fragments handed down by literary quotations, such as the formulas of the *carmina Saliorum*, which according to Quintilian were obscure even to the sacerdotal college in charge of reciting them. The fragments of the XII Tables (about 450 BCE according to the historical sources) are also problematic: one may wonder whether their linguistic features reliably correspond to their original redaction, or whether they were adapted to make them understandable in order to favor their oral transmission (according to Cicero).

In classical Rome the practice of renovating archaic texts proceeded in parallel with the artificial “aging” or archaizing trend, exerted in epigraphic texts as well as in literary style (for example, in Sallust). In this process, however, the threshold of archaism did not go beyond the beginning of Roman literature. This clearly shows that such a threshold not only was a barrier to a linguistic understanding but also represented a cultural starting point for the history of the Latin language from a Roman’s perspective. Indeed, in the Roman meta-linguistic awareness, the archaic documentation was consigned to oblivion by the Gaulish invasion in 390 BCE, which symbolically marked the loss of the main written documents that represented historical memory of archaic Latium. This makes it still difficult to establish the sources (written and/or oral) upon which the late republican Romans could reconstruct their archaic history.

The beginning of the Latin literary age coincides with a different relationship established between orality and writing also in other documentary aspects. Partially, this was an effect of social events, such as the subtraction of the right to power of the *Pontifices*, which brings about the great tradition of writing juridical texts (from the age of Appius Claudius). Partially it was an outcome of widespread Hellenism throughout Roman society (from aristocracy to middle-class to slaves), that inserted Rome in a wider circulation on a Mediterranean scale. The adherence to Hellenistic models conditions various literary genres of Latin literature, such as theatre, prose, and poetry. This produced more and more deep contacts between Greek and Latin, which became the official languages of the Roman world. A widespread bilingualism as well as interlinguistic experiences (translations, calques, borrowings) strongly affected the oral and written structures of Latin and accelerated the process of convergence between both languages, whose effects are ultimately reflected by the Romance languages on the one hand and by medieval and Modern Greek on the other.

Thus, the literary age changed text typologies on both the oral and written levels. Also writing materials changed, as shown by stone inscriptions of public acts, which were previously recorded by *Pontifices* on ephemeral or perishable material (*tabula dealbata*). Similarly, sepulchral or celebratory epigraphy replaces the genre of the archaic *elogia*, which were previously assigned to orality on the occasion of funerals or commemorations. New theatre genres performed on a Greek model represent the written correspondences to the indigenous scenic forms (such as *Atellanae*, *Fescennini*), which were devoted to improvisation or to enactment of plots transmitted by oral learning.

Latin epigraphy, forming the most extensive corpus in the ancient world with its more than 400,000 inscriptions, reflects a language in uniformity with the literary standard. Latin inscriptions were mainly produced by a professional class in charge of the writing but also became a powerful vehicle for the diffusion of literacy in the Roman domain, which owing to this vehicle possessed one of the highest rates of alphabetization in the ancient world. The variations of epigraphic Latin are very limited in the different parts of the Empire, while the taxonomy of texts in the various regions (milestones, military diplomas, laws) is heterogeneous, for different reasons.

The Faliscan documentation, geographically concentrated in a small territory around the settlement of *Falerii* situated in the middle basin of the Tiber, is chronologically divided into three groups: a) early Faliscan (7th–5th BCE); b) middle Faliscan (400–240 BCE); c) late Faliscan (after 240 BCE). In this year the ancient site of *Falerii* was destroyed by the Romans, which compelled the inhabitants to move to a neighbouring place, named *Falerii novi* (as distinct from *Falerii veteres*) (Bakkum 2009).

Faliscan together with Latin provides the most ancient documentation of IE languages in Italy (around the middle of the 7th century BCE): evidence for Sabellic languages started about a century later. Faliscan epigraphic materials, amounting to around 500 texts, consist mostly of pottery inscriptions and sepulchral inscriptions, few votive texts and very few public documents.

The Faliscan area is diachronically and synchronically crossed by different alphabetic models and linguistic contacts. Faliscan of the latest period adopted the Latin alphabet and is strongly influenced by Roman Latin, a circumstance that makes it difficult to distinguish texts in Faliscan mingled with Latin from those in dialectal Latin. Faliscan of the archaic period is alphabetically and linguistically influenced by its Etruscan neighbourhood. The origin of the archaic Faliscan alphabet poses problems analogous to that of Latin: a direct derivation from a Greek alphabet with Etruscan influences. But from an early period Faliscan writing presents its own identity by creating a special sign for /f/. Contacts are also revealed by Etruscan sepulchral inscriptions attested in Faliscan necropolises and Faliscan inscriptions found in Sabine (that is, Sabellic-speaking) areas (Cristofani 1988).

The documentation of the Sabellic languages encompasses a quite extended territory of the Peninsula, within the high course of the Tiber, the Ionian Sea, and the Straits of Sicily (except Apulia). This documentation refers to a system of populations with shared cultural features, as well as with heavy differences among them. Latin and Greek sources identified them under various regional and sub-regional designations, such as, for the archaic age (7th–5th centuries BCE), *Ausones*, *Opici*, *Oenotri*, Sabines, Picenians, Umbrians, and, for a more recent age (4th–1st centuries BCE), *Samnites*, *Campani*, *Lucani*, *Bruttii*, *Marsi*, *Paeligni*, *Vestini*, *Marrucini*.

Until a few decades ago the documentary knowledge of these languages was limited to their recent phase, conventionally labelled Oscan and Umbrian. This phase, starting from the middle of the 4th century BCE, came to an end in the middle of the 1st century BCE, when the writing of the Sabellic languages definitively stopped, at least in the formal and official stage. Oscan and other Sabellic varieties continued to be understood and, probably, were used in colloquial contexts until the Augustan age, as shown by some graffiti in Pompei and Herculaneum during the Roman period. The disappearance of these languages from the written record is related to the political and administrative Romanization of Italy after the *Bellum Sociale* (90 BC), while the process of Latinization

had already begun in the early 3rd century BCE. The end of their oral use coincides with the diffusion of Latin affected by a condition of bilingualism (or rather diglossia), characterizing generally the history of “submerged” Latin.

The re-edition of a corpus of texts improperly defined as “South-Picene” (with a generic reference to their provenance) and newly discovered texts from other regions enabled the scholarly world to recognize a more archaic level of the language, which can be defined as “Paleo-Sabellic” (Marinetti 1985). These documents present no standardized language even in texts sharing the same typology (monumental and commemorative stone inscriptions) and the same alphabetic shape (South-Picene corpus). The recently acquired Paleo-Sabellic documentation (Lazzarini and Poccetti 2001) permitted scholars to establish that 1) the distinction between two alphabetic areas in the domain of the Sabellic languages, which was well known in later documentation of the Oscan Language, is to be dated back to archaic times: one uses the same Greek alphabet as the colonies of Magna Graecia (specifically the alphabet of the Achaean colonies); the other uses a transformation of a Greek alphabet, introduced throughout the course of the Tiber; 2) there were contacts between the two alphabetic areas, in particular shown by the introduction of the sign /f/ from the Central (South-Picene) alphabetic area to the Southern (Greek) alphabetic area; 3) an independent script tradition emerges in Pre-Samnite Campania, showing important links between Central alphabets and Southern alphabets (Poccetti 2010).

Oscan and Umbrian documentation is divided into two main alphabetic systems: a) Etruscan based alphabets (with different adaptations in Oscan and Umbrian); b) a Greek (Hellenistic) alphabet adapted to the Oscan language. Later, with the diffusion of Latin, the Latin alphabet was (alternatively and not unitarily) adopted for writing Sabellic languages.

In spite of changing alphabetic models, from Paleo-Sabellic down to Oscan and Umbrian, Sabellic languages show an uninterrupted writing tradition in the respective alphabetic areas; and the various Sabellic communities of the centre and the south of the Peninsula show a continuity of linguistic contacts as well as a practice of writing, when passing from one alphabet to another one. Nevertheless, archaic differences between northern Sabellic documentation and the southern one allow us to conclude that the dialectal distinction between a Sabellic group of central Italy (“Umbrian”) and one of southern Italy (“Oscan”) is earlier than the appearance of their documentation. Therefore the formation of the dialectal varieties of the individual communities, and especially of the macro-distinction between the Oscan and the Umbrian groups, is not the result of recent migrations, but rather an outcome of sociolinguistic facts and of contacts with other languages in historical times. Actually, Oscan innovations in comparison with Umbrian, whose oldest documents preserve archaic features of the Paleo-Sabellic phase, result from its connections with a Mediterranean linguistic circulation and, particularly from Greek and Latin influences.

On the other hand, our knowledge of lexical and syntactic-grammatical structures of the Sabellic languages is mainly determined by their most extended text, the Tables of Iguvium, a religious ritual (in a partly Etruscan and partly Latin alphabet); the most prominent editions are those of Poultney (1958) and Prodocimi (1984, 2015 with largest commentary). Because of its intrinsic character, this important text presents linguistic features more conservative and more similar to the “Paleo-Sabellic” corpus, in compari-

son with Umbrian minor inscriptions (about 25: Rocca 1996; Agostiniani, Colderini, and Massarelli 2011), which reveal a strong mixture with Latin.

In contrast, the two most significant Oscan texts (the *Tabula Bantina*, a public law in the late phase of Romanization, and the *Cippus Abellanus*, a treaty between two Campanian communities) are variously affected by formulas and by syntactic structures of their corresponding Greek and Roman textual types. These foreign influences were combined into a tradition of civil and religious normative texts which the Oscan culture independently possessed, as shown by the recently discovered fragment of a Lucanian law chronologically ranging from the late 4th and early 3rd century BCE. (Gualtieri and Poccetti 2001) that opens new perspectives on the formation of legal style in both Sabellic and Latin languages (Poccetti 2009).

The name “Oscan” is assigned by the Romans to the language of the regions located immediately to the south of Latium. Thus, this language must have appeared relatively unitary to their eyes. The same impression is offered by the public epigraphy, which presents a remarkable form of language standardization. This condition is probably related to the existence of literary genres in the Oscan language, such as the theatre of the *fabulae Atellanae*, which, according to Strabo, were still performed in their original language in Imperial Rome. Moreover, the existence of an Oscan literary language in contexts of bilingualism with Greek and/or Latin is indirectly confirmed by the personalities of the earliest Latin literary figures, who came from regions where Oscan was spoken (e.g. *Naevius*, *Lucilius*, *Accius*, *Pacuvius*) or claimed to possess competence in the Oscan language (such as *Ennius*) (Adams 2003). The presence of Oscan words (or terms of Sabellic origin, more broadly speaking) in Latin literature is a further clue to literary bilingualism and represents another channel of knowledge of these languages: the glosses. These glosses, provided with lexical-semantic information, are transmitted by heterogeneous sources that certainly had different interests in assembling these materials as well as an indirect knowledge of their specific language. For example, the glosses ascribed to Oscan number thirteen in Festus, three in Varro, one in Servius, and two in others (Poccetti 2004). Varro transmits fourteen glosses which he assigns to Sabine, while assigning another three to Oscan. Festus, too, includes seven Sabine glosses, while Servius preserves six. Umbrian is represented by only three glosses in Festus (for distinctions among “Sabine” glosses, see Negri 1993). Finally, the epigraphic documentation of Oscan, which is distributed between an alphabet of Etruscan origin (used in Campania and in Samnium) and a Greek alphabet (of a Hellenistic type, used in Lucania, Brutium, and Messina), comprises over 600 inscriptions of various dimensions. The outstanding collections of epigraphic documents of Sabellic languages are Vetter (1953) and Rix (2002). The very recent and up-to-date Crawford (2011) is provided with photographs, drawings and archeological details of each text.

References

- Adams, James N.
 2003 *Bilingualism and Latin Language*. Cambridge: Cambridge University Press.
- Adams, James N.
 2007 *The Regional Diversification of Latin 200 BC–AD 600*. Cambridge: Cambridge University Press.

- Adams, James N.
2013 *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Agostiniani, Luciano, Alberto Calderini, and Riccardo Massarelli (eds.)
2011 *Screhto est: lingua e scrittura degli antichi umbri*. Perugia: Univerità degli studi.
- Bakkum, Gabriël Cornelis Leonides Maria
2009 *The Latin Dialect of the Ager Faliscus. 150 Years of Scholarship*. 2 vols. Ph.D. dissertation, Amsterdam University. Amsterdam: University Press.
- Calboli, Gualtiero
1984 Il latino merovingico fra latino volgare e latino medievale. In: Edoardo Vineis (ed.), *Latino volgare, latino medievale, lingue romanze. Atti del convegno della Società Italiana di Glottologia*. Pisa: Giardini, 63–81.
- Crawford, Michael H. (ed.)
2011 *Imagines Italicae. A Corpus of Italic Inscriptions*. 3 vols. In collaboration with William M. Broadhead, James Clackson, Federico Santangelo, S. Thompson, Margaret Watmough with computing by Errietta M. A. Bissa and Gabriel Boddard. London: Institute of Classical Studies, University of London.
- Cristofani, Mauro
1988 Etruschi nell'agro falisco. *Papers of the British School at Rome* 56: 13–24. [Reprinted 2001 in *Scripta Selecta*, Vol. 1. Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali, 199–214.]
- De Bernardo Stempel, Patrizia
2000 Kernitalisch, Latein, Venetisch: Ein Etappenmodell. In: Michaela Ofitsch and Christian Zinko (eds.), *125 Jahre Indogermanistik in Graz*. Graz: Leykam, 47–70.
- Devoto, Giacomo
1929 Italo-greco e italo-celtico. In: Giacomo Devoto and Benvenuto Terracini (eds.), *Silloge linguistica dedicata alla memoria di G. I. Ascoli*. Turin: Chiantore, 200–239. [Reprinted 1958 in *Scritti Minori*. Florence: Le Monnier, 129–154.]
- Franchi de Bellis, Annalisa
2007 La fibula di Numasio e la coppa dei Veturii. *Quaderni dell'Istituto di Linguistica dell'Università di Urbino* 12: 65–142.
- Franchi de Bellis, Annalisa
2011 La fibula prenestina. Margherita Guarducci e Wolfgang Helbig, presunto falsario. In: Simo Örmä and Kaj Sandberg (eds.), *Wolfgang Helbig e la scienza dell'antichità del suo tempo. Atti del convegno internazionale. Roma 2. 2. 2009*. (Acta Instituti Romani Finlandiae 37). Rome: Institutum Romanum Finlandiae, 181–215.
- Gualtieri, Maurizio and Paolo Poccetti
2001 *La Tabula Bronzea di Roccagloriosa*. In: *Roccagloriosa II. L'oppidum Lucano e il territorio*. (Collection du Centre Jean Bérard). Naples: Centre Jean Bérard.
- Hartmann, Markus
2005 *Die frühlateinischen Inschriften und ihre Datierung*. Bremen: Hempen.
- Herman, József
1990 *Du latin aux langues romanes. Étude de linguistique historique*. Tübingen: Niemeyer.
- Lazzarini, Maria L. and Paolo Poccetti
2001 *L'iscrizione paleoitalica di Tortora*. In: Maurizio Bugno and Concetta Masseria (eds.), *Il mondo Enotrio tra VI e V secolo a.C. Atti dei seminari napoletani (1996–1998)*. (Quaderni di Ostraka 1, 2). Naples: Loffredo.
- Lejeune, Michel
1945 La position du latin sur le domaine indo-européen. In: *Mémorial des Études Latines*. Paris: Les Belles-Lettres, 7–31.
- Lejeune, Michel
1950 Notes de linguistique italique, I. Latin et chronologie «italique». *Revue des Études Latines* 28: 97–104.

- Marinetti, Anna
1985 *Le iscrizioni sud-picene*. Florence: Olschki.
- Mayrhofer, Manfred
1980 *Zur Gestaltung des etymologischen Wörterbuchs einer "Grosscorpusprache"*. (Sitzungsberichte der Österreichischen Akademie der Wissenschaften, 368.11). Vienna: Österreichische Akademie der Wissenschaften.
- Meillet, Antoine
1928 *Esquisse d'une histoire de la langue latine*. Paris: Klincksieck.
- Negri, Mario
1993 Il sabino di Roma. Considerazioni linguistiche. In: Helmut Rix (ed.), *Oskisch-Umbrisch. Texte und Grammatik. (Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia, vom 25.–28. September 1991 in Freiburg)*. Wiesbaden: Reichert, 196–206.
- Pisani, Vittore
1954 Zur Sprachgeschichte Altitaliens. *Rheinisches Museum* 97: 47–68. [Reprinted 1959 in *Saggi di linguistica storica*. Turin: Rosenberg & Sellier, 160–180.]
- Pocchetti, Paolo
2001 Perspectives actuelles dans le domaine des langues italiques dans le sillage de Michel Lejeune. *Comptes Rendus de l'Académie des Inscriptions et Belles-Lettres*: 185–196.
- Pocchetti, Paolo
2004 Servio come fonte di documentazioni linguistiche ed etnografiche dell'Italia antica tra tradizioni indigene e filtri allogloti. In: Carlo Santini and Fabio Stok (eds.), *Hinc Italae gentes. Geopolitica ed etnografia dell'Italia nel commento di Servio all'Eneide*. Pisa: ETS, 259–306.
- Pocchetti, Paolo
2009 Lineamenti di tradizioni non romane di testi normativi. In: Augusto Ancillotti and Alberto Calderini (eds.), *L'umbro e le altre lingue dell'Italia mediana antica*. Perugia: Jama, 165–248.
- Pocchetti, Paolo
2010 Intorno ai nuovi documenti di area sorrentina: riflessioni sul novum e sul notum. In: Felice Senatore and Mario Russo (eds.), *Sorrento e la Penisola Sorrentina tra Italici, Etruschi e Greci nel contesto della Campania antica. Atti della giornata di studio in omaggio a Paola Zancani Montuoro (1901–1987), Sorrento, 19 maggio 2007*. (Quaderni di OEBALUS 1). Rome: Scienze e Lettere, 65–102.
- Poultney, James Wilson
1958 *The Bronze Tables of Iguvium*. Baltimore: American Philological Association.
- Prodocimi, Aldo Luigi
1984 *Le Tavole Iguvine*. Florence: Olschki.
- Prodocimi, Aldo Luigi
2015 *Le Tavole Iguvine. II. Preliminari all'interpretazione. La testualità: fatti e metodi*. Firenze: Leo S. Olschki.
- Rix, Helmut
1994 Latein und Sabellisch. Stammbaum und/oder Sprachbund. *Incontri Linguistici* 17: 13–29.
- Rix, Helmut
2002 *Sabellische Texte. Die Texte des Oskischen, Umbrischen und Südpikenischen*. Heidelberg: Winter.
- Rocca, Giovanna
1996 *Iscrizioni umbre minori*. Florence: Olschki.
- Triantafyllis, Elena
2005 Il concetto di coiné italica dal 1930 ad oggi. *Atti dell'Istituto Veneto di Scienze, Lettere ed Arti* 163: 603–683.

Untermann, Jürgen

1980 *Trümmersprachen zwischen Grammatik und Geschichte. (Rheinisch-Westfälische Akademie der Wissenschaften, Vorträge G 245)*. Opladen: Westdeutscher.

Untermann, Jürgen

1989 Zu den Begriffen “Restsprache” und “Trümmersprache”. In: Heinrich Beck (ed.), *Germanische Rest- und Trümmersprachen*. Berlin: De Gruyter, 15–19.

Vetter, Emil

1953 *Handbuch der italischen Dialekte*. Heidelberg: Winter.

Walde, Alois

1917 *Über älteste sprachliche Beziehungen zwischen Kelten und Italikern*. Innsbruck: Rektoratschrift.

Paolo Poccetti, Rome (Italy)

47. The phonology of Italic

- | | |
|---|---|
| 1. Defining Italic | 6. Proto-Sabellic innovations |
| 2. Specific Proto-Italic sound changes | 7. Developments common to Latino-Faliscan and Umbrian |
| 3. Other Proto-Italic sound changes | 8. References |
| 4. Phonological inventory of Proto-Italic | |
| 5. Common Italic developments | |

The aim of this chapter is to outline the major phonological developments which occurred in the history of the Italic languages. The main attention, therefore, is given to features occurring in all or at least in a number of Italic languages, as well as to their relative chronology. To be sure, terms like “Proto-Italic”, “Proto-Sabellic” etc. clearly presuppose the language tree model (assumed by the author) as a schematic rendering of the prehistory of languages. A less literal understanding of these terms at least implies a relative chronology of the sound changes in question.

Epigraphical quotations in local scripts (Etruscan, Oscan alphabet, etc.) are indicated by bold type, those in the Latin alphabet by small capitals. Etruscan examples are quoted from Rix and Meiser (1991).

1. Defining Italic

The Italic branch of the Indo-European language family is characterized by a set of specific phonological and morphological developments which produce the typical shape of the Italic languages and separate them as a group from other branches of Indo-European like Celtic, Germanic, or Greek as well as from the Indo-European Protolanguage. These innovations took place in the prehistory of all Italic languages. Obviously they go back to a hypothetical Proto-Italic (PI) period and antedate developments which are

particular to a single language or subgroup of the Italic branch like L(atino)-F(aliscan) or Sabellic. PI sound changes include, e.g., the spirantization of the PIE voiced aspirates (Mediae aspiratae; MA) and the vocalization of the syllabic liquids (2.1–2.2). On the other hand, innovations common to all Italic languages chronologically preceded by language specific developments belong to the post-PI period and are therefore classified as Common Italic (CI). An example is the loss of short vowels in word internal syllables (5.2), which took place probably about 500 BCE. To be sure, due to the scarce attestation of the Sabellic languages, the attribution of a sound change to one of these chronological layers is not always clear. Nevertheless, in sections 2–7. Italic sound changes are arranged according to their relative chronology and diffusion. Thus, (2) treats sound changes specific to PI and (3) PI sound changes which also occur in other (mainly western) IE branches. The hypothetical phonemic inventory of PI, which originated from these developments, is given in (4). The last three sections present developments of the post-PI period: sound changes common to all Italic languages but posterior to PI (5); the main innovations of the Sabellic languages (“Proto-Sabellic”) (6), and coincidences between LF and Umbrian (7).

2. Specific Proto-Italic sound changes

2.1. Undoubtedly the spirantization of the PIE Mediae Aspiratae $*g^h / *g^h *g^{wh} *b^h *d^h$ is the most “prominent” specifically Italic – or more precisely: Italo-Venetic – feature. Generally speaking, in word-initial position they become voiceless fricatives (1), in word internal position, however, at least originally voiced fricatives (2). A further peculiarity is the merger of the outcomes of word-initial PIE $*g^{wh} *b^h *d^h$ in *f*.

- a) $*d^h$: Lat. *faciō*, O. **fakiad**, U. **façia** 3.sg.subj. ‘make’ (PIE $*d^h_{1}k-$ ‘put’, cf. Phryg. *ad-daket*) – $*b^h$: Lat. *frater*, O. **fratrúm**, U. **FRATROM** g.pl. ‘brother’ < $*b^hrātr-$ (PIE $*b^hreh_2-tr-$, cf. OInd. *bhrātara-*) – $*g^{wh}$: Lat. *faveō* ‘I am propitious to’ < $*g^{wh}ow-eh_1-ye/o-$, U. **FONS** ‘propitious’ n.sg. < $*g^{wh}ow-ni-$ (PIE $*g^{wh}ew-$ ‘regard, respect’, cf. OCS *gověti* ‘to worship’) – $*g^h / *g^h$: Lat. *hortāri* ‘to urge on’, O. **HEREST**, U. **HERIEST** 3.sg.fut. ‘like’ (PIE $*g^her-$ ‘like’, cf. OInd. *háryati* ‘enjoys’). Apparently the same changes, including the merger of word initial $*g^{wh}$, $*b^h$, $*d^h$, occur in Ven(etic), which makes us suppose a Proto-Italo-Venetic unity preceding the PI period: $*d^h$: Ven. **vhaxsto** /fagsto/ ‘offered’ 3.sg.prt. – $*b^h$: Ven. **fouxontia** /fougont-/ [proper name] (probably from PIE $*b^hewg-$ ‘flee’, cf. Gk. $\phi\epsilon\upsilon\gamma\omega$) – $*g^h / *g^h$: Ven. **horeionte** n.du. [pl.?] m.part. prs. ‘willingly’ (PIE $*g^her-$ ‘like’).
- b) The details of the representation of PIE Media Aspiratae in word-internal position are less clear. In some cases it is difficult to establish the precise phonological value of the relevant letters used in various Italic alphabets. Does <C> in Fal(iscan) *lece* ‘lies’ < $*leg^h-ē-ye-ti$ represent a voiced stop /g/ or a (voiced) fricative /ɣ/? In the second case the Fal. development would be similar to that of the Sabellic languages, in which voiced fricatives are the normal outcomes of word-internal PIE MA. Peculiar to the Sabellic languages is the merger of word-internal $*g^{wh}$, $*b^h$, $*d^h$, (see however 7.2.), whereas the change of $*b^h$ and $*d^h$ to voiced stops – obviously through an intermediate stage $*\beta / *d$ – is peculiar to Latin.

d^h*: Lat. *medius*, O. **mefiáí 1.sg.f. ‘the middle one’ (PIE **med^hyo-*, cf. OInd. *mád-hya-*) – **b^h*: Lat. *albus*, U. ALFIR (d.pl.n.) ‘white’ < **h₂alb^ho-* – **g^{wh}*: Lat. *voveō* ‘I promise (to a god)’, U. **vufru** /wofrom/ ‘votive’ < **wog^{wh-ro-}* (PIE **weg^{wh-}* ‘preach’) – **g^h/g^h*: Lat. *vehō* ‘I convey’ (PIE **weg^{h-}*, cf. OInd. *váhati*); O. **feihúss** acc.pl. ‘walls’ < **d^heyg^ho-* (PIE **d^heyg^{h-}* ‘form’, cf. Gk. *τείχος* n. ‘wall’). – Again Ven. shows a development of internal MAs similar to that of the Italic languages, esp. to Latin, cf. **louzerofos** /louderobos/ d.pl. ‘children’ (PIE **h₁lewd^herob^h(y)os* d.pl. ‘free’, cf. Gk. *ἐλεύθερος*, Lat. *liber*, O. LOUFIR). The precise value of the relevant letters <z> and <f>, however, is not entirely clear (/d/ or /ð/? /b/ or /β/?).

In addition, the following sound changes are specifically Italic and seem to date back to the PI period:

2.2. Vocalization of **r*, **l* to *or* (*ur*), *ol* (*ul*) before consonant, to *ar*, *al* before a (sequence of laryngeal plus) vowel: Lat. *mortuus*, Ven. *murtuvoi* d.sg. ‘dead’ (PIE **m₁rtwo-*, cf. OCS *m₁r₂tv₂*) – Lat. *currō* ‘I run’ < **k₁rs-e/o-* – Pael. FORTE(S) g.sg. ‘luck’ (**b^hrti-*, cf. Lat. *fortūna* < **b^hrtu-hn-ah₂*) – U. TRAHUORFI ‘crosswise’ < *(*trāns-*) *worssēd* < **w₁rt-tēd* (ppp. of PIE **wert-* ‘turn’, cf. OInd. *v₁rttá-*) – Lat. *fulg(e)ō* ‘I flash’ < **b^hlg(-eh₁-y)-e/o-*, cf. Gk. *φλέγω* (PIE **b^hleg-* ‘shine’) – O. **kulupu** /kolopom/ g.pl. ‘thieves’ < **kolp-om* < **k₁lp-om* vs. O. *κλοπουστ* 3.sg.fut.pfct. ‘steal’ (PIE **k₁lep-* ‘steal’, cf. Gk. *κλώψ* ‘thief’) – Lat. *carō* ‘meat’ (originally ‘portion’), U. **karu** ‘part’ < **k₁ron/n-* (PIE *(*s*)*ker-* ‘cut’) – Lat. *caleō* ‘I am warm’ < **k₁eh₁-ye/o-* (PIE **kēl-*, cf. Lit. *šiltas* ‘warm’ < **k₁l-to-*).

2.3. Lengthening of vowels preceding (former) spirants: Lat. *sānctus*, O. **saah₁túm** / *sāhtom*/ n.sg.n., U. *sahatam* /*sātam*/ acc.sg.f. ‘holy’ < **sānxto-* < **sankto-* (PIE **sak-* ‘sanctify’).

2.4. Fronting of **ūy* (*-uhy-) to *īy* (‘pius-law’): Lat. *pius*, O. **pihiúí** d.sg. ‘pious’ < **p₁yo-* < **puh₂yo-* ‘pure’, cf. Lat. *pūrus* – Lat. *inciēns* ‘pregnant’ < **-kuh₁-eye-nt-* (PIE **k₁weh₁₋* ‘swell’, cf. Gk. *κυέω* ‘I am pregnant’)

2.5. Lowering of **ow* to *aw* before vowels (“Thurneysen-Havet’s law”; preceding the PI rounding of *ew* to *ow* [3.2.]): Lat. *caueō* ‘I am wary of’ < **kowh₁-eye/o-*, (PIE *(*s*)*kewh₁₋* ‘perceive’), cf. Gk. *κοέω* ‘I perceive’.

2.6. Assimilation of word-internal **gy* (and also **dy*?) to *yy*: Lat. *aiō* /*ayyō*/ ‘I say so’ < **ag-ye/o-*, U. **aiu** /*ayyā*/ n.pl. (‘oracular’) responses’ < **h₂ag-yo-* (PIE **h₂ag^h-* ‘say’) – Lat. *peius* ‘worse’ < **ped-yos*.

2.7. Word-internal PIE **sw* > *rw*: Lat. *Minerva* < **menes-wah₂₋*. The Etruscan tradition (**menerva** Ve 3.10, 6th c. BCE) demonstrates that the development is not due to the rhotacism of the 4th c. BCE. The U. form **menerra* (with regular **rw* > *rr*) is attested indirectly by Etr. **mera** Vs 7.29 (with Etr. syncope **menerra* > **menra* and assimilation).

2.8. Voicing of **-t* in word-final position, attested by the 3.sg.prt. and subj. endings: OL *feced* ‘made’, O. **deded** ‘gave’ < **-et*.

3. Other Proto-Italic sound changes

The following sound changes can also be ascribed to the PI period. In contrast with the innovations mentioned above, these are not “typically” Italic, but occur also in other western IE branches:

3.1. Shortening of long vowels preceding a sequence of resonant/semivowel + consonant (“Osthoff’s law”): Lat. *ventus* ‘wind’ < **wēnto-* < **h₂wéh₁-nt-o-* (cf. OInd. *váta-/váata-*). The formation is based on the part. prs. of PIE **h₂weh₁-* ‘blow’. Owing to the defective orthography it is difficult to prove this change in the Sabellic languages. A possible outcome of shortening of a long diphthong before consonant is the d./l./abl.pl. ending of the *o*-stems, O. -*ois -úís* /-oys/ < **-ōys*, attested more than 20 times (as the direct reflex of **-ōys* we would expect O. *-uis*). The once-attested O. **veruís** ‘gates’ is certainly an error of the scribe). Counter-examples are O. **stahínt**, S(outh) P(icene). **praistaínt** 3.pl.prs. ‘they stand’ < **staent* (with *e* < *ē*), which would reflect PI **sta(y)ēnt*, but here *e* may have been introduced by analogy from the 3.sg., cf. O. **stait**, SP **praistaít**.

3.2. Rounding of **e* to *o* before *w*: Lat. *novus* ‘new’ < PIE **néwo-*, cf. Gk. νέος – Lat. *ūrō* ‘I burn’ (PIE **h₁ews-* ‘burn’, cf. Gk. εὔω ‘I singe’) – O. LOUFIR ‘free’, cf. 2.1b.

3.3. Shortening of long post-PIE vowels preceding a sequence of liquid + stressed vowel (“Dybo’s law”): Lat. *vir*, U. UEIRO /wērǎ/ acc.pl. ‘men’ < PI **wiro-* < PIE **wih-ró-*, cf. OInd. *vīrá-* ‘man, hero’.

3.4. Vocalization of a group consisting of word-initial laryngeal and syllabic resonant (cf. “Rix’s law” in Greek): Lat. *amb-*, U. *amb-* ‘around, about’ < PI **ambi* < PIE **h₂m̥b^hi-*, cf. Gk. ἀμφί, OInd. *abhí*, OIr. *imb-*, OHG *umbi* – Lat. *umbilicus* ‘navel’ < PI **ombel-* < **h₃m̥bhel-*, cf. Gk. ὀμφαλός, OIr. *imbliu*.

3.5. Vocalization of syllabic resonants preceding a sequence of laryngeal + consonant: *CR_hHC* > *CRāC*: Lat. (*g*)*nātus*, Pael. CNATOIS d.pl. ‘born; children’ < **ǵnh₁to-* (PIE **ǵenh₁-* ‘be born, give birth’).

3.6. Loss of laryngeals in all positions, after vowels with compensatory lengthening before consonants a). Laryngeals in syllabic position, *h₁*, change to *a* b):

- a) Lat. *fēstus* ‘festive’, O. **fíisnú** /fēsno/ n.sg. ‘sanctuary’ < **d^heh_{1s}-to-* / *-nah₂₋* (PIE **d^heh_{1s}-* ‘divine’, cf. Gk. θεός ‘god’ < **d^hh_{1s}-o-*)
- b) Lat. *datus*, O. **datas** g.sg.f. ‘given’ (PIE **deh₃₋* ‘give’, cf. Gk. δοτός < **d_h3-tó-*).

3.7. The syllabification of **y* after “heavy” syllables (“Sievers’ law”: e.g. *..CCy* > *..CCiy*) caused the split of the **ye/o*-present stems into two classes: IIIB Lat. *faciō facis*, U. **façiu** /fašom/ inf.prs., O. FACTUD inv.II ‘make’ < **fakye-/faki-* vs. IV Lat. *serviō servīs* ‘serve’, U. SERITU < **serwītōd* inv.II ‘observe, protect’ < **serwiye-/serwī-* < **serwye-*.

3.8. Loss of intervocalic *y: Lat. *trēs*, O. **trís** /trēs/ ‘three’ < PIE *tréyes, cf. OInd. *tráyas*.

3.9. Assibilation of dental geminates PIE *tʰt < *t-t, *d-t to ss: Lat. *sessus* ppp. ‘sit’ < *sed-to- (PIE *sed-) – O. *ῥεσσορει* d.sg. (epithet of Iuppiter) < *wert-tor- (PIE *wert- ‘turn’).

3.10. Voicing of *s to z in intervocalic position or adjacent to liquids (with further development z > r – “rhotacism” – in Latin, Fal(iscan) and Umbrian, cf. 7.1): Fal. **carefo** ~ Lat. *carēbō* 1.sg.fut. ‘I will lack’ < *kazēbō (PIE *kēs/kās ‘cut’, cf. Lat. *castus* ‘chaste’) – g.pl. ending of ā-stems: Lat. *mēns-ārum*, O. EGM-AZUM ‘rerum’, U. PRACAT-ARUM ‘?’ < *āzom < *āsom. In non-rhotacizing dialects using the local alphabets there is no indication of the voiced quality of the sibilant (i.e., it is written <S>).

4. Phonological inventory of Proto-Italic

For the late PI period we can reconstruct the following phonemic inventory:

4.1. Vowels and diphthongs:

short vowels:	<i>i</i>	<i>e</i>	<i>a</i>	<i>o</i>	<i>u</i>
long vowels:	<i>ī</i>	<i>ē</i>	<i>ā</i>	<i>ō</i>	<i>ū</i>
nasalized vowel:		<i>ẽ</i>			
short diphthongs:	<i>ey</i>	<i>ay</i>	<i>oy</i>	<i>aw</i>	<i>ow</i>
long diphthongs:		<i>āy</i>	<i>ōy</i>		

4.2. Consonants:

voiceless stops:	<i>k</i>	<i>kʷ</i>	<i>p</i>	<i>t</i>
voiced stops:	<i>g</i>	<i>gʷ</i>	<i>b</i>	<i>d</i>
voiceless fricatives:	<i>x</i>	<i>xʷ</i>	<i>f</i>	
voiced fricatives:	<i>ɣ</i>	<i>ɣʷ</i>	<i>β</i>	<i>δ</i>
sibilant:	<i>s</i> [z]			
resonants:	<i>r</i>	<i>l</i>	<i>m</i>	<i>n</i>
semivowels:	<i>y</i>	<i>w</i>		

Notes:

- The reconstruction of a nasalized vowel *ẽ* is based on the fact that the outcomes of PIE **ŋ* and **ŋ̃* are different in Latin and Sabellic, cf. 6.5. The assumption of a vowel quality similar to French *en* [ɛ̃] best explains the two Sabellic reflexes *an* (in initial, i.e. stressed) syllables and *en* in internal or final (i.e. unstressed) syllables.
- The voiced fricatives **ɣ*, **ɣʷ*, **β* could be categorized as word-internal allophones of their word-initial voiceless counterparts (just as *z* can be classified as an allophone of *s*, at least in the PI period). But this is not possible in the case of **δ*, which lacks a voiceless parallel.

5. Common Italic developments (cf. 1)

The Italic sub-branches – Latino-Faliscan and Sabellic – differ in some specific innovations in the same way as the Italic branch is differentiated from other IE branches. In this section, however, we will treat features common to all Italic languages, but dating from the post-PI period.

5.1. The sound change *tl > kl* is common to all Italic languages: Lat. *pōculum* ‘bowl’ < **pōtlom* (PIE **peh₃-* ‘drink’) – O. **sakaraklúm** ‘sanctuary’ < **sakrā-klo-* < **-tlo-*. An argument against a PI date for this change is Etr. **putlumza** Ta 2.31 ‘little bowl’ (ca. 300 BCE; *-za* is the usual Etr. diminutive suffix), apparently borrowed from Italic **pōtlom*. It is improbable that the contact between speakers of an Italic language and Etruscans should have taken place already in PI times. In Venetic the original group is preserved, cf. **maxetlon** /magetlon/.

5.2. Fixation of the accent on the first syllable of a word (also in Etruscan). The chronology is unclear; to be sure, it precedes the syncope of word-internal short vowels (see next section). An argument for at least a Proto-Sabellic date is the fact that in Oscan as well as in Umbrian the outcome of PI **ē* (< PIE **η*) is different, depending on its position within a word (cf. 6.5).

5.3. Loss of word-internal short vowels in open syllables (vowel syncope). This tendency is characteristic of all central-Italic languages, including Etruscan: Lat. *pergō* ‘I continue’ < **per-regō*, O. ACTUD inv.II ‘treat’ < **agetōd*, Etr. **larke** (Sp 2.72; beginning of 3rd c. BCE) < **lareke** (Fe 3.2; 5th c. BCE). This syncope is preceded by the specific Umbrian palatalization of velars: U. **struhšla** /strūšlam/ acc.sg. ‘cake’ < **struwi-kelā-* (dimin. suffix **-kelo-*) attested again by the Etruscan tradition already in the beginning of the 5th c. BCE, cf. *luzies* ‘Lucii’ Cm 2.53 (1st half of 5th c. BCE).

5.4. The vowel syncope was responsible also for the origin of “secondary liquids/nasals” *r̥, l̥, ŋ* which later on were vocalized as *er, el, en /in*: Lat. *agellus* ‘small plot of land’ < **agerlo-* < **agrlō-* < **agrolō-*. Oscan and Umbrian show the same development *r̥ > er*, cf. O. **agerllúd** abl.sg.

5.5. Vowel syncope in final syllables: O. **húrz** n.sg. ‘sacred grove’ ~ Lat. *hortus* ‘garden’ < PI. **xortos*, U. **setums** (570 BCE) < **sehtumos* < **seftemos* (proper name ‘Septimus’). In Latin syncope occurs only in a few cases after *t* and *r* (regularly in the nom.sg. of *i*-stems, if the preceding syllable is long, cf. *pars* < **part-i-s* ‘part’ vs. *sitis* ‘thirst’), but its scarce attestation may be the effect of analogical restoration in the majority of cases. The post-PI date of this sound change is proven not so much by SAKROS ESED ‘sacer erit’ from the Forum Cippus – theoretically the ending *-os* in *sakros* could be reintroduced by analogy (cf. *ferus*). A better argument is the different outcome of original **-ns > f* (e.g. in the ending of the acc.pl., cf. U. TORUF ‘bulls’ < **tawrōns*) and secondary *-ns*, originating from syncope (O. BANTINS ‘citizen of Bantiae’ < **bantinos*). Indeed, we can ascribe the vowel syncope in final syllables to the Proto-Sabellic period (6.). By “samprasāraṇa” a final group **Cyos* developed to *Cis*, cf. O. **lúvkis** < **lowkyos* ‘Lucius’. Inferring from the Oscan orthography, we can assume a closed pronunciation of *i* – and,

by the way, the origin of a “new” phonemic *i*, the PI vowel **i* having been lowered to *i* / *e* (cf. 6.2).

6. Proto-Sabellic innovations

6.1. Certainly the most typical feature of the Sabellic languages is the labialization of labiovelars: PIE, PI **k^w* > *p*; PIE, PI **g^w* > *b*; PIE **g^{wh}*, PI **x^w* / **g^w* > *f* / *b* (cf. 2.1): O. **pís** ‘who?’ ~ Lat. *quis* (PIE **k^wi-s*) – O. **bivus** n.pl. ~ Lat. *vīvus* ‘alive’ (PIE **g^wih₃wo-*) – U. **vufru** /wofrom/ ‘votive’ < **wog^{wh}ro-* (see above 2.1b).

6.2. According to a general tendency common to all Sabellic languages, short vowels tend to be articulated openly, whereas long vowels have a closed pronunciation. Thus, the PI correlation of vowel quality (cf. 4.1) was abolished. Because the “vowel letters” of the traditional alphabets were not sufficient to render all nuances of quality, new letters were invented in order to convey them, e.g. **í** (besides **e i**) and **ú** (besides **o u**) in South Picene, **í** in Oscan (besides **e i**). Another solution was the use of digraphs, e.g. **ei** in some sections of the text of the Umbrian bronze tables of Iguvium, or **ei** in Oscan (written in the Greek alphabet). Moreover, frequent alternations in spelling can indicate a special quality of the vocalic phoneme in question. – For the later periods of Oscan and Umbrian we can assume the restoration of a vowel correlation, which differentiated one more quality degree than the PI system.

- *i* > **í** O. <**í**>, cf. O. **pís** < **k^wis* (5.1)
- *ī* > **ī**, O. <**ii**, **i**>, cf. O. **liimitú[m]** g.pl. ‘border lines’ (Lat. *līmes*)
- *e* > **e**, O. <**e**>, cf. O. **meſiaí** l.sg.f. (2.1b)
- *ē* > **ē**, O. <**í**, **íí**>, cf. O. **fiísnú** (3.6)
- *a* > **a**, O. <**a**>, cf. O. **fakiiad**, U. **fačia** (2.1a)
- *ā* > **ā**, O. <**a**, **aa**>, cf. O. **saahútum**, U. SAHATAM (2.3)
- *o* > **o**, O. <**ú**>, cf. O. **púd**, POD ‘what’ < **k^wod*, U. PORS-E /poře/< **k^wod-i*
- *ō* > **ō**, SP <**ú**> / <**o**>, cf. SP **dúnoh** d.sg. (?) ‘gift’ < **dōnōy* ~ Lat. *dōnō*. In Oscan and Umbrian **ō** changed to **ū**, cf. O. **dunúm** /dūnom/, **duunated** ‘dōnāvit’, U. PIHATU inv.II. ‘expiate’ /p̄eatu/ < **pī-ā-tōd* (from **pīyo-*, cf. 2.4).
- *u* > **u**, O. <**u**>, cf. O. **puklui** d.sg. ‘son’, SP. **puqloh** < PIE **putlo-*, cf. OInd. *putrá-* ‘son’
- *ū* > **ū**, O. <**u**>, cf. O. **fruktatiuf** ‘usufruct’, cf. Lat. *frūgēs* ‘fruit’

6.3. Final PIE *-ā* changes to *-ā* (> *-o*?). In O. the vowel is rendered by <**ú/o**>, in Neo-Umbrian (Lat. alphabet) by <**o**>, in more archaic sections of the Iguvine Tables by <**a**> or <**u**>, cf. O. TOUTO n.sg. ‘people’, SP **toúta** (cf. Got. *þiuda*) – O. MOLTO, U. **muta**, **mutu** n.sg. ‘penalty’ ~ Lat. *multa*.

6.4. Spirantization of stops preceding *t*: PI **kt* > *χt*, PI **pt* > *ft*: In Umbrian *ft* changes to *xt*, finally *x* is lost (with compensatory lengthening): O. **ehtrad** ‘outwards’ < **ektrād*, cf. Lat. *extrā* – O. SCRIPTAS n.pl.f., U. SCREIHTOR /skrēhtâr/ n.pl.n. ‘written’ < **skripto-* (**skreyb^h-* ‘write’) – U. **speturie** /spēturyay/ d.sg.f. ‘spectoriae’ / ‘augural’ < **spextōriyo-* < **spekī*. In some Umbrian dialects the loss of *x* occurs very early, cf. **setums** ‘Septimus’ (proper name; 570 BCE) < **sehtemos* < **septemos*, cf. 5.4.

6.5. Development of \tilde{e} in initial syllables to *an*, in internal or final syllables to *en*: O. **fangvam**, FANCUA acc.sg. ‘tongue’ < * $f\tilde{e}gw\tilde{a}$ - < PI * $d\eta\eta w\tilde{a}$ - < * $d\eta\tilde{g}^h wah_2$ -, the proto-form also for Lat. *lingua* < OLat. *dingua* < * $dengw\tilde{a}$ -, Goth. *tuggo* /*tungō*/ – O. **anter**, U. ANDER ‘between’, Lat. *inter*, OHG *untar* < * η -*ter* – U. **iveka** IUENGA /*iwengaf*/ acc.pl. ‘heifers’ ~ Lat. *iuvenca* < PIE * $h_2yu-h\eta-k\tilde{o}$ -, cf. OInd. *yuvaśá*-.

7. Developments common to Latino-Faliscan and Umbrian

Latin (Latino-Faliscan) and Umbrian share some features which are alien to Oscan or South Picene. The most prominent is the rhotacism of intervocalic *s* (i.e. /z/, cf. 3.10), others are the weakening (and hence orthographic neglect) of final *-m* and *-s*.

7.1. Rhotacism: For examples see 3.10. The Latin rhotacism traditionally is dated to the 4th c. BCE. As a pre-rhotacized form of Umbrian we may cite **heruses** (Um 3, end of the 7th c. BCE) – or **heduses**? However the meaning and morphological analysis of this word are not clear.

7.2. Clusters with nasals and voiced fricatives: After a nasal a voiced fricative develops to a voiced stop in Latin and Umbrian, but not in Oscan (see below):

- *nd^h*: Lat. *offendīcēs* ‘knots of the straps fastening a priest’s cap’ (PIE * $b^h end^h$ - ‘bind’) – U. PREUENDU imv.II ‘turn against’ < * $wend^h-e-t\tilde{o}d$ (PIE * $wend^h$ - ‘turn, wind’)
- *mb^h*: Lat., U. *amb*-, cf. 3.4.
- *ng^h*: Lat. *fingō* ‘I form’, U. **fiktu** /*fęngtōd*/ imv.II < * $d^h ing^h-e/o$ - (PIE * $d^h ey\tilde{g}^h$ - ‘form’) – U. CRINGATRO /*kręngatrom*/ acc.sg. ‘stole’ (?), * $kreng^h$ -, cf. OIcel. *hringr*, OE *hring*
- *ng^{wh}*: Lat. *ninguit* ‘snows’ < * $sning^{wh-e/o}$ - (PIE * $sneig^{wh-}$) – no U. example. The sound change probably occurred independently in Latin and Umbrian. It is preceded by the specific Umbrian assimilation of a voiced labial stop to a preceding nasal, cf. U. **umen** /*ommen*/ ‘ungent’ < * $omben$ < * ong^{wen} , cf. Lat. *unguen* < * h_3ong^{w-n} , and it proves also that in the post-PI (i.e. early Oscan and Umbrian) period the outcomes of the PI voiced fricatives δ (< * d^h) and β (< * b^h) were distinct from each other and that their merger in intervocalic position (cf. 2.1b) is only common Sabellic, not Proto-Sabellic. – In Oscan there is no trace of this sound change. O. **anafrīss** d.pl. ‘(deities of) rain’ < * $\eta\beta ri$ -, cf. Lat. *imber* (PIE * neb^h - ‘wet’), retains the fricative quality of the consonant. In **fangvam** < * $d\eta\tilde{g}^h wah_2$ - (6.5) the fricative feature of the velar has been transferred to the word initial-consonant, in **amnud** ‘because of’ if from * $amb-no$ - < * h_3mb^h-no - (cf. 3.4) *b* may be ousted by the adjacent nasals.

8. References

Buck, Carl Darling

1928 [1904] *A Grammar of Oscan and Umbrian*. Boston: Ginn.

Fogolari, Giulia and Aldo Luigi Prosdocimi (eds.)

1988 *I Veneti antichi. Lingua e cultura*. Padova: Studio Editoriale Programma.

- Leumann, Manu
1977 [1928] *Lateinische Laut- und Formenlehre*. Munich: Beck. [Göttingen: Vandenhoeck & Ruprecht.]
- Meiser, Gerhard
1986 *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut der Sprachwissenschaft der Universität.
- Meiser, Gerhard
1998 *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Rix, Helmut
2002 *Sabellische Texte. Die Texte des Oskischen, Umbrischen und Süd-pikenischen*. Heidelberg: Winter.
- Rix, Helmut and Gerhard Meiser
1991 *Etruskische Texte. Editio minor*, 2 vols. Tübingen: Narr.
- Schrijver, Peter
1991 *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam: Rodopi.
- Sommer, Ferdinand
1914 [1902] *Handbuch der lateinischen Laut- und Formenlehre*. 2nd and 3rd edn. Heidelberg: Winter.
- Stuart-Smith, Jane
2004 *Phonetics and Philology. Sound Change in Italic*. Oxford: Oxford University Press.
- Untermann, Jürgen
2000 *Wörterbuch des Oskisch-Umbrischen*. (Handbuch der italischen Dialekte 3). Heidelberg: Winter.
- Walde, Alois and J. B. Hofmann
1938–1956 *Lateinisches etymologisches Wörterbuch*, 2 vols. Heidelberg: Winter.

Gerhard Meiser, Halle a. Saale (Germany)

48. The morphology of Italic

- | | |
|-----------------|----------------------------|
| 0. Introduction | 5. Numerals |
| 1. Nouns | 6. Derivational morphology |
| 2. Adjectives | 7. Verbs |
| 3. Adverbs | 8. References |
| 4. Pronouns | |

O. Introduction

“Italic morphology” as construed here concerns the morphological systems visible in and (insofar as they can be reconstructed) ancestral to Latin (from its earliest stages through the Classical period) and Faliscan (the representatives of the “Latino-Faliscan”

branch of the family), as well as the Sabellic languages (Oscan, Umbrian, South Picene, Pre-Samnite, and the so-called “minor dialects” of central Italy, such as Paelignian, Marsian, Volscian, and others). Although Venetic, Sicel, and possibly Lusitanian may belong to an “Italic” family (in some sense), they are not treated here.

Inevitably, the bulk of the data cited below comes from Latin, given the sparse and often fragmentary attestation of the Sabellic languages and of Faliscan. Forms unidentified as to language are Latin, sometimes specified as to chronology, for which this chapter follows a version of the periodization in Weiss (2011: 23–24): “VOL” = “Very Old Latin” (7th/6th to 4th/3rd centuries BCE), “OLat.” = “Old Latin” (3rd/2nd centuries BCE), “Class.” = “Classical Latin” (1st century BCE to 3rd/4th centuries CE), plus occasional reference to Late Latin. Forms are generally cited without textual reference; many standard resources are available for locating Latin material (such as the *Thesaurus Linguae Latinae* and its online counterpart), and Sabellic data may be located via the indexes in Rix (2002) and Crawford (2011). Following standard notational conventions, Latin material in SMALL CAPS is inscriptional; **boldfaced** forms in Faliscan and Sabellic languages are drawn from texts that use the native (or “national”) alphabets of those languages, while Sabellic forms in *italics* come from texts that use a version of the Roman alphabet.

Coverage of Italic morphology may be found in the standard handbooks of Latin historical grammar, such as Ernout (1953) and the morphological portions of Leumann (1977), Sihler (1995), Meiser (1998), and Weiss (2011), along with the concise survey by Clackson (2011). For Sabellic, see Buck (1928: 113–194) and Clackson (2015); for Faliscan, see Bakkmum (2009: 117–176). (More specialized works are cited below under individual morphological categories. For developments in Latin subsequent to the Classical period, see Väänänen 1981; Weiss 2011: 503–535; and Adams 2013.)

From the perspective of Proto-Indo-European (PIE) morphology, Italic presents a mixed picture: nominal and pronominal morphology are relatively conservative (note, for example, the maintenance of seven distinct cases in the noun, including a robust locative in Sabellic, as against the more limited system of post-Mycenaean Greek), whereas the verbal system presents numerous innovations, with marked divergences separating Latino-Faliscan and Sabellic. Nevertheless, most features of PIE verbal morphology survive in some form. Similarly, derivational morphology (which can be treated only briefly here) preserves many archaic features typical of the older-attested IE languages (such as root nouns), alongside many innovative formations.

1. Nouns

Bibliography: in addition to the handbook material cited above, note Klingenschmitt (1992) and Gerschner (2002) for Latin, Tikkanen (2011) for Sabellic.

Italic nominal declension (comprising both nouns and adjectives) preserves most features of PIE, with some simplification. Nouns and adjectives are inflected for number, but singular and plural only. (See, however, 1.1.2 [“Nom. pl. masc.”] and 5.1.1 [“2”] for dual relic-forms.) The PIE array of eight cases (as in Sanskrit) has been reduced to seven (nominative, vocative, accusative, dative, ablative, genitive, and locative, this last with marginal usage in Latin), with the Italic “ablative” continuing both PIE ablative and

instrumental forms, as well as some PIE locatives. The various PIE stem-classes devolved into five descriptive nominal categories, traditionally (in Latin grammar) called “declensions”, organized as follows if considered historically:

- “1st declension”: PIE *eh₂*-stems (or “*ā*-stems”);
- “2nd declension”: PIE thematic formations (or “*o*-stems”);
- “3rd declension”: PIE *i*-stems and consonant stems;
- “4th declension”: PIE *u*-stems;
- “5th declension”: descriptively “*ē*-stems”, with controversial PIE status (see 1.5).

(For the distribution of grammatical gender within the declension classes, see the individual treatments below.) Reductions in ablaut variation in the original athematic classes have largely obscured the PIE accent-ablaut categories (“acrostatic”, “proterokinetic”, etc.), as these have been conceived in traditional accounts of PIE grammar, as well as patterns of so-called “internal derivation” (see e.g. Fortson 2010: 119–122).

It will be most convenient to begin with the 2nd declension.

1.1. The second declension

The Italic “2nd declension” continues PIE *o*-stems, including both plain *o*-stems and forms built with complex thematic suffixes (nouns and adjectives in **-ro-*, instrument nouns in **-tro-*, etc.; see 6.1 on nominal suffixal derivation). The “*o*” of the term “*o*-stems” refers to one of the two variants of the PIE stem vowel known as the “thematic vowel” (**-o-* ~ **-e-*), which is present in all case forms except gen. sg., preceding a desinence. This original structure is sometimes visible on the surface (e.g. voc. sg. *-e*, with thematic vowel **-e* plus “zero ending”, or OLat. nom. sg. masc. *-os* < **-o-s*), but is more often obscured by phonological (and sometimes analogical) developments (see 1.1.2 for some details). Nouns in this category are mostly masc. and neut., the only systematic exception (apart from episodic cases: e.g. *humus* fem. ‘earth’, an original root noun, cf. Gk. *χθών* ‘id.’ fem.) being tree and plant names, which are regularly feminine (e.g. *fāgus* fem. ‘beech’, cf. Gk. *φηγός* ‘id.’ fem.); see Weiss (2011: 226–227).

1.1.1. Sample Latin paradigm (*lupus* ‘wolf’, *Corinthus* ‘Corinth’ for loc. sg., *Carseolī* ‘Carseoli [mod. Carsoli]’ for loc. pl.)

	SINGULAR	PLURAL
NOM.	<i>lupus</i>	<i>lupī</i>
VOC.	<i>lupe</i>	<i>lupī</i>
ACC.	<i>lupum</i>	<i>lupōs</i>
DAT.	<i>lupō</i>	<i>lupīs</i>
ABL.	<i>lupō</i>	<i>lupīs</i>
GEN.	<i>lupī</i>	<i>lupōrum</i>
LOC.	<i>Corinthī</i>	<i>Carseolīs</i>

For neuters (e.g. *iugum* ‘yoke’): as above, except nom./voc./acc. sg. *iugum*, nom./voc./acc. pl. *iuga*. (This pattern of formal case identity is a regular feature of neuter inflection, inherited from PIE.)

1.2.2. Notes on the case endings

Nom. sg. masc.: OLat. *-os* (and *-os* after [w] in Class. Lat.: *servos* ‘slave’) < PIE **-o-s*. Nouns and adjectives in **-ros* undergo phonological developments leading to nom. sg. *-er* (*ager* ‘field’ < pre-Lat. **agros*; cf. 2.1 below on *sacer*, *līber*).

Nom./voc./acc. sg. neut.: a few neut. forms show nom./voc./acc. in *-us*, e.g. *pelagus* ‘sea’ (borrowed from Gk. πέλαγος neut.), *vīrus* ‘poison’ (with complex PIE background, probably involving a root noun or consonant stem).

Voc. sg.: < PIE **-e* (cf. Gk. ἄδελφε ‘O brother’), i.e. the bare thematic stem, with **-e* + “zero ending”.

Acc. sg.: OLat. *-om* (and *-om* after [w] in Class. Lat.: *servom* ‘slave’) < PIE **-o-m*.

Dat. sg.: VOL and Fal. /-ōi/ (VOL ΔΒΕΝΟΙ ‘bonō’; Fal. **tītoi** [personal name], perhaps with shortening to /-oi/) < PIE **-ōi* (contracted from ***-o-ei*) = Gk. -ω, Ved. *-āy-* (in dat. sg. *-āya*); monophthongized to /-ō/ by OLat. (Fal. **tīto** is probably an error for **tīto<i>**; Bakkum 2009: 126–127.)

Abl. sg.: OLat. *-ōd* (mainly inscriptional), cf. Ved. *-ād* (and similar material in Balto-Slavic); the PIE reconstruction is controversial, see Weiss (2011: 202).

Gen. sg.: < PIE **-iH*, replacing the thematic vowel (as in Celtic and elsewhere), perhaps related to the so-called *vrkī* suffix (‘belonging/appertaining to X’); VOL (and OFal.) also *-osio* (cf. Ved. *-asya*, Hom. *-oio* < PIE **-osio*); see Weiss (2011: 203–204, 222) and Bakkum (2009: 129–130) for further details and references. There is limited evidence for a Faliscan genitive spelled <-oi> (of unclear analogical origin, if the forms in question are really genitives); see Bakkum (2009: 131–132) for discussion.

Loc. sg.: OLat. *-ei* < PIE **-ei* (as in Sabellic, cf. below) or **-oi* (cf. Gk. οἶκοι ‘at home’), with regular monophthongization to *-ī*; used freely only in place names and a few locational nouns and fixed expressions (e.g. *domī duellique* ‘at home and in war’).

Nom. pl. masc.: < PIE **-oi* (nom. pl. of the pronominal declension, cf. 4), with regular monophthongization. The synchronic nom. pl. *-ī* in a few forms (including some neuters) may continue the PIE *o*-stem neut. dual **-oih₁*, e.g. *frēnī* ‘reins’ (sg. *frēnum*); see Weiss (2011: 195 n. 9). An alternate nom. pl. masc. in *-eis* (also spelled *-es*, *-is*) is well attested inscriptionally (mainly in OLat.) but virtually unattested in literary texts (apart from pronominal forms). It is functionally restricted (frequent onomastic use and in official titles, e.g. *magistreis* ‘public officials’), and its origin is obscure; see Wachter (1987: 253–254), Vine (1993: 215–239), Bakkum (1994), Adams (2003: 677–678), Dupraz (2004), Adams (2007: 417).

Nom./voc./acc. pl. neut.: *-ā* is an innovation (cf. Ved. *yugā* ‘yokes’, with ending **-eh₂*), though it is not clear whether this is a morphological development based on consonant stems (**-C-h₂* > **-C-ā*) or a phonological shortening (cf. 1.2.2 on nom. sg. fem. **-eh₂*).

Acc. pl.: < PIE **-o-ns* (perhaps originally ***-o-m-s*, cf. acc. sg.), although the details of the phonological development to *-ōs* are disputed.

Dat. pl. and abl. pl.: VOL -OIS, OLat. -EIS < PIE instr. pl. **-ōis*; forms in *-ibus* (cf. 1.2.2 on *-ābus*) are either analogical (e.g. *generibus* ‘sons-in-law’ [Accius] after *patribus* ‘fathers’; *diibus* ‘gods’ [Petronius] after *deābus* ‘goddesses’) or reflect post-Classical developments (thus *FILIBVS* et sim. on late inscriptions, Ernout 1953: 34).

Gen. pl.: *-ōrum* is based on the fem. gen. pl. pronominal ending (see 1.2.2 on *-ārum*), replacing **-ōm* (< **-oHom*), maintained (with regular shortening) as a variant for some forms (e.g. *deum* ‘of the gods’) and in some fixed expressions (see also 1.1.3 on Sabellic).

Loc. pl.: < PIE **-oisu* (Ved. *-eṣu*), with regular monophthongization; the loss of the final *-u* may be phonologically regular. For further details on the case endings, see Weiss (2011: 220–225).

1.1.3. Notes on Sabellic

The most salient morphological differences (apart from purely phonological developments, which often obscure fundamental relationships with Latin) are the following:

Gen. sg.: *-eis*, i.e. the original *i*-stem ending (O. **sakarakleís** ‘shrine’).

Loc. sg.: *-ei* remains productive (vs. the restricted behavior of Latin); also, *-ei* + postposition **-en* leads (by yod-loss and contraction in **-e_i-en*) to a new ending “*-ēn*”, especially in noun + adj. expressions (O. **húrtín Kerriiún** ‘in the grove of Ceres’, SPi. **ombriiēn akren** ‘in Umbrian territory’).

Nom. pl. masc.: PIE **-ōs* (< ***o-es*, with early contraction) is retained (O. **Núvlanús** ‘inhabitants of Nola’).

Nom./voc./acc. pl. neut.: unshortened **-ā* (< PIE **-eh₂*, cf. 1.1.2) (O. **prúftú** ‘[things] put forth’).

Gen. pl.: **-ōm* is maintained (with a secondary shortening; Weiss 1998) (U. **pihaklu** ‘purification rites’, SPi. **raeliom** ‘the Raelii’).

1.2. The first declension

The Italic “1st declension” continues PIE feminine formations (“*ā*-stems”) built with an invariable suffix **-eh₂(-)* (the view followed here) or (according to a different conception, as in Beekes 2011: 199–201) an ablauting suffix **-eh₂(-) ~ *-h₂(-)*. These forms provide both substantives and the feminine forms of *o*-stem adjectives. The rare Latin masculine nouns reflect personalizations of original feminine abstracts, mostly compounds (e.g. *agricola* ‘field work’ → ‘farmer’), or appear in borrowings (*nauta* ‘sailor’, cf. Gk. ναύτης) and some onomastic forms (in some cases of Etruscan origin).

1.2.1. Sample Latin paradigm (*fēmina* ‘woman’, *Rōma* ‘Rome’ for loc. sg.,
Athēnae ‘Athens’ for loc. pl.)

	SINGULAR	PLURAL
NOM.	<i>fēmina</i>	<i>fēminae</i>
VOC.	<i>fēmina</i>	<i>fēminae</i>
ACC.	<i>fēminam</i>	<i>fēminās</i>
DAT.	<i>fēminae</i>	<i>fēminīs</i>
ABL.	<i>fēminā</i>	<i>fēminīs</i>
GEN.	<i>fēminae</i>	<i>fēminārum</i>
LOC.	<i>Rōmae</i>	<i>Athēnīs</i>

1.2.2. Notes on the case endings

Nom. sg.: The expected outcome /-ā/ (< *-eh₂) survives in Sabellic (see 1.2.3). The Latin ending *-a* probably reflects the original vocative (also from *-eh₂, but with the laryngeal lost *in pausa* by “Kuiper’s Law”).

Acc. sg.: The expected outcome /-ām/ (by “Stang’s Law”) underwent a regular shortening to /-am/ (but cf. on Sabellic below).

Dat. sg.: PIE *-eh₂-eī > *-āī, OLat. -AI (well attested inscriptionally), later *-ae*. (See Adams 2007: 46–50, 78–88 and Weiss 2011: 233 n.5 on the inscriptionally-attested monophthongizations -A and -E, especially frequent outside Rome.)

Abl. sg.: < Ital. *-ād (well attested inscriptionally in OLat. as -AD), analogically formed after the *o*-stem abl. sg. (in PIE, the eh₂-stem abl. sg. was identical to gen. sg.); but see also 3.1 on abl. *viā/U. vea, uia* ‘along the way’ (perhaps an old instrumental).

Gen. sg.: PIE *-eh₂-es > /-ās/, retained in Sabellic (see below) and in a few relic forms in Latin (e.g. *familiās* in *pater familiās* ‘head of the household’), but replaced with *-āī* (at first disyllabic, as often in Ennius, Plautus, Lucretius) on the basis of 2nd declension gen. sg. *-ī*, eventually developing to *-ae*.

Loc. sg.: PIE *-eh₂-i > /-āī/ (perhaps generalizing the pre-consonantal sandhi outcome), OLat. -AI.

Nom. (and voc.) pl.: PIE *-eh₂-es > /-ās/, preserved in Sabellic (see below) but replaced in Latin with -AI, Class. *-ae* (modeled on *o*-stem /-oi/, see 1.1.2); the rare literary and inscriptional examples of Lat. *-ās* have been variously interpreted, but are probably not archaisms (see Weiss 2011: 235).

Dat.-Abl. pl.: Lat. *-īs* < *-ēs* < *-eīs* (e.g. AASTVTIEIS ‘cleverness’) < *-aiś < *-āīś, replacing expected *-ābus by analogy with *o*-stem *-ōīś; attested forms in *-ābus* are secondary creations to distinguish fem. forms from their masc. counterparts (thus *deābus* ‘goddesses’ vs. ambiguous *d[e]īś* ‘gods’ or ‘goddesses’, LIBERTABVS ‘freed women’ beside *libertīs* ‘freedmen’ or ‘freed women’), which later spread to some fem. nouns (e.g. FEMINABVS ‘women’) (Weiss 2011: 236).

Gen. pl.: *-ārum* < **-āsōm* (cf. Myc. *-a-o* /-āhōn/, Hom. *-ᾱων*) drawn from pronominal declension (replacing expected **-ōm*), and the model for *o*-stem *-ōrum* (1.1.2). For further details, see Weiss (2011: 232–237).

1.2.3. Notes on Sabellic

Nom. sg.: /-ā/ (< **-eh₂*) survives (vs. Latin), though with regular rounding in Oscan and Umbrian, usually spelled: O. *touto*, U. **tuta** ‘people’, SPi. **qora** ‘statue’.

Acc. sg.: O. **paam** (relative pronoun, cf. Lat. *quam*) may indicate that Sabellic retained /-ām/ without shortening.

Dat. sg.: Paelignian forms like *Minerva* ‘to Minerva’ probably reflect influence from comparable forms in dialectal Latin (see 1.2.2, “Dat. sg.”).

Gen. sg.: /-ās/ is retained: O. *eituas* ‘money’, U. **tutas** ‘people’.

Nom. pl.: /-ās/ is retained: O. **aasas** ‘altars’, U. **urtas** ‘standing up’.

Gen. pl.: unrhotacized in Oscan (*egmazum* ‘things’) and SPi. (**fitiasom** ‘deeds’), but with rhotacism in Umbrian (**hapinaru** ‘lambs’), as in Latin.

1.3. The third declension

The Italic “3rd declension” continues PIE root nouns (Untermann 1992) and athematic consonant stems (*s*-stems, *t*-stems, *n*-stems, etc.), including *i*-stems (**-i-* ~ **-ei-* stems, in PIE terms), but not *u*-stems (1.4). All three genders are represented, with restricted behavior in some classes following expected PIE patterns (e.g. heteroclitic *r/n*-stems are regularly neuter). Most root nouns and consonant stems are substantival (with exceptions, e.g. *āk*-stems and some other *k*-stems are adjectival), and both substantives and adjectives are represented among *i*-stems. (See 2.1 below on the *i*-stem inflection of adjectives generally.) Particularly important for the historical development is a complex interplay between consonant-stem endings and the *i*-stem suffix-plus-ending conglomerate (i.e. **-i-* ~ **-ei-* plus ending), whence ultimately a mixed inflection for the category as a whole. Thus, for example, Lat. nom. pl. animate *-ēs* (for both consonant stems and *i*-stems) reflects *i*-stem **-ei-es* (vs. consonant-stem **-ēs*, unattested in Latin but regular in Sabellic), while gen. sg. *-is* (for both categories) reflects consonant-stem **-es* (vs. *i*-stem **-ei-s*, preserved in Sabellic and used for both *i*-stems and consonant stems). A pivotal event for this history may have been the reduction (by haplology) of *i*-stem dat. sg. **-ei-ej* to **-ej*, which thus became identical to consonant-stem dat. sg. **-ej* (Klingenschmitt 1992: 105–107).

1.3.1. Sample Latin consonant-stem paradigm illustrating animate nouns (masc. *dux* ‘leader’, *Carthāgō* ‘Carthage’ for loc. sg., *Calēs* ‘Cales’ for loc. pl.) (see directly below for neuter nouns and *i*-stems)

	SINGULAR	PLURAL
NOM.	<i>dux</i>	<i>ducēs</i>
VOC.	<i>dux</i>	<i>ducēs</i>
ACC.	<i>ducem</i>	<i>ducēs</i>
DAT.	<i>ducī</i>	<i>ducibus</i>
ABL.	<i>duce</i>	<i>ducibus</i>
GEN.	<i>ducis</i>	<i>ducum</i>
LOC.	<i>Carthāgine, -ī</i>	<i>Calibus</i>

Neuters: As usual, neuters of all categories have the same form for nom./voc./acc., e.g. *s*-stem sg. *genus* ‘origin, race’, pl. *genera*; *r/n*-stem sg. *femur* ‘thigh’, pl. *femina*. (Neut. nom./voc./acc. pl. *-a* is regular for all classes, but the nom./voc./acc. sg. form varies by stem-class.)

i-stem nouns: Regularly with gen. pl. *-ium* (vs. consonant-stem *-um*), neut. nom./voc./acc. pl. with *-ia*; in other parts of the paradigm where distinctive *i*-stem endings are found (details below), they generally alternate with consonant-stem forms – thus (for *turris* ‘tower’) acc. sg. *turrim* ~ *turrem*, abl. sg. *turrī* ~ *turre*, acc. pl. *turrīs* ~ *turrēs* (but see 2.1 on adjectival declension).

1.3.2. Notes on the case endings

Nom. sg. (masc., fem.): PIE animate **-s* (in consonant stems) is maintained in some phonological contexts (cf. *dux* = /duk-s/, 1.3.1), but has been lost in others, in some cases as early as PIE itself (thus *pater* ‘father’ < PIE **ph₂tér* < pre-PIE ***ph₂-tér-s*); there is thus no uniform nom. sg. animate desinence visible on the surface for consonant stems. In some categories, however, a nom. sg. marker /-s/ was added secondarily to an asigmatic PIE nom. sg.: thus some 3rd declension nouns in nom. sg. *-ēs* may derive from PIE hysterokinetic forms in nom. sg. **-éj* or **-én* (> pre-Lat. **-ē*), e.g. *verrēs* ‘boar’ (original *n*-stem; cf. *fidēs*, 1.5). For *i*-stems: **-i-s* > *-is* (Lat. *turris*, 1.3.1), except when suppressed phonologically, e.g. **-V:tis* > nom. sg. *-V:s* (*Larīnās* ‘inhabitant of Larinum’).

Nom./voc./acc. sg. (neut.): Prominent categories among consonant-stem substantives include *s*-stems, *men*-stems, and *r/n*-stems; *i*-stems are also represented (e.g. *mare* ‘sea’ < **mór-i-Ø*, with adjusted vocalism [Vine 2011: 264–265]).

Acc. sg.: *i*-stem **-i-m* maintained to some extent in OLat. (and regularly in adverbs in *-tim* < **-ti-m*), but largely replaced by *-em* from consonant stems (i.e. **C-ṃ* > Lat. /C-em/: **ped-ṃ* > *pedem* ‘foot’).

Dat. sg.: OLat. *-ei* (< PIE **-eī*) > Class. *-ī* (consonant stems); see above (1.3, introduction) on *i*-stem **-eī-eī* > **-eī* by haplology, whence also *-ī*.

Abl. sg.: Class. Lat. *-e* (consonant stems) < PIE loc. sg. **-i* (the few OLat. forms of the type [c]OSOLED ‘consul’, LEGED ‘law’ are hyperarchaizing artificial creations, see Weiss 2011: 238 n. 1, with reference); *i*-stem *-ī* (best attested for neuter nouns and regular for adjectives, see further 2.1) < *-īd*, a Proto-Italic innovation (analogical to *o*-stem abl.

sg. *-ōd*), e.g. OLat. *LOVCARID* ‘grove’, Class. *ignī* ‘fire’ (generally replaced by consonant-stem *-e*, but sometimes infiltrating consonant stems: e.g. OLat. *BOVID* ‘ox’, Class. *bove*).

Gen. sg.: Class. Lat. *-is* (both consonant stems and *i*-stems) < PIE (consonant-stem) **-es* (OLat. *-ES*), but OLat. also *-os* and (with regular raising) *-vs* (< PIE **-os*) (*DIIVOS* ‘Iovis’, *HONORVS* ‘honōris’, etc.); Faliscan, however, may have maintained *-os* (consonant stems) and *-is* (*i*-stems, spelled <*-e*>) distinct (see Bakkum 2009: 143–145).

Loc. sg.: *-e* = consonant-stem abl. sg., *-ī* is analogical to *o*-stem loc. sg.; see above (“Abl. sg.”) for the PIE (consonant-stem) loc. sg. (the PIE *i*-stem loc. sg. does not survive).

Nom./voc. pl. (animate): *i*-stem **-ej-es* > Lat. *-ēs*, extended to consonant stems; an *i*-stem variant *-īs* (well-attested in MSS of Plautus, inscriptionally, and in the grammatical tradition, and also appearing in some consonant stems) may be analogical to the consonant-stem pattern nom. pl. *-ēs* ~ acc. pl. *-ēs* (cf. *i*-stem acc. pl. *-īs*); differently Nyman (1990) and Vine (2012: 565–567), with references (cf. also 1.3.3 on Sabellic).

Acc. pl.: consonant-stem **-ns*, thus **C-ns* > pre-Lat. **C-ens* > Lat. *C-ēs*; *i*-stem **-i-ns* > Lat. *-īs*, but undergoing replacement by *-ēs* during Class. Lat.

Dat.-Abl. pl.: *i*-stem **-i-b^hos* > Lat. *-ibus*, extended to consonant stems (thus *rēgibus* ‘kings’ [root noun], not †*rēgbus*).

Gen. pl.: consonant-stem **-oHom* > pre-Lat. **-ōm* > OLat. *-OM* (*POIMILIONOM* ‘dwarves’), Class. *-um*; *i*-stem *-ium*, but with variation (e.g. *mēnsium/mēsum* ‘months’) (cf. also 2.1 on adjectives/participles).

Loc. pl.: PIE forms do not survive (Latin uses the dat.-abl. pl.). For further details, see Weiss (2011: 198, 243–246).

1.3.3. Notes on Sabellic

In general, the consonant-stem vs. *i*-stem distinctions are more faithfully preserved in Sabellic than in Latin. Some of the more salient morphological differences are as follows:

Gen. sg.: *i*-stem **-ej-s* is extended to consonant stems (O. **medikeís** ‘public official’).

Acc. sg.: *i*-stem **-im* is regular in Sabellic (though recessive in Latin), but *o*-stem **-o-m* is used for consonant stems (O. *tanginom* ‘decision’, *n*-stem); problematic is SPi. **dikdeintem** (vs. **áudaqum**), cf. Fortson (2016: 23).

Abl. sg.: *o*-stem **-ōd* is used for consonant stems in Oscan (O. **tanginúd** ‘decision’), vs. loc. sg. **-i* (as in Latin) in Umbrian (**kapiře** ‘[sacrificial] bowl’); *i*-stem *-īd* like Latin (SPi. **aríth** ‘with skill’).

Loc. sg.: U. *ocre* ‘mountain’ (*i*-stem) with **-ej* or **-ēj* (no Oscan example); **-i* in U. consonant stems (e.g. **ferine** ‘platter (?)’); O. *κενσορατη* ‘office of censor’ with **-ej* from *o*- or *i*-stems).

Nom. pl. (animate): PIE **-es* is preserved in consonant stems (*n*-stem O. **humuns** ‘people’, with regular syncope < **...n-es*); *i*-stem **-ej-es* (O. **trís** ‘three’, U. *pacrer*/Pael., Marruc. *pacris* ‘propitious’); a few forms with *-īs*/ may also be attested, as in Latin (see the references cited in 1.3.2 “Nom./voc. pl.”).

Dat./Abl. pl.: Oscan consonant stems and *i*-stems, as well as Umbrian *i*-stems, use the *i*-stem form (as in Latin); but Umbrian consonant stems use a form based on *u*-stems, e.g. *fratrus* ‘brethren’ (-*us* < *-*u-b^hos*, see 1.4.3).

1.4. The fourth declension

The Italic “4th declension” continues PIE *u*-stems, both plain *u*-stems (with all three genders represented in Italic) and complex formations, among which *tu*-stem verbal nouns (masc. only; see also 7.3.1.5 [“supines”]) are best attested.

1.4.1. Sample Latin paradigms (animate: fem. *tribus* ‘division of the people’; neut. *cornū* ‘horn’)

	SINGULAR	
	ANIMATE	NEUTER
NOM.	<i>tribus</i>	<i>cornū</i>
VOC.	<i>tribus</i>	<i>cornū</i>
ACC.	<i>tribum</i>	<i>cornū</i>
DAT.	<i>tribuī, tribū</i>	<i>cornū</i>
ABL.	<i>tribū</i>	<i>cornū</i>
GEN.	<i>tribūs</i>	<i>cornūs</i>

	PLURAL	
	ANIMATE	NEUTER
NOM.	<i>tribūs</i>	<i>cornua</i>
VOC.	<i>tribūs</i>	<i>cornua</i>
ACC.	<i>tribūs</i>	<i>cornua</i>
DAT.	<i>tribibus</i>	<i>cornibus</i>
ABL.	<i>tribibus</i>	<i>cornibus</i>
GEN.	<i>tribuum</i>	<i>cornuum</i>

1.4.2. Notes on the case endings

Nom. sg.: neut. -*ū* (for expected -*ǔ*, cf. Hitt. *tāru*, Ved. *dāru*, Gk. δόρυ ‘wood’) is secondary, perhaps based on the old neut. pl. (or collective) in *-*uh₂* > -*ū* (see Nom./voc./acc. pl. neut. below).

Dat. sg.: animate $-ū$ < $*-eu-ej$; the variant in $-ū$ (anim. and neut.) is analogical to i -stems; Fal. **mercui** perhaps with analogical $/-uḡ/$ or $/-ūḡ/$ after \bar{a} -stem $/-āḡ/$ and o -stem $/-ōḡ/$ (Bakkum 2009: 146–147).

Abl. sg.: < $*-ūd$ (OLat. CASTVD ‘abstinence’), analogical to 2nd declension abl. sg. (although some plain $-ū$ forms may derive from instr. sg. $*-u-h_1$, cf. 6.1 on *astū*).

Gen. sg.: $-ūs$ < $*-eu-s$; alternate OLat. $/-u(ṽ)os/$, $/-u(ṽ)is/$ (SENATVOS [cf. Fal. **zenatvo**] ‘senate’, *fructuis* ‘fruit’) either from old ablaut variant $*-u-os/es$ or analogical after 3rd declension; regular in Plautus and frequent in OLat. is yet another ending $-ī$ (e.g. SENATI), by analogy with 2nd declension masculines.

Loc. sg.: only *diū* ‘by day’ < $*dijēu$ (whence analogically *noctū* ‘at night’); the lengthened-grade desinence $*-ēu$ (Ved. *śātrau* ‘enemy’) is not preserved.

Nom. pl.: animate $-ūs$ < $*-eu-es$.

Acc. pl.: animate $-ūs$ < $*-u-ns$.

Nom./voc./acc. pl. neut.: $-ua$ analogically (cf. 3rd declension: 1.3.1 *genera, femina*, cf. 1.1.2) for expected $-ū$ < $*-u-h_2$.

Dat.-abl. pl.: $-ibus$ < $*-u-b^hos$ (with regular vowel-weakening); Class. Lat. forms in $-ubus$ (e.g. *artubus* ‘joints’) are artificial (cf. OLat. TREBIBOS ‘tribibus’ [1.4.1]). For further details, see Weiss (2011: 250–253).

1.4.3. Notes on Sabellic

In general, u -stems are poorly attested in Sabellic.

Acc. sg.: U. **trifu** = Lat. *tribum*; O. *manim* is an artificial transfer to i -stem inflection, analogical to abl. sg. $*manid$ (with regular treatment $*-ūd > /-īd/$, cf. O. *castrid* ‘?’, U. **mani** ‘hand’).

Dat. sg.: U. *trifo* with the alternate ending as in Lat. $-ū$.

Loc. sg.: U. **manuv-e** ‘hand’ (with postposition $-e = /-en/$), *maronato* (term for a public office) < $*-ou < *-eu$; U. *maronatei* shows a transfer to o -stem inflection (cf. OLat. gen. sg. SENATI above).

Abl. pl.: U. **berus** ‘spits (for roasting)’, SPi. **manus** ‘hands’, regularly < $*-u-b^hos$ (> $*-ufos > *-ufs > *-uss > /-us/$).

1.5. The fifth declension

This category of descriptive “ \bar{e} -stems” is beset with difficulties, beginning from uncertainty as to whether there was an Italic 5th declension at all, rather than a purely Latin one. Apart from U. **ri/re** (cf. Lat. *rēs* ‘thing’), hardly any Sabellic forms can be identified with certainty as \bar{e} -stems of the Latin type (see Tikkanen 2011: 42–43); and it is even possible that the Umbrian word is a borrowing from Latin (see Untermann 2000: 635, s. v.). Not surprisingly, therefore, the Indo-European background of the category is also in question. It is generally agreed that the two core forms (Lat. *rēs* and *diēs* ‘day’; for the latter, cf. also Fal. **foied** ‘today’ < $*g^h\bar{o}-dīēd$) arose secondarily, through regular phonological developments and subsequent analogies; thus for *diēs*: PIE acc. sg. $*dīeu-m$ (more specifically the “Lindeman-variant” $*dīēu-m$) > $*dīēm$ (by Stang’s Law; later > Class. Lat. *diem*, with regular shortening), whence analogical nom. sg. $*dīēs$ (= Lat.

diēs and secondarily the rest of its 5th declension paradigm, apart from relic forms from the original paradigm, such as voc. sg. **djēu* > *Iū-* in *Iūpiter/Iuppiter*). (Similarly, on the phonological development of *rēs* < PIE **reh₁i-*, see Weiss 2011: 254.) Other 5th declension forms may have individual explanations: e.g. *fidēs* ‘faith’ may continue a hysterokinetic *i*-stem (as if PIE **b^hid^h-ēj* [+s]) related (by “internal derivation”) to the amphikinetic *i*-stem in Gk. *πειθώ* ‘persuasion’ (Hamp 1999), or may be influenced by a synonymous **crēdēs*, cf. *crēdō* ‘trust’ and Ved. *śrād + dād^hāti* (Ernout-Meillet 1985 s. v. *fidēs*). Other prominent forms pose complex etymological and/or morphological problems: see e.g. Klingenschmitt (1992: 127) on *plēbēs* ‘common people’; for *famēs* ‘hunger’, with disputed etymology, see Vine (2013); on *spēs* ‘hope’, see Nussbaum (2011). Thus, according to one view, “there was no *ē*-declension in PIE” (Fortson 2010: 209, similarly Weiss 2011: 253); but others (e.g. Beekes 2011: 199), broadly following Pedersen (1926), argue for PIE “*-(*e*)*h₁-* stems”, partly on the basis of alleged evidence from Baltic. A further 5th declension problem concerns nouns (mostly abstracts) built with a formant *-iēs* (e.g. *aciēs* ‘sharp edge’, *maciēs* ‘leanness’), including some that show an alternation with *-ia* (e.g. *luxuriēs/luxuria* ‘luxury’). The details are complex, but probably involve developments associated with the so-called “*devī*” and “*vṛkī-*” formations of PIE (Piwowarczyk ms.).

Latin 5th declension nouns are feminine except for *diēs* (and its compound *merīdiēs* ‘mid-day’), which is regularly masculine except in the meaning ‘appointed day’.

1.5.1. Sample Lat. paradigm (*diēs* ‘day’)

	SINGULAR	PLURAL
NOM.	<i>diēs</i>	<i>diēs</i>
VOC.	<i>diēs</i>	<i>diēs</i>
ACC.	<i>diem</i>	<i>diēs</i>
DAT.	<i>diē, diēi, diei</i>	<i>diēbus</i>
ABL.	<i>diē</i>	<i>diēbus</i>
GEN.	<i>diēi, diei, diei</i>	<i>diērum</i>

1.5.2. In general, given the (likely) innovatory status of this category, the case forms are all analogical (mainly based on 1st declension models), and show complex variation in the historical record. (Thus for ‘day’: also e.g. OLat. gen. sg. *diēs*, cf. OLat. 1st declension gen. sg. *-ās* [1.2.2].) For many details see Weiss (2011: 253–255), and for the OLat. inflection, see especially Gerschner (2002: 151–163), Ernout (1953: 70), Leumann (1977: 445–446). For many details on the forms belonging to ‘day’ and ‘Jupiter’ in both Latin and Sabellic, see Rix (2004).

1.5.3. Notes on Sabellic

Apart from U. dat. **ri**, abl. **ri/re** (cf. 1.5), the only clear \bar{e} -stem form is O. dat. **kerrí**, **keri** ‘to Ceres’ (originally an *s*-stem). All other alleged “5th declension” forms are problematic.

2. Adjectives

2.1. Basic patterns

Adjectives are inflected according to one of two types, i.e. “1st and 2nd declension adjectives” or “3rd declension adjectives”. In the first type, the feminine is supplied by 1st declension forms (corresponding to 1st declension nominal inflection), and the masculine and neuter are supplied by 2nd declension forms (corresponding to 2nd declension masculine and neuter nominal inflection). 3rd declension adjectives are inflected like 3rd declension *i*-stems. (There are no “4th declension adjectives” or “5th declension adjectives”.) Perfect passive, future active, and future passive participles are inflected as 1st/2nd declension adjectives; present active participles are inflected as 3rd declension adjectives (with a complication in abl. sg. forms, explained below). (The formations of the participles are treated in 7.3.1.)

In dictionaries and grammars, adjectives are conventionally cited in the order “[nom. sg.] masc., fem., neut.,” e.g. *laetus, laeta, laetum* ‘happy’. In the 1st and 2nd declension type, 2nd declension forms exhibit the same formal variation as in nouns, i.e. nom. sg. *-us/-um* for most forms, but also nom. sg. masc. *-er* for many forms (whence acc. sg. and neut. nom./acc. *-rum* or *-erum*): thus *sacer* (still VOL SAKROS), *sacra, sacrum* ‘sacred’ (like *ager*, gen. *agrī* ‘field’) or *liber, libera, liberum* ‘free’ (like *vesper*, gen. *vesperī* ‘evening’). In the majority of 3rd declension adjectives, masc. and fem. forms are identical, with nom. sg. in *-is*, while the neut. nom. sg. ends in *-e* (so-called “two-ending” 3rd declension adjectives, e.g. masc./fem. *facilis*, neut. *facile* ‘easy’). A number of other 3rd declension adjectives (mostly ending in *-x* or *-ns*, the latter including present active participles) have only a single form for nom. sg. masc., fem., and neut. (e.g. *audāx* ‘bold’, *audēns* ‘daring’). A few “three-ending” 3rd declension adjectives (all with suffixal *-ri-*) have distinct nom. sg. forms for each gender: masc. *ācer*, fem. *ācris*, neut. *ācre* ‘sharp’. The OLat. situation, however, is more complex; further details in Leumann (1977: 432–433). A unique relic of a feminine adj. in (original) **-ih₂* (the so-called “deví suffix”) may survive in *Laurentis* (Ennius), if this reflects *Laurentis* (Nussbaum 1973). On adjectival declension in Sabellic (largely the same as in Latin, insofar as this can be determined), see Buck (1928: 133–134).

There is, finally, considerable variation in 3rd declension forms with regard to consonant stem vs. *i*-stem inflection, comparable formally to the same variation in 3rd declension nouns (abl. sg. *-e* vs. *-ī*, gen. pl. *-um* vs. *-ium*, etc.; cf. 1.3), but with different distribution: e.g. *nt*-participles use abl. sg. *-e* when used as nouns or in abl. absolute constructions, vs. *-ī* in attributive usage; most adjectives favor (masc./fem.) acc. pl. *-īs*, but comparative adjectives (see next section) favor *-ēs*; etc.

2.2. Comparison

2.2.1. The comparative

In PIE (see Rau 2014), the primary comparative suffix **-ios-* (zero-grade **-is-*) – probably in origin an elative marker (‘quite X, rather X’) – was added directly to the root, as in Lat. *magnus* ‘big’ (root **mag-* plus suffix **-no-*), but comparative (masc., fem.) *maior* (< **mag-ios-*), (original neut.) *magis* adv. ‘more’ (< **mag-is*). But this pattern was largely given up: in the productive comparative formation, the suffix is added to the stem of the positive form of the adjective, whether or not this is characterized by suffixal material; thus *facilis* ‘easy’ (< pre-Lat. **fak-li-*), compar. *facilior* ‘easier’ (not †*facior*). PIE **-ios-* inflected as an amphikinetic *s*-stem (animate nom. sg. **-iōs*, acc. sg. **-ios-ŋ*, gen. sg. **-is-és*, loc. sg. **-iés ± i*, etc.), but the nom. sg. form of the suffix has been generalized throughout the paradigm (thus gen. sg. *faciliōris* < **-iōs-es*, with rhotacism; cf. the pre-rhotacism form *meliosem* ‘better’ [acc. sg.], preserved in the Roman grammatical tradition). The neut. nom./acc. sg. has the form **-ios* (*facilius*), beside relic forms with zero grade **-is* (cf. the adv. *magis* above, vs. the regular neut. comparative adj. form *maius*; in Sabellic, e.g. O. *mais* ‘more’ if from **meh₂-is*, although the form is ambiguous). Traces of *e*-grade **-ies-* may survive in a few forms, such as *mulier* ‘woman’ (< **m_l-ies-*, generally assumed to belong originally with *melior* ‘better’; this connection is rendered uncertain, however, by VOL MVLIAR[] and the stem *maies-* of *maiestās* ‘greatness’. The neut. (nom./)acc. form of the adjective supplies the regular comparative adverb (*facilius* ‘easier [thing]’ and ‘more easily’), the type of comparative form best attested in Sabellic (O. *fortis* ‘fortius’, with regular final-syllable syncope). (See 3 for more on adverbial formations.)

The PIE suffix **-tero-*, although used to make regular comparative adjectives elsewhere (as partly in Greek and Sanskrit), was not in origin a comparative suffix as such, but had a contrastive or oppositional function, well-represented in Italic: Lat. *dexter*, U. *destrame* ‘right (as opposed to left)’ (cf. Gk. δεξιτερός). (Differently on **-tero-*: Baldi and Cuzzolin 2010; but see the critical remarks, with further references, by Martzloff 2013: 118–119.)

Synthetic comparison (i.e. phrasal expressions like *magis idōneus* ‘more useful’, *maximē egregius* ‘most outstanding’) is attested from the earliest Latin literature, and becomes widespread in later stages of the language (see Väänänen 1981: 118–119). Apart from a tendency for this pattern to appear, in the Classical language, with adjectives in *-eus* and *-ius*, this has more to do with syntax (and sociolinguistics) than morphology, and will be left aside here.

2.2.2. The superlative

The superlative suffix **-is-ŋmo-* (a significant isogloss with Celtic, see Cowgill 1970) consists of the zero-grade of the comparative suffix followed by a formant **-ŋmo-*, otherwise used to derive superlatives from adverbial bases (**sup-ŋmo-* ‘highest’ > **sup-omo-* > [with syncope and assimilation] Lat. *summus*, U. *sume*). This suffix is visible only indirectly (having been obscured by the application of regular phonological process-

es) in three classes of forms, namely (i) adjectives in *-er* (whether 1st/2nd declension or 3rd declension: *liberrimus* ‘most free’, *ācerrimus* ‘most sharp’), (ii) adjectives in *-lis* (*facillimus* ‘most easy’), and (iii) in some suppletive superlative forms (2.2.3). Other adjectives use a suffix *-issimus* (the synchronically regular superlative suffix), apparently a version of **-is-ḡmo-* with “affective” or “expressive” gemination of *-s-*. (See, however, Gunkel 2012 for a more nuanced approach to the gemination.) Oppositional adjectives in **-tero-* use a suffix **-tḡmo-* (*dextimus* ‘rightmost’), sometimes also found (instead of plain **-ḡmo-*) with adverbial bases, as probably in *optimus* ‘best’ (cf. Lat. *ob*, O. *úp* ‘at’; Weiss 2011: 361). (On the irregular syllabification in **-ḡmo-* and **-tḡmo-*, see Rau 2014: 331.)

The PIE superlative suffix **-isto-* (Ved. *-iṣṭha-*, Gk. *-ιστο-*) – again built with comparative **-is-* plus another formant – may survive in a few relic forms, such as Lat. *sōlistimus* ‘very favorable’ (secondarily remade with *-imus*, cf. OE *sēlest* ‘best’ < **sōl-isto-*; Dieu 2009).

2.2.3. Suppletive comparison

In a small number of frequent and semantically basic adjectives (‘good’, ‘bad’, ‘small’, ‘many’), the root form of the comparative and superlative differs from that of the positive form of the adjective (a widespread pattern in Indo-European, cf. Dieu 2011):

bonus ‘good’, but *melior* ‘better’ (cf. 2.2.1 on *mulier*) and *optimus* ‘best’ (2.2.2).

malus ‘bad’, but *peior* ‘worse’, *pessimus* ‘worst’ (based on a root **ped-*, perhaps the same as **ped-* ‘foot’ > ‘(at) bottom’, hence ‘worse, worst’).

parvus ‘small’, but *minor* ‘smaller’, *minimus* ‘smallest’ (cf. Gk. *μείων* ‘smaller’, OCS *mǫnjii* ‘id.’, among other comparanda).

multus ‘much, pl. many’, but *plūs/plūrēs* ‘more’ (neut. *plūs* mainly in partitive use, with genitive: *plūs pecūniae* ‘more [of] money’, etc.), *plūrimī* ‘very many, most’ (based on PIE **pleh₁-* ‘fill’; original comparative **pleh₁-ios-* perhaps continued in VOL PLEORIS/PLEORES [Carmen Arvale]. The inscriptional and grammatical records provide a great variety of forms, and the background of this material, as for the suppletive formations generally, is particularly complex; see Weiss 2011: 359–361 for details.)

3. Adverbs

3.1. Productive patterns

Each adjectival type (2.1) has its own productive adverbial formation: 1st/2nd declension adjectives make regular adverbs in *-ē*, OLat. *-ēd* (OLat. *FACILVMED* ‘very easily’, Lat.-Fal. *RECTED* = U. *rehte*, Lat. *improbē* ‘improperly’ = O. *amprufid*; the frequent forms *bene* ‘well’, *male* ‘badly’, with short *-ē*, result from the application of iambic shortening; cf. below on *modo* ‘only’), and 3rd declension adjectives make regular adverbs in *-iter* (*fortiter* ‘strongly’), which can surface as *-ter* as a result of syncope (*audāx* ‘bold’ → **audākiter* > *audācter* ‘boldly’) or as *-er* as a result of haplology (*diligēns*, stem *diligent-* ‘careful’ → **diligentiter* > *diligenter* ‘carefully’). In OLat. (and to a lesser extent in

Class. Lat.) usage, the *-iter* formation spread secondarily to 1st/2nd declension adjectives (*pūrus* ‘pure’ → *pūriter* [Cato, Pomponius, Catullus], beside *pūrē* [Cicero, Livy, etc.]); and the regular adverb of *alius* ‘other’ is *aliter* ‘otherwise’. A correspondent to the *-iter* formation is not attested in Sabellic; but cf. O. **akrid**, which may mean ‘ācritter’ and which has the form of an abl. sg. to the 3rd declension adjective corresponding to Lat. *ācer* ‘sharp’.

The backgrounds of both formations are to some extent controversial. Lat. *-ē(d)* is generally considered to reflect an *e*-grade variant of the *o*-stem abl. sg. ending *-ō(d)*, partly on the strength of attested *-ēd* forms with *-d*, and also because of a presumed parallel with adverbial forms in both *-ō(d)* itself (e.g. *MERITOD/meritō* ‘deservedly’, O. *contrud* ‘against’; the adv. *modo* ‘only, just now’, with *-ō*, results from iambic shortening, cf. above on *bene*, *male*) and *-ā(d)* (e.g. *eā* ‘that way’ in *praetereā* ‘besides’, cf. OLat. *ARVORSVM EAD* ‘against that’; O. [s]úllad ‘wholly’), which are also taken to continue old abl. sg. forms. But there is little evidence otherwise for a thematic abl. sg. ending *-ē(d)*, and an alternative explanation that assumes old instrumental forms in **-eh₁* > Ital. **-ē* (secondarily outfitted with “ablative *-d*”) may be preferable. (On the instrumental solution, which also accounts for a series of directional adverbs in *-ā* [**-eh₂-eh₁*] unlikely to have a source in ablatives [*hāc* ‘this way, over here’, etc.], see 4.2 below, with references; similarly perhaps *viā/U. vea*, *uia* ‘along the road’, Vine 2010: 136. See also 4.2 on the directional adverbs in *-ū* [*hūc* ‘to this place’, etc.]) The suffix *-iter* may be based on a reinterpretation of the nom. sg. masc. of the oppositional suffix **-tero-* (cf. *alter* ‘the other [of two]’ and *aliter* ‘otherwise’), or may involve an extension of the *-ter* seen in some “compound prepositions” of the type *praeter* ‘beside’ (cf. *prae* ‘in front’), which are in turn analogical to the inherited adverbial form (> Lat. preposition) seen in *inter* ‘between’ (cf. Ved. *antár*).

As noted above (2.2.1), the neut. acc. sg. of the comparative adjective functions as the comparative adverb; but the process is more general. The “adverbial accusative” (see in general Ernout-Thomas 1953: 28–29) is not fully productive, but is well attested in terms referring to quantity and extent, as well as in pronominal forms that often develop as conjunctions (e.g. *quod* ‘because’). The pattern is used for some common adjectives of both declensional types, e.g. *multum* ‘much’ (to *multus* ‘much, pl. many’, 1st/2nd declension adj.), *facile* ‘easily’ (to *facilis* ‘easy’, 3rd declension adj.); similarly in Sabellic: U. *promom* ‘first’ (adj. stem **promo-*), cf. (with a version of the same root but a different suffixal formation) Lat. *primum* ‘first’ (adj. *prīmus*). The adverbial suffix *-tim* originated in the acc. sg. of *ti*-stem nouns (e.g. *partim* ‘partly’, *statim* ‘immediately’), but then, on the basis of reinterpretations like *statim* ← *status* (participle to *stāre* ‘to stand’), came to be formed from perf. passive participles (e.g. *cursim* ‘on the double’ ← participle *cursus*, to *currō* ‘run’), whence forms in *-ātim* (*nōminātim* ‘by name’), which in turn led to the productive use of *-ātim* with nominal stems (*generātim* ‘by classes’), among other secondary developments. (See Leumann 1977: 501–502 for further details.) An acc. sg. origin may also underlie some adverbs in *-am* (e.g. *clam* ‘secretly’, *palam* ‘openly’). But the *-im* of adverbs like *interim* ‘meanwhile’, *ōlim* ‘at a distant time’ may go back to a metanalyzed suffixal *-im* abstracted from *i*-stem forms that preserve an archaic ablative-instrumental formant **-m*, cf. *hinc* ‘from here’ (< **g^hi-m-k[e]*), Gk. *πρίν* ‘before’ (< **pri-m*). (On adverbial *-[i]m*, see Dunkel 1997 and 2014: 1.137–148, developing an insight of Delbrück.)

3.2. More isolated patterns

The inherited adverbial suffix **-tos* (Dunkel 2014: 1.191–194), originally used with local adverbs (*intus* ‘from within’, cf. Gk. ἐντός), has an ablational sense. It spread secondarily to nouns (*funditus* ‘from the ground up’, cf. *fundus* ‘ground, bottom’), just as occurred in Sanskrit (Ved. *parītaḥ* ‘round about’, but also *hṛttāḥ* ‘from the heart’). A series of adverbs in *-per* (mostly time adverbs, e.g. *semper* ‘always’, *topper* ‘quickly’) probably show a postpositional usage of the preposition *per-* (Baldi and Cuzzolin 2009). Some prepositions (a category derived mainly from PIE adverbial particles) also function as self-standing adverbs (e.g. *ante*: prep. ‘before, in front of’ and adv. ‘beforehand’). Some isolated adverbs based on nouns have arisen from case forms other than acc. sg. or abl./instr. sg., e.g. *nox* ‘by night’ (old nom. sg. or gen. sg.).

See 4.2 on some pronominal adverbs and 5.3 for multiplicative adverbs.

4. Pronouns

Bibliography: in addition to the standard lexica, see relevant entries in Dunkel (2014, Vol. 2); for the demonstrative pronouns in Sabellic, see Penney (2002), Tikkanen (2011: 44–48), Dupraz (2012).

For the gendered pronouns (represented in Italic by the deictic and anaphoric pronouns [4.1], demonstrative pronouns [4.2], and relative and interrogative pronouns [4.3]), PIE displayed a special “pronominal inflection”, which nevertheless shared a number of features with thematic (*o*-stem and *ā*-stem) inflection. (The gendered pronouns have the same inflectional properties as nouns, except that there are no special vocative forms.) Some of the features of PIE pronominal inflection are retained in Italic pronouns and pronominal adjectives, and will be pointed out below. Salient features of pronominal inflection include the following: (i) suppletive stems (two or more different stems within a single paradigm); (ii) lack of **-s* in some nom. sg. masc. forms; (iii) final **-d* in some nom./acc. sg. neut. forms; (iv) gen. sg. forms with **-sġ-* (cf. 1.1.2 on *o*-stem gen. sg. **-osġo*); (v) dat., abl., loc. sg. masc./neut. forms with **-sm-*; (vi) some pl. forms with a diphthongal element, e.g. nom. pl. masc. **-oġ-* (cf. 1.1.2 on 2nd declension nom. pl. masc.) and medial **-oġ-* in some other endings. Elements of pronominal declension are also found in a series of semantically basic adjectives (4.4).

Personal pronouns (4.5) are still more anomalous, with forms that vary widely across the IE languages; thus the reconstruction of the personal pronouns remains unclear in many respects. As often, Italic retains some archaic features, along with a number of innovations.

4.1. Deictic and anaphoric pronouns

The PIE **so/*to-* deictic pronoun (Ved. *sá/tá-*; Gk. ὁ, neut. τό; TB *se*; Go. *sa*, neut. *þata*, etc.) survives for the most part only in traces, especially in frozen adverbial forms (e.g. Lat. *tum* ‘then’, *tam* ‘so’) and in pronominal forms (and some adverbs) largely restricted to OLat., such as forms built to an innovated paradigm in **so-* (e.g. OLat. masc. acc.

sg. *sum*, fem. *sam*, masc. acc. pl. *sōs*; here also loc. sg. **sei(-ke)* > OLat. *SEI*, Lat. *sī*, Volsc. *se* ‘if’ and Lat. *sīc* ‘thus’). But the inherited fem. nom. sg. **seh₂* (Ved. *sá*, Gk. *ή*, Go. *so*) may survive as an adnominal enclitic in South Picene, with definite reference or demonstrative function: **praistakla-sa** ‘the standing object [= statue]’ or ‘this statue’ (Dupraz 2012: 256–258, with references).

The pronominal stem **kō-* (developed from the near-deictic particle **ke/*kī*) does not survive as such, with the probable exception of SPi. **oidom** ‘here, in this place’ (grammaticalized as an adverb from neut. nom./acc. sg. **kīd-om*; see Dupraz 2012: 252–256 and 4[iii] on the *-d* ending).

The anaphoric pronoun (used as a third person pronoun) is expressed by the suppletive stem (or stems; see Dunkel 2014 s.vv. **e-* and **i-*) **i-/*e₂jo-/*e-* (Go. *is*, acc. pl. *ins*; Ved. *ayám*, *asyá*, etc.), sometimes reconstructed with initial **h₁-*, on which see Dunkel (2014: 2.363). The Class. Lat. forms – provided as a sample that also illustrates essential points of the remaining gendered pronoun forms – are as follows (with commentary below):

SINGULAR			
	MASC.	FEM.	NEUT.
NOM.	<i>is</i>	<i>ea</i>	<i>id</i>
ACC.	<i>eum</i>	<i>eam</i>	<i>id</i>
DAT.	<i>eī</i>	<i>eī</i>	<i>eī</i>
ABL.	<i>eō</i>	<i>eā</i>	<i>eō</i>
GEN.	<i>eius</i>	<i>eius</i>	<i>eius</i>

PLURAL			
	MASC.	FEM.	NEUT.
NOM.	<i>eī, ī, iī</i>	<i>ae</i>	<i>ea</i>
ACC.	<i>eōs</i>	<i>eās</i>	<i>ea</i>
DAT.	<i>īs, iīs, eīs</i>	<i>īs, iīs, eīs</i>	<i>īs, iīs, eīs</i>
ABL.	<i>īs, iīs, eīs</i>	<i>īs, iīs, eīs</i>	<i>īs, iīs, eīs</i>
GEN.	<i>eōrum</i>	<i>eārum</i>	<i>eōrum</i>

Various archaic features of pronominal inflection are readily apparent, including:

4(i): suppletive stems **i-* (masc. nom. sg. *is*, neut. nom./acc. sg. *id*, OLat. masc. acc. sg. *im*) ~ **e₂jo-* in most other forms (masc. acc. sg. *eum*, abl. *eō*, fem. nom. sg. *ea*, acc. sg. *eam*, etc.) ~ **e-* in gender-indifferent gen. sg. *eius*, with gender-indifferent *eī* (< **e₂jeī*) built to this stem;

4(iii): neut. nom./acc. sg. in *-d* (*id*);

4(iv): gen. sg. in **-s₂-* (*eius* < **es₂jo+s*, cf. Ved. *asyá*).

Beyond the variation indicated for the plural forms above (partly phonological [the monosyllabic forms result from contractions] and partly analogical [the *e-* forms in nom. pl. masc. and dat./abl. pl. are analogical to the other *e-* forms]), there is considerable

further variation, especially in OLat. (e.g. gen. pl. *eum*, analogical to *o*-stem nominal gen. pl.; *i*-stem dat./abl. pl. *ībus*, etc.); see Ernout (1953: 83–84) for details.

The PIE stem-form **e-sm-* (in dat./abl./loc., cf. Ved. masc. dat. *asmaí*, abl. *asmát*, loc. *asmín* and 4[v]) does not survive in Latin but appears in some Sabellic forms (e.g. loc. sg. U. *esmei*, SPi. **esmín** < **esmej-en*, with postposition; but see 4.2 below on **esmo-* in Sabellic). Two additional Sabellic features are worth noting: “stem-doubling” (cf. de Vaan 2015) via a postposed particle *-id-* (originally the neut. nom./acc. sg. = Lat. *id*), and the frequent appearance of the postposed deictic particle **-ke* (in Italic terms, and generally apocopated in these forms) < PIE near-deictic **kē*; thus (with both features) O. nom. sg. masc. *izic* (< **is-id-k[e]*) and neut. **idik** (< **id-id-k[e]*), further (showing the second feature) O. masc. acc. sg. *ionc* (as if Lat. “*eunc*”), fem. nom. sg. *ioc* (as if Lat. “*eac*”), etc.

A suffixed version of the anaphoric pronoun produces a pronoun of identity, as in Lat. *īdem*, *eadem*, *idem* ‘the same’, and with a related suffix O. nom. sg. masc. **ísidum** ‘id.’, vs. U. nom. sg. masc. *eront* ‘id.’, with a different suffix *-(h)ont*. The historical analysis of these suffixes is complex and subject to differing interpretations; see Buck (1928: 147), Weiss (2011: 342), Dupraz (2012: 232–234), Dunkel (2014: 2.596 with n. 6). Some adverbial forms based on the anaphoric pronoun are noted in 4.2.

4.2. Demonstrative pronouns

The demonstrative pronouns in Italic display considerable formal (and to some extent syntactic/semantic) divergence between Latino-Faliscan and Sabellic. Cognates elsewhere in IE are limited and the Italic formations are largely innovated, apart from the maintenance of features of archaic pronominal inflection, much as with the anaphoric pronoun.

In Latin, there is a three-way contrast involving pronouns that are near-deictic (*hic* ‘this [by me]’, also indirectly attested in Faliscan), distal with second-person reference (*iste* ‘that [by you]’), and far-deictic (*ille* ‘that [over there]’). As with the anaphoric pronoun, all forms have gender-indifferent gen. sg. (*huius*, *istius*, *illius*, cf. 4.1 *eius*) and dat. sg. (*huic*, *istī*, *illī*, cf. 4.1 *eī*) and neut. nom./acc. sg. in *-d* (*illud*, *istud*, *hoc* < **hod-k[e]*, cf. 4.1 *id*). *Iste* and *ille* have so-called “reinforced” forms (i.e. with the deictic element *-c*) in OLat. and in spoken varieties (e.g. neut. nom./acc. sg. *istuc[c]* [< **istud-k(e)*] vs. Class. *istud*, masc. acc. sg. *illunc* vs. Class. *illum*); see Adams (2013: 454–459) on the facts of attestation and their sociolinguistic interpretation. Some further details on each of the three Latin demonstrative pronouns and their counterparts in Sabellic:

hic, *haec*, *hoc*. PIE **g^ho(-)* is innovative as a pronoun in Italic, but perhaps related to the Ved. particle *ghā/ha* and the particle *-go* appended to pronominal genitives in Slavic (Dunkel 2014: 2.283–288). Lat. *hic* < (endingless) pre-Lat. **ho-k(e)* (cf. PIE **so* [4.1] and 4[ii]) and phrasal vowel reduction, whence *hic*, OLat. **hec** (see Dunkel 2014: 2.285 n. 15). Fem. nom. sg. and neut. nom./acc. pl. *haec* have a particle **-i-* or **-ī-* preceding the deictic element. Neut. nom./acc. sg. *hoc* is underlyingly /hocc/ < **hod-k(e)*, thus always scanned heavy in poetry (whence analogical heavy scansions for nom. sg. masc. *hic*, as if /hicc/). OLat. inflected *eccum* ‘here he is’, *eccam* ‘here she is’, etc. may

contain this pronoun (i.e. < **ekke-hom*, **ekke-ham*, etc.), as does the unverbated form *hodiē*, Fal. **foied** ‘today’ (with hypercorrect *f*-), probably with (secondarily shortened) abl. sg. **hō(d)* + abl. sg. *diē(d)* ‘day’; a similar unverbation may underlie Lat. *hōrnus* ‘this year’s’, unless it is a compound with uninflected first member **ho-*. There is no corresponding Sabellic form (see below on the proximal pronouns in Sabellic).

iste, ista, istud. The background of this form is controversial; for competing theories and possible extra-Italic comparanda, see Weiss (2011: 345). Part of the problem: how to evaluate, vis-à-vis Lat. *iste* (with /i-/), the vocalism of the Sabellic forms, with initial /e-/ , as in U. **este, estu** etc., SPi. **estas** etc. (This pronoun is not found in Oscan, with the partial exception of **estam** in Pre-Samnite, a language closely related to Oscan.)

ille, illa, illud. OLat. (and in archaizing Class. Lat. contexts) *olle*, with the original vocalism (cf. O. **ulas**, and extra-Italic comparative material pointing to **ol-no-*, see Dunkel 2014: 2.592–593), cf. also OLat. *uls* ‘on the other side’, Class. *ultrā* ‘beyond’ (compar. *ulterior*, superl. *ultimus*), and a lengthened-grade form in the adv. *ōlim* ‘at one time’ (cf. Ved. *ārē* ‘in the distance’, *ārāt* ‘from a distance’). The *i*-vocalism of *ille* must be secondary (e.g. after *is, iste, ipse*).

All three have related adverbial forms, e.g. locative in *-ei* (HEIC, *hīc* ‘here’ [Fal. **hec, fe**], similarly *istīc* ‘there (by you)’, *illīc* ‘there’); directional forms in *-ō* and *-ā*, perhaps old instrumentals (García Ramón 1997; Vine 2010: 127–128), or (for *-ō-*) from the PIE “directive” (*hōc* ‘to this place’, *eō* ‘to that place’; *hāc* ‘over here’, *aliā* ‘in another direction’, *eā* ‘that way’ in *intereā* ‘meanwhile’); directive forms in *-ūc* (i.e. *-ū*, of unclear origin, plus **k[e]*: *hūc* ‘to this place’, *illūc* ‘to that place’), and others (see Weiss 2011: 354 for details).

As just seen, Lat. *iste* and *ille* have cognates in Sabellic; but the Sabellic systems are in many ways quite different, and also involve marked dialectal distributions within Sabellic. Thus the Sabellic proximal pronoun stem **esto-* (direct cases) appears to be in a suppletive relationship with **esmo-* (oblique cases), as shown by the Umbrian and South Picene data (Dupraz 2012: 29–60); and the Sabellic distal pronouns **ollo-* (Oscan) and **ōlo-* (Umbrian: U. **ulu/ulo**; cf. Lat. *ōlim*) are poorly attested but show interesting commonalities (Dupraz 2012: 117–127). Much better attested, finally, are proximal pronoun stems **eko-* and **ekso-* (in Oscan and Umbrian), but with a complex distribution: **ekso-* (Umbrian) vs. suppletive **eko-* (direct cases) and **ekso-* (oblique cases) in Oscan (Dupraz 2012: 63–115).

There is also a so-called “emphatic” (or focalizing) pronoun *ipse* ‘himself’ in Latin (with the same features of pronominal inflection as the anaphoric and demonstrative pronouns, except neut. nom./acc. sg. *ipsum*). Although *ipse* is often thought to have arisen within Latin (e.g. Weiss 2011: 346–347, based in part on forms of the type OLat. fem. nom. sg. *eapse* vs. Class. *ipsa*), comparison with formally and functionally similar O. **essuf/esuf**, U. **esuf** (Dupraz 2012: 239–247) complicates the question. The background of all of these forms is obscure, in part given their formal divergence (Lat. *i-* vs. Sab. *e-*, Sab. formant *-ōn-*).

4.3. Relative and indefinite/interrogative pronouns

The PIE relative pronoun *(H)*ǵjo-* (as in Indo-Iranian, Greek, Phrygian, Balto-Slavic) was not maintained in Italic. Rather, relative function was taken over by the PIE indefinite/

interrogative pronoun **k^wi-* (*i*-stem) ~ **k^wo-* (*o*-stem and *ā*-stem), as in Hittite, Tocharian, and some other Indo-European traditions; and the same forms continue to be used with indefinite and interrogative functions.

The Class. Lat. forms of the relative pronouns are as follows:

	SINGULAR			PLURAL		
	MASC.	FEM.	NEUT.	MASC.	FEM.	NEUT.
NOM.	<i>quī</i>	<i>quae</i>	<i>quod</i>	<i>quī</i>	<i>quae</i>	<i>quae</i>
ACC.	<i>quem</i>	<i>quam</i>	<i>quod</i>	<i>quōs</i>	<i>quās</i>	<i>quae</i>
DAT.	<i>cui</i>	<i>cui</i>	<i>cui</i>	<i>quibus</i>	<i>quibus</i>	<i>quibus</i>
ABL.	<i>quō</i>	<i>quā</i>	<i>quō</i>	<i>quibus</i>	<i>quibus</i>	<i>quibus</i>
GEN.	<i>cuius</i>	<i>cuius</i>	<i>cuius</i>	<i>quōrum</i>	<i>quārum</i>	<i>quōrum</i>

Note gender-indifferent dat. sg. *cui* and gen. sg. *cuius* (cf. 4.2: *eī*, *huic* etc.; *eius*, *huius* etc.), as well as the neut. nom./acc. sg. in *-d* (cf. 4[iii], 4.1 *id*, 4.2 *illud* etc.). The forms *quem* and *quibus* are *i*-stem forms and the rest are *o*-stem or *ā*-stem forms. In Class. Lat. usage, indefinite/interrogative function calls for the distinct *i*-stem forms *quis* (animate nom. sg.) and *quid* (neut. nom./acc. sg.), including compound pronominal forms like *aliquis* ‘someone’, *nesciōquis* ‘someone or other’ (and others; on the full set, see Weiss 2011: 352–353). But at earlier periods (and in Sabellic), and in archaizing usages in the Classical language, there is considerably more variety, with *i*-stem forms attested more prominently, such as anim. nom. pl. QVES/*quēs* and (so-called) abl. sg. *quī* (also in the postpositional phrase *quicum* and the complementizer *quīn*), perhaps actually (or partly) an old instr. sg.; conversely, the old dat./abl. pl. *o*-stem form appears in OLat. *quīs*. Other forms survive as frozen relics outside the regular relative and indefinite/interrogative paradigms, such as *o*-stem masc. acc. sg. OLat. *quom* > Class. *cum* ‘when’ or *i*-stem neut. (nom./)acc. pl. *quia* ‘because’.

The Sabellic picture is similar (with the regular development of PIE **k^w-* to Sab. *p-*), as in Oscan *o*-stem/*ā*-stem forms like masc. nom. sg. **pui** (cf. Lat. *quī*), fem. **paī/pae** (cf. Lat. *quae*), neut. **púd** (cf. Lat. *quod*), fem. acc. sg. **paam** (cf. Lat. *quam*), masc. acc. pl. Pael. *puus* (cf. Lat. *quōs*), neut. nom. pl. **paí** (cf. Lat. *quae*), and (with non-Latin nom. pl. forms, corresponding to standard Sab. *o*-stem/*ā*-stem declension, cf. 1.1.3 and 1.2.3) masc. nom. pl. O. **pús**, fem. **pas**. For *i*-stem forms: e.g. anim. nom. sg. O. **pis/pis**, neut. **píd** (cf. Lat. *quid*), anim. acc. sg. SPi. **pim** (cf. Lat. *quem*, replacing **quim*), anim. nom. pl. O. **piís** (cf. OLat. *quēs*), acc. pl. U. *pifí* (< **k^wi-ns*, plus a particle). Distinctive, however, are various oblique *sm*-forms (cf. 4[v]), such as dat. sg. U. **pusme**, SPi. **pos-múí**, Pre-Samn. *πυσμοι*. But there is again considerable variety on the Sabellic side: e.g. masc. gen. sg. O. **púieh** (for ***púieis**), which does not match Lat. *cuius*, is the gen. sg. of a possessive adj. (with normal Oscan *o*-stem gen. sg. ending, i.e. the original *i*-stem ending [1.1.3]), this being otherwise attested in O. **púiiu** ‘cuia’. For the pattern, however, cf. Lat. pronominal gen. sg. forms like *nostrī*, *vestrī*, likewise based on the corresponding possessive adjectives (4.5.1 below).

There is, finally, a series of adverbial forms based on a stem **k^wu-* (cf. Ved. *kútra* ‘where?’, *kútas* ‘from where?’, etc.), e.g. O. **puf**, U. *pufe* ‘where’, cf. Lat. *-cubi* (in

alicubi ‘anywhere’). The corresponding forms in Latin lack /k-/ in initial position (thus, in addition to Lat. *ubi* ‘where’, e.g. the particle *ut, utī*, which belongs with O. **puz**, U. **puze/puse**, and the pronominal adjective *uter* ‘which of two?’ [cf. 4.4]), but it is unclear whether this is a phonological or an analogical development (see Weiss 2011: 78 with n. 54).

4.4. Pronominal adjectives

Elements of pronominal inflection (in particular: gender-indifferent gen. sg. in *-ūs* and dat. sg. in *-ī*) are also found in a series of semantically basic adjectives (otherwise of the normal 1st and 2nd declension type), including *ūnus* ‘one’ (and related forms: *ūllus* ‘any’ < **oinelo-*, *nūllus* ‘none’ < **ne-oinelo-*), *alius* ‘other’ (and *alter* ‘one or another of two’, based on the stem of *alius*), *uter* ‘which of two?’ (and *neuter* ‘neither’), *sōlus* ‘alone’, and *tōtus* ‘all’. To some extent, this behavior is inherited: thus for *alius*, with neut. nom./acc. sg. *aliud* (4[*iii*]), the only form in this set with that feature: cf. Gk. nom./acc. sg. neut. ἄλλο < **aliod* and Ved. *anyád* ‘id.’ (but Lat. neut. *alid* [Lucretius, Catullus] is a secondary innovation); more generally, cf. pronominal inflection for Ved. *éka-* ‘one’ (e.g. loc. sg. *ékasmin*) < *(*h₁*)oi-ko-, cf. Lat. *ūnus* < *(*h₁*)oi-no-, and for Ved. *visva-* and *sárva-* ‘all’ (e.g. loc. sg. *visvasmin*, dat. sg. *sárvasmai*), cf. Lat. *tōtus*.

4.5. Personal pronouns

The personal pronouns in Italic (very poorly attested in Faliscan and Sabellic, given the nature of the text types in those languages) preserve much of the archaic suppletion and other irregularity characteristic of the PIE personal pronouns. For the reasons discussed in 4, reconstruction is exceedingly difficult; thus this material has more to do, in some respects, with lexicon (and lexical reconstruction) as opposed to morphology proper. The following discussion selectively treats some of the more salient morphological points, with only limited attention to etymology and extra-Italic comparanda. For further detail, see Weiss (2011: 325–334); also Buck (1928: 139–140) for Sabellic; and Bakkum (2009: 148–151) for Faliscan.

4.5.1. The personal pronouns proper lack gender (as elsewhere in IE), and have no special vocative forms (the nominative is used in vocative contexts); see 4.5.2 for the personal possessives (“possessive adjectives”). The Italic personal pronouns show three major reductions in inflection, as compared with PIE, namely (i) a reduced case system (nom., acc., dat., abl., gen., i.e. no continuation of instr. and loc. forms); (ii) no dual forms (unlike Indo-Iranian, Greek, Gothic, etc.); (iii) no systematic contrast between accented and enclitic forms (unlike Indo-Iranian, Greek, Anatolian, etc.), though some enclitic forms survive as relics and served as derivational bases for new formations.

The Latin forms are given first, with commentary below, including some additional Latin material. (The treatment is restricted to the 1st and 2nd person pronouns and the reflexive pronoun; see 4.1 for the anaphoric pronoun, used as a 3rd person pronoun. The

reflexive pronoun had no nominative, and the sg. and pl. forms are identical, treated here with the 1st and 2nd person sg. forms.)

SINGULAR AND REFLEXIVE FORMS	1 st person	2 nd person	reflexive
NOM.	<i>egō, ego</i>	<i>tū</i>	-----
ACC.	<i>mē(d)</i>	<i>tē(d)</i>	<i>sē(d)</i>
DAT.	<i>mihī, mihi</i>	<i>tibī, tibi</i>	<i>sibī, sibi</i>
ABL.	<i>mē(d)</i>	<i>tē(d)</i>	<i>sē(d)</i>
GEN.	<i>meī</i>	<i>tuī</i>	<i>suī</i>

1st person nom. Lat. *ego* via iambic shortening; /egō/ also in Faliscan (OFal. **eqo/eko**, later **eco**) and South Picene (**ekū**). For the form, cf. Gk. ἐγώ, vs. Ved. *ahám* (with added particle *-om).

2nd person nom. Like Lat. *tū* also O. **tiú** and (with added particle *-om) **tiium**, cf. elsewhere e.g. Hom. τύνη, Av. *tū*, OCS *ty*.

1/2 acc. and refl. acc. OLat. *mēd* (Fal. **med/met**), *tēd*, *sēd* vs. Class. *mē*, *tē*, *sē*, but the reconstruction of the dental suffix is controversial; with a different particle in Sabellic, cf. Paleo-U. **miom**, U. *tiom*, O. *siom*; SPi. **tiom**, in an unclear context, may be acc. or nom.

1/2 dat. and refl. dat. OLat. MIHEI, TIBEI, SIBEI; Class. *tibi*, *mihi*, *sibi* via iambic shortening; mostly the same forms in Sabellic (U. *mehe*, SPi. **tefeí**, U. **tefe**, O. **tfēi**, O. **sifeí**, Pael. *sefeí*), also showing that Lat. /i/ in the first syllable is from enclitic weakening; but U. refl. *seso* probably < **soi* or **sei* followed by unclear material (see Untermann 2000: 682).

1/2 abl. and refl. abl. Cf. Ved. *mát*, *tvát*; the formal equivalence with acc. sg. is secondary (details in Meiser 1998: 157–158).

1/2 gen. and refl. gen. These forms belong formally to the gen. sg. of the corresponding pronominal adjectives (4.5.2). OLat. *tīs* (Plautus) continues the gen./dat. enclitic form **toi* plus a genitive *s*-marker from the nominal system; a corresponding OLat. *mīs*, though claimed in Roman grammatical literature, may not have existed (Weiss 2011: 327 n. 7).

PLURAL FORMS	1 st person	2 nd person
NOM.	<i>nōs</i>	<i>vōs</i>
ACC.	<i>nōs</i>	<i>vōs</i>
DAT.	<i>nōbīs</i>	<i>vōbīs</i>
ABL.	<i>nōbīs</i>	<i>vōbīs</i>
GEN.	<i>nostrum, nostrī</i>	<i>vestrum, vestrī</i>

1/2 nom. These are the original acc. forms.

1/2 acc. These (also 2 pl. Pael. *uus*) continue long-vowel versions of the enclitic forms, cf. Ved. *naḥ*, *vaḥ*. Problematic, however, is Fal. **ves** (see Vine 1993: 179; Katz 1998: 69; Bakum 2009: 150–151).

1/2 *dat./abl.* OLat. NOBEIS, VOBEIS; these reflect a complex Italic innovation (see Meiser 1998: 158–159 and Weiss 2011: 330 for details).

1/2 *gen.* OLat. *vostrum*; the two forms are functionally distinct (*-um* mainly for partitive use, *-ī* mainly for objective gen.); as in the sg. (and refl.), these are derived from the corresponding possessive adjectives (4.5.2), *-um* from gen. pl. and *-ī* from gen. sg. (In Plautus and Terence, masc. *nostrōrum/vostrōrum* and fem. *nostrārum/vostrārum* are also used as gen. pl. pronominal forms.)

4.5.2. For the possessive pronouns (i.e., adjectives based on pronominal stems), the sg. and pl. have distinct formations.

The 1/2 sg. forms reflect thematizations of old pronominal forms: enclitic gen. /dat. **mei* (beside **moi*, cf. Gk. μοι) → **mejos* > Lat. *meus*, *-a*, *-um* (but voc. *mī* from unthematized enclitic **moi* or **mei*); old orthotonic gen. **téye* (cf. Ved. *táva*) → **teyo-* (cf. Hom. τε[Ϝ]ός) > Ital. **toyo-* > U. *touer*, O. **tuvaī**, Lat. *tuus* (with weakened initial vowel). Similarly refl. **seyo-* (Hom. ἐ[Ϝ]ός) > Ital. **soyo-* > OLat. SOVEIS, O. **súvad**, SPi. **súais** (again with Lat. *suus* having undergone a vowel-weakening in unstressed positions; but see Weiss 2011: 334 n.25 for an alternative explanation). The inherited version **suo-* (cf. Ved. *svá-*, Hom. [Ϝ]ός) is also marginally attested in OLat. *sīs* ‘suīs’ (Ennius, Lucretius), *sam* ‘suam’ (Paulus ex Festo).

In the 1/2 pl. forms, the enclitic forms (comparable to Ved. *nah*, *vah*, cf. 4.5.1) were suffixed with oppositional **-tero-* (cf. 2.2.1 and Gk. ἡμέτερος ‘our’, ὑμέτερος ‘your [pl.]’), hence Pr.-Ital. **nostero-* (> Lat. *noster*, *nostra*, *nostrum*) and **uestero-* (cf. U. *uestra*, as well as Fal. **ves** above) → OLat. *voster* (after *noster*) > Class. *vester* by regular sound change. (Oppositional **-tero-* also in [*ne*]uter, 4.3–4.4)

5. Numerals

The study of numerals has as much to do with lexicon as morphology, and involves many special developments. The following sections therefore focus selectively on the morphological features of numerals in Italic, again (as with pronouns) with only limited comparative material. (For details, see Coleman 1992 and handbook treatments like Ernout 1953: 104–112; Leumann 1977: 484–495; Meiser 1998: 170–177; Weiss 2011: 364–376. For recent discussion of Italic numeral forms, including many that cannot be included here, see Prósper 2014–15.)

5.1. Cardinal numbers

5.1.1. ‘1’ through ‘10’

In PIE, the numerals ‘1’ through ‘4’ were fully declined, with ‘5’ through ‘10’ indeclinable. Italic retains inflected forms for ‘1’, ‘2’, and ‘3’ (with inflectional relics for ‘4’):

‘1’: PIE **(h₁)oi-no-*, perhaps originally ‘sole, unique’ (also used in Celtic, Germanic, Balto-Slavic, and marginally in Greek; Dunkel 2014: 2.588–589) > Lat. *ūnus*, *-a*, *-um* (cf. 4.4 on inflection). PIE **sem-*, used as the cardinal numeral ‘1’ in some traditions

(Greek, Armenian, Tocharian), appears in Latin in adverbial usage and in derived adjectives (combined with other elements), as in forms like *semel* ‘once’, *singulī* ‘one each/at a time’; but relics of the cardinal usage may survive in unverbated expressions, such as *mille* ‘one thousand’ (5.1.2 below), OLat. *simītū* ‘at the same time’ (< pre-Lat. **sem* [e] *eitū* [d] ‘at one go’; Vine forthcoming), and possibly *semper* (Dunkel 2014: 2.673).

‘2’: Lat. nom. masc. and neut. *duo*, fem. *duae* (with direct and indirect survival of PIE dual inflection in the nom. forms, and *duo* by iambic shortening from earlier **duō*; cf. Ved. *dvāu/dvā*, Gk. *δύω*, *δύο*, etc.). The remaining Lat. forms (mostly remade on the basis of thematic plural inflection, apart from athematic dat./abl. pl.) are acc. masc. *duōs*, fem. *duās*, neut. *duo*; dat./abl. masc./neut. *duōbus*, fem. *duābus*; gen. masc./neut. *duōrum*, fem. *duārum* (but OLat. *duom*, *duum*, preserved in the technical term *duomvir*, *duumvir* ‘member of a board of two’). Sabellic forms are attested in Umbrian, which shows no trace of dual inflection even in the nom., and has remade the inflection entirely after thematic plurals: nom. masc. *dur* (with regular *o*-stem ending **-ōs* [1.1.3]); further acc. **tuf** (though possibly feminine, see Untermann 2000: 193), dat./abl. **tuves/duir**, acc. neut. **tuva**.

‘3’: PIE masc. **tréjes* > Lat. masc./fem. *trēs*, neut. **trih₂* → Lat. *tria* (remodeled from expected **trī*, cf. 5.1.2 on ‘30’), inflected as an *i*-stem: cf. gen. pl. *trium*. (Cf. O. nom. masc./fem. **trís**, U. acc. masc./fem. *trif*, nom./acc. neut **triia**, etc.). (The PIE feminine **t[r]i-sr-és* does not survive in Italic.)

‘4’ through ‘10’ are indeclinable: Lat. *quattuor* ‘4’ (O. **pettiur**; but cf. neut. pl. *petiro*-with postposition [5.3.2]), and a trace of inflection may be preserved in the gloss form O. *pitora*, v.l. *petora* [Festus]; see Buck 1928: 138, Prósper 2014–15: 10–12, 35 n. 59), *quīnque* ‘5’, *sex* ‘6’, *septem* ‘7’, *octō* ‘8’, *novem* ‘9’, *decem* ‘10’. (‘5’ through ‘10’ are not directly attested in Sabellic, apart from derived forms and compounds, including onomastic usage.)

5.1.2. Cardinals above ‘10’

‘11’ through ‘17’ are dvandva compounds consisting of the uninflected unit numeral followed by ‘10’ (like Ved. *dvādaśa*, Gk. *δώδεκα* ‘12’, literally ‘two-ten’). The combinations have in most cases undergone phonological developments affecting the unit numerals, as well as a special development affecting ‘10’ (in which the vowels in the expected vowel-weakening result */-dikem/* < *decem* have been metathesized); thus, for example, Lat. *undecim* ‘11’, with late Osthoff-shortening in **ūndecim*, from syncopated **oino-dekem*, with dissimilatory *n*-loss from **oinon-dekem*, itself with nasal assimilation from **oinom-dekem*. See the handbooks for details on the other forms, which are *duodecim* ‘12’, *trēdecim* ‘13’ (or *tredecim*; the vowel quantity in the initial syllable is uncertain), *quattuordecim* ‘14’, *quīndecim* ‘15’, *sēdecim* ‘16’, *septemdecim* ‘17’. The only Sabellic form in this series is U. *desenduf* ‘12’ (inflected; masc. acc. pl.), with a reversal of the Latin pattern (literally ‘ten-two’).

‘18’ and ‘19’ use subtractive expressions: *duodēvīgintī* (literally ‘2 from 20’) and *ūndēvīgintī* (literally ‘1 from 20’). Subtractive numeral expressions are well-attested in Etruscan, which is often thought to be the source of this usage in Latin (see Meiser

1998: 172); but this is disputed, given subtractive formats in Greek, Vedic, Germanic, and elsewhere (see Weiss 2011: 371 with n. 41, 485).

‘20’, ‘30’, etc.: the decad terms were formed in PIE with the neut. pl. of the unit numeral followed by a form based on ‘10’, probably **-d̥k̥omth₂* (cf. Gk. -κοντα); but the resulting forms have undergone many special developments, especially via contaminations (e.g. influence from adjacent terms) and allegro reductions or lenitions, as is typical for counting systems. Again, see the handbooks for details. Interesting as an inherited term is Lat. *vīgintī* ‘20’ (cf. Ved. *viṃśati-*, Gk. εἴκοσι, etc.) < **d̥ui(h₁) d̥k̥mtih₁* ‘two decads’. This should have led to Lat. *†vīcentī*; the attested form shows a vowel-harmonic assimilation (*/ī ... e/ > /ī ... i/*) and an allegro lenition of the intervocalic voiceless stop. The other Lat. terms are *trīgintā* ‘30’ (with *trī-* perhaps preserving the expected outcome of PIE neut. **trih₂* ‘3’, cf. 5.1.1), *quadrāgintā* ‘40’, *quīnquāgintā* ‘50’, *sexāgintā* ‘60’, *septuāgintā* ‘70’, *octōgintā* ‘80’, *nōnāgintā* ‘90’. Note (among many other details) the spread of a “connecting vowel” */-ā-/* from *quadrāgintā* to all the other decads except ‘80’.

‘100’ and the hundreds: Lat. *centum* ‘100’ < PIE **d̥k̥mtóm* (cf. Ved. *śatám*, Gk. ἑκατόν, etc.). (A trace of the Sabellic version **kantom* may survive in the onomastic form *CANTOVIO*, in an OLat. inscription in Marsian territory.) The hundreds are formed with the unit numbers as first members of compounds, with second member *-centī* (for ‘200’, ‘300’, ‘600’) and *-gentī* (for the rest) inflected as 1st/2nd declension adjectives (thus *-centī*, *-centae*, *-centa* etc.); the forms, which again show many special features, are *ducentī* ‘200’, *trecentī* ‘300’, *quadringentī* ‘400’, *quīngentī* ‘500’, *sescentī* ‘600’, *septingentī* ‘700’, *octingentī* ‘800’, *nōngentī* ‘900’. (In an alternate OLat. usage, the neut. sg. form of the number accompanies a genitive of the term being counted.)

‘1000’: Lat. *mīlle* is a sg. based on unverbated **smih₂ ḡ^heslih₂* ‘one thousand’, whence pl. *mīlia* (cf. Ved. *sahásra-* < **sm-ḡ^heslo-* and, with a derivative of the same **ḡ^heslo-*, Gk. [Ion.] χεῖλιοι).

5.2. Ordinal numbers

The ordinals are inflected as 1st/2nd declension adjectives. (The lower ordinals through ‘10th’, and especially ‘5th’ through ‘10th’, are prominent in onomastic usage.)

5.2.1. 1st through 10th

As in many IE traditions, ‘1st’ and ‘2nd’ are unrelated to cardinal ‘1’ and ‘2’: Lat. *prīmus* ‘1st’ (Pael. *Prisma*, woman’s name), originally the superlative (< **pri-ismo-*) of the adverb OLat. *pri* ‘prae’ (Festus), cf. comparative *prior* ‘earlier, former’; different formations (though based on related adverbial roots and with a different superlative suffix) in Fal. **pramo** (< **p̥rh₂-mo-*) and U. *promom* (< **pro-mo-* [3.1]). Lat. *secundus* ‘2nd’ (literally ‘following’) is originally the gerund (7.3.1.4) of *sequor* ‘follow’.

‘3rd’: Lat. *tertius* (U. *tertīm*) < **tri-ti̯o-* (cf. Av. *θritiia-*, MW *trydyd*), one of several versions in other traditions (e.g. **tri-to-* in Gk. τρίτος, **tj-t-* in Ved. *tṛtīya-*, etc.).

‘4th’, ‘5th’, ‘6th’: These are formed with a suffix **-to-*. ‘4th’: Lat. *quārtus* (with complex and partly unexplained developments from **k^watuyr-to-* vel sim.; see recently Prósper 2014–2015: 7–10), cf. Sabellic **peturto-* (in the name ΠΕΤΥΡΤΙΥΣ) < **k^weturto-* (cf. Ved. *caturthá-*). ‘5th’: Lat. *quīntus* (cf. Gk. πέμπτος, OHG *fimfto* etc.) (see recently Prósper 2014–2015: 24–26 on complex phonological issues raised by the Latin form), cf. O. *pomtis* ‘5 times’ and onomastic forms like πομπτιεσ (gentilicium). ‘6th’: Lat. *sextus* (cf. Gk. ἕκτος, Li. *šēštas* etc.), cf. U. *sestentasiaru* ‘bi-monthly’ (i.e. in cycles of one-sixth of a year), with a basis in this ordinal form.

‘7th’ through ‘10th’: These adjectives are derived by thematizing the cardinals with a suffix **-o-*. ‘7th’: Lat. *septimus* (cf. Ved. *saptamá-*), Paleo-U. *setums* ‘Septimus’. ‘8th’: Lat. *octāvus*, cf. O. *ūhtavis* ‘Octavius’ (with a development **-ō- > /-ā-/*). ‘9th’: Lat. *nōnus* (< pre-Lat. **nouen-o-*, with the original final dental nasal; cardinal *novem* with *-m* after *decem*), cf. Ved. *navamá-* (with *-m-* after *daśamá-* ‘10th’). ‘10th’: Lat. *decimus* (O. δεκμας).

5.2.2. ‘11th’ through ‘19th’

‘11th’ and ‘12th’: based on the cardinals and *decimus* ‘10th’, i.e. *undecimus*, *duodecimus*.

‘13th’ through ‘17th’: both the unit and the decimal are declined as ordinals, thus *tertius decimus* ‘13th’, etc.

‘18th’ and ‘19th’: with the subtractive pattern as in the cardinals (cf. also ‘20th’ below), i.e. *duodēvīcēsīmus* and *undēvīcēsīmus*.

5.2.3. ‘20th’ through ‘100th’ (and higher)

‘20th’: *vīcē(n)simus* (VICENSVMAM) < **ūīkmt-tṃmo-*, serving as model for the rest, e.g. *trīcē(n)simus* ‘30th’, *quadrāgē(n)simus* ‘40th’, *septuāgē(n)simus* ‘70th’ (cf. the decad terms in 5.1.2) and *centē(n)simus* ‘100th’. The same pattern is used for ordinals based on the cardinal terms for ‘hundreds’ (5.1.2), thus *ducentē(n)simus* ‘200th’, *quīngentē(n)simus* ‘500th’, etc., up to *mīllē(n)simus* ‘1,000th’ (based on *mīlle* ‘1000’).

5.3. Other number forms

There are many isolated forms derived from (or related to) numerals, which cannot be treated here; already mentioned above were derivatives of **sem-* (5.1.1), to which can be added the first-compound-member term for ‘one-half-’, i.e. Lat. *sēmi-* (cf. Gk. ἡμι-), based on the loc. sg. of the root noun **sem-* (Dunkel 2014: 2.679). Note further, for example, various forms based on **(h₁)oi-no-* ‘1’ itself: Lat. *ūnicus* ‘the one (and only)’ (cf. Go. *ainahs** ‘only’, similarly [but with *e*-grade in the root] OCS *inokū* ‘monk’, and with similar suffixation Ved. *ekaká-* ‘alone’), and the pronominal adjectives *ūllus* ‘any-one’ (< **oinelo-*) and *nūllus* ‘no one’ (< **ne-oinelo-*) (4.4), as well as the negative particle *nōn*, OLat. *noenum* (< **ne-oinom*; differently Dunkel 2014: 2.533); and with

different primary suffixation in **-ko-* (like Ved. *eka-*) as well as different secondary suffixation, U. *ecla* ‘every’ (< **eiko-lo-*), *ekvi* ‘one (time?)’ (< **eiko-ūio-*). There are, however, several types of morphological behavior connected with numerals that can conveniently be sketched here (although some of this material belongs to “derivational morphology”).

5.3.1. First compound members

Apart from predictable stem-forms based on cardinals (such as **[h₁]oi-no-* > OLat. OINO-, Class. *ūni-* ‘1-’, *quīnqu-/quīnc-* ‘5-’, or pre-consonantal *sē-* ‘6’ with regular phonological reduction of /seks-C/), the following are noteworthy:

‘2-’: PIE **dwi-* (cf. Ved. *dvi-*, Gk. δ[ɸ]ι-) > OLat. *dvi-* (Festus), Class. *bi-* (Lat. gloss forms in *di-* [*diennium* ‘period of two years’ etc.] may be Grecisms); the vowel length in some forms is regular via contraction (e.g. *bīmus* ‘two years old’ < **dwi-him-o-* ‘having two winters’), but must be analogical in *bīduum* ‘period of two days’ (second member **-diuom*), although the source is uncertain. A form *du-* is also used, e.g. *duplex* ‘two-fold’ (U. **tuplak** ‘two-fold [cloth]’), perhaps influenced by the cardinal (cf. also *ducentum/ducentī* ‘200’), but see Weiss (2011: 367) for a possible phonological explanation.

‘3-’: PIE **tri-* (cf. Ved. *tri-*, Gk. τρι-) > Lat. *tri-*, but *ter-* (cf. *tertius* ‘3rd’, *ter* ‘3 times’) spreads in Imperial Lat.: *trigeminus* ‘triplet, threefold’ (Plautus+) but *tergeminus* (ps.-Tibullus+); also *trē-* and (with secondary shortening) *tre-* in some numerals (see 5.1.2 on ‘13’, ‘300’) < *trēs* ‘3’.

‘4-’: PIE **k^wtru-* (cf. **k^wetru-* in Av. *caθru-*, Gaul. *Petru-*) > Lat. *quadru-* (with irregular voicing of the **-t-*, as in other forms: *quadrāgintā* ‘40’, *quadrāre* ‘to square’) and *quadri-*; most of the attested *quadru-* forms precede a labial (e.g. *quadrupēs* ‘4-footed’), so the /u/ is ambiguous and could reflect **u*, **o*, or **i*; differently U. *peturpur-sus* ‘4-footed’, based on the cardinal (like Ved. *catur-*).

‘7-’, ‘8-’, ‘10-’, ‘100-’: In addition to *septem-*, also *Septi-* (in the ancient topographical term *Septimontium*) and *septu-* before vowels (*septuennis* ‘7 years old/long’), perhaps analogical to *quadri-/quadru-* (but see Meiser 1998: 172 for a phonological explanation); similarly, in addition to *Octō-* (*Octōber*) also *octi-/octu-* (*octipēs* ‘8-footed’), perhaps from shortened **octō-*; and besides *centum-* also *centi-* (Horace+). Note also the derivational basis **deku-* ‘10’ required by some forms in both Latin (*decuria* ‘group of 10 men’) and Sabellic (O. **dekkviarim**, U. **tekvias**, of uncertain meaning but probably related to ‘10’), cf. Prósper (2014–2015: 35 n. 59).

5.3.2. Multiplicative forms

‘Once’: The background of Lat. *semel* (apart from its basis in **sem-* [5.1.1]) is controversial; see Meiser (1998: 176) and Dunkel (2014: 2.491).

‘Twice’, ‘3 times’, ‘4 times’: Following an inherited pattern, a suffix **-s* (Dunkel 2014: 1.169) is added to the compositional form of the cardinal, thus Lat. *bis* ‘twice’ < **dwi-s* (cf. Ved. *dviḥ*, Gk. δ[ɸ]ίς, OHG *zwir*), *ter* ‘three times’ < **tri-s* (cf. Ved. *triḥ*, Gk. τρίς), *quater* ‘4 times’ < pre-Lat. **quatrus* (cf. Av. *caθruš*). O. *pomtis* ‘5 times’, based

on the ordinal, has apparently extracted a suffix */-is/* from **dūi-s* and **tri-s*. (But a different strategy in U. **triiuper** ‘3 times’, O. *petiropert* ‘4 times’, via a postposition *-pert* with neut. pl., literally ‘up to’; Untermann 2000: 545–546 s. v. *pert*.)

‘5 times’ etc. in Latin: Cardinal plus a suffix *-iē(n)s* (*quīnquiē[n]s* ‘5 times’ etc., also with shortened form *centiē[n]s* ‘100 times’, cf. *centi-* [5.3.1], and haplological shortening in other forms, e.g. *vīciē[n]s* ‘20 times’ for †*vīcentiē[n]s*); the formation arises from a false segmentation of *totiēns* ‘so many times’, *quotiēns* ‘how many times?’ < **toti* ‘so many’ (> Lat. *tot*), **k^woti* ‘how many?’ (> Lat. *quot*, cf. Ved. *kāti*) plus a suffix **-ent-*, as in Ved. *īy-ant-* ‘so great’, *kīy-ant-* ‘how great?’.

5.3.3. Distributive forms

Apart from the isolated *singulī* ‘one each/at a time’ based on **sem-* (5.1.1), a series of so-called distributive adjectives (‘two each’ etc., inflected as 1st and 2nd declension plural adjectives) was originally made by suffixing **-no-* to the multiplicative adverbs, thus **dūis-no-* > *bīnī* ‘two each’, **tris-no-* > *ternī* ‘three each’. But a false segmentation based on such forms yielded a suffix variant **-sno-*, used for ‘five each’ (*quīnquēnī*) and above, in some cases added to shortened versions of the cardinals (e.g. *dēnī* < **dek-sno-*, cf. *dec-iē[n]s* ‘10 times’). See also Weiss (2011: 375–376) on homophonic collective numeral adjectives, which may have had a different origin.

6. Derivational morphology

Bibliography: extensive treatments of Latin material in Leumann (1977: 273–403) and Weiss (2011: 266–324) (see also the bibliography cited at 2011: 270 n. 20); for Sabellic, note Heidermanns (1996) and (in less detail) Buck (1928: 182–194), as well as Poultney (1959: 84–97) (for Umbrian).

Italic displays a fairly standard profile of primary and secondary derivation, comparable to other early-attested IE languages and branches, such as Indo-Iranian or Greek. Some formations are of course better represented or more productive than others, or have undergone special developments, displaying a characteristic Italic “touch”. For example, deverbial or deradical neuter *s*-stems of the inherited (originally proterokinetic) type are well represented (*genus*, gen. *generis* ‘birth, race, kind’, cf. Ved. *jānas-*, Gk. γένος), but animate (originally amphikinetic) *s*-stems (in Italic almost exclusively masc.) are very well developed, especially as part of a derivational system associated with 2nd-conjugation stative verbs (7.1.1), also encompassing adjectives in *-idus*: *timor* ‘fear’, gen. *timōris* (*/-r-/* < **-s-* via rhotacism) ~ *timeō* ‘be afraid’ ~ *timidus* ‘afraid’.

There is no space for anything like full treatment here; it is possible, however, to sketch some developments that are particularly characteristic of Italic (beyond those already mentioned above, such as Lat. *-iēs/-ia* [1.5], the Italo-Celtic superlative morpheme [2.2.2], etc.).

6.1. Nominal derivation: suffix patterns

The following points concern nominal suffixation – in Italic, as in PIE, the primary modality for derivational morphology.

Original *u*-stem adjectives do not survive as such, but show an extension in *-i-* (thus appearing as “*ui*-stem” adjectives, e.g. Lat. *gravis* ‘heavy’ vs. Ved. *gurú-*, Gk. βαρύς) – in this case, an innovation that may be restricted to Latin (though see Heidermanns 1996: 156–157 for a possible trace in Sabellic). In general, the most characteristic pattern of Italic innovation involves the development of suffix “conglomerates”, with a prominent concentration of new abstract and adjective formations, although other types (such as agent nouns) are also represented.

In some cases, the pattern involves athematic material only, as in the “*ui*-stems” just mentioned, or the *k*-extended feminine agent nouns in **-trī-k-* of the type Lat. *genetrīx* ‘mother’ (vs. the *k*-less “*devī*-formations” in Ved. *jānitṛī*, Gk. γενέτειρα), Marruc. *sacra-crix* ‘priestess’. In other cases, the suffixal extension involves (plain) thematization of an athematic suffix, as in adjectival forms in *-ācus* (e.g. *merācus* ‘undiluted’), perhaps in origin reflecting thematizations of *āk*-stems (themselves well represented: *audāx* ‘bold’ and many more in Latin; with *audāx* probably SPi. **áudaqum**, cf. Fortson 2016: 20–23); still more frequent is thematization via a complex thematic formant, thus in addition to *men*-stems (originally deverbal nouns, well represented as such: e.g. *sēmen* ‘seed’, cf. *serō* ‘sow’), there are deverbal nouns in *-mentum* (*alimentum* ‘nourishment’, cf. *alō* ‘nourish’; already VOL IOVXMENTA in the Forum Inscription, Class. *iūmenta* ‘yoked teams [of animals]’, cf. *iungō* ‘join, yoke’) < **-mṇ-to-*, perhaps originally substantivized possessive derivatives in **-to-* based on *men*-stems (Weiss 2011: 313–314, with comparative material). Similarly, new adjectival formations with PIE **-ijō-* (frequently substantivized) added to athematic bases are typical: e.g. *mōn*-stems → *-mōnium/-mōnia* (*alimōnium/-ia* ‘nourishment’), pres. participial stems in *-ent-* (7.3.1.1) → *-entio-* (*silentium* ‘silence’, cf. *sileō* ‘be silent’, ptcple. *silēns*), etc.

A highly characteristic subtype of the general case just mentioned (athematic base plus complex thematic suffix) appears in so-called “deinstrumental” adjectival formations: thus *astū* ‘cleverly’ (*u*-stem instr. sg. **-u-h₁*, i.e. ‘with cleverness’) → *astūtus* ‘clever’, similarly *-ītus* adjectives based on *i*-stem instr. sg. **-i-h₁* (*aurītus* ‘having ears’, cf. *auris* ‘ear’), *-ōtus* adjectives based on *o*-stem instr. sg. **-o-h₁* (*aegrōtus* ‘sick’, cf. *aegrum* ‘distress’); likewise, with instr. sg. plus **-no-*, adjectival suffix conglomerates *-īnus*, *-ūnus*, *-ōnus* etc. (cf. Weiss 2011: 290, 293, and further 7.3.1.3 on *futūrus*). In some such forms, however, the underlying derivational process may be different: thus some *-īnus* adjectives may go back to thematic gen. sg. noun plus **-no-* (*vīcīnus* ‘neighboring’, cf. *vīcus* ‘district’; Weiss 2011: 288). A formation of this general type that became productive (though not exclusively based on instr. sg. forms) is *-ānus* (also well represented in Sabellic), originally based on *ā*-stems (place name *Abella* → adj. *Abellānus*, O. dat. sg. **abellanúí**) but extended to other stem types (e.g. *i*-stem *urbs* ‘city’ → *urbānus* ‘urban’). Some similar thematic formations that became integrated into the verbal system are treated below (e.g. verbal adjectives in **-tō-*, as in Lat. *-āto-*, *-ito-* [< **-eto-*], *-īto-*, 7.3.1.2).

Some complex athematic formations are characteristic of Italic but also have extra-Italic comparanda, even if questions remain about their morphological background. Thus the frequent abstract suffix *-tāt-* (OLat. *aevitās*, Class. *aetās* ‘lifetime’, O. acc. sg. **aita-**

tum) has cognates in Indo-Iranian and Greek (see Pike 2011), and the abstract suffix *-tūt-* (Lat. *senectūs* ‘old age’) has cognates in Celtic and Germanic; their internal structure is uncertain, but may again involve a basis in instrumental forms (Weiss 2011: 304). Though not frequent, the complex *s*-stem suffix **-n-es-* appears in a series of culturally important terms relating to commerce/exchange and social obligation (e.g. *fēnus* ‘profit’, *mūnus* ‘duty, service’) and has cognates elsewhere that display a similar semantic profile. In other cases, complex athematic suffix formations are Italic innovations, particularly frequent with secondary *n*-stems, as in the abstracts in *-iō*, gen. *-iōnis* (Lat. *legiō* ‘legion’, O. acc. sg. **leginum**) and (much more productively, based on verbal adjectives in *-tus* and *-sus* [7.3.1.2]) *-tiō* and *-siō* (Lat. *nātiō* ‘birth; race, nation’, U. abl. sg. **natine**; in this and the preceding Oscan form, note the discrepancy between the Latin and Sabellic inflectional behavior: Lat. oblique *-iōn-* vs. Sab. *-in-*), and in a series of complex abstract suffixes with dental or velar preceding the *n*-stem suffix proper, e.g. *-ēdō*, *-īdō*, *-ūdō* and *-tūdō* (gen. *-ēdinis* etc.), *-īgō*, *-āgō*, *-ūgō* (gen. *-īginis* etc.). In one prominent case, a borrowed source has been claimed for a suffix, i.e. (secondary) *-ti-*, *-āti-*, *-ūti-* in ethnonyms based on place names (*Saepinās*, gen. *-ātis* ‘inhabitant of Saepinum’, O. nom. sg. **saipinaz**), with a possible Etruscan background (see the discussion, with references, by Penney 2009: 92–93); but the pattern is also attested in Celtic and in some native Italic vocabulary (e.g. *quoiātis* [Plautus] ‘of what country?’ > Class. *cuiās*, cf. *cuius* ‘whose’), which may suggest, at least in part, a native background.

Innovations are also found among thematic suffixal material. Thus Italic inherited a diminutive formation in **-elo-* (OFal. **arcentelom** ‘a little silver’) > Lat. *-ulo-*, e.g. *catulus* ‘puppy’ = U. **katel** (see Leumann 1977: 309 for possible comparative material); but the Italic diminutive conglomerate **-kelo-* (> Lat. *-culo-*, cf. U. **struhçla** ‘offering cake’, with <ç> the palatalization result of **k* before the **e* was syncopated) is an innovation, based on the productive suffix **-ko-*, in Italic terms (Fruyt 1986), reflecting PIE **-ko-* and **-kō-* and with interesting correspondences to Indo-Iranian **-ka-* (Jamison 2009). In other cases, however, the background of a complex thematic suffix is controversial, as in the highly characteristic Lat. suffixes *-ōsus* and *-ulentus* (forming possessive adjectives); see Weiss (2011: 296–297) for details.

6.2. Compounds

The repeated complaints of Lucretius about the poverty of Latin expression (*De Rerum Natura* 1.832, 1.136–139, 3.260), especially in technical vocabulary, probably referred at least partly (if not largely) to the less extensive use of compounding in Latin, as compared with Greek (Kenney 1971: 110 ad 3.260). But compound formations of all types represented otherwise in Sanskrit and Greek (and generally reconstructed for PIE) are not rare in Italic (cf. on *agricola*, 1.2), even if less prominent than in those languages. In addition to handbook treatments like Leumann (1977: 383–403) and Weiss (2011: 262–265), see the bibliography cited at (Weiss 2011: 262 n. 14), to which can be added Moussy (2005), Lindner and Oniga (2005), and (for Sabellic and Proto-Italic) Heidermanns (1996: 307–320 and 2002).

6.3. Verbal derivation

The pattern *timor/timeō/timidus* (6) straddles both nominal and verbal derivation. But some regular processes of verbal derivation are also found. These cannot be surveyed in detail, but prominent examples include (i) factitive 1st-conjugation verbs based on adjectives (*sānus* ‘healthy’ → *sānāre* ‘heal, restore to health’, *novus* ‘new’ → *renovāre* ‘renew’), an inherited pattern (cf. Hitt. *newahh-* ‘renew’ and 7.1.1 on athematic factitives in **-eh₂-*); (ii) repetitive and frequentative verbs, synchronically associated with the perfect passive participle (see Weiss 2011: 401–402 for details).

7. Verbs

Bibliography: important monograph-length treatments (mainly on Latin) include Meiser (2003; cf. Schrijver 2006), de Melo (2007), Garnier (2010), Seldeslachts (2001); for Sabellic: García Castillero (2000).

The major Italic innovation was the reorganization of the PIE present, aorist, and perfect tense/aspect categories (as well as secondary categories like iterative-causative, stative, and desiderative) into a two-part (mainly) tense-based system, opposing for each verb an *inflectum* or “present system” (7.1) (with four “conjugation classes” and all forms based on a “present stem”) to a *perfectum* or “perfect system” (7.2) (with all forms based on a “perfect stem”). The Italic present system is the repository of the PIE present types (including secondary categories: iterative-causative, etc.), while the perfect system reflects a merger between the PIE aorist and perfect. The Latin and Sabellic present systems largely coincide, but there are marked divergences between the Latin and Sabellic perfect systems. Finite verbs are inflected for singular and plural, the dual having been lost (apart from possible traces in some synchronic plural endings).

For finite forms: in the indicative, the present system has three tenses (present, imperfect, and future) and the perfect system has three tenses (perfect, pluperfect, and future perfect); in the subjunctive, the present system makes a present and an imperfect subjunctive, while the perfect system makes a perfect and a pluperfect subjunctive. (The PIE optative has been lost as a category, but survives in some synchronic subjunctives.) The present imperative and the future imperative are based on the present stem. Indicative and subjunctive have both active and passive forms (with the passive continuing the PIE “oppositional middle”); but some verbs (so-called “deponents”; see Flobert 1975) have only passive forms, with non-passive meaning (continuing PIE *media tantum* verbs). In a few cases, so-called “semi-deponent” verbs have active inflection in the present system but passive inflection in the perfect system. (For an overview of PIE tense, aspect, voice, and mood and their development in Italic, see Weiss 2011: 377–384.)

The non-finite forms (infinitives, participles, verbal nouns) are treated separately (7.3).

7.1. Present system

7.1.1. Regular verbs

Each of the four regular conjugation classes has a characteristic stem vowel, as follows: I *-ā-*; II *-ē-*; III *-e-* (but in Latin normally *-i-*, as a result of vowel-weakening), including a “III-*iō*” subtype (with some forms similar to IV); IV *-ī-*. To some extent, the reorganization into these classes was driven by purely phonological factors, especially (for I and II) the loss of intervocalic yod and subsequent vowel contractions. Thus for I, forms based on PIE **-eh₂-* (e.g. athematic verbs to roots ending in **...eh₂-*, athematic factitives in **-eh₂-*) or **-eh₂-je/o-* (e.g. denominatives to *eh₂-*stems or thematized versions of the preceding) naturally merged into a new “*ā*-conjugation”; for II, the PIE iterative-causatives (in **-éje/o-*, whence **-ee-* and contraction to **-ē-*) and statives (in **-éh₁-* and **-éh₁-je/o-*, whence **-ē-* for the first and **-ēje/o- > *-ēe- > *-ē-* for the second) naturally merged into a new “*ē*-conjugation”, joined by PIE athematic (or thematized) verbs with roots ending in **...eh₁-*. Class III continues PIE plain thematic verbs; and classes III-*iō* and IV both continue PIE **je/o-*verbs, with a secondary differentiation that is again phonologically driven, in this case by syncope processes. (For those details, see Weiss 2011: 122. An alternative view, however, considers III-*iō* verbs to descend from an inherited class of athematic *i*-presents; see Schrijver 2003, de Vaan 2011.) Some monograph-length works on the individual conjugation classes are available: I Steinbauer (1989; with special attention to denominatives); II Hocquard (1981), Vernet i Pons (2008); III Bock (2008); III-*iō* and IV: Martzloff (2006); and for primary presents in Sabellic, see García Castillero (2000).

7.1.1.1. Sample present active indicative paradigms for Latin verbs belonging to each of the conjugation classes are as follows (I *amō* ‘love’, II *doceō* ‘teach’, III *agō* ‘lead’, III-*iō* *capiō* ‘take’, IV *audiō* ‘hear’):

	SINGULAR				
	I	II	III	III- <i>iō</i>	IV
1	<i>amō</i>	<i>doceō</i>	<i>agō</i>	<i>capiō</i>	<i>audiō</i>
2	<i>amās</i>	<i>docēs</i>	<i>agis</i>	<i>capis</i>	<i>audīs</i>
3	<i>amat</i>	<i>docet</i>	<i>agit</i>	<i>capit</i>	<i>audit</i>

	PLURAL				
	I	II	III	III- <i>iō</i>	IV
1	<i>amāmus</i>	<i>docēmus</i>	<i>agimus</i>	<i>capimus</i>	<i>audīmus</i>
2	<i>amātis</i>	<i>docētis</i>	<i>agitis</i>	<i>capitis</i>	<i>audītis</i>
3	<i>amant</i>	<i>docent</i>	<i>agunt</i>	<i>capiunt</i>	<i>audiunt</i>

All of these forms continue PIE thematic inflectional endings. The most salient innovation is the loss by apocope of the “hic et nunc” *-i* on the 2 sg., 3 sg., and 3 pl. primary

thematic endings (PIE **-si*, **-ti*, **-nti*), although it is widely assumed (see Sarullo 2014: 210–215) that VOL *tremonti* ‘they tremble’, in a fragment from the extremely archaic Carmen Saliare, preserves the unapocopated 3 pl. ending (cf. Dor. Gk. -ovτι, Ved. -anti). (The thematic endings have been adopted even by regular verbs with a background in athematic categories, e.g. nasal presents: *linquunt* ‘they leave’ [cf. Ved. *riñcánti* < **li-n-k^w-énti*] with **-o-nti*, not †*linquent*.) The most serious indeterminacy concerns the 2 pl. ending, where Lat. *-tis* (< **-tes*) and SPi. **-tas** (in *videtas* ‘you see’) may continue old dual endings (Weiss 2011: 386; Clackson 2015: 17–18).

In general, and apart from purely phonological differences (e.g. non-contraction of 1 sg. */-ā-ō/* < **-ājō* in 1st-conjugation forms [e.g. U. *subocau* ‘I invoke’], or final stop loss, as is characteristic of Umbrian [e.g. U. *habe* ‘he has’]), the Sabellic forms match those of Latin, insofar as this can be judged from the attestations (1 pl. pres. indic. forms are lacking, and most second person forms are imperatives). The major exception is the 3 pl., with */-ent/* (from the original athematic ending **-ent[i]*) instead of thematic **-o-nt(i)*; cf. O. **fiiet** (with suppression of the nasal) ~ Lat. *fiunt* ‘they become’. A representative selection of additional forms (beyond those just cited) includes: I 3 sg. O. **faamat** ‘he declares’, SPi. **qumat** ‘he lies’ (: Lat. *cubat*, Fal. **cupat**), 3 pl. SPi. **persukant** ‘they declare (?)’; II 1 sg. SPi. **kduíú** ‘I am famed’ (: Lat. *clueō*), 3 sg. O. **kasit** ‘it is fitting’ (: Lat. *caret* ‘it is lacking’); III 1 sg. U. **sestu** ‘I set up’ (: Lat. *sistō*), 2 sg. U. **seste** ‘you set up’ (: Lat. *sistis*, with final *s*-loss in Umbrian), 3 sg. Vest. *didet* ‘he gives’; III-*iō*: see 7.1.1.4 for some non-indicative forms. As these examples indicate, the Sab. present stem formations for individual verbs generally correspond to those of Latin (and Faliscan); but a notable exception is ‘give’, regularly a 3rd-conjugation reduplicated verb in Sab., cf. Vest. *didet* above (as if Lat. †*didō*, *didere*) vs. Lat. *dō*, *dare* (7.1.2).

7.1.1.2. The PIE imperfect indicative (with augment and secondary endings) does not survive in Italic (but see 7.2.1 [item (iv) in the paragraph on “long-vowel perfects”] for a possible exception), where the imperfect of regular verbs has an innovative formation consisting, synchronically, of the present stem vowel followed by a formant *-bā-* (thus Lat. 1 sg. *-bam* [note the secondary active ending **-m*], 2 sg. *-bās*, 3 sg. *-bat*, etc., with regular pre-desinential shortening in some forms). The stem vowels for classes I and II, and partly for IV, correspond to the present stem vowels (thus I *laudābam*, II *monēbam*, IV [mainly OLat.] *audībam*), but III is unexpected, i.e. *agēbam* (likewise III-*io* *capiēbam*, IV [Class. Lat.] *audiēbam*), not †*agibam* < **agēbam* (cf. pres. infin. *agere*). O. **fufans** may be an athematic formation of the same type (‘they were’), but has also been interpreted as a pluperfect, at least historically (Meiser 1998: 197–198). The origin of the formation is disputed, beyond the widely accepted assumption that it involves a univerbation of a nominal form (such as a present participle or the instrumental sg. of a root noun) and the verb **b^huH-* ‘be(come)’, followed by a tense/mood marker “*-ā-*” (< **-eh₂-*), found elsewhere in the Latin verb system (see below on imperfect *erā-* [7.1.2] and the pluperfect [7.2.3]), as well as in other branches (Jasanoff 1978: 121–122, 1983). For a survey of alternative approaches, with bibliography, see Willi (2016: 89 n. 57).

7.1.1.3. The Latin/Faliscan and Sabellic future formations (both innovative, given the absence of a category “future” in PIE) diverge. For the regular Latin future, classes I and II show a formation similar to the imperfect, with a *b*-suffix that again involves a univerbation with **-b^huH-*, here specifically in the form **-b^huH-e/o-* (subjunctive of the

root aorist, cf. 7.1.1.4 and 7.2 on other aorist-based forms), whence thematic-appearing forms of the type I 1 sg. *-ā-bō*, 2 sg. *-ā-bis*, II 3 sg. *-ē-bit*, 3 pl. *-ē-bunt*, etc., < Ital. **-ā-φυ-e* < **-ā-φυυ-e* < (Transponat) **-eh₂-b^huH-e-*, etc. Faliscan forms of this type are attested in **pafo/pipafo** ‘I will drink’ and **carefo** ‘I will lack’. Classes III/III-iō and IV, in contrast, continue the PIE thematic subjunctive, with **-ē-* generalized (III 2 sg. *agēs*, III-iō 1 pl. *capiēmus*, IV 3 pl. *audient*, etc.) and with a substitution of *-am* endings in 1 sg. forms (*agam*, *capiam* etc.). In OLat., however, IV futures in *-ībō* (analogical to the I and II pattern) are not uncommon.

Sabellic shows an *s*-future (cf. the PIE desiderative and the *s*-futures of Greek, Old Irish, and other IE languages, although the Italic futures are based on present stems, not the root itself): thus e.g. O. **fust** ‘it will be’, *deiuast* ‘he will swear’, U. **ferest** ‘he will carry’. The (mainly) OLat. *s(s)*-formation *faxō*, *amassō* is either the subjunctive of such an *s*-formation (Weiss 2011: 419–420) or, according to another view (e.g. de Melo 2007), continues the subjunctive of an *s*-aorist (see Weiss 2011: 420 n. 16 for discussion). Finally, some relics of the PIE desiderative have entered the present system: *quaesō* ‘seek’ (< pre-Lat. **kwais-s-e/o-*, vs. **kwais-e/o-* > *quaerō* ‘ask’), *vīsō* ‘visit’ (< **weid-s-e/o-*); and the particle *dumtaxat* ‘provided that’ is based on a subjunctive form (related to *tangō* ‘touch’) with a similar background.

7.1.1.4. The present subjunctive of regular verbs displays two formants synchronically, *-ē-* for the 1st conjugation and *-ā-* for the rest. As in the imperfect indicative, endings corresponding to the PIE secondary endings are used, hence with 1 sg. *-m* in Latin (and probably likewise in U. *aseriaia* ‘I may observe’, with *m*-loss). Thus for the Lat. verbs in 7.1.1.1: I *amem*, *amēs*, *amet* etc. (with contraction of **amā-ē-* to *amē-*), vs. II *doceam*, *doceās*, *doceat* etc., III *agam*, *agās*, *agat* etc., III-iō *capiam*, *capiās*, *capiat* etc., IV *audiam*, *audiās*, *audiat* etc. The Sabellic behavior is the same (apart from the lack of contraction in **-ā-ē-*; cf. also 7.2.4 on the secondary ending *-ens*): I 3 sg. O. *deiuaid* ‘let him swear’, U. **kuraia** ‘he shall take care of’ (< **-ā-ē-ā-d*, i.e. in Umbrian with recharacterization by analogy with the subjunctives of the other classes); II 3 sg. U. **habia** ‘he shall have’ (: Lat. *habeat*), 3 pl. O. **putians** ‘may they be able’ (as if Lat. *†poteant* to a verb *†potēre*, cf. Late Lat. *potēre* > Ital. *potere* [Väänänen 1981: 136] vs. OLat. *potis sum* > Class. *possum*); III 3 sg. Pael. *dida* ‘may he give’, 3 pl. O. *deicans* ‘they may say’ (: Lat. *dīcant*); III-iō 3 sg. U. **façia**, O. **fakiad** ‘he shall do/make’ (: Lat. *faciat*).

The *ē*-subjunctive has two possible sources, namely the long-vowel thematic subjunctive of **-ā-je/o-* verbs (as in the frequent denominatives of this type), with the *ē*-version generalized; and the strong form of the athematic optative suffix (i.e. **-ieh₁-* > Ital. **-iē-*, with regular *i*-loss in Ital. sequences **-āiē-*), appropriate for original athematic stems in *-ā-*, such as factitives in **-eh₂-* (cf. 7.1.1). (There is no trace in Italic of the PIE thematic optative; for the doubtful interpretation of VOL OPETOIT [Duenos Inscription] in these terms, see Meiser 1998: 201 and Weiss 2011: 417, with references, as well as Tichy 2004 for a more attractive alternative analysis.) The source of the *ā*-subjunctive is more controversial; for a theory based on developments with roots in final **-h₂*, see Meiser (1998: 200–201), and for comparison with the Celtic *ā*-subjunctive, see Weiss (2011: 418, 466 n. 8). A further point connected with the *ā*-subjunctive is its appearance in OLat. forms based on historical aorist or perfect stems, rather than present stems: thus OLat. *ēvenat* (i.e. *-ven-ā-*, cf. root aorist forms of **g^wem-* ‘come’: Ved. *āgan*, Arm. *ekn*,

O. **kúmbened** [7.2.2]) vs. Class. Lat. regular pres. subj. *ēveniat* (based on the present stem of IV *veniō* ‘come’).

7.1.1.5. Latin and Sabellic share an innovative imperfect subjunctive marker **-s(-)ē-* (appended to the present stem, thus surfacing mainly as *-rē-* in Latin owing to rhotacism): e.g. Lat. 2 sg. *amārēs* (cf. pres. infin. *amāre* ‘to love’), O. 3 sg. **fusid** = Lat. *foret* (suppletive imperf. subj. to *sum*, cf. 2 sg. imperf. subj. *essēs*). The origin of the formation is uncertain (see the references at Weiss 2011: 420 n. 20).

7.1.1.6. The oppositional passive forms of regular (and some irregular) verbs, as well as the finite forms of deponent verbs, use a special set of passive endings, descended from the PIE middle endings. As in the primary middle forms of Celtic, Tocharian, and Hittite, the Latin passive endings are characterized by an added *-r* (in all forms except 2 sg. and 2 pl.): thus for *amō* ‘love’, pres. indic. *amor*, *amāris/-e*, *amātur*, *amāmur*, *amāminī*, *amantur*, and similarly for other present system forms: pres. subj. *amer*, *amēris/-e*, *amētur* etc., imperf. *amābar*, *amābāris/-e*, *amābātur* etc., fut. *amābor*, *amāberis/-e*, *amābitur* etc., imperf. subj. *amārer*, *amārēris/-e*, *amārētur* etc.; and likewise for the other conjugation classes, with their characteristic stem vowels. (For the alternate 2 sg. endings *-ris* and *-re*: the latter is more common in OLat.; in Class. Lat., *-ris* prevails in indicative forms, *-re* in the rest and in the indicative of deponent verbs.) The historical analysis of these endings is replete with difficulties, and there are marked divergences with Sabellic; only a few salient points can be selected for comment here. (See Meiser 1998: 218–219 and Weiss 2011: 387–391 for details.) The non-*r* forms (2 sg./pl.): For 2 sg., the PIE primary and secondary endings beginning in **-th₂-* (with traces in Hittite, Indic, and Old Irish) were replaced by an ending **-so*, cf. 2 sg. active *-s* (similar developments elsewhere, cf. Gk. 2 sg. imperf. mid.-pass. ἐφέρου ‘you were carried’ < Pr.-Gk. **e-p^here-ho*, with **-ho* < **-so*); for stems ending in a vowel (as in all regular classes), rhotacism and final vowel weakening produced Lat. *-re*. Early recharacterization by **-s* (before weakening of final **-o* in **-so*) produced an ending **-sos*, whence **-ros* with rhotacism and attested *-rus*: this archaism is found a little over a dozen times, mainly in late Republican and early Imperial inscriptions associated with a lower social dialect (Adams 2007: 445–450). A later recharacterization by **-s* (i.e. *-re* > **-res*) produced the alternate form *-ris* in standard varieties of Latin. 2 pl.: there is disagreement as to whether Lat. *-minī* reflects a complex series of developments with starting point in original **-d^hye* (see Meiser 1998: 219) or, perhaps more likely (Weiss 2011: 391), a wholesale replacement of the original ending, based on nom. pl. forms of the middle participle in **-mh₁no-* (7.3.1.1; e.g. nom. pl. masc. > Pr.-Ital. **-manoi* > Lat. *-minī*), whether in a periphrastic construction with *estis* ‘you (pl.) are’ (as in Greek: e.g. *sequiminī* [estis] ‘you (pl.) are following’ ~ ἐπόμειναι ἔστε) or in some other construction. For a possible 2 pl. mid. *m-* ending in Sabellic, see under 7.1.1.7 below (imperatives).

3 sg./pl. and Lat. vs. Sabellic: Umbrian distinguishes primary 3 sg. *-ter*, 3 pl. *-nter* from secondary 3 pl. *-ntur* (with 3 sg. unattested in this category). In contrast, Oscan shows only 3 sg. *-ter*, 3 pl. *-nter*, while Latin has only 3 sg. *-tur*, 3 pl. *-ntur*. (SPi. **qolofitūr** ‘is erected (?)’ [Vine 2006] could indicate a pattern of the Latin type.) Umbrian may preserve the Proto-Italic situation (although how this system arose from PIE remains unclear), with different generalizations having occurred in Oscan and Latin (and perhaps South Picene). While the Lat. forms and the Umbrian secondary forms could descend

from PIE 3 sg. mid. **-tor*, 3 pl. **-ntor*, the forms with *e*-vocalism reflect a formation of the type **(n)trV* (e.g. 3 sg. **-tro*, 3 pl. **-ntro*), reminiscent of the endings to be reconstructed for Celtic deponent verbs. (See Weiss 2011: 390–391, with reference, and 465 n.3. Synchronically the Umbrian secondary forms, as well as the South Picene form, could reflect **(n)tōr*; but this could be analogical to 1 sg. **-ōr*, see Zair 2014: 378.)

7.1.1.7. The Italic imperative distinguishes a “present imperative” (or “Imperative I”) from a “future imperative” (or “Imperative II”).

The present imperative has only second person forms; in the active, the 2 sg. imperative uses the bare stem (thus for Lat. regular verbs: I *amā*, II *docē*, III *age*, III-*iō cape*, IV *audī*) – the essential comparandum being III *-e* (the bare thematic vowel) = Gk. *-ε*, Ved. *-a* – while the 2 pl. uses the secondary ending *-te* (I *amāte*, II *docēte*, Fal. **salvete**, etc.), cf. Gk. *-τε*, Ved. *-ta*. (Sabellic imperatives of this kind are attested very rarely: e.g. U. *aserio* ‘observe!’, a 2 sg. *ā*-conjugation form.) Irregular verbs (i.e., for the most part, historically athematic verbs [7.1.2]) operate the same way: thus sg. *es*, pl. *este* ‘be!’, sg. *ī*, pl. *īte* (= Pael. *eite*) ‘go!’, etc. (The PIE 2 sg. imperative marker **-d^{hi}* [Gk. *ἴθι*, Ved. *ihī* ‘go!’ etc.] does not survive in Italic as such, apart from reanalyzed relics of the type **g^{wh}h-d^{hi}* [Ved. *jahī*] ‘strike!’ > pre-Lat. **fende* → Lat. *-fendere* in *dēfendere* ‘ward off’, *offendere* ‘strike’; see Rix 2001: 219 s.v. **g^{wh}hen-*, n.4.) For the 2 sg. forms: some simplex III and III-*iō* verbs use an apocopated form in the Classical language (*dīc* ‘say!’, *dūc* ‘lead!’, *fac* ‘do!’; but prefixed *addīce*, *addūce*, etc.), with some full forms (*dīce* etc.) attested in OLat. (see Weiss 2011: 422 n.22 for details). For *fer* ‘carry!’, however, it is not clear whether this is a truncated thematic form or an athematic relic (cf. again 7.1.2). The discourse particle *em* is traditionally interpreted as a truncated thematic 2 sg. imperative, based on the root of *emō* ‘buy’, originally ‘take’; but it may instead be a root aorist injunctive (thus endingless) < **h₁ém* (Meiser 2003: 62), as in some other relics of this kind: e.g. *-do* in *cedo* ‘give [it] here!’ < **déh₃*, OLat. FV ‘be(come)!’ (Carmen Arvale) < **b^húH*.

The 2 sg. passive and deponent pres. imperative uses the 2 sg. middle secondary ending **-so* (cf. 7.1.1.6 on the 2 sg. finite pres. passive/deponent alternate form), thus *amāre* ‘love!’, *sequere* ‘follow!’, etc. The fragments of the archaic Carmen Saliare may preserve a form *orieso* (= Class. Lat. *orīre* ‘arise!’; see Sarullo 2014: 167). The 2 pl. passive/deponent imperative is the same as the finite present form (see 7.1.1.6 on *-minī*).

Unlike the present imperative, the future imperative is very well attested in Sabellic (given the prescriptive nature of some Sabellic texts), and differs in some respects from Latin. The Latin forms (by Classical times mainly restricted to legal formulations and similar contexts) show an ending *-tō*, OLat. *-TOD*, used for both 2 sg. and 3 sg. (cf. Ved. *-tāt*, Gk. [3 sg.] *-τω*), added to the present stem of regular verbs or to the stem of irregular verbs, and the same formation is frequent in Sabellic: I *amātō*, O. *deiuatud* ‘he shall swear’; II Lat. LICETOD ‘it shall be permitted’ = O. *licitud*, Fal. **salvetod** (and **salveto** apparently as 3 pl., see Bakum 2009: 171); III *agitō*, O. *actud* (with syncope), etc.; irreg. *estō* = O. *estud*. Latin has innovated both a 2 pl. in *-tōte* (on the analogy of the finite 2 pl. pres.) and a 3 pl. in *-untō* (on the analogy of the finite 3 pl. pres.): 2 pl. *agitōte*, *estōte* etc., 3 pl. *aguntō*, *suntō* etc. Similarly, Umbrian has innovated a 2/3 pl. form in */-tōtā/*, apparently built on a 2 pl. finite ending */-tā/*, perhaps historically a dual ending (Clackson 2015: 18, with references): **fututo** ‘estote’, **habetutu** ‘habento’. In the Classical language, Latin passives and deponents use endings 2/3 sg. *-tor* and 3 pl. *-ntor*

(with a passivizing use of *-r* to mark *-tō* and *-ntō*); but OLat. uses the active endings, e.g. *ūtītō*, VTVNTO (to deponent *ūtor* ‘use’). Also found in OLat. (and occasionally in Republican Latin) is a deponent 2/3 sg. imperative in *-minō* (*antestāminō* ‘he shall call as witness’, XII Tables), analogically formed on 2 pl. *-minī* after the ending of 2/3 sg. *-tō*. An independent counterpart to this development is found in Sabellic: thus U. **persni-mu** ‘you/he shall pray’ (deponent verb) shows an ending */-mō/*, apparently built to an (unattested) 2 pl. passive/deponent ending characterized by *-m-* (cf. Lat. *-minī*); in Oscan, such imperative forms were secondarily “passivized” via *-r* (O. *censamur* ‘he shall be counted’). Umbrian also developed a plural version, with a suffix */-mā/* (cf. active */-tā/* above): *persnimumo* ‘they shall pray’.

7.1.2. Irregular verbs

These are mainly PIE athematic presents that preserve traces of athematic inflection in at least some parts of the paradigm, while other forms have undergone remodeling (usually based on thematic inflection). Some typical examples:

**h₁es-* ‘be’ (see Rix 2001: 241 s. v.; Ved. 3 sg. *ásti*, 3 pl. *sánti*, etc.)

2 sg. pres. Lat. *es* < **h₁ési*, regularly < ***h₁és-si* (*es[s]* in Plautus < recharacterized **essi*).

3 sg. pres. Lat. *est*, O. **est**, U. **est** < **h₁és-ti*.

3 pl. O. **sent**, U. *sent* < **h₁s-énti*, but Lat. SONT/*sunt*, Fal. **zot** remodeled after thematic 3 pl. forms in *-ont*, Class. *-unt* (cf. 7.1.1.1 on *linquunt* for †*linquent*).

**h₁ej-* ‘go’ (see Rix 2001: 232 s. v.; Ved. 3 sg. *éti*, 3 pl. *yánti*, etc.)

2 sg. pres. Lat. *īs* < **h₁éj-si*.

3 sg. pres. Lat. *it* < **h₁éj-ti*, but the full grade has been generalized to the plural: thus e.g. 1 pl. Lat. *īmus* for expected †*imus* (cf. Ved. *imás[i]*); and 1 sg. and 3 pl. show thematized forms (thus 3 pl. *eunt* as if from **ej-ont[i]*), for expected †*ient*, i.e. /*ient/*, cf. on *sunt*, *linquunt* above).

For many details concerning these verbs and others (principally ‘eat’ [Lat. *edō*], ‘carry’ [Lat. *ferō*], ‘give’ [Lat. *dō*], ‘wish’ [Lat. *volō*], ‘become, be made’ [Lat. *fīō*]), see e.g. Leumann (1977: 521–531), Meiser (1998: 221–224), Weiss (2011: 425–435). In some cases, the historical background is controversial, as with the synchronic “semi-thematic” inflection of Lat. *ferō* ‘carry’ (see Jasanoff 1998 and 2003: 224–227 for assumptions involving original athematic inflection, vs. Meiser 1998: 224 on forms like Lat. 2 sg. pres. *fers*, 3 sg. *fert* as syncopated allegro-forms, cf. Marrucin. 3 sg. pres. **feret**).

In a second type of irregularity, “defective verbs” (Weiss 2011: 435–436) present a restricted inflectional profile. Thus e.g. *aiō* ‘I say’ and its other limited forms (3 sg. pres. *ait*, imperf. *aiēbat* etc.) are restricted to the present system in the Classical language, while e.g. *meminī* ‘I remember’ has only perfect forms (though with present meaning, cf. 7.2.1).

In the imperfect indicative, an isolated formation appears in the stem Lat. *erā-* (1 sg. *eram* ‘I was’, etc.), i.e. **h₁es-* with the bare suffix **-ā-* (PIE **-eh₂-*), without preceding *-b-* (< PIE **-b^huH-*, cf. 7.1.1.3). Likewise, the future of *sum* (1 sg. *erō*, 2 sg. *eris*, 3 sg. *erit*, etc.) does not follow any of the regular future formations (7.1.1.4), but continues

the inherited pres. subjunctive of **h₁es-*, i.e. **h₁és-e-* (3 sg. **h₁és-e-ti*, etc., cf. Ved. *ásati*).

The pres. subjunctive of some irregular verbs is remarkable in preserving traces of the PIE athematic optative; the OLat. paradigm

	SINGULAR	PLURAL
1	<i>siem</i>	<i>sīmus</i>
2	<i>siēs</i>	<i>sītis</i>
3	<i>siet</i>	<i>sient</i>

closely matches that of Gk. opt. εἶην and Ved. opt. *syām*, showing both full grade **-ieh₁-* and zero grade **-ih₁-* of the optative suffix in their original distribution; but the forms are leveled to a stem *sī-* (1 sg. *sim*, 2 sg. *sīs* etc.) in the Classical language, as has also occurred in Umbrian (2 sg. *sir*, 3 sg. **si**, 3 pl. **sis/sins**). Similar forms (but always with *-ī-*, from the zero grade stem) appear for ‘wish’ (*velim* etc.), ‘eat’ (*edim* etc.), and ‘give’ (in the alternate stem-form seen in OLat. subjunct. *duim*), as well as in *faxim* and similar *s*-forms based on forms of the *faxō* type (7.1.1.4).

7.2. Perfect system

The Italic “perfect” system, from the synchronic perspectives of Latino-Faliscan and Sabellic, reflects a merger of the PIE perfect and the PIE aorist. Historically, however, the *disiecta membra* of the Italic perfect make clear that both PIE categories remained distinct at the stage of Proto-Italic. (For a full treatment of the material according to this conception, with a focus on the Latin data, see Meiser 2003.) Despite certain parallel developments, such as an innovative “future perfect” category, the details of the breakdown of the PIE perfect and aorist varied considerably in the two branches of Italic, thereby accounting for the most significant area of divergence between Latino-Faliscan and Sabellic morphology. Here we survey the major features and developments, organized in terms of stem-formation (7.2.1 [Latino-Faliscan], 7.2.2 [Sabellic]), innovative pluperfect and future perfect categories (7.2.3, including perfect and pluperfect subjunctive), and endings (7.2.4). (This section treats the perfect active; for the perfect passive, see 7.3.1.2.)

7.2.1. The PIE primary perfect (with *e*-reduplication, *o*-grade root in the singular and zero grade in the plural) is continued in Latin and Faliscan by reduplicated perfects, which nevertheless vary from the PIE model. Root ablaut is no longer found: either the zero grade is generalized (*pungō* ‘pierce’, root **peug-*: perf. *pupugī*) or the present-stem vocalism is copied (*pariō* ‘give birth, produce’, perf. *peperī*, OFal. **pepara[i]**). There are also a number of formal innovations involving the reduplication: (i) the vocalism of the reduplicating syllable copies the root vocalism of verbs with pres. stem in /i, u, o/, although variation is found (cf. *pupugī* just cited, the normal Classical form, vs. *pepugī* cited by Aulus Gellius); (ii) reduplication is lost in compound forms, whence in some

cases a synchronic dereduplicated simplex (OLat. *tetulī* ‘I carried’, but *rettulī*, *attulī* and Class. simplex *tulī*), and some other perfect stems might have arisen via dereduplication (Meiser 2003: 162); (iii) roots with initial *sT-* clusters show full reduplication but loss of the internal /-s-/ (*scindō* ‘cut’, perf. *scicidī*). Semantically, the resultative meaning of the PIE perfect is retained as an archaism in a few forms (cf. the “defective” verb *meminī* ‘I remember’ < *‘I have called to mind’ [7.1.2]); but by and large, Italic perfects denote simple past actions or (less frequently) a tense value similar to the English “present perfect” (see Weiss 2011: 452–455, with additional details).

The PIE aorist is continued in Italic by perfect stems descended from PIE root aorists and *s*-aorists (cf. also 7.2.4 below on endings derived from the thematic aorist). Forms descended from root aorists typically show full grade: Lat. *līquī* ‘I left’ (*līqu-* < **leik^w-*, cf. Ved. 3 sg. *āraik*, 2 sg. mid. *rikthās* and Gk. [thematized] ἔλιπον). In sigmatic perfects (based on *s*-aorists), the original lengthened grade is sometimes seen (Lat. pres. *vehō* but perf. *vēxī* ‘I carried’, cf. Ved. *āvākṣam*) but more often lost via Osthoff’s Law (Lat. *dīxī* ‘I said’ < pre-Lat. **dejk-s-*, cf. Gk. ἔδειξα) or other secondary patterns; and in some cases a sigmatic perfect has replaced a different PIE preterite formation or reflects an Italic innovation. (See in detail Meiser 2003: 107–146.) Traditionally, some reduplicated perfects have been thought to continue PIE reduplicated aorists (e.g. Lat. *tetigī* ‘I touched’, cf. Hom. ptcple. τεταγών), but this is far from certain (see Meiser 2003: 147–150 contra).

Descriptively, Latin shows a number of primary perfects with long stem vowel; these have a range of sources, beyond the *s*-aorist type just mentioned (Lat. *vēxī*), including the following: (i) regular development of preconsonantal vowel plus laryngeal, e.g. *fēcī* ‘I did, made’ (< *k*-extended root aorist stem **d^heh₁-k-*, cf. Gk. ἔθηκα and, without the *k*-extension, Ved. *ádhatī*) or reduplicated *HeC-* stems with zero grade root, i.e. *He-HC-* (e.g. *ēmī* ‘I bought’ < **h₁e-h₁m-*); (ii) other phonological effects associated with reduplication, e.g. *sēdī* ‘I sat’ < **se-sd-* (pres. *sedeō*); (iii) much older phonological effects, as perhaps in Lat. *vēnī* ‘I came’ (pres. *veniō*), where the original root aorist (cf. 7.1.1.4 on Lat. *ēvenat*) could have developed a long stem vowel in part of its paradigm as a result of Stang’s Law and Szemerényi’s Law (see Weiss 2011: 412 n. 12, with reference); (iv) continuation of the lengthened-grade imperfect stem of some acrostatic (“Narten present”) verbs, as has been suggested for *lēgī* ‘I collected’ and a series of other verbs (see Weiss 2011: 412–413 and Jasanoff 2012). The historical analysis of many such forms, however, is a matter of controversy, along with questions about the possible relationship of some forms to long-vowel preterites in Germanic (e.g. Lat. perf. *scābī* ‘I scraped’ to pres. *scabō*, cf. ON pret. *skóf* to pres. *skafa* ‘id.’), sometimes involving “deep” reduplications with subsequent consonant loss and compensatory lengthening (e.g. *Te-TT-* > *TēT-*), followed in some cases by adjustments of vocalism – see e.g. Meiser (2003: 156) on **ske-skb^h-* > **skēb^h-*, replaced by **skāb^h-*, and more generally Schumacher (2005; and Jasanoff 2012 contra). For detailed treatments of Latin (and some Sabellic) long vowel perfects, see Meiser (2003: 152–158) and Garnier (2010).

The most important morphological innovation of Latin is the so-called “*v*-perfect” or “*u/v*-perfect” (no such forms are attested in Faliscan, perhaps by chance), the regular formation for all vowel-final perfect stems, including (i) the “secondary conjugations” (i.e. denominatives belonging to both the 1st and 4th conjugations, as well as the iterative-causatives and statives belonging to the 2nd conjugation), whence the productive perfect formations I (stem vowel *-ā-*) perf. *-āvī*, II (stem vowels *-ē-* [stative] and *-ē-* [iterative-

causative, i.e. *-e- of *-e(i/e/o)-] perf. -ēvī and -uī (< *-e-u-ai), IV (stem vowel -ī-) perf. -īvī, and (ii) perfects based on laryngeal-final roots, whether those resulting in a long vowel (of the type *ḡneh₃- ‘know’, perf. [g]nōvī) or those with root-final vocalized laryngeal (of the type *sekH- ‘cut’, perf. secuī < *seka-u-ai; see recently Martzloff 2015 on this type). The source of the formation is disputed, but may involve a morphologization of the secondary glide in the pre-Lat. perfect *fūyai (to *b^huH- ‘be[come]’) > OLat. fūī, Class. fuī (Seldeslachts 2001: Ch. I, assuming an old root aorist; similarly Willi 2009, but on the basis of a reduplicated perfect *fuβuai > *fūyai). For a different theory, involving a phonological development associated with laryngeal-final roots (cf. nōvī ‘I know’, root *ḡneh₃-), see Weiss (2011: 411). For further discussion, see Meiser (2003: 220–224).

The absence of *v*-perfect forms in Faliscan has already been noted. The sparse documentation of perfect forms shows convergences with Latin (cf. above on Fal. **pepara**[i], further **keset** ‘gessit’), but also some divergences: Fal. **fifiked/f[if]iqod** ‘fashioned’ (reduplicated) vs. Lat. *finxī* (sigmatic), Fal. **faced/facet** (dereduplicated? cf. VOL **vhevhak-ed**, O. *fefacid*) vs. Class. Lat. *fēcī* (above).

7.2.2. The overall structure of the Sabellic perfect system is similar to that described in 7.2.1 for Latin (and Faliscan), but many of the details differ strikingly. Thus, for primary perfects derived from PIE perfect and aorist stems, some verbs show the same (or nearly the same) development in both branches: e.g. for *prek̑- ‘pray’: Lat. *poposcī* ~ U. **pe-purkurent** (both with reduplicated perfect stem); and for *g^wem- ‘come’: Lat. *vēnī*, though descriptively a long-vowel perfect, may descend from the PIE root aorist (7.2.1), while O. **kūmbened** and U. *benust* (3 sg. fut. perf.) may derive from a short-vowel form within the same root aorist paradigm. But most verbs differ: O. **fifikus** (reduplicated, like Fal. **fifiked/f[if]iqod** above) vs. Lat. *finxī* (sigmatic), O. (ptcple.) *sipus* ‘knowing’ (*sip-* < *sēp-; long-vowel perfect) vs. Lat. *sapuī* ‘I knew’ (*v*-perfect), etc. It is this predominant distinctive patterning that points most clearly to the survival of both (PIE) perfect and aorist into Proto-Italic.

Similarly, for the perfects of secondary conjugations (denominatives, etc.), sometimes referred to as “weak perfects” (vs. the “strong perfects” based on PIE primary perfect and aorist stems): the *v*-perfect (as in Latin) is unattested in Sabellic, which instead displays an elaborate series of innovative perfect formations of its own (largely diverging even within the Sabellic languages). While descriptively suffixal, some may reflect univerbations of periphrastic constructions. Many questions remain about these formations, which cannot be treated here in full. Most important are the following: (i) the “*tt*-perfect” (Oscan and related dialects only, including Pre-Samnite), e.g. O. **prūfatted** ‘he approved’ (cf. Lat. *probāvit*, with *v*-perfect) (see Clackson 2015: 25–26 and especially Willi 2016 for survey and critique of previous approaches, and a new theory based on a periphrastic construction with pres. ptcple. + STAND); (ii) the “*nkī*-perfect” (Umbrian only), e.g. *com-bifiansiust* (and other spellings) ‘he will have communicated’ (see Willi 2010 for survey and critique of previous approaches, and a new theory based on a periphrastic construction with acc. sg. noun + DO, MAKE); (iii) the “*ō*-perfect” (South Picene only as such, but cf. 7.2.3 below on the Sabellic future perfect), e.g. **opsút** ‘he made’ (see Clackson 2015: 19 and especially Zair 2014: 377–382 for survey and critique of previous approaches, and a new theory based on developments involving the inherited *o*-grade perfects of laryngeal-final roots). Other formations are more sparsely attested, such as a

possible “*k*-perfect” in Oscan (e.g. **kellaked** ‘he enclosed [?]’), on which see Willi (2010: 12–13).

7.2.3. Latin shows an innovated pluperfect formation, but there is no evidence for pluperfect forms in Sabellic (or virtually no evidence, cf. 7.1.1.2 on O. **fufans**), given the nature of those texts; and both Latin and Sabellic developed innovative future perfect forms. (Faliscan is omitted here, given the absence of both pluperfect and future perfect forms. This section treats active forms; see 7.3.1.2 for pluperfect and future perfect passive and deponent forms.)

The Latin pluperfect and future perfect show a marker **-is-* (preconsonantal /-is-/ but intervocalic /-er-/ via weakening and rhotacism) following the perfect stem and preceding personal endings, namely the “*ā*-endings” (as in the imperfect of *sum*: *eram*, *erās* etc. [7.1.2]) for the pluperfect (thus 1 sg. *amāv-er-am* ‘I had loved’, 2 sg. *amāv-er-ās* ‘you had loved’, etc., and the thematic endings for the future perfect (1 sg. *amāv-er-ō* ‘I will have loved’, 2 sg. *amāv-er-is* ‘you will have loved’, etc., except 3 pl. *amāv-er-int*, which uses the subjunctive ending). The historical analysis of these formations, however, is very controversial; discussion is reserved for 7.2.4, since the characteristic element **-is-* is also a recurring feature of the perfect endings. Also built with this formant, at least descriptively, is the perfect subjunctive (**-is-* plus **-ī-* [*<* athematic optative **-ih₁-*] plus secondary endings, thus *amāv-er-i-m*, *amāv-er-ī-s*, etc.) and the pluperfect subjunctive (**-is-* plus **-sē-* [suffix of the imperfect subjunctive, 7.1.1.5] plus secondary endings, thus *amāv-is-se-m*, *amāv-is-sē-s*, etc.). But see 7.2.4 for a different conception.

The Sabellic future perfect (which is very well attested) appears as a suffix /-us-/ (preceding personal endings) on the surface (e.g. U. *benust*, *combifiansiust* [7.2.2]), but its background remains controversial. See Clackson (2015: 24 n. 70) and especially Zair (2014, with detailed survey of previous approaches), arguing for historical **-ō-* (identified with the South Picene *ō*-perfect [7.2.2]) plus the future suffix **-s-* (7.1.1.3).

7.2.4. Not surprisingly, the endings of the Italic perfect system have their sources in PIE perfect and aorist endings; these have nevertheless undergone a series of innovations and display considerable variation, especially in third person forms. (In addition to the handbook treatments, see recently Clackson 2015: 30–31.)

A characteristic innovation of Latin and Faliscan consisted in the addition of the *hic et nunc* particle **-i* (as in the primary active endings) to the PIE perfect endings: thus 1 sg. **-h₂e* (or **-h₂a*, with laryngeal coloring) → **-h₂a-i* > OFal. (and pre-Lat.) /-ai/, whence regularly OLat. -EI and Class. Lat. -ī (OFal. **pepara[i]** = Lat. *peperī*, 7.2.1); similarly 2 sg. **-th₂e* (**-th₂a*) → /-tai/, cf. OLat. -TEI and Class. Lat. -tī in -ISTEI, -istī (with additional element *-is-*, see below) and 3 pl. **-ēr* (*<* **-ers*) → **-ēr-i* > Lat. -ēre. (Unique is VOL STETERAI ‘they have set up’, with **-ēr* plus **-ai* from 1/2 sg.) Likewise, 3 sg. **-e* → **-e-i*, but this was renewed by adding the general 3 sg. marker /-t/, whence (with monophthongization, and long scansion retained in Plautus) OLat. -īt. (1 pl. *-imus* and 2 pl. *-istis* were remade after the primary active endings, but *-istis* again shows the characteristic *-is-* element.) Also well-attested inscriptionally, however, is the 3 sg. secondary ending -ED (e.g. Lat. FECED, Fal. **faced** ‘[s]he did, made’), originally from the thematic aorist; this was in turn remade with final /-t/, whence OLat. -ET (cf. Fal. **keset** ‘gessit’) and (with final vowel-weakening) already OLat. -īt. Still more variation is found in the 3 pl.: beside archaic -ēre (also characteristic of Latin poetry and some

prose genres), literary Latin (especially OLat.) attests, relatively rarely, an ending *-ĕrunt* (perhaps also seen in some inscriptional forms of the type *DEDRON*, with syncope or syllabic notation <D> = /dĕ/), which survived into Romance (Väänänen 1981: 141); this form is built with the marker **-is-* plus 3 pl. aorist **-ont*, while the regular Class. Lat. ending *-ĕrunt* combines *-ĕr(e)* and *-ont*. Faliscan diverges further: OFal. 3 pl. perf. **fifiqod** (7.2.1) shows /-ond/, with the original version of the 3 pl. thematic aorist secondary ending. (On the 3 sg. and 3 pl. variation and its significance for Proto-Italic, see Kümmel 2007.) The element **-is-* referred to above (Lat. *-istī, -istis, -ĕrunt*) is presumably the same as in the Latin pluperfect and future perfect, and the perfect and pluperfect subjunctives (7.2.3), as well as the perfect active infinitive in *-is-se* (7.3.2.3). Its source is controversial. See Leumann (1977: 609, with older literature and critique of earlier accounts involving the Vedic *iṣ*-aorist); for an elaborate theory involving a haplogologized version of an original periphrastic expression based on the perfect active participle (which does not survive in Latin as such [7.3.1.2]), see Meiser (1998: 215). Finally for Latin, a special feature connected with the forms of perfect endings is the existence of “short forms” (or “contracted forms”) of the perfect, well attested for both *s*-perfects and *v*-perfects: e.g. 2 sg. *dīxtī* (for *dīxistī*), infin. *dīxe* (for *dīxisse*), pluperf. subj. *dīxem* (for *dīxissem*) etc.; similarly *audīstī, audīsse, audīssem* etc. (for *audīvistī* etc.). These have their sources in phonological reductions (thus *audīstī* and the like with regular glide loss and contraction for /u/ between like vowels), but then induce analogical behavior in other contexts, thus *amāstī* etc. for *amāvistī* etc. (where glide loss should not have occurred). The behavior of these forms in the texts is complex; for many details (including other types of shortened perfects), see Leumann (1977: 598–602) and Weiss (2011: 411–412).

The Sabellic endings show a similar picture, although they differ in detail. Thus O. **manafum** ‘I entrusted’ shows a 1 sg. thematic aorist ending (< **-o-m*), cf. the Latin/Faliscan 3 sg. type *FECED/faced*, itself well-attested in Sabellic (O. **deded**, U. **dede** ‘he gave’, also with occasional replacement of final /-d/ by /-t/ in South Oscan, as in Latin). The 3 pl. ending /-ens/ of Oscan and Umbrian (O. **prūfattens** ‘they approved’) ultimately reflects developments based on athematic secondary **-ent*, while SPI. *ō*-perfects show 3 sg. *-t* (**opsút** [7.2.2]), secondarily based on primary **-ti*, and 3 pl. **-h** (**adstaiúh** ‘they set up’) < secondary *-nd* (see Zair 2014: 378 n. 39, with references). Finally, if Pael. *lexe* ‘you (pl.) have read’ continues a sigmatic perfect /lek-s-e/ (to a verb comparable to Lat. *legō* ‘read’), it may preserve a remarkably archaic 2 pl. ending **-e* comparable to Ved. *-á* (Weiss 2011: 392–393).

7.3. Non-finite forms

7.3.1. Participles and supines

7.3.1.1. As in PIE, the active participle belonging to the present stem inflects with a suffix *-nt-*, generalized as such (apart from a few irregular forms) in Latin (vs. ablaut variants elsewhere, e.g. *-ont-* in Greek for thematic verbs): thus synchronically *-ent-* (< post-consonantal *-nt-*) for III and III-iō/IV conj. verbs (III *agent-*, III-iō *capient-*, IV *audient-*), and *-nt-* following the stem of I conj. verbs (I *amant-*) and II conj. verbs

(II *sedent-* ‘sitting’ = U. **zeřef/serse**); the original I and II stems (**amānt-*, **sedēnt-*) underwent regular Osthoff-shortening of the stem vowel. (See 2.1 on the variable inflection of the abl. sg. form.)

Some irregular forms show a suffix **-ont-*, as in *eunt-* ‘going’ (beside nom. sg. *iēns*; see Weiss 2011: 429 on the complex background of this paradigm), *sōns* ‘guilty’ (the original participle of *sum*, otherwise seen as *-sent-* in prefixed forms: *praesēns*, cf. O. abl. sg. *praesentid*), and the **uel(H)-ont-* ‘wishing’ that underlies *voluntās* ‘desire’.

The PIE aorist active participle, also made with this suffix, is not a part of the Latin verb system as such, but may survive in some lexicalized forms, such as *cliēns* ‘dependent’, *trāns* ‘across’ (Meiser 2003: 46; Vine 2008: 20–21 [*trāns*]).

Similarly, the PIE middle participle in **-mh₁no-* survives only in the 2 sg. passive ending *-minī* (if that is the correct analysis, cf. 7.1.1.6) and in lexicalized relics, e.g. *alumnus* ‘nursling’ (cf. *alō* ‘nourish’), *fēmīna* ‘woman’ (among other forms based on PIE **d^heh₁-[i-]* ‘suck[le]’, e.g. *fēcundus* ‘fertile’, *fētus* ‘offspring’, etc.).

7.3.1.2. The so-called “perfect passive participle” of the Italic verb is based on a PIE verbal adjective with suffix **-tó-*, originally with root in the zero grade, and in PIE probably with neither exclusively passive nor preterital value (see Weiss 2011: 437). Formally, there are many archaic examples: thus *dictus* ‘said’ (< zero grade **dik-tó-*) vis-à-vis the thematic present stem *dīcō*, OLat. DEIC- ‘say’ (< e-grade **dējk-e/o-*), or O. **prúftú** ‘placed’ < pre-O. (i.e. pre-syncope) **pro-fa-to-* (**-fa-to-* < **-d^hh₁-to-*, cf. Lat. *prōditus* ‘put forth’). The original system, however, has undergone considerable adjustment, including patterns such as the following:

- (i) most I conj. verbs attach **-to-* (> Lat. *-tus*, *-a*, *-um*) to the present stem, thus *amāre* (pres. stem *amā-*) → *amātus* ‘loved’, U. **pihaz/pihos** ‘purified’;
- (ii) similarly, some II conj. verbs attach the suffix to the present stem, e.g. *dēlēre* ‘destroy’ → *dēlētus* ‘destroyed’;
- (iii) many II conj. verbs display a format “root + **-eto-*” (based on an analogy that originated in iterative-causative forms with suffix **-éje/o-*), e.g. *tacēre* ‘be silent’ → *tacitus*, U. **tačez** ‘quiet’;
- (iv) the root form of the Italic pres. indic. has often spread to the participle, and the outcome of the suffixation is often disrupted by phonological processes; e.g. for the above pattern: *docēre* ‘teach’ → pre-Lat. **dok-eto-* > (via syncope) *doctus* ‘taught, learned’;
- (v) some II conj. verbs with *s*-perfect indicative have created perf. pass. ptcples. in *-sus*, e.g. *haereō* ‘stick’, perf. *haesī* → *haesus* ‘stuck’;
- (vi) some III conj. verbs with original nasal infix have spread the pres. stem (with nasal) to the participle: e.g. *iungō* ‘join’, but *iūctus* ‘joined’ (for expected †*iuctus*), vs. the archaic distribution preserved in cases like *vincō* ‘conquer’, ptcple. *victus* ‘conquered’;
- (vii) most IV conj. verbs add *-tus* to the pres. stem, thus *audīre* → *audītus* ‘heard’, though many build the ptcple. directly to the root (e.g. *sarcīre* ‘mend’, but ptcple. *sartus*, with regular cluster reduction < **sarktos*); differently U. *sarsite* (adv.) ‘altogether’ (vel sim.), as if Lat. †*sarcītē*;
- (viii) *-itus* has spread to a number of III/III-*iō* verbs with perf. in *-ivī*, e.g. *quaerō* ‘seek, ask’ (perf. *quaesivī*) → *quaesītus* ‘sought’;

- (ix) some verbs display suppletive behavior, e.g. *ferō* ‘carry’ (root **b^her-*) but pteple. *lātus* (root **telh₂-*);

and still other patterns; for many additional details, see Ernout (1953: 220–228), Weiss (2011: 437–443). (For possible traces of the perfect passive participle in Faliscan, see Bakkum 2009: 173. The possible South Picene forms remain unclear in some respects; see e.g. Vine 1998: 21 n. 46 on SPi. **deiktam** ‘?’, with unexpected full grade and possible **-eto-*, with syncope.)

With regard to function: apart from typical (and some special) participial usages, what is important for morphology is that this form is used, together with forms of *sum* ‘be’, in a periphrastic formation that provides the regular perfect passive (including pluperfect and future perfect passive, perfect and pluperfect subjunctive passive), and the perfect-system forms of deponent verbs: i.e. perfects with present *sum* (*amātus sum* ‘I was loved’, *amāta est* ‘she was loved’, etc.), pluperfects with imperfect *eram* (*amātī erāmus* ‘we had been loved’, etc.), future perfects with future *erō* (*amātus eris* ‘you will have been loved’, etc.); similarly for deponents, e.g. *sequor* ‘follow’ → *secūtus sum* ‘I followed’, *secūta erat* ‘she had followed’, etc. (The same usage is documented for Sabellic perfects and future perfects, though the auxiliary stem *fu-* is also used: O. *scriptas set*, U. *screihtor sent* ‘they were written’, U. **pihaz fust** ‘it will have been purified’, etc.; but differently O. *comparascuster* ‘it will have been decided’, with the active fut. perf. form passivized.) Variant forms with the perfect of the auxiliary verb are also found in Latin (*amātus fūi/fueram/fuerō*), but the details of usage and attestation are left aside here.

The PIE perf. act. pteple. in **-uos-/*-ues-/*-us-* (Skt. *-vāms-/-uṣ-*, Myc. */-woh-/*, etc.) does not survive as such. In addition to the possibility of its appearing in the Sabellic future perfect (but see 7.2.3), it has been taken to appear in SPi. **vepses** ‘?’ (if from pre-SPi. ***vep-us-**, with syncope), but the meaning and interpretation of the form remain controversial (see Nishimura [forthcoming], with references); and it may appear in a few lexicalized relics, such as Lat. *apud* ‘among’, *cadāver* ‘corpse’, O. *sipus* and Volsc. *sepu* ‘knowing’, this word perhaps borrowed into Latin as *sībus* ‘callidus sive acutus’ (Festus) (see de Vaan 2008 and Untermann 2000 s.vv.).

7.3.1.3. An innovated future active participle (attested for Latin only) shows a formant *-ūrus* (i.e. *-ūrus*, *-a*, *-um*), normally added to the stem of the perf. pass. pteple. (thus *ductus* ‘having been led’ → *ductūrus* ‘about to lead’), but in some cases formed to the present stem (e.g. *morior* ‘die’ → *moritūrus* ‘about to die’, cf. perf. pass. pteple. *mortuus*). The starting point for the formation may have been *futūrus* ‘about to be’ (the fut. act. pteple. to *sum*), ultimately a *ro*-adjective based on the instr. sg. of an abstract, i.e. pre-Lat. **futū* ‘with futurity’ (Fortson 2007; cf. 6.1 on deinstrumental derivation).

7.3.1.4. The gerundive (or “future passive participle”), attested in both Latin and Sabellic, is an innovative verbal adjective that expresses necessity or obligation. Descriptively, a thematic suffix that appears as *-ndo-* in Latin is attached to the present stem of I conj. and II conj. verbs (*amandus* ‘to be loved’, *docendus* ‘to be taught’, in both cases with Osthoff-shortening), while in III/III-iō and IV conj. verbs the form varies between *-endo-* (*dūcendus* ‘to be led’, *faciendus* ‘to be done’, *audiendus* ‘to be heard’) and *-undo-* < **-ondo-* (*faciendus* ‘to be done’), the latter especially in OLat. and archaizing contexts,

and also appearing in a series of isolated lexicalized forms (e.g. *secundus* ‘second’ < *‘following’ [5.2.1], *rotundus* ‘round’ < *‘rolling’, and others). (The Latin “gerund” – a pure verbal noun used in non-nominative singular case forms – has the same formation.) The Sabellic picture is similar, except that the element corresponding to Lat. *-nd-* appears as *-n(n)-* (O. **úpsannúm** ‘to be done’, U. *pihaner* ‘to be purified’) and thematic verbs show only the *e*-grade form (U. *anferener* ‘lustration to be performed’). There are many theories about the background of the formation, none entirely successful; see the discussion by Weiss (2011: 443–444) and the extensive bibliography cited at Weiss (2011: 444 n. 75) and Meiser (1998: 228).

7.3.1.5. There are two so-called “supine” forms, which in origin are case forms of verbal *tu*-stem nouns. In PIE terms, old *tu*-stems would be expected to show full grade of the root, and there may be some relics of such forms (e.g. *genitum* to *genō/gignō* ‘engender’). Synchronically, however, most of these forms are derived from the past participle stem (for *genō/gignō* cf. [*gnātus* ‘born’, synchronically associated with *nāscor* ‘be born’).

The supine in *-um* (i.e. < acc. sg. **-tum*), attested in Umbrian in addition to Latin, is mainly used to express purpose after verbs of motion (U. **avēf anzeriatu etu** ‘he shall go to observe the birds’), an inherited pattern with good parallels in Indic and Balto-Slavic.

The supine in (Class.) Lat. *-ū* (in Plautus also *-uī*), which is not found outside Latin, appears in phrases introduced by adjectives meaning ‘good’, ‘easy’, ‘useful’ and the like (e.g. *facile factū* ‘an easy thing to do’) and by *fās est* ‘it is proper’ and *opus est* ‘there is a need’. Although the (Class. Lat.) form looks synchronically like an abl. sg., it may well be dat. sg. (cf. 1.4.2 and Plautine *-uī*).

7.3.2. Infinitives

7.3.2.1. The Latin present active infinitive is marked synchronically with a suffix */-se/*, surfacing as such after consonants (e.g. *es-se* ‘to be’) but appearing as *-re* after vowel stems, owing to rhotacism (as in the regular conjugation classes: I *amā-re*, II *docē-re*, III *age-re*/III-*iō cape-re*, IV *audī-re*), and undergoing assimilation in some post-consonantal contexts (e.g. *fer-re* ‘to carry’, *vel-le* ‘to wish’). (Rare apocopated forms are also found: TANGER ‘to touch’, Class. *tangere*.) The suffix was probably abstracted from reanalyzed loc. sg. forms of neuter *s*-stems that seemed to match III conj. stems, of the type (pre-Lat.) loc. sg. **gen-es-i* ‘in birth, in bearing’ (cf. Lat. *genus* [1.3.1, 6]) → thematic infin. **gen-e-si* ‘to bear’, cf. OLat. *genō*, infin. *genere* ‘to bear’. (Archaic instances of the original formant */-si/* may appear in OFal. **menes{e}i** [i.e. /menesi/] ‘to remain’, cf. Gk. μένω for the plain thematic form, vs. Lat. *manēre* [discussion in Dupraz 2006: 329–330; for a different interpretation of the Faliscan form, see Weiss 2011: 446 n. 81], and perhaps in VOL οἰσι ‘to bring’ [see Tichy 2004].) The suffix, once abstracted, then spread to other present classes, and beyond: cf. *fore* ‘about to be’ < pre-Lat. **fu-se* (alternate future active infin. to *esse*, beside *futūrum esse* [7.3.1.3, 7.3.2.5]), and further 7.3.2.3 (perfect active infin.).

The Sabellic ending is **-om* (added to the present stem), e.g. I O. *moltaum* ‘to fine’, II O. **fatium** ‘to speak’, III O. **deikum/deicum** ‘to say’, III-*iō* U. **façiu** ‘to do’, cf. also

O. **ezum**, U. *erom*, **eru** ‘to be’. The appearance of the ending as a thematic acc. sg. desinence is deceptive: more likely (see Fortson 2012: 76 n.5, with references) this reflects an original athematic acc. sg., in the first instance appearing in root nouns (cf. the infinitival use of root-noun accusatives in Vedic). As seen above (1.3.3), the consonant-stem acc. sg. ending was replaced in Sabellic by the thematic ending.

7.3.2.2. The present passive (and deponent) infinitive appears as follows in the Classical language: III and III-iō conj. verbs add *-ī* to the root (thus III *ag-ī* ‘to be led’, III-iō *cap-ī* ‘to be taken’), while the other classes add *-rī* to the present stem (I *amā-rī* ‘to be loved’, II *docē-rī* ‘to be taught’, IV *audī-rī* ‘to be heard’). The III/III-iō pattern probably reflects an original root-noun dat. sg. in **-eġ* (cf. infinitival use of dat. sg. root nouns in Vedic), while the I/II/IV type (*-rī* < **-seġ*) may be a contamination of this with the **-si* of the present active infinitive. (Alternatively, it could reflect an *s*-stem dat. sg. in **(e)s-eġ*, reanalyzed as **-seġ*; cf. 7.3.2.1 on loc. sg. **-es-i* → thematic infin. **-e-si*.)

In OLat., however, III/III-iō forms show *-ier* (III *agier*), the other classes *-rier* (I *amārier*, etc.). According to a proposal by Meiser (2003: 57–58), supported with detailed arguments by Fortson (2012), here it is the *-rier* group that is primary, reflecting an Italic **-dīēr* (later remodeled, within Latin, to **-ziēr*, whence *-rier* with rhotacism, after vowel stems). This form is cognate with the Sabellic present passive infinitive in **-fēr* (O. **sakrafir** ‘to be consecrated’), which in turn shows a secondary addition of passivizing *-r* to a passive infin. in **-fē* (U. *pihafēi* ‘to be propitiated’). This *r*-less form, finally, is ultimately cognate with the Indo-Iranian infinitives in **-d^hiāi* (Ved. *-dhyai*, Av. *-diiāi*) < dat. sg. **-d^hiōi*, while the Italic forms may continue an *e*-grade version **-d^hiēi* (Fortson 2013). For many additional details on the Italic developments, see Fortson (2012: 85–92, 106–107).

7.3.2.3. The perfect active infinitive is formed by adding *-isse* to the perfect stem, consisting of the formant *-is-* (7.2.3, 7.2.4) and the infinitive marker *-se* (7.3.2.1), thus *amāvisse* ‘to have loved’ (1 sg. perf. *amāvī*), *docuisse* ‘to have taught’ (1 sg. perf. *docuī*), etc.

7.3.2.4. The perfect passive (and deponent) infinitive is a periphrastic formation consisting of the perfect passive participle (7.3.1.2) with *esse* (rarely *fuisse*). This form is attested in Sabellic, e.g. U. **kuratu eru** ‘to have been taken care of’ (Lat. *cūrātum esse*).

7.3.2.5. The future active infinitive is a periphrastic formation consisting of the future active participle (7.3.1.3) with *esse* (infectum) or (more rarely) *fuisse* (perfectum), although these auxiliaries are often omitted. In Classical Latin, the participle agrees in number and gender with the subject of the infinitive clause. But especially in OLat., an invariant *-tūrum/-sūrum* is found. Although sometimes considered to be a genuine archaism, it is more likely innovative (see e.g. Ernout 1953: 230; Fortson 2007: 84).

7.3.2.6. The rarely used future passive infinitive is a periphrastic construction with supine in *-um* (7.3.1.5) followed by the (impersonal) present passive infinitive of *eō* ‘go’, i.e. *īrī* (OLat. *īrier*). As shown by a passage in Cato (*Orat.* 176 Malcovati = Cugusi and Sblendorio Cugusi 171), this is an infinitivization of the impersonal passive indicative of *eō* used with the supine in *-um* (Ernout 1953: 232).

8. References

- Adams, J. N.
2003 *Bilingualism and the Latin Language*. Cambridge: Cambridge University Press.
- Adams, J. N.
2007 *The Regional Diversification of Latin 200 BC–AD 600*. Cambridge: Cambridge University Press.
- Adams, J. N.
2013 *Social Variation and the Latin Language*. Cambridge: Cambridge University Press.
- Bakkum, Gabriël Cornelis Leonides Maria
1994 The second declension nominative plural in *-eis, -es, is*, and the first declension nominative plural in *-as*. In: Jozseph Herman (ed.), *Linguistic Studies on Latin: Selected Papers from the 6th International Colloquium on Latin Linguistics*. Amsterdam: Benjamins, 19–39.
- Bakkum, Gabriël Cornelis Leonides Maria
2009 *The Latin Dialect of the Ager Faliscus: 150 Years of Scholarship*. Amsterdam: Amsterdam University Press.
- Baldi, Philip and Pierluigi Cuzzolin
2009 The Latin temporal adverb *semper* and its relatives. *Alessandria* 3: 13–26.
- Baldi, Philip and Pierluigi Cuzzolin
2010 Osservazioni sulla morfologia del comparativo in greco antico. In: Ignazio Putzu, Giulio Paulis, Gianfranco Nieddu and Pierluigi Cuzzolin (eds.), *La morfologia del greco tra tipologia e diacronia*. Milan: FrancoAngeli Edizioni, 46–60.
- Beekes, Robert S. P.
2011 *Comparative Indo-European Linguistics: An Introduction²*. Amsterdam: Benjamins.
- Bock, Bettina
2008 *Die einfach thematischen Präsentien in der dritten Konjugation des Lateinischen*. Graz: Leykam.
- Buck, Carl Darling
1928 *A Grammar of Oscan and Umbrian²*. Boston: Ginn. [Reprint Hildesheim: Olms.]
- Clackson, James
2011 The forms of Latin: Inflectional morphology. In: James Clackson (ed.), *A Companion to the Latin Language*. Malden (MA): Wiley-Blackwell, 105–117.
- Clackson, James
2015 Subgrouping in the Sabellian branch of Indo-European. *Transactions of the Philological Society* 113: 4–37.
- Coleman, Robert
1992 Italic. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: De Gruyter, 389–445.
- Cowgill, Warren
1970 Italic and Celtic superlatives and the dialects of Indo-European. In: George Cardona, Henry M. Hoenigswald, and Alfred Senn (eds.), *Indo-European and Indo-Europeans: Papers Presented at the Third Indo-European Conference at the University of Pennsylvania*. Philadelphia: University of Pennsylvania Press, 113–153. [Reprinted in Jared S. Klein (ed.). 2006. *The Collected Writings of Warren Cowgill*. Ann Arbor: Beech Stave, 191–226.]
- Crespo, Emilio and José Luis García Ramón (eds.)
1997 *Berthold Delbrück y la sintaxis indoeuropea hoy*. Madrid: Ediciones de la UAM; Wiesbaden: Reichert.
- Crawford, M. H.
2011 *Imagines Italicæ: A Corpus of Italic Inscriptions*. London: Institute of Classical Studies, University of London.

- Dieu, Éric
2011 *Le supplétisme dans les formes de gradation en grec ancien et dans les langues indo-européennes*. Geneva: Droz.
- Dieu, Éric
2009 Les formes de gradation vieil-anglaises *sēlra* ‘meilleur’, *sēlest* ‘le meilleur’ et le superlatif latin *sōlistimus* / *sollistimus* ‘très favorable’. *Historische Sprachforschung* 122: 31–38.
- Dunkel, George E.
1997 B. Delbrück and the instrumental-ablative in *-m. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy*. Madrid: Ediciones de la UAM; Wiesbaden: Reichert, 63–83.
- Dunkel, George E.
2014 *Lexikon der indogermanischen Partikeln und Pronominalstämme*. 2 vols. Heidelberg: Winter.
- Dupraz, Emmanuel
2004 Les nominatifs masculins pluriels thématiques en -es du latin républicain. *Revue de philologie* 78: 239–255.
- Dupraz, Emmanuel
2006 Problèmes falisque. *Res Antiquae* 3: 319–336.
- Dupraz, Emmanuel
2012 *Sabellian Demonstratives: Forms and Functions*. Leiden: Brill.
- Ernout, Alfred
1953 *Morphologie historique du latin*. Paris: Klincksieck.
- Ernout, Alfred and François Thomas
1953 *Syntaxe latine²*. Paris: Klincksieck.
- Flobert, Pierre
1975 *Les verbes déponents latins des origines à Charlemagne*. Paris: Belles Lettres.
- Fortson, Benjamin W. IV.
2007 The origin of the Latin future active participle. In: Alan J. Nussbaum (ed.), *Verba Docenti: Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues, and Friends*. Ann Arbor: Beech Stave, 83–95.
- Fortson, Benjamin W. IV.
2010 *Indo-European Language and Culture: An Introduction²*. Malden, MA: Wiley-Blackwell.
- Fortson, Benjamin W. IV.
2012 Latin -rier and its Indo-Iranian congeners. *Indogermanische Forschungen* 117: 75–118.
- Fortson, Benjamin W. IV.
2013 Pre-Italic *-dhjē (*-dhjeh₁) versus Pre-Indo-Iranian *-dhjōi: *Bridging the Gap*. In: Adam I. Cooper, Jeremy Rau, and Michael Weiss (eds.), *Multi Nominis Grammaticus: Studies in Classical and Indo-European Linguistics in Honor of Alan J. Nussbaum on the Occasion of His Sixty-Fifth Birthday*. Ann Arbor: Beech Stave, 50–60.
- Fortson, Benjamin W. IV.
2016 Latin *audēre* and related forms. In: Stephanie W. Jamison, H. Craig Melchert, and Brent Vine (eds.), *Proceedings of the 26th Annual UCLA Indo-European Conference*. Bremen: Hempen, 13–26.
- Fruyt, Michèle
1986 *Problèmes méthodologiques de dérivation à propos des suffixes latins en ...-cus*. Paris: Klincksieck.
- García Castellero, Carlos
2000 *La formación del tema de presente primario osco-umbro*. Vitoria-Gasteiz: Universidad del País Vasco, Servicio Editorial/Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua.

- García Ramón, José Luis
 1997 Adverbios de dirección e instrumental indoeuropeo. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy*. Madrid: Ediciones de la UAM; Wiesbaden: Reichert, 113–141.
- Garnier, Romain
 2010 *Sur le vocalisme du verbe latin: Étude synchronique et diachronique*. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Gerschner, Robert
 2002 *Die Deklination der Nomina bei Plautus*. Heidelberg: Winter.
- Gunkel, Dieter
 2012 Zur Grammatikalizierung expressiver Dehnung in gradmodifizierenden Suffixen: lateinisch *-issimus*, vedisch *-(ī)yas*, attisch *-(ī)yon-*. Paper presented at the Institut für Sprachwissenschaft, Universität Wien.
- Hamp, Eric P.
 1999 *Fidēs* and the Latin 5th declension. *Glotta* 74: 54–56.
- Heidermanns, Frank
 1996 *Sabellische Nominalbildung: Untersuchungen zur nominalen Wortbildung im Oskisch-Umbrischen*. Habilitationsschrift, Universität zu Köln.
- Heidermanns, Frank
 2002 Nominal composition in Sabellic and Proto-Italic. *Transactions of the Philological Society* 100: 185–202.
- Hocquard, Marianne
 1981 *Les verbes d'état en -ē- du latin*. Lille: Atelier de Reproduction de Thèses de Lille; Paris: Champion.
- Jamison, Stephanie W.
 2009 Sociolinguistic remarks on the Indo-Iranian *-ka-suffix: A marker of colloquial register. *Indo-Iranian Journal* 52: 311–329.
- Jasanoff, Jay H.
 1978 *Stative and Middle in Indo-European*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Jasanoff, Jay H.
 1983 The IE. “*ā*-preterite” and related forms. *Indogermanische Forschungen* 88: 54–83.
- Jasanoff, Jay H.
 1998 The thematic conjugation revisited. In: Jay H. Jasanoff, Craig Melchert, and Lisi Oliver (eds.), *Mir Curad: Studies in Honor of Calvert Watkins*. Innsbruck: Institut für Sprachwissenschaft der Universität, 301–316.
- Jasanoff, Jay H.
 2003 *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.
- Katz, Joshua T.
 1998 *Topics in Indo-European Personal Pronouns*. Ph.D. dissertation, Harvard University.
- Kenny, E. J. (ed.)
 1971 *Lucretius, De Rerum Natura* (Book III). Cambridge: Cambridge University Press.
- Klingenschmitt, Gert
 1992 Die lateinische Nominalflexion. In: Oswald Panagl and Thomas Krisch (eds.), *Latein und Indogermanisch: Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg 23.–26. September 1986*. Innsbruck: Institut für Sprachwissenschaft der Universität, 89–135.
- Kümmel, Martin J.
 2007 The third person endings of the Old Latin perfect and the fate of final *-d* in Latin. In: Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe, and Miriam Robbins Dexter (eds.), *Proceedings of the 18th Annual UCLA Indo-European Conference*. Washington, DC: Institute for the Study of Man, 89–101.

- Leumann, Manu
1977 *Lateinische Laut- und Formenlehre*. Munich: Beck.
- Lindner, Thomas and Renato Oniga
2005 Zur Forschungsgeschichte der lateinischen Nominalkomposition. In: Gualtiero Calboli (ed.), *Lingua Latina! Proceedings of the 12th International Colloquium on Latin Linguistics, Bologna 2003*. Rome: Herder, 149–160.
- Martzloff, Vincent
2006 *Les thèmes de présent dans l'épigraphie italique et en latin archaïque*. Ph.D. dissertation, Université Lumière-Lyon II.
- Martzloff, Vincent
2013 Review of Philip Baldi and Pierluigi Cuzzolin (eds.). 2011. *New Perspectives on Historical Latin Syntax*, Vol. 4. *Kratylos* 58: 115–129.
- Martzloff, Vincent
2015 *Domū, secuī, cubū*: Sur les verbes latins en °Cāre formant un parfait en °Cuī. In: Gerd V. M. Haverling (ed.), *Latin Linguistics in the Early 21st Century: Acts of the 16th International Colloquium on Latin Linguistics, Uppsala, June 6–11, 2011*. Uppsala: Acta Universitatis Upsaliensis, 59–72.
- Meiser, Gerhard
1998 *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Meiser, Gerhard
2003 *Veni Vidi Vici: Die Vorgeschichte des lateinischen Perfektsystems*. Munich: Beck.
- de Melo, Wolfgang
2007 *The Early Latin Verb System: Archaic Forms in Plautus, Terence, and Beyond*. Oxford: Oxford University Press.
- Moussy, Claude (ed.)
2005 *La composition et la préverbaton en latin*. Paris: Presses Université Paris-Sorbonne.
- Nishimura, Kanehiro
Forthcoming On syncope of *u*-vocalism in Sabellic: Syllable structure, glide treatment, and other phonological issues. To appear in *Indogermanische Forschungen*.
- Nussbaum, Alan J.
1973 *Ennian Laurentis Terra*. *Harvard Studies in Classical Philology* 77: 207–215.
- Nussbaum, Alan J.
1999 *JOCIDUS: An account of the Latin adjectives in *-idus*. In: Heiner Eichner and Hans Christian Luschützky (eds.), *Compositiones indogermanicae in memoriam Jochem Schindler*. Prague: Enigma, 377–419.
- Nussbaum, Alan J.
ms A nominal contribution: *Spēs* exploration. Paper presented at the 30th East Coast Indo-European Conference. Harvard University, June 7–10, 2011.
- Nyman, Martti
1990 Latin *-īs* 'nom. pl.' as an Indo-European reflex. *Glotta* 68: 216–229.
- Pedersen, Holger
1926 *La cinquième déclinaison latine*. Copenhagen: Höst.
- Penney, John
2002 Notes on some Sabellic demonstratives. *Oxford University Working Papers in Linguistics, Philology and Phonetics* 7: 131–142.
- Penney, John
2009 The Etruscan language and its Italic context. In: Judith Swaddling and Philip Perkins (eds.), *Etruscan by Definition: The Cultural, Regional and Personal Identity of the Etruscans* (Papers in honour of Sybille Haynes, MBE). London: The British Museum, 88–94.

- Pike, Moss
 2011 *Latin -tās and Related Forms*. Ph.D. dissertation, University of California, Los Angeles.
- Piowarczyk, Dariusz
 ms Latin *māteriēs* and the derivational history of the *-iēs/-ia* abstracts. Paper presented at the 14th Fachtagung of the Indogermanische Gesellschaft. Copenhagen, 17–22 September, 2012.
- Poultney, James Wilson
 1959 *The Bronze Tables of Iguvium*. Baltimore: American Philological Association.
- Prósper, Blanca María
 2014–2015 The Indo-European ordinal numerals ‘fourth’ and ‘fifth’ and the reconstruction of the Celtic and Italic numeral systems. *Die Sprache* 51: 1–50.
- Rau, Jeremy
 2014 The history of the Indo-European primary comparative. In: Norbert Oettinger and Thomas Steer (eds.), *Das Nomen im Indogermanischen*. Wiesbaden: Reichert, 327–341.
- Rix, Helmut (ed.)
 2001 *Lexikon der indogermanischen Verben*². Wiesbaden: Reichert.
- Rix, Helmut
 2002 *Sabellische Texte: Die Texte des Oskischen, Umbrischen, und Süd-pikenischen*. Heidelberg: Winter.
- Rix, Helmut
 2004 Ein neuer oskisch-griechischer Göttername in Pompei. In: Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson, and Thomas Olander (eds.), *Per Aspera ad Asteriscos: Studia Indogermanica in honorem Jens Elmegård Rasmussen sexagenarii Idibus Martiis anno MMIV*. Innsbruck: Institut für Sprachwissenschaft der Universität, 491–505.
- Sarullo, Giulia
 2014 *Il Carmen Saliare: Indagini filologiche e riflessioni linguistiche*. Berlin: De Gruyter.
- Schrijver, Peter
 2003 Athematic *i*-presents: The Italic and Celtic evidence. *Incontri Linguistici* 26: 59–80.
- Schrijver, Peter
 2006 Review of Meiser 2003. *Kratylos* 51: 46–64.
- Schumacher, Stefan
 2005 ‘Langvokalische Perfekta’ in indogermanischen Einzelsprachen und ihr grundsprachlicher Hintergrund. In: Gerhard Meiser and Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17–23. September 2000*. Wiesbaden: Reichert, 591–626.
- Seldeslachts, Herman
 2001 *Études de morphologie historique du verbe latin et indo-européen*. Louvain: Peeters.
- Sihler, Andrew L.
 1995 *New Comparative Grammar of Greek and Latin*. New York: Oxford University Press.
- Steinbauer, Dieter
 1989 *Untersuchungen zu den bei Plautus belegten Verben der lateinischen ersten Konjugation. Unter besonderer Berücksichtigung der Denominative*. Unpublished Regensburg dissertation.
- Thesaurus Linguae Latinae*
 1900– Stuttgart, Leipzig et al.: Teubner, De Gruyter.
- Tichy, Eva
 2004 Gr. οἴσομαι, lat. *ūtī* und die Mittelzeile der Duenos-Inschrift. *Glotta* 78: 179–202.
- Tikkanen, Karin
 2011 *A Sabellian Case Grammar*. Heidelberg: Winter.

Untermann, Jürgen

- 1992 Wurzelnomina im Lateinischen. In: Oswald Panagl and Thomas Krisch (eds.), *Latein und Indogermanisch: Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg 23.–26. September 1986*. Innsbruck: Institut für Sprachwissenschaft der Universität, 137–153.

Untermann, Jürgen

- 2000 *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.

de Vaan, Michiel

- 2008 *Etymological Dictionary of Latin and the Other Italic Languages*. Leiden: Brill.

de Vaan, Michiel

- 2011 PIE *i*-presents, *s*-presents, and their reflexes in Latin. *Glotta* 87: 23–36.

de Vaan, Michiel

- 2015 Reduplicated demonstratives in ancient Indo-European languages. *Transactions of the Philological Society* 113: 38–52.

Väänänen, Veiko

- 1981 *Introduction au latin vulgaire*³. Paris: Klincksieck.

Vernet i Pons, Mariona

- 2008 *La segona conjugació verbal llatina: Estudi etimològic i comparatiu sobre l'origen protoindoeuropeu de la formació dels seus temes verbals*. Barcelona: Institut d'Estudis Montjuïc.

Vine, Brent

- 1993 *Studies in Archaic Latin Inscriptions*. Innsbruck: Institut für Sprachwissenschaft der Universität.

Vine, Brent

- 1998 *Aeolic ὄρπετον and Deverbative *-etó- in Greek and Indo-European*. (Innsbrucker Beiträge zur Sprachwissenschaft, Vorträge und Kleinere Schriften 71). Innsbruck: Institut für Sprachwissenschaft der Universität.

Vine, Brent

- 2006 Autour de sud-picénien **qolofitúr**: Étymologie et poétique. In: Georges-Jean Pinault and Daniel Petit (eds.), *La langue poétique indo-européenne. Actes du Colloque de travail de la Société des Études Indo-Européennes (Indogermanische Gesellschaft/Society for Indo-European Studies) Paris, 22–14 octobre 2003*. Leuven: Peeters, 499–515.

Vine, Brent

- 2008 On the etymology of Latin *tranquillus* 'calm'. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 5: 1–24.

Vine, Brent

- 2010 Latin *aliās* 'at another time'. In: B. Richard Page and Aaron D. Rubin (eds.), *Studies in Classical Linguistics in Honor of Philip Baldi*. Leiden: Brill, 123–140.

Vine, Brent

- 2011 Initial **mo-* in Latin and Italic. *Münchener Studien zur Sprachwissenschaft* 65: 261–286.

Vine, Brent

- 2012 PIE mobile accent in Italic: Further evidence. In: Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen, and Jens Elmegård Rasmussen (eds.), *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics*. Copenhagen: Museum Tusulanum, 545–575.

Vine, Brent

- ms Taking the fifth: Latin *famēs* 'hunger' and the 5th declension. Paper presented at the 144th Annual Meeting of the American Philological Association. Seattle, January 3–6, 2013.

Vine, Brent

- Forthcoming Latin *simītū*, *simītūr*, SEIMITVM 'at the same time, together'. To appear in *Glotta*.

- Wachter, Rudolf
 1987 *Altlateinische Inschriften: Sprachliche und epigraphische Untersuchungen zu den Dokumenten bis etwa 150 v. Chr.* Bern: Lang.
- Weiss, Michael
 1998 On some problems of final syllables in South Picene. In: Jay H. Jasanoff, Craig Melchert, and Lisi Oliver (eds.), *Mir Curad: Studies in Honor of Calvert Watkins*. Innsbruck: Institut für Sprachwissenschaft der Universität, 703–715.
- Weiss, Michael
 2011 *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave.
- Willi, Andreas
 2009 To be or not to be: The Latin perfect in *-v-*. *Historische Sprachforschung* 122: 228–247.
- Willi, Andreas
 2010 The Umbrian perfect in *-nç/-nš-*. *Transactions of the Philological Society* 108: 114.
- Willi, Andreas
 2016 The Oscan perfect in *-tt-*. *Transactions of the Philological Society* 114: 75–94.
- Zair, Nicholas
 2014 The future perfect in Oscan and Umbrian, and the *ō*-perfect in South Picene. *Transactions of the Philological Society* 112: 367–385.

Brent Vine, Los Angeles, CA (USA)

49. The syntax of Italic

- | | |
|----------------------|-------------------|
| 1. Introduction | 4. Clausal syntax |
| 2. Nominal syntax | 5. References |
| 3. Verbal categories | |

1. Introduction

By the term “Italic” we mean to designate all those languages, including Latin, which are covered under the “Proto-Italic” node on a tree diagram of Indo-European (IE) languages, as e.g. in Fortson (2010a: 10). This usage differs, for example, from that found in a relatively recent compendium (Ramat and Ramat 1998), which treats Latin separately from “The Italic Languages”. It also excludes Venetic (on which see Wallace, this handbook) and, of course, the non-IE Etruscan. Within Italic we distinguish two primary branches: the Latin-Faliscan group (Latin and Faliscan); and the Sabellic group (Oscan, Umbrian, and South Picene). Latin is extensively documented, but Faliscan, while reasonably well-known, has no diagnostic syntactic structures which are not also found in Latin. Within Sabellic we recognize three subdivisions, namely the Oscan branch, comprising Oscan, Paelignian, Marrucinian, Vestinian, and Hernican; the Umbrian branch, which includes Umbrian, Aequian, Marsian, and Volscian; and the Picene branch, made up of South Picene and pre-Samnite (see Wallace 2007). The literature on the Italic languages is immense, and I make no attempt in this essay to attain bibliographical

comprehensiveness. The best approach is to invite the reader to pursue pertinent literature mentioned in the more recent general sources, such as Rix 1993; Untermann 2000; Caiazza 2006; Wallace 2007; Clackson and Horrocks 2007; Bakum 2009; Tikkanen 2009; and Weiss 2009, 2010).

Because of the fragmentary nature of all the Sabellic languages except Oscan and Umbrian, this chapter will emphasize data from the latter two and Latin, with supplementary material from other languages as needed. This emphasis is not only necessitated by the patterns of usable material for syntactic comparison and analysis, but is further justified by the representative nature of the data provided by these languages for Italic syntax.

1.1. The general typological profile of the Italic languages

As members of the Indo-European family of languages, representatives of the Italic branch display the classic paradigmatic and syntagmatic characteristics of the family as a whole. The Italic languages are of the fusional (synthetic, inflecting) type, which is to say that they express many basic syntactic relations morphologically. Nouns are organized into structural groups called “declensions”, which are characterized by their common stem segment: *o-* stems, *ā-* stems, and so on, as seen in the common stem vowel *-a-* before the inflectional ending *-m* in Lat. *viam*, Osc. **víam**, S. Pic. **viam** (acc.) ‘road’, as opposed to the vowel *-u-/-o-* (< **o*) seen in the accusative forms Lat. *populum* ‘people’, Umbr. **puplum**, *puplom* ‘army’, Osc. **húrtúm** ‘enclosure’, S. Picene **meitimúm** ‘gift’. Nouns are inflected in six or seven cases (nominative, genitive, dative, accusative, ablative, vocative, and locative), with the locative more general in Oscan and Umbrian; two numbers (singular and plural, with vestiges of an ancient dual); and three grammatical genders (masculine, feminine, and neuter). There are five nominal declensions in Latin and an equal number in Sabellic, though the patterns of occurrence and productivity between the two groups differ widely, especially in the 4th and 5th declensions. Adjectives follow the declensional pattern of a few noun classes, and agree with their head noun in gender, number, and case. Both Sabellic and Latin show two adjectival types, one based on the first and second declensions, the other based on the third; adjectives are “compared” in the positive-comparative-superlative degrees. There are complete sets of numerals (ordinals, cardinals, distributives, and so on), and a full set of inherited pronouns, including personal, possessive, demonstrative, interrogative, indefinite, and relative pronouns, all with more or less the same functional distribution in the two groups.

Verbs are assigned to structural groups called “conjugations”, of which there are four in both Latin and Sabellic (with some internal differences of organization). The labeling of a given conjugation is based on a characteristic vowel (*-ā-*, *-ē-* and so on), as in the first conjugation (*-ā-*) verbs Lat. *am-ā-re* ‘to love’, Osc. **dadik-a-tted** ‘he dedicated’. The following categories are expressed inflectionally within the verbal complex: aspect (completed vs. non-completed action), tense (past, present, and future); mood (indicative, subjunctive, and imperative); voice (active, passive, and some residual middle usages); number (singular, plural); and person (1st, 2nd, 3rd). The passive of the completed aspect system is compounded with the perfect passive participle and the verb *be*; otherwise the Italic verb is a complex of bound morphemes in the order root-stem/mood-tense-ending.

Verbs are indexed to their subjects by rules of agreement, by which person and number are reflected in the verbal endings. There are many non-finite verb forms as well, including participles, gerundials, infinitives, and supines, though the distribution of these forms differs between the two groups.

Finally, there are the so-called “invariable words”, that is word-classes whose function is not reflected in paradigmatic variation. Included here are conjunctions of the familiar type (coordinating, disjunctive, etc.), with many derived from Indo-European pronouns; adverbs, largely frozen case-forms; and prepositions, particles, and postpositions, most derived from Indo-European particles. From a typological point of view the pre-/postpositions are of special importance because they have diagnostic value for word-order. While postpositions are found residually in the dominantly prepositional Latin (*tēcum* ‘with you’ instead of *cum tē*), postpositions are much more common in Sabellic (cf. Umbr. *poplu-per* ‘for the people’). The standard word order of the Italic languages is SOV, though every permutation of this order is found in Sabellic and Latin, depending on pragmatic factors, text type and so on. Impersonal constructions are found everywhere in Italic, and subordination of all types (adverbial clauses, relative clauses, noun clauses, and so on) is equally common.

2. Nominal syntax

2.1. Agreement

The Italic languages, like many of their Indo-European relatives, require that every nominal form be in some case, which must be motivated by syntactic or semantic criteria. There is no “root” form of the noun that appears in non-syntactic contexts, as for example *girl* in English. Furthermore, nominal constituents are subject to rules of agreement (or concord) which stipulate that the form of a head word be matched by the corresponding form of its modifiers. In Latin these formal requirements are met by the categories of gender, number, and case. For example, Lat, *rēgī bōnō* ‘to the good king’ (dat.) matches the gender (masc.) number (sing.) and case (dat.) of the head noun *rēgī* and the modifier *bōnō*. Parallel textual examples from Umbrian, Oscan, and Latin establish this feature as Common Italic, as in (1)–(3):

- (1) **rupinie: e: tre: purka: rufra: ute: peia: fetu: ... (Umbrian Um 1, ib, 27)**
 ‘at Rubinia sacrifice three pigs, red or black ...’
purka acc.pl.fem., modified by **rufra** and **peia**
- (2) **[e]ísai. víai. mefiái. terem[|n]iú. staiét. (Oscan Cm 1 = CA, b, 31–32)**
 ‘in the middle of that road stand the boundary markers’
víai loc.sg.fem., modified by **[e]ísai** and **mefiái**
- (3) *Veterem atque antiquom quaestum maiorum meum servo atque obtineo et magna cum cura colo* ‘The ancient and venerable vocation of my ancestors I preserve and maintain and cultivate with great care’ (Plaut. *Persa* 53–54)
quaestum acc.sg.masc., modified by *veterem* and *antiquom*; *maiorum* gen.pl.masc., modified by *meum* (for *meorum*, a usage found elsewhere in Plautus as well); *cura* abl.sg.fem., modified by *magna*

2.2. Heads and modifiers

The nominal phrase consists of a head and a number of modifiers (attributes). The head of the phrase is usually a noun, though pronouns, infinitives, and substantival participles and adjectives can also be heads.

2.2.1. Heads

Latin

- (4) (noun) *vir videt* ‘the man sees’
- (5) (pronoun) *qui videt* ‘(he) who sees’
- (6) (infinitive) *errare est humanum* ‘to err is human’
- (7) (substantival participle) *basiata ridet* ‘the one who has been kissed is laughing’.
- (8) (substantival adjective) *bona occurrunt* ‘good things are happening’

Sabellic

- (9) (noun) *ocar pihos. fust.* ‘the mount will be purified’ (**Umbrian Um 1 vib, 46–47**)
- (10) (pronoun) **ařfertur: pisi: pumpe: fust:** ‘whoever shall be ařfertor’ (**Umbrian Um 1, va, 3–4**)
- (11) (infinitive) no examples available
- (12) (substantival participle) *eso. naratu. uesteis.* ‘let the one who is pouring speak thus’ (**Umbrian UM 1, via, 22**)
- (13) (substantival adjective) *esisco. esoneir.* ‘for these sacred (rights)’ (**Umbrian Um 1, via, 18**)

Heads can be modified by adjectives (the usual case), nouns in the genitive or some other case, pronominal adjectives, prepositional phrases, and certain types of subordinate clauses.

2.2.2. Modifiers (attributes)

Latin

- (14) *vir bonus* ‘the good man’ (noun plus adjective)
- (15) *ille vir* ‘that man’ (noun plus pronominal adjective)
- (16) *talentum auri* ‘a talent of gold’ (noun plus genitive)
- (17) *quam ignavus ac sine animo miles* ‘what a weak and spiritless soldier’ (Cic. *Att.* 1, 18,5) (noun plus prepositional phrase) [this construction is disfavored]

- (18) *rumor regem venire* ‘the rumor that the king is coming’ (noun plus subordinate clause)

Sabellic

- (19) **aasaí. purasiaí.** ‘on the fiery altar’ (noun plus adjective) (**Oscan Osc Sa 1, a 16**)
- (20) **ekak. víam.** ‘this road (acc.)’ (noun plus pronominal adjective) (**Oscan Osc Po 1**)
- (21) **medíkeís. púmpaiianeís serevkiđ.** ‘by the authority of the Pompeian meddix’ (noun plus genitive) (**Oscan Osc Po 1**)
- (22) (noun plus prepositional phrase) no examples found
- (23) **v. aadirans.v. eítivam. paam vereiiaí. púmpaiianaí. trístaa|mentud. deded** ‘the money that V. Adiranis, (son) of Vibius, gave to the community (?) of Pompeii from his will’ (noun plus subordinate clause) (**Oscan Osc Po 3**)

2.3. The cases

Case is a grammatical category used in the analysis of word classes to identify the syntactic relationship between words in a sentence (Crystal 2007: 63). PIE is traditionally reconstructed with eight cases (nominative, genitive, dative, accusative, ablative, vocative, locative, and instrumental); to these is sometimes added the directive (marking direction *to*). We assume seven cases for Common Italic, with the PIE instrumental occurring only vestigially in Sabellic, as is the case with the instrumental and locative in Latin.

2.3.1. Nominative

The primary function of the nominative is to mark the subject of a sentence.

- (24) **aasas. ekask. eestínt. húrtúí** ‘these altars stand in the grove’ (**Oscan Sa 1, b, 1–2**)
- (25) **ma kuprí koram opsút ani{ni}s rakinelíş pomp[úne]í** ‘Annius Rakinelius made (this) statue very well for Pompo’ (**South Picenē SP AQ 2**)
- (26) **setums: míom | face** ‘Septimus made me’ (**Umbrian Um 4**)
- (27) *puer litteras scribit* ‘the boy is writing a letter’ (**Latin**)

2.3.2. Genitive

The genitive is above all else an adnominal case, used to indicate a relation between two nouns. Common relations expressed by the genitive are possession, family relationship, and part-whole relationships; secondary functions are the genitive of time, the

objective and subjective genitive, and some highly specific uses such as the genitive of the penalty, the genitive of the price, the genitive of quality, and others.

- (28) *deiuast. maimas. carneis. senateis. tanginud* ‘he shall take an oath with the agreement of the majority of the senate’ (**Oscan Lu 1 = TB, 3**)
- (29) ... **safinúm: nerf: persukant: ...** ‘(they) call the leaders of the Sabines’ (**South Picene SP TE 6**)
- (30) ... *fetu. popluper. totar. iiouinar. ...* ‘let him do it for the people of the Iguvine state’ (**Umbrian Um 1, vib, 43**)
- (31) *magna pars hominum* ‘a great part of mankind’ (**Latin**)

2.3.3. Dative

The primary function of the dative is to mark the indirect object, or beneficiary, of the verb. Other common uses are the dative of possession (*liber est puellae* ‘the girl has a book’, literally ‘a book is to/at the girl’), a very old construction inherited from PIE, the dative of purpose, the dative of agent, and others.

- (32) *pa. ui. pacuies. medis | uesune. dunom. ded(e)* ‘Pacius Pacuius, (son) of Vibius, the meddix, gave (this) as a gift to Vesona’ (**Marsian VM 3, 1–2**)
- (33) **ma kuprí koram opsút ani{ni}s rakineliš pomp[úne]í** ‘Annius Rakinelius made (this) statue very well for Pompo’ (**South Picene SP AQ 2**)
- (34) **iuve: krapuvi: tre buf: fetu: ...** ‘sacrifice three cows to Jupiter Grabovius’ (**Umbrian Um 1, ia, 3**)
- (35) *magister puellae librum dedit* ‘the teacher gave a book to the girl’ (**Latin**)

2.3.4. Accusative

The original function of the accusative was to mark the goal of the verbal action, which is expressed grammatically by the direct object. This is the primary function of the accusative; others include the expression of the extent of time and space, the cognate accusative, and a few others.

- (36) **este: persklum: aves: anzeriates: enetu:** ‘after the birds have been observed, begin this ceremony’ (**Umbrian Um 1, ia, 1**)
- (37) *izic. eizeic. zicele[i] comono[.] ni. hipid.* ‘this (man) shall not hold an assembly on this day’ (**Oscan Lu 1 = TB 7–8**)
- (38) **postin: viam:** ‘along the road’ (**South Picene SP TE 2**)
- (39) *lego librum* ‘I am reading the book’ (**Latin**)

2.3.5. Ablative

The ablative case has many uses, especially in Latin, where it absorbed the functions of the PIE locative and instrumental. In Sabellic it is used primarily to mark separation and instrument, with many subcategories. The ablative also marks the ablative absolute construction, location and time, and a number of functions inherited from other cases in PIE.

- (40) **este: persklum: aves: anzeriates: enetu:** ‘after the birds have been observed, begin this ceremony’ (**Umbrian Um 1, ia, 1**)
- (41) **eítiuvad múltasíkad[.] kúmbennieís. tangi(nud) aamanaffed** ‘he built this (sundial) from money raised from fines in accordance with a decree of the senate’ (**Oscan Osc Po 4**)
- (42) *curis liberatus* ‘free from cares’ (**Latin**)
- (43) *nullo adversante regnum obtinuit* ‘with no one opposing (him), he secured the throne’ (**Latin**)

2.3.6. Vocative

The vocative case marks nouns used in direct address. Ordinary in Latin, the vocative is relatively rare in the Sabellic corpus. Its use is quite unremarkable.

- (44) **statie silie. salavs s** ‘O Statius Silius, (I hope that) you are secure’ (**Oscan Osc Cm 18**)
- (45) *dei grabouí.* ‘O Jupiter Grabovius’ (**Umbrian Um 1, via, 23**)
- (46) *mihí credite, iudices* ‘believe me, judges’ (**Latin Cic. In Verr. II 4,133**).

2.3.7. Locative

The locative marks stationary location and time-when in Sabellic, while in Latin the locative exists in only a few nouns, and in the names of cities, towns, and small islands, its functions having been taken over by the ablative.

- (47) **[e]ísaí. víaí. mefiaí. teremén[|n]iú. staíet.** (**Oscan Cm 1 = CA, b, 31–32**) ‘in the middle of that road stand the boundary markers’
- (48) *izic. eizeic. zicele[i] comono[.] ní. hipid.* ‘this (man) shall not hold an assembly on this day’ (**Oscan Lu 1 = TB 7–8**)
- (49) **postin: viam: videtas: tetis: tokam: alies: esmen: vepses: vepeten** ‘along the road you see the memorial stele of Titus Allius (who is) buried(?) in this tomb’ (**South Picene, SP TE 2**)

(50) *domī* ‘at home’, *humī* ‘on the ground’, *Romae* ‘at Rome’ (**Latin**)

2.4. Clitics and particles

2.4.1. Clitics

A clitic is a bound form which cannot stand on its own since it is phonologically dependent on a host. Clitics are similar to affixes: like affixes, they do not possess syntactic freedom, though unlike affixes, they can be attached to inflected forms. There are three types of clitics identified for PIE (Fortson 2010a: 162–163), though only two occur in Italic. The first type contains word-level clitics, which are paired with a single word or constituent, usually in an emphatic or contrastive way, and are placed directly after the word or constituent they modify, as in Lat. *hoc quoque maleficium* ‘this evil deed also’ (Cic. *Q. Rosc.* 40). Also of this type originally was the *-que* of Lat. *quisque* ‘each one’, *uterque* ‘each of two’, and the final **-píd** and **-pe** of Osc. **pútúrús-píd**, Umbr. **putres-pe** ‘both’. However, over time these became lexicalized together with their hosts, so that synchronically these univerbations must be considered independent words. Included as well in the word-level group are pronominal clitics such as Lat. *-met*, *-pte*, and *-pe* as in *egomet* ‘I myself’, *eampte* ‘herself’ or *quippe* ‘of course’, and possibly the *-tu* of forms such as Osc. **estu** ‘this one’. The second type in Italic is made up of sentence-connective and phrase-level clitics, which conjoin or disjoin clauses or phrases (which may minimally consist of single words), e.g. Lat. *-que*, *-ve*, as in **SIBEI ET POSTEREISQV SVEIS VIVOS FECIT** (*CIL* I². 1613: lines 3–4) ‘While alive, he made (this tomb/monument) for himself and his descendants’ and *quos Elea domum reducit palma caelestis pugilemum equumum* (Hor. *Carm.* 4.2.18) ‘whom, (whether) horse or boxer, the celestial palm of Elis leads home’, as well as Osc. **ínim (íní)** ‘and’, seen in **Ekass víass íní vía Iúvíia íní Dekkviarím ... uupsens** ‘They built these roads and the road of Jupiter and the Decurialis ...’. Lat. *autem* ‘but, however’, *enim* ‘for, indeed’, and the *ne* of ‘certain indirect interrogative types’ are also sentence-connective clitics, as are Osc. **avt** ‘but’, Umbr. **et** ‘and’, and **ene** ‘then’. Many older IE languages have rich sets of pronominal enclitics (Hittite, Vedic, Greek, Celtic, Avestan) which exist alongside full pronominal forms, but clitics in the Italic languages can be difficult to detect because of the conservative writing practices of those recording classical texts and inscriptions. This is especially true of the weak forms of pronouns. Sentence clitics are governed by Wackernagel’s Law, a general feature of IE languages, which states that unstressed clitic elements appear second in their clause after the first stressed element, as in Lat. *ego enim vocari iussi* ‘I indeed ordered (you) to be called’ (Plaut. *Cas.* 262).

2.4.2. Particles

Closely related to, and often inseparable from clitics, particles are invariable elements which perform a variety of grammatical functions, and appear in many grammatical categories. Among the categories instantiated by particles are prepositions and postposi-

tions, modality markers, certain kinds of adverbs, negators, and others. We can illustrate some particle uses in the Italic languages with the following:

- (51) *cupio ne veniat* ‘I desire that he not come’ (*nē* ‘that not’) (**Latin**)
- (52) **nep. deikum. nep. fatium. púfiad nep. memnim. nep. úlam. sífeí. heriiad (Oscan Osc Cp 36, 8–9; translation Wallace 2007: 60)**
 ‘may he be able neither to speak nor give testimony, nor may he wish (to speak of) his recollection (?) or this (accusation) for his own benefit’ (**nep** ‘that not’)
- (53) *neip. mugatu. nep. arsir. andersistu* ‘let him not make a noise nor let another sit in between’ (*neip, nep* ‘not’) (**Umbrian Um 1 via, 6**)

2.5. Deixis and anaphora

By the term *deixis* we mean those features of language which refer directly to the personal, temporal, or locational characteristics of the situation in which an utterance takes place, and whose meaning is relative to that situation (*now/then, I/you, this/that*; Crystal 2007: 127). *Anaphora*, on the other hand, is used to describe the process or result of a linguistic unit deriving its interpretation from some previously expressed unit or meaning (the antecedent; see Crystal 2007: 24). For example, as a *pro-drop* language, Latin does not require a subject pronoun in unmarked contexts, and a full description of the interaction between the tendency to leave pronouns out and the need to leave them in would far exceed the limits of this chapter. Furthermore, abstract phenomena such as these have not been systematically addressed for the Sabellic languages, though research on Latin is quite advanced.

2.5.1. Deixis

According to Pieroni (2010: 405): “Among Latin demonstratives, only three may be traced back to Proto-Indo-European stems with certainty: *is, sum* [a rare Old Latin demonstrative – PB] and (the inflectional part of) *iste*, the first from the stem **eǵ-/i-*, the second and the third from **so-/to-*. Tripartite deictic systems, which are clearly attested in some languages, for instance Armenian, Greek, and possibly Sanskrit, in addition to Latin, may themselves be considered as the result of independent developments or as indicative of an original structure. As for Latin, the three lexical forms marking the deictic function, attested since the archaic period, are *hic, iste, ille*”. The Sabellic situation is sufficiently controversial that, with the exception of a few well understood forms cited in (57) and (58) below, we will avoid the citation of forms and speculation on their number and organization. It is possible that Sabellic had a three-way system like Lat. *hic, ille* and *iste*, but concrete evidence is difficult to establish.

Latin

- (54) *Dux his verbis epistulam misit* ‘the leader sent a letter in these words’ (*hic*)

- (55) *illud intellego, omnium ora in me conversa esse* ‘this I understand, that the faces of all are turned toward me’ (*ille*)
- (56) *de isto cane hoc scio, pueros mordere* ‘I know this about that dog of yours, that it bites boys’ (*iste*)

Oscan

- (57) **ekas: iúvilas. iúveí. flagiúi stahínt** ‘these memorials are established for Jupiter Flagius’ (**Oscan Osc Cp 25, 2–3**)

Umbrian

- (58) **este: persklum: aves: anzeriates: enetu:** ‘begin this ceremony after the birds have been observed’ (**Umbrian Um 1 ia, 1**)

2.5.2. Anaphora

We will illustrate anaphora in Italic with the relative clause construction (Pompei 2011), which is headed by the relative pronouns (Lat.) *qui*, (Sab.) *po-*, *pa-* and *pi-*, both continuing PIE *k^wi-, *k^wo-. Relative clauses typically follow, but can also precede their “antecedents”, and are found with both subjunctive and indicative mood (mostly indicative in Sabellic). Preposed relatives are archaic retentions from PIE, and are consistent with adjective-noun word order.

Latin

- (59) *semperque istam quam nunc habes aetatulam optinebis* ‘and you will always keep that youthful age which you have now’ (Plaut. *Cas.* 48)
- (60) *quisquis | ama valia | peria quin | osci amare | bis [t]anti pe | ria quisqu | is amare | vota* ‘whoever is in love, may he prosper. May he perish who does not know how to be in love. Twice as much may he perish again who forbids himself to be in love’ (**POMPEII, CIL 4.4091**)

Oscan

- (61) **v. aadirans. v. eítivam. paam vereiiaí. púmpaiianaí. trístaa|mentud. deded. eísak. eítivad ...** ‘which money V. Adiranis, (son) of Vibius, gave to the community(?) of Pompeii from his will, that money ...’ (**Oscan Osc Po 3**)
- (62) **íním. iúk. tríba|rakkiuf. pam. núvlanús tríbarakattuset.** ‘and this building, which the citizens of Nola will have built’ (**Oscan Osc Cm1 = CA b, 11–12**)

2.5.3. Reflexives

The Italic languages continue the PIE reflexive form *se- morphologically, though with different syntax. The PIE reflexive form (itself a late replacement of the middle voice) was used for all persons and numbers, and like other personal pronouns, was indifferent

to gender. Italic languages maintain the genderless and numberless feature, but restrict the reflexive form to the 3rd person. The number of attested reflexive forms in Sabellic is quite small (about four), so it is difficult to say much about the syntax of the reflexive beyond the obvious, namely that it stood in the same clause as its antecedent (subject), as in English (here in [63] in an indirect statement construction, where *sjom* is the accusative subject of the infinitive *deicum*):

- (63) *tanginom. deicans. sjom. dat. eiza(i)sc. idic. tangineis. deicum.* ‘that they pronounce judgment that they themselves are rendering judgment about these things’
(Oscan Lu 1 = TB 9–10)

Latin, on the other hand, evidences not only the standard clause-mated reflexive construction (*se amant* ‘they love themselves’, *bona sibi faciunt* ‘they are doing good things for themselves’), but also has the “long distance reflexive” (traditionally known as the “indirect reflexive”), in which the reflexive pronoun is mated with the subject (topic) of a higher clause, as if in English *John asked Mary to drive himself to the store*.

- (64) *Iccius ... nuntium ad eum mittit nisi subsidium sibi submittatur ...* ‘Iccius sends him a message that unless relief is sent to him (i.e. Iccius) ...’ (Caes. Gall. 2,6,7)

2.6. Adpositions

Comparative evidence from some of the older IE branches, such as Indo-Iranian (Vedic) and Anatolian (Hittite), strongly suggests that PIE was postpositional, that is, that the closed set of particles which combine with nominals to create phrases follow the noun with which they combine (prepositions precede). When, however, these elements combine with verbs, they precede their heads and are referred to as preverbs. Examples are Lat. *ad-moneo*, *com-mitto*, Osc. **aa-manaffed**, Umbr. *aha-uendu*. Except for a few pronominal relics such as *mēcum* and *quoad*, and phrasal combinations such as *quā dē causā*, Latin is prepositional (*dē puellā, in tantīs rēbus*), but this is not the case in Sabellic, especially Umbrian, which is one of the IE groups that points to an original postpositional pattern for PIE (on this issue see Fortson 2010b). Latin prepositions occur with two cases, the accusative and the ablative, while adpositions in Sabellic occur with the accusative, ablative, or locative. A sample of adpositions from IE languages follows, with Latin and Sabellic representatives. Despite the unitary glosses provided, the meanings of these forms vary widely.

- PIE *kom ‘with’: Lat. *cum*, Umbr. *com*, *-com*, **-ku**, **-kum**, Osc. *com*, *con*, OIr. *co-*, Goth. *ga-*
PIE *ad ‘to’: Lat. *ad*, Goth. *at*, Umbr. **-ař**, **-a**, Osc. **az**, OIr. *ad-con-darc*
PIE *uper ‘above’: Ved. *upāri*, Gk. *ὑπέρ*, Lat. *s-uper*, OIr. *for*, Umbr. **s-uper**

The postpositional Umbrian, plus the archaic nature of postpositions in Oscan and Latin, provide significant evidence for a postpositional Proto-Italic. It further suggests verb-final unmarked word-order, which we will discuss below.

3. Verbal categories

The finite verb in Italic encodes, overtly or covertly, the categories of tense, aspect, mood, voice, person, and number. Verbs are organized by stem-class, with four major groupings called conjugations. Latin and Sabellic are essentially the same in the categories they represent, the differences being largely accidental, or matters of detail. For example, there are no examples of the pluperfect attested in Sabellic (an accident, most likely), nor is there a gerund (a detail).

3.1. Agreement

Verbal agreement in Italic is of the standard type for nominative-accusative languages, namely, a finite verb agrees with its (nominative) subject in number and person. In compound verbal forms containing a participle, there is also agreement in gender. Examples of agreement phenomena are basically superfluous, so only one from each language is given here.

- (65) **inuk ukar: pihaz: fust:** ‘then the mound will have been purified’ (**Umbrian, Umbr 1, ib 7**)
pihaz fust periphrastic 3rd sg. fut. perf. pass. indic.; nom. sg. masc. of part. **pihaz** plus 3rd sg. fut. **fust** [agrees with **ukar**]
- (66) **vibieisen: beriiéis: anei: upsatuh: sent: tiianeí:** ‘(These ceramic vessels) were manufactured in the workshop (?) of Vibius Berrius at Teanum’. (**Oscan, Osc Si 5**; translation Wallace 2007: 58–59)
upsatuh sent periphrastic 3rd pl. perf. pass. indic.; nom. pl. masc. of part. **upsatuh** plus 3rd pl. pres. **sent** [agrees with unexpressed pl. subj.]
- (67) **NEI TED ENDO COSMIS VIRCO SIED ...** ‘if the girl is not friendly toward you ...’ (Duenos Inscription; see Baldi 2002: 197–200)
SIED an archaic form of *sit*, 3rd sg. pres. act. subj. [agrees with **VIRCO**]

3.2. Verbal constituent structure

In both branches of Italic the only required member of a predication is a verb, as in Lat. *currit* ‘he runs’, *pluit* ‘it is raining’, Osc. **Karanter** ‘they eat’, Umbr. *purdito fust* ‘the offering shall take place’. Transitive verbs take object complements, either nominal (Lat. *amat patrem* ‘he loves his father’), or with a complementary infinitive (Lat. *vult patrem vidēre* ‘he wishes to see his father’). Similarly in Sabellic: **iuve: krapuvi: tre buf: fetu:** ‘sacrifice three cows to Jupiter Grabovius’ (**Umbrian, Um 1, ia, 3**): **pune: puplum: aferum: heries:** ‘when you will wish to purify the people’ (**Umbrian Um 1, ib, 10**).

In addition, predicates can be expanded by a number of adjunct types. The following schema comes from Pinkster (1990), slightly modified.

Tab. 49.1: Verbal adjunct types

A	(i) <i>Manner</i> (ii) <i>Instrument</i> (iii) <i>Degree</i>	<i>eloquenter</i> ‘eloquently’, <i>summa audacia</i> ‘with the greatest courage’, <i>gladio aliquem necare</i> ‘to kill someone <u>with a sword</u> ’ <i>prosperere</i> ‘with positive outcome’
B i	<i>There is a relation of “co-involvement” between satellite and nuclear predication</i>	
	Beneficiary	<i>ut maioribus natu adsurgatur</i> ‘that one gets up <u>for older persons</u> ’, Cic. <i>inv.</i> 1,48
	Involved party	<i>hic tibi rostra Cato advolat</i> ‘then, would you believe it, Cato ran to the stand’, Cic. <i>Att.</i> 1,14,5 – the so-called “ethical dative”
	Companion	<i>degrediente eo magna prosequentium multitudine</i> ‘as he was leaving <u>with a great throng</u> of followers’, Tac. <i>ann.</i> 13,14,1
ii	<i>Location in time</i>	
	Time Position	<i>in illa tempestate</i> ‘in those times’, <i>feriis Latinis</i> ‘during the <u>Feriae Latinae</u> ’
	Time Duration	<i>diem unum supplicatio fuit</i> ‘for <u>one day</u> there was public thanksgiving’
	Time within which	<i>tribus mensibus villam suam aedificavit</i> ‘ <u>in three months</u> he built his villa’
iii	<i>Location in space</i>	
	Place	<i>terra marique</i> ‘on land and at sea’, <i>in locis idoneis</i> ‘in suitable places’
	Route along which	<i>illo ascensu Haeduos mittit</i> ‘he sent the Haedui <u>along that slope</u> ’
	Place to which	<i>in mensam manum porrigit</i> ‘he stretches out his hand <u>to the table</u> ’
	Place from which	<i>Roma venire</i> ‘to come <u>from Rome</u> ’
iv	Circumstances, conditions, etc.	
	Accompanying circumstances	<i>degrediente eo</i> ‘as he was leaving’ – the so-called ablative absolute; <i>qui potuisset assensu omnium dicere Ennius</i> ‘how could Ennius say <u>with the assent</u> of all?’, Cic. <i>nat. deor.</i> 2,4

Tab. 49.1: (continued)

Cause	<i>aetate</i> in the following example
Motive	<i>ei vel aetate vel curae similitudine patres appellabantur</i> ‘these were called “fathers”, either <u>because of their age</u> , or <u>because their task resembled that of a father</u> ’, Sall. <i>Catil.</i> 6,6
Purpose	<i>esse oportet ut vivas, non vivere ut edas</i> ‘one should eat <u>in order to live</u> , not live <u>in order to eat</u> ’ (<i>Rhet. Her.</i> 4,28,39); <i>admonitum venimus te</i> ‘we come <u>to remind you</u> ’, Cic. <i>de orat.</i> 3,17 – the so-called supine
Result	<i>si quando non pluet, ut terra sitiāt</i> ‘If at some point it does not rain, <u>so that</u> , as a result, the land is <u>dry</u> ’, Cato <i>agr.</i> 151,4

Some Sabellic Examples:

- (68) *uia. auiecta. esonome etuto.* ‘they shall go by the augural way to the sacrifice’ (**Umbrian Um1, vib, 52**) [ablative of the route, probably an old instrumental]
- (69) **sakriss. sa|krafiṛ aṽt ūltiumam ker|ssnaís** ‘to be consecrated with sacrifices, but the last with banquets’ (**Oscan Osc Cp 31, 9–12**) [ablative of means]
- (70) **menzne: kurčlasiu:** ‘in the last(?) month’ (**Umbrian Um 1 iia, 17**) [ablative of time]
- (71) **kupifiatu: rupiname: erus: teṛa:** ‘he shall give the command to Rubinia that one distribute the erus’ (**Umbrian Um 1, ib, 35**) [indirect command]
- (72) **postin: viam: videtas: tetis: tokam: alies:** ‘along the road you see the memorial stele of Titus Allius (**South Picene, SP TE 2**) [accusative with prep. **postin**]

4. Clausal syntax

4.1. Main clauses

Main clauses in the Italic languages occur primarily in the indicative, imperative, and interrogative modalities. The indicative mood is the unmarked mood, and it is commonplace in expressions of fact, both in Latin and Sabellic. Examples are plentiful, and superfluous. Interrogatives have no special mood marking; rather, interrogativity is signalled either by question words such as ‘who’, ‘when’, ‘where’ for “wh-questions”, or by question particles for “yes-no” questions. Imperatives are abundant as well, especially in textual genres which are aimed at giving instructions, as so many of the Sabellic inscriptions are. We illustrate the imperative clause type with the XII Tables for Latin, and some selections from the Iguvine Tables for Sabellic.

4.1.1. Interrogatives

The “yes-no” interrogative type is marked by sentence-level interrogative adverbs/clitics. These particles are found in Latin, but owing to the nature of the texts, are absent from the Sabellic corpus. It is possible that there was some intonational accompaniment to a yes-no question, but what we find in the texts are the following particles (this section relies on Panhuis 2006):

- a. The particle *-ne* is attached to the first word or phrase of the question when neither a positive nor a negative response is anticipated:
 - (73) *Putasne satis esse Romanis haec omnia?* ‘Do you think all this is sufficient for the Romans?’ (Macr. *Sat.* 2,2,2)
- b. *Nōnne* occurs at the beginning of a sentence; it anticipates a positive response:
 - (74) *Nonne emori per uirtutem praestat?* ‘Isn’t it indeed preferable to die with courage?’ (Sall. *Catil.* 20,9)
- c. *Num* occurs at the beginning of the sentence, and anticipates a negative answer:
 - (75) *Num negare audes?* ‘You don’t dare deny it, do you?’ (Cic. *Catil.* 1,8)

The ‘wh-’ type interrogative is amply attested in Latin, scantily in Sabellic. ‘Wh-’ elements can be interrogative pronouns such as Latin *quis* ‘who?’, *quid* ‘what?’, *quantus* ‘how much?’, and so on, along with Osc. **pís**, **píd** ‘who?, what?’.

- (76) **pís: tiú:** ‘who are you?’ (Oscan Sa 31)
- (77) **peṛḱium. púieh súm** ‘whose am I? (I belong to) the Perkii’ (Oscan Cp 41)
- (78) *quid ais tu* ‘what do you say?’ (Plaut. *Aul.* 717)
- (79) *ubinam gentium sumus* ‘where in the world are we?’ (Cic. *Catil.* 1,9)

4.1.2. Imperatives

PIE had both 2nd and 3rd person imperatives, singular and plural. The Italic situation is mixed. In Latin there are two imperative forms, present and future, occurring in both the active and passive. Negative commands are expressed in a variety of ways, canonically with *nōlī* plus the infinitive, but also with *nē* and in various subjunctive constructions. Sabellic has a parallel imperative mood formation, with a 2nd and 3rd person, singular and plural, active and passive (with many gaps). Given the subject matter of much of the Sabellic corpus, it is not surprising that imperatives number in the hundreds. This aligns Sabellic with Early Latin, where imperative forms abound in such texts as the Twelve Tables or the *Lex agraria*, which corresponds to the Oscan *Lex Bantina*. There is also a subjunctive imperative construction in Sabellic, though Oscan and Umbrian split in the details: Umbrian and Oscan both employ the imperative subjunctive construction in positive constructions, but Oscan normally uses the subjunctive in prohibitions.

- (80) *si in ius vocat, <ito>. Ni it, antestamino. Igitur em capito.* ‘if he (the plaintiff) calls him (the defendant) into court, he shall go. If he does not go, the plaintiff shall call a witness. Then let him seize the defendant’ (*XII Tables* I,1)
- (81) *hoc plus ne facito. Rogum ascea ne polito.* ‘one must not do more than this. One must not smooth a funeral pile with an ax’ (*XII Tables* X,2)
- (82) **ene: tra: sahta: kupifiaia:** ‘then let him announce to Trans Sanctam’ (**Umbrian Um 1, ib, 35**)
- (83) *este. persclo. aueis. aseriatier. enetu.* ‘after the birds have been observed, let him begin that sacrifice’ (**Umbrian Um1, vi, a, 1**)
- (84) **nep. abel|lanús. nep. núvlanús. pí dum. tríbarakat{.}tíns.** ‘let neither the Abellani nor the Nolani build anything’ (**Oscan Cm 1 = CA, b, 20–22**)

4.2. Word order

As with fusional languages generally, word order in the Italic languages does not serve a strictly syntactic function, but rather a highlighting or stylistic one. That is to say, the speaker or writer of Latin discourse has the option to package information in certain ways by ordering the words accordingly. The sentence is divided into two parts: the *theme*, which is what the sentence is about (typically the subject of the sentence); and the *rheme*, which is what is said about the theme (typically the predicate of the sentence). Theme and rheme are also known as *topic* and *comment*. In unmarked sentences, the theme occurs at the beginning of the sentence. The rheme is primarily represented by the verb. The place of the verb in the Latin sentence varies widely according to author and text type. There is first a strong tendency for the verb to occur finally in certain authors such as Caesar and Sallust (Panhuis 2006: 194); this is the classical pattern. In poetry it can vary according to the way the author wishes to present his information.

Latin

With these general guidelines in mind, we note the main orders for words in Archaic Latin, bearing in mind that other orders are found as well, even in the same document. This is especially true for Genitive-Noun and Noun-Adjective constructions. In fact, few languages, including Latin, exhibit consistent word order. By the time of the earliest literature, word order in Latin is highly variable, even in non-emphatic contexts, making a neat description problematic. Particularly problematic is the frequent occurrence of discontinuity between the head and its modifier, as in a complex noun phrase such as *quo maiore faciant animo* (Caes. *Gall.* 77, 66, 6). This topic is too complicated for inclusion here. For an illuminating treatment see Bolkestein (2001); further Devine and Stephens (2006); Bauer (2009); Spevak (2010); and Baldi and Cuzzolin (2011):

- (85) *caseus allium olit.* ‘the cheese stinks of garlic’ (Lucil. 14, 481; Warmington 1967, 3: 152) [Subject-Object-Verb]
- (86) *hoce in tabolam ahenam inceideretis* ‘that you engrave this on a bronze tablet’ (S.C. *de Bacch.* 26) [Noun-Adjective]

- (87) *de pr. urbani senatusque sententiad* ‘by authorization of the urban praetor and the senate’ (S.C. *de Bacch.* 21) [Genitive-Noun]
- (88) *sei ques esent quei sabei deicerent* ‘if there were any who say that they...’ (S.C. *de Bacch.* 3–4) [Head noun-Relative]
- (89) *dum ne minus senatoribus C adesent* ‘as long as no less than 100 senators are present’ (S.C. *de Bacch.* 17–18) [Ablative of comparison, without particle; see Cuzzolin 2011]

Sabellic

According to Wallace (2007: 48), the dominant order of words in the typical Sabellic sentence was SOV, as seen in the following excerpts from inscriptions. But as in Latin, word order is highly variable:

- (90) *suae pis prumeddixud. altrei. castrous. auti. eituas. zicolom. dicust* ‘if anyone shall have publicly (lit. before the magistracy) designated the day for another (in a suit) involving the death penalty or money’ (**Oscan Lu 1 = TB 13–14**) [Subject-Object-Verb]
- (91) **ahal trutitis dunum dede** ‘A. Trutitis gave (me) as a gift’ (**Umbrian Um 16**) [Subject-Object-Verb]
- (92) **... saffínúm: nerf: persukant: ...** ‘(they) call the leaders of the Sabines’ (**South Picene SP TE 6**) [Subject-Object-Verb]
- (93) **lígatúís. abellaj[úís] íním. lígatúís. Núvlanúís** ‘to the Abellian envoys and the Nolanian envoys’ (**Oscan Cm 1 = CA, a, 6–7**) [Noun-Adjective]
- (94) *eine angluto somo.* ‘and from the highest angle’ (**Umbrian Um 1, vi a, 10**) [Noun-Adjective]
- (95) **herekleís. físnú.** ‘temple of Herakles’ (**Oscan Cm 1 = CA, b, 4**) [Genitive-Noun]
- (96) *cubrar. matrer. bio. eso* ‘this shrine of The Good Mother’ (**Umbrian Um 7**) [Genitive-Noun]
- (97) **ařfertur: pisi: pumpe: fust: eikvasese: atiiērier:** ‘the *ařfertor*, whoever will be in the Atiedian brotherhood’ (**Umbrian Um1, va, 3–4**) [Head noun-Relative]
- (98) **íním. íúk. tríba|rakkiuf. pam. núvlanús tríbarakattuset.** ‘and this building which the Nolani will have built’ (**Oscan Osc Cm 1 = CA b, 11–12**) [Head noun-Relative]
- (99) *mais. zicolois. X. nesimois.* ‘more than the ten following days’ (**Oscan Osc Lu 1 = TB 25**) [Comparative construction; ablative of comparison, without particle]

4.3. Pro-drop

Since the Italic languages encode the subject in the verbal endings, it is not grammatically necessary to mention pronominal subjects except for contrast, topicalization, or as an

anaphoric marker. In this regard Italic languages differ from modern languages like English or French, which require subject pronouns whatever their pragmatic status in the sentence. Contrast the Latin examples (100a) and (100b), and the Sabellic (101a) and (101b):

(100a) *filios amo* ‘I love my sons’

(100b) *ego scribo sed tu legis* ‘I am writing but you are reading’

(101a) *sacre. stahu* ‘I stand inviolable’ (**Umbrian Um 10, 7**)

(101b) **svai: neip: avt svai tiium: idik fifikus pust** ‘if not, or if you will decree/make/fashion it after’ (**Oscan Osc Cp 37, 5**)

4.4. Negation

Simple negation (i.e. negation which does not involve quantifiers, indefinites, and the like) is a fairly straightforward process in Italic. Negative particles such as Lat. *nōn, nē, neque*, Umbr. *neip*, Osc. *ne, nei, neip* combine with items to be negated (verbs for the most part), usually preceding the item over which they have scope:

(102) *non eo* ‘I do not go’ (**Latin**)

(103) *pusei. neip. heritu.* ‘that it not be as intended’ (**Umbrian Um 1 vi, a, 27**)

(104) *nei. suae. q(uaestor). fust. nep. censtur. fuid. nei. suae. pr(aetur). fust.* ‘unless he will be the quaestor, unless he might have been the censor, unless he will be the praetor’ (**Oscan Lu 1 = TB, 28**).

4.5. Movement rules

As noted earlier, the Italic languages have a flexible word order, in which certain elements can be highlighted by putting them in a non-canonical pattern. In Latin and Sabellic, adjectives typically follow their head noun, but they can be moved to the left of the noun for emphasis. Thus Umbr. *ocriper Fisiu, totaper Iiouina* ‘pro monte Fisiu, pro civitate Iguvina’ [for the Fisian mount, for the Iguvine state], but *destruco persi, nertruco persi* ‘ad dextrum (sinistrum) pedem’ [at the right foot, at the left foot] (Buck [1904] 1928: 224). Question words (*pís, píđ*, etc.) assume the first position in their clause, and any word in a sentence may be dislocated to topic (theme) position for emphasis. However, questions are extremely rare in the Sabellic corpus. A few Latin examples follow:

(105a) *quis venit?* ‘who is coming?’

(105b) *quem vides?* ‘whom do you see?’

Relative clauses may precede their head noun for emphasis as well, which is a reversion to the archaic pattern inherited from PIE.

4.6. Coordination

Italic languages follow the PIE pattern of coordinating both phrasal and clausal elements. In Latin the usual coordinator is *et*, though postpositional *-que* is equally common, usually for conjoining like members of a constituent like *terrā marīque* ‘by land and by sea’. Postpositional *-que* dominated in the archaic period, since it was inherited as a conjunction from PIE, whereas *et* was an old adverbial particle. Disjunction is expressed by *sed* ‘but’, as in the phrase *nōn solum, sed etiam* ‘not only, but also’. Choice was expressed by *aut...aut* ‘either...or’, as well as by *vel ... vel*, with subtle contrast. In Oscan coordination of phrases was expressed by **inim** ‘and’, while disjunction was marked by **avt** ‘but’. Choices were marked by either **avt/aut(i)** or *loufir*. For Umbrian the coordinating function was indicated by **et** and **ene**, choice by **heri(s) ... heri(s)**, disjunction by **avt**, and alternatives by **ute**, *ote*. The very old type of coordination via asyndeton seen in the common quotation *vēnī, vīdī, vīcī* attributed to Caesar by Suetonius (*Iul.* 37,4) is found extensively in Oscan and especially in Umbrian: **arvia ustentu: vatuva ferine: feitu: heris: vinu: heri puni: ... feitu: ... kutef: pesnimu ařepes arves**: ‘present offerings of grain, place the ribs on a tray, sacrifice ... with wine or mead, pray silently with (offerings of) grain and fat’ (**Umbrian Um 1, i a, 3–6**)

4.7. Subordination

Subordinate clauses are sentences contained within other sentences. There are many ways to classify subordinate clauses in Italic languages. Limitations of space make it impossible to do more than sketch some basic types. The following treatment tries to bring out the major distinctions, in the case of Sabellic following in the main Wallace (2007: 42–46).

1. Is the verb of the subordinate clause finite or non-finite?
2. If finite, is the verb of the subordinate clause indicative or subjunctive?

The first distinction, finite or non-finite, allows us to separate nominal clauses from adverbial and relative clauses.

4.7.1. Nominal clauses

Nominal clauses are subordinate clauses which function as the complement of verbs. Sentences of the type *Mary knows that Marica stole the car*, in which the clause *Marica stole the car* is the object of *knows*, have a special character in the Italic languages. This is the *Accusativus cum Infinitivo* (AcI) construction, also known as indirect statement, or indirect speech. This construction consists of an accusative plus infinitive in both

Latin and Sabellic, which we illustrate here with the English *I consider [him to be my friend]*. In Italic languages all nominal clauses expressing indirect statement require the AcI construction except indirect questions like *I ask you how much this costs*, or indirect commands of the type *I demand that you arrest Marica*, both of which take finite subjunctive clauses. From a syntactic point of view, the AcI is the “signature” construction of Italic syntax.

- (106) *deiuatud. sipus. comenei. perum. dolom. mallom. siom. ioc. comono. mais. egm[as touti] cas. amnud. pan. pieisum. brateis. auti. cadeis. amnud. ... pertumum*. ‘let him swear knowingly in the assembly without criminal intent that he prevented this assembly more on account of public business than on account of favor or enmity’. (**Oscan Osc Lu 1 = TB 5–7**; trans. Wallace 2007: 46).
- (107) *is ubi malam rem scit se meruisse...* ‘when he knew that he had deserved a bad thing ...’ (Plaut. *Cas.* 160)

4.7.2. Adverbial clauses

Adverbial clauses are subordinate clauses which are typically introduced by a subordinating construction; the verb of the subordinate clause is in the subjunctive. Included in this category are clauses indicating time, purpose, result, cause, indirect questions, indirect commands, and conditions. As in Latin, the Sabellic adverbial clauses of time are often in the indicative mood, while the rest tend to be subjunctive. This suggests a different underlying syntax for the temporal clauses. We illustrate this category with a result clause, an indirect command, and an indirect question.

4.7.3. Result clauses

Latin result clauses are introduced by *ut*, while those in Sabellic are introduced by (Umbrian) *pusi* ‘so that’. The subjunctive is the subordinate mood in both.

- (108) *eo. iso. ostendu. pusi. pir. pureto. cehefi. dia*. ‘let him present them in this way so that it is possible for fire to be taken from fire’ (**Umbrian Um 1, via, 20**; trans. Wallace 2007: 43).
- (109) *eloquere utrumque ... ut nos sciamus* ‘tell us two things ... so that we might understand’ (Plaut. *Cas.* 56–57).

4.7.4. Indirect questions

- (110) **ehvelklu: feia: ... panta: muta: arferture: si:** ‘he shall take a vote as to how great a fine should be for the priest’ (**Umbrian, Um 1, vb 1–3**)
feia 3rd sg. pres. act. subj.; **si** 3rd sg. pres. act. subj.

- (111) *haec sat scio quam me habeat male* ‘I know well enough how badly she has it in for me’ (Plaut. *Most.* 708)
habeat 3rd sg. pres. act. subj.

4.7.5. Indirect commands

- (112) **kupifiatu: rupiname: erus: teřa** ‘he shall give the command to Rubinia that one distribute the *erus*’ (**Umbrian Um 1, ib, 35**)
teřa 3rd sg. pres. act. subj.
- (113) *orat ut suadeam Philolacheti* ‘he begs me to persuade Philolaches’ (Plaut. *Most.* 798)
suadeam 1st sg. pres. act. subj.

4.7.6. Relative clauses

The final subordination type is the relative clause. Relative clauses are adjectival in that they modify nouns, and can precede (the older order) or follow the noun which they modify. Verbs are typically in the indicative.

- (114) **íním. íúk. triba|rakkiuf. pam. núvlanús tríbarakattuset. íním úittiuf. núvlnúm. estud** ‘and this building which the Nolani will have built and let it be of use to the Nolani’ (**Oscan, Osc CM 1=Ca b, 11–13**)
- (115) *quem agrum eos uendere heredemque sequi licet, is ager uectigal nei siet* ‘the field which (lit. which field) they are allowed to sell and pass on to an heir, that field may not be taxable’ (Sent. *Minuc. CIL I² 584, 5*); example from Fortson (2010a: 164).

4.8. Impersonal constructions

Impersonals are those 3rd pers. sg. constructions in which the verb is construed without person contrast, that is without a subject phrase. Such constructions are of several types. The first is equivalent to English constructions such as ‘it is snowing’ or ‘it is raining’: Lat. *ninguit, pluit*.

A second group comprises verbs of personal evaluation, such as Lat. *piget* ‘it disgusts’, *miseret* ‘it grieves’; Sabellic examples are Umbr. **herter** ‘it is desirable, desired’ (= Lat. *oportet*), Osc. *loufir* ‘or’ (lit. ‘it pleases’, cf. Lat. *libet*).

The final type is the most widespread, expressing unspecified agency in such expressions as Eng. ‘one says’, Fr. *on dit*, Ital. *si dice*, and so on. In both Latin and Sabellic such expressions are typically in the 3rd sg pass., even with intransitive verbs such as ‘come’ and ‘go’. We find Umbr. *ferar* ‘one carries’, Osc. **sakarater** ‘a sacrifice is made’, Umbr. **herter** ‘it is desired’; Lat. *itur* ‘one goes’, *perventum est* ‘there was an arrival’, and so on.

4.9. Sequence of tenses

Italic languages have a fairly complex system of rules governing the tense relations between main clauses and dependent clauses. Latin with its somewhat more elaborate tense inventory is more complex than Sabellic, which attests relatively few examples. Simply put, **primary tenses** (i.e. those expressing non-past action) in the main clause are followed by the present or perfect subjunctive in the dependent clause. The secondary sequence requires that a **secondary tense** (i.e. one expressing past action) in the main clause be followed by an imperfect or pluperfect subjunctive in the dependent clause. In Sabellic only a few imperfect subjunctives are found, and there is no pluperfect indicative or subjunctive.

4.10. Primary tenses

Latin

- (116a) *Marcus scribit ut Caesarem moneat* ‘Marcus writes so that he might warn Caesar’
scribit 3rd sg. pres. act. indic.; *moneat* 3rd sg. pres. act. subj.
- (116b) *Marcus scribet ut Caesarem moneat* ‘Marcus will write so that he might warn Caesar’
scribet 3rd sg. fut. act. indic.; *moneat* 3rd sg. pres. act. subj.
- (116c) *Marcus rogat quid fecerimus* ‘Marcus is asking what we were doing’ (present-perfect)
rogat 3rd sg. pres. act. indic.; *fecerimus* 1st pl. perf. act. subj.

Sabellic

- (117) **kupifiatu: rupiname: erus: teřa:** ‘he shall give the command to Rubinia that one distribute the *erus*’ (**Umbrian Um 1, ib, 35**)
kupifiatu 3rd sg. pres. act. imper.; **teřa** 3rd sg. pres. act. subj.
- (118) **ehvelklu: feia: ... panta: muta: ařferture: si:** ‘he shall take a vote as to how much of a fine there should be for the priest’ (indirect question) (**Umbrian Um 1, vb, 1–3**)
feia 3rd sg. pres. act. subj.; **si** 3rd sg. pres. act. subj.

4.11. Secondary tenses

Latin

- (119a) *Marcus scripsit ut Caesarem moneret* ‘Marcus wrote so that he might warn Caesar’
sripsit 3rd sg. perf. act. indic.; *moneret* 3rd sg. imperf. act. subj.
- (119b) *Marcus scribebat ut Caesarem moneret* ‘Marcus was writing so that he might warn Caesar’
scribebat 3rd sg. imperf. act. indic.; *moneret* 3rd sg. imperf. act. subj.

- (119c) *Marcus rogavit quid fecissimus* ‘Marcus asked what we had done’
rogavit 3rd sg. perf. act. indic.; *fecissimus* 1st pl. pluperf. act. subj.

Sabellic

- (120) **ekss. kúmbened. ... puz. ídík. sakara[klúm] íním. ídík. terúm. múíník[úm] múíníkeí. tereí. fusíd.** ‘thus it was agreed ... that this sanctuary and this property should be held in common on common property’ (**Oscan Cm 1 = CA a, 10, 17–18**)
kúmbened 3rd sg. perf. act.; **fusíd** 3rd sg. imperf. act. subj.
- (121) **kúmbened. ... puz. ... the|saurúm. ... pún. patensíns. múíníkad. ta[n]ínúd. patensíns.** ‘it was agreed that, when they open the treasury, they should open it by common agreement’ (**Oscan Cm 1 = CA a, 10, b 23–25**)
kúmbened 3rd sg. perf. act.; **patensíns** 3rd pl. imperf. act. subj.

Acknowledgment

I am grateful to Brent Vine, Paul Harvey, and Pierluigi Cuzzolin, whose comments on an earlier draft of this chapter resulted in many improvements. Needless to say, they are blameless for any errors of fact, interpretation, or presentation which may remain.

5. References

- Bakkum, Gabriel Cornelis Leonides Maria
 2009 *The Latin dialect of the Ager Faliscus: 150 years of scholarship*. 2 vols. Ph.D. dissertation, Amsterdam University. Amsterdam: University Press.
- Baldi, Philip
 2002 *The foundations of Latin*. Berlin: Mouton de Gruyter.
- Baldi, Philip and Pierluigi Cuzzolin (eds.)
 2009–2011 *New perspectives on historical Latin syntax*. Volume 1 (2009): *Syntax of the sentence*, Volume 2 (2010): *Constituent syntax, Adverbial phrases, adverbs, mood, tense*, Volume 3 (2010): *Constituent syntax: quantification, numerals, possession, anaphora*, Volume 4 (2011): *Complex sentences, grammaticalization, typology*. Berlin: Mouton de Gruyter.
- Baldi, Philip and Pierluigi Cuzzolin
 2011 Syntactic Change in the History of Latin: Do New Perspectives Lead to New Results? In: Baldi and Cuzzolin (eds.), Vol. 4, 865–893.
- Bauer, Brigitte
 2009 Word order. In: Baldi and Cuzzolin (eds.), Vol. 1, 241–316.
- Bolkestein, Machtelt
 2001 Random scrambling? Constraints on discontinuity in Latin noun phrases. In: Claude Moussy (ed.), *De lingua latina novae quaestiones. Actes du X^{ème} colloque internationale de linguistique latine. Paris-Sèvres, 19–23 avril, 1999*. Leuven: Peeters, 245–258.

- Buck, Carl D.
1928 [1904] *A grammar of Oscan and Umbrian*. 2nd edn. Boston: Ginn.
- Caiazza, Domenico (ed.)
2006 *Studi in onore di Aldo Prosdocimi per il premio "I Samniti"*. 2 vols. Piedimonte Matese: Arti Grafiche Grillo.
- Clackson, James and Geoffrey Horrocks
2007 *The Blackwell history of the Latin language*. Oxford: Blackwell.
- Crystal, David
2007 *A dictionary of linguistics and phonetics*. 5th edn. Oxford: Blackwell.
- Cuzzolin, Pierluigi
2011 Comparative and superlative. In: Baldi and Cuzzolin (eds.), Vol. 4, 549–659.
- Devine, Andrew M. and Laurence D. Stephens
2006 *Latin word order: structured meaning and information*. New York: Oxford University Press.
- Fortson, Benjamin W. IV.
2010a *Indo-European language and culture: an introduction*. 2nd edn. Oxford: Blackwell.
- Fortson, Benjamin W. IV.
2010b Reconsidering the history of Latin and Sabellic adpositional morphosyntax. *American Journal of Philology* 131: 121–154.
- Panhuis, Dirk
2006 *Latin grammar*. Ann Arbor: University of Michigan Press.
- Pieroni, Silvia
2010 Deixis and anaphora. In: Baldi and Cuzzolin (eds.), Vol. 3, 389–501.
- Pinkster, Harm
1990 *Latin syntax and semantics*. [Revised and updated English translation of 1984 Dutch original]. Amsterdam: Grüner. (web version at <http://perseus.uchicago.edu/cgi-bin/philologic/navigate.pl?NewPerseusMonographs.19>) [Last access May 30, 2017].
- Pompei, Anna
2011 Relative clauses. In: Baldi and Cuzzolin (eds.), Vol. 4, 427–547.
- Ramat, Anna Giacalone and Paolo Ramat (eds.)
1998 *The Indo-European Languages*. London: Routledge.
- Rix, Helmut (ed.)
1993 *Oskisch-Umbrisch. Texte und Grammatik. Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia vom 25. bis 28. September 1991 in Freiburg*. Wiesbaden: Reichert.
- Rix, Helmut
2002 *Sabellische Texte. Die Texte des Oskischen, Umbrischen und Süd-pikenischen*. Heidelberg: Winter.
- Spevak, Olga
2010 *Constituent Order in Classical Latin Prose [SLCS 117]*. Amsterdam: Benjamins.
- Tikkanen, Karin
2009 *A comparative grammar of Latin and the Sabellic languages. The system of case syntax*. Unpublished Ph.D. dissertation, Uppsala University.
- Untermann, Jürgen
2000 *Wörterbuch des Oskisch-Umbrischen*. (Handbuch der italischen Dialekte 3). Heidelberg: Winter.
- Wallace, Rex
2007 *The Sabellic languages of ancient Italy*. Munich: LINCOM.EUROPA.
- Warmington, E. H.
1967 *Remains of Old Latin, III*. (Loeb Classical Library 329). Cambridge, MA: Harvard University Press.

Weiss, Michael

2009 *Outline of the historical and comparative grammar of Latin*. Ann Arbor: Beech Stave.

Weiss, Michael

2010 *Language and Ritual in Sabellic Italy. The ritual complex of the third and fourth Tabulae Iguvinae*. Leiden: Brill.

Philip Baldi, Boalsburg, PA (USA)

50. The lexicon of Italic

- | | |
|--|-------------------|
| 1. Preliminary remarks | 4. Word-formation |
| 2. The structure of the Italic lexicon | 5. References |
| 3. Loanwords | |

1. Preliminary remarks

This article deals with the lexicon of the Italic languages and is divided into the following main sections: 1) the structure of the Italic lexicon with regard to its Indo-European patrimony and Italic peculiarities, 2) loanwords, and 3) word-formation. It goes without saying that, due to the limited space, an exhaustive treatment is impossible and only tendencies can be delineated and illustrated with a few examples, but with ample bibliographical references. The term *Italic* is employed here in its genealogical value to encompass Latino-Faliscan (L.-F.) as one sub-branch and Sabellic – i.e. Oscan-Umbrian (O.-U.) and minor dialects – as the other. For a general discussion of the term *Italic*, see Poccetti, this volume. A comprehensive survey of all the languages of ancient Italy, both Indo-European (IE) and non-IE (e.g. Venetic, Messapic, and Etruscan, among others) is provided by Prosdocimi (1978a).

2. The structure of the Italic lexicon

Since Latin alone among the ancient languages of Italy is a *Großkorpussprache* with ample documentation, our overview of the lexicon of the other Italic languages is fragmentary and presents many problems. The inadequacy of our knowledge of these languages is owing to the fact that they are all attested in small, epigraphic corpora restricted both thematically and lexically. Thematically, they deal with administration and politics, and religion and rites, with corresponding restrictions in lexicon. Another well represented category comprises epitaphs, which are by their very nature limited to certain phraseological types. Only Latin, with its robust literary tradition dating to the late 3rd century BCE, bears witness to a complete range of semantic fields and documents the lexicon in an exhaustive manner, from the literary refinements of Cicero to the everyday conver-

sation of average people as reflected in Plautus. This level of exhaustive documentation is completely lacking in Faliscan and Sabellic.

2.1. To get an impression of the basic lexicon of Italic, a small list of common words inherited from Indo-European will suffice (cf. Meillet [1903] 1934: 378 ff.). The following items of the Indo-European inherited lexicon (body parts, family members, numerals, core verbs) are attested in both Latino-Faliscan and Sabellic (all are cited with some of their most evident cognates and their full attestation in *WOU*):

- IE **ped-/pod-* ‘foot’: L. *pēs, pedis*, O. **pedú** (‘measure of length’), U. **peři, persi; peřum** (< **ped-o-* ‘ground’), cf. L. *oppidum* ‘town’ < **op-pedo-m* (*WOU* 522 ff.)
- IE **mar-/man-* (heteroclitic) ‘hand’: **manu-* (*u*-stem) as Italic innovation: L. *manus*, O. *manim*, U. **manuve, mani/mani, manf** (cons. stem), SP. **manus** (*WOU* 450 f.)
- IE **ph₂ter-* ‘father’: L. *pater*, O. **patir**, SP. **patereih**, Mars. *patre* (*WOU* 518 f.)
- IE **meh₂ter-* ‘mother’: L. *māter*, O. **maatreis**, U. **matres, matrer**, SP. **matereih** (*WOU* 441)
- IE **b^hreh₂ter-* ‘brother’: L. *frater*, O. **fratrúm**, U. **frater/frater** etc. (*WOU* 293 f.)
- numerals (‘two, three, four’): L. *duō*, U. **dur, tuf, duir**; comp. L. *bi-* (**dwi-*), O.-U. *du-* (*bipēs, dupursus*) (*WOU* 192 f.); L. *trēs*, O. **trís**, U. **trif, tris** (*WOU* 767 f.); L. *quattuor, quadri-/quadru-*, O. **pettiur**, U. *petur-* (*WOU* 550 f.)
- ‘be’, IE **h₁es-*, **b^huH-*: L. *sum, est, esse* etc., O. **súm, íst, est**, U. **eru, erom, est/est**, SP. **esum**; L. *fuit* etc., O. **fufens**, U. **fust** etc. (*WOU* 245 ff.)
- ‘go’, IE **h₁ei-*: L. *ire*, U. **etu/etu** etc.; cf. O. **eitiuvam** (*WOU* 207 ff.)
- ‘bear’, IE **b^her-*: L. *ferō, ferre*, U. **ferest, fertu/fertu**, Volsc. *ferom*, Marr. *feret*; O. (with preverb) **amfret** (*WOU* 275 ff.)
- ‘show, point to, say’, IE **deik’-*: L. *dicō*, O. **deikum/deicum**, U. **teitu/deitu** (*WOU* 159 f.)
- ‘turn’, IE **uert-*, **uend^h-*: L. only *vertere*, O.-U. both roots, cf. U. **vurtus, kuvertu/covertu**, O. ῥερσορει ($\delta\iota\upsilon\text{ῥε}\text{ῖ}\text{ῥε}\text{ρσορει} \sim \text{Ζεὺς Τροπαῖος}$); U. *-uendu* (*WOU* 835 f., 844 f., 864 f.)

2.2. It is worth noting that there are a few basic vocabulary items of the inherited Indo-European lexicon that are attested in Sabellic but not in Latin or Faliscan. We must assume that the Latino-Faliscan branch lost these items, replacing them with others. The following are significant lexical isoglosses separating the two main branches of Italic:

- IE **wed-ōr/*ud-en-* ‘water’: U. **utur, une** (cf. L. *unda*), but L. *aqua* (cf. O. **aapam**) (*WOU* 815 f.)
- IE **peh₂-w₁/*peh₂-un-* ‘fire’: U. **pir/pir**, O. *pur-* in **purasiaí** ‘pertaining to fire’, but L. *ignis* (*WOU* 557 f., 612)
- IE **d^hugh₂ter-* ‘daughter’: O. **futír**, but L. *fília* (*WOU* 306 f.)
- ‘son’, IE **pu-tlo-*: O. **puklum**, Mars. *pucl[e]s*, SP. **puqloh**, Pael. *puclouis* (cf. L. *puer*), but L. *filius* (*WOU* 599 f.); there are no traces of **sunu-* in Italic.
- ‘civitas, people, community’, (Western) IE **teutā*: O. *touto*, $\tau\omega\text{ῤ}\tau\omega$, u. **tuta/tota(m)**, SP. **toúta**, Marr. *toutai*, but L. *civitās, populus* (see below), *rēs pública* (*WOU* 779 f.)
- IE **derk-* ‘see’: U. *terkantur*, but L. *vidēre* (sceptical *WOU* 747 f.)
- IE **h₂ner-* ‘man’: O. **nír**, SP. **nír**, U. *nerf* ‘man’, but L. *vir* (cf. personal name *Nerō*) (*WOU* 495 ff.)

- O.-U. **medos* (O. **meddíss**, U. **meřs/mers**, *mersto*) ‘right, law’, but L. *iūs* (< **iouis*), *iūstus* (archaic *iovestōd*), cf. *WOU* 459.
- O.-U. *ais-* ‘god’, probably from Etruscan (*WOU* 68 ff.), see below.

2.3. Finally, there is an important set of words shared by the Italic dialects that either have no cognates elsewhere in IE or have particular meanings in Italic lacking in cognates elsewhere. The following list represents some lexical isoglosses of Italic (in both branches). Typically, we have specific words connected with religion and ritual practices (cf. Porzio Gernia 1961: 97 ff.), but there are a few examples also from more common lexical domains (cf. Untermann 1993: 96 f.).

- ‘have, seize’: L. *habēre*, U. **habia**, **habetu** etc. (*hab-ē-* < **g^hab-ē-*, *WOU* 311 ff.), cf. U. **hahtu/hatu** (< **g^hab/p-je-tōd*), O. *hafiest* (*haf-io-* < **g^hab^h-io-*), *hipid*, *hipust* (*hip-* < **g^hēp-*) (*WOU* 315 f.)
- ‘use’: Italic **oit-*: L. *ūtī*, Pael. *oisa* (part.), O. *úttiuuf* (subst., *n*-stem) (*WOU* 790 f.)
- ‘(pro-)cure’: L. *cūrāre*, Pael. *coisatens*, U. **kuraia**, **kuratu** (*WOU* 407); from Italic **koiśā* (Lat. *cūra*)
- ‘lie (*scil.* buried)’: L. (*in*)*cubat*, F. **cupat**, SP. **qupat**, Marr. *cibat*, Pael. *incubat* (*WOU* 418 f.)
- ‘write’: Latin *scrībere*, O. *scriftas*, U. *screhto*, *screihtor* (*WOU* 685 f.; cf. Untermann 1993: 95).
- ‘family’: Italic **fameliā*: L. *familia*, O. *famelo*, U. **fameřias** (cf. **famelos* ‘slave’: *famulus*, *famel*) (*WOU* 262 ff.; on the terminology of slavery in Italic, see Rix 1994, especially 35 ff.)
- ‘people, exercitus’: L. *populus*, U. **puplum**/*poplom* (*WOU* 610 f.; cf. **toutā*, O. *comono/comenei*, **kúmbennieis**)
- ‘peace’: Italic **pāk-*: L. *pāx*, U. *pase* (*pāse*), cf. U. *pacer* etc. (< **pāk-ri-* ‘propitius’) (*WOU* 508 ff.)
- ‘meal’: L. *cēna* (old *cesna*), U. *śesna*, O. **kerssnaís**, **kersnu** (*WOU* 392 ff.)
- ‘(religious) feast’: L. *fēriae*, O. **fīśiaís** (cf. ‘temple’: O.-U. **fīśnú**, **fīśnam**; **fesnafe/fesnere** < **fēsñā*, L. *fānum* < **fāsnom*) (*WOU* 281 ff.)
- ‘urban architecture’: L. *forum*, U. **furu**/*furo* ‘forum’ (*WOU* 305); L. *via*, O. **víú**, U. **via**/*uia* ‘street’, SP. **víam** (*WOU* 860 f.)
- ‘holy, consecrated’: Italic **sakro-/*sākri-*: L. *sacer*, F. **sacru**, **sacra**, O. **σακορο**, U. **sacru**/*sacra* etc. (*WOU* 647 ff.), cf. L. *sanctum* (*sancīre*), O. **sahtúm**, U. (**tra**) **sahta** (top.), all belonging to the root **sak-* ‘to consecrate’ (*WOU* 640 f.)
- Italic **pī-jo-* ‘pious etc.’: L. *pius*, *piāre*, Volsc. *pihom*, Marr. *peai*, O. **pihiúí**, U. *pihatu*, *pihaclu*/**pihaklu** (cf. lat. *piaculum*) (*WOU* 552–555)
- the pronouns L. *ille* (old *olle*)/O. **úlleis**/*olu*, L. *alter*/O. **alttram**/*altrud*, preverbs *prai*, *pri*, *re*, superlative L. *ultimus*/O. **últiumam** etc.

2.4. The seminal article discussing the structure, stratification, and hierarchies within the different domains of the Italic lexicon is by Campanile (1967: 106 ff.); the criteria for *Urverwandtschaft* and possible later exchange among Common Italic and the Italic dialects are debated by Untermann (1993: 97 ff.); and the various isoglosses between Italic and the other branches of IE are discussed by Porzig (1954: 97 ff., 131 ff.). Molinari

(1965) provides a special study of the Italic and Germanic isoglosses, while Haas (1960) concentrates on substrate elements in the Italic languages. A representative bibliography of work on the Italic lexicon can be found in Heidermanns (2005: 561 ff.).

2.4.1. The different aspects of the Latin lexicon are beyond the scope of this article, especially the literary influences of Greek and the postclassical development (cf. Hofmann and Szantyr 1965: 37*ff.; on Latin lexicology and lexicography, see Hofmann and Szantyr 1965: 74*ff. as well as the exhaustive bibliographies in Cousin 1951: 272 ff., Cupaiuolo 1993: 443 ff., and Heidermanns 2005: 573 ff.). A very concise overview concerning the stratification and changes of the Latin lexicon – from its IE background to Vulgar Latin and Proto-Romance – is provided by Morani (2000: 301–325).

2.4.2. The Faliscan lexicon (with its limited remains of personal names, theonyms, and appellatives) is discussed in its entirety by Giacomelli (1963: 169–259).

2.4.3. The Sabellic lexicon is enregistered completely (and alphabetically) in *WOU* with ample semantic and etymological remarks. This dictionary replaces older works and glossaries, such as von Planta (1897: 674–765), Muller (1926), and Vetter (1953: 379 ff.).

2.5. A special field within the Italic lexicon is constituted by anthroponymy. Most of the languages of central Italy gave up the IE system of nomenclature (usually a compound name, e.g. **H₂uesu-kleues-* ‘He of good renown’) and introduced a pluripartite system of names, consisting of an obligatory *gentile* (‘family name’) deriving from a patronym as well as a language-specific ordering of the whole name set: Latin with *praenomen* + *nomen gentile* and *cognomen* (and affiliation), e.g. *Lucius Furius (L.f. Ouf.) Crassipes*, Faliscan: *Uoltio Uecineo Maxomo Iuneo*, Umbrian with *praenomen* + patronymic + *nomen gentile*, e.g. *Vuvśis Titis Teteies*, Oscan with *praenomen* + *nomen gentile* + patronymic (and a facultative cognomen), e.g. *Niumsis Heirennis Niumsiels*. Cf. also Etruscan with *praenomen* + *nomen gentile* + *cognomen* + affiliation: *Vel Tite Meluta Arnθal*. For an extensive account of the origin and various complications of central Italic (i.e. Etruscan, Latino-Faliscan, Sabellic) and peripheral (i.e. Venetic, Messapic, etc.) nomenclature, see Rix (1995a, 1995b), Untermann (1995), and Wallace (2007: 49–53) with further literature.

3. Loanwords

The directions of loans are manifold in the Italic lexicon. The two major external sources are Greek and Etruscan, with loans sometimes passing from the first of these to the second. But also internally, cultural terms may pass from Latin into Sabellic and, much less frequently, *vice versa*.

Cf. Greek to Italic: *χοῖνιξ* → O. **kúiniks** ‘grain measure’, Etruscan to Italic: *ais, ais* → O. *aiso-* etc. ‘god’ (cf. Wallace 2007: 56), Greek to Etruscan to Italic: *κυλίχνη* → Etr. *culiχna* → O. **culchna** ‘goblet, cup’, Latin to Sabellic: *cisterna* → U. *cisterno* ‘cistern’, Greek to Latin to Sabellic: *θησαυρός* → L. *thēsauros* → O. **thesavrúm** ‘treasure’. Cf. L. *turris* ~ O. **tiurri** ‘tower’ (from Greek *τύρσις*? But perhaps from a mediter-

ranean substrate). – L. *vīnum*, F. **uino** ~ U. **vinu**, Volsc. *uinu* ‘wine’ (mediterranean loanword).

3.1. A chronological layer can be discerned within the loanwords. The earliest borrowings come from Greek (often via Etruscan) and comprise names of gods and cultural items.

– Theonyms: O. **herekleís**, Vest. *herclo*, L. *Herculēs*, (*me*)*hercle*, Etr. *hercle* ← Ἡρακλῆς (*WOU* 318 f.); O. **appelluneís**, (L. *Apollō*), Etr. *aplu(ne)* ← Ἀπόλλων (Dor. Ἀπέλλων) (*WOU* 115 f.); (more recent) Pael. *perseponas*, *urantias*. – For cultural borrowings, see above.

3.2. Later on, in the third and second centuries BCE, Latin political and legal vocabulary began to penetrate into the Sabellic lexicon. This shows the growing influence of the Roman expansion all over Italy. Interestingly, the Sabellic loanwords in Latin are restricted to a more rural and agricultural sphere (cf. Ernout [1909] 1928: *passim*).

Cf. O. **kvaísstur**, κραιστρ, U. **kvestur** ← L. *quaestor*, O. **keenzstur/censtur**, **kenzsur** ← L. *cēnsor* (cf. Untermann 1993: 100 f.), O. **áidil** ← L. *aedīlis*; O. **lígatúís** (*lígatús**) ← L. *legātus*, O. **senateís** ← L. *senātus* (genuinely O. **kúmbennieís**, **kú]mparakineís**), O. *ceus* ← L. *cīvis* ‘citizen’, O. **embratur** ← L. *imperātor* (cf. Pael. *empratois*). Cf. *WOU*, s.vv. – O. **meddíss**, μεδδειξ, etc. (**med(o)-dik-*) → L. *med(d)ix* (citation word), which is formally equivalent to *iūdex* (cf. *WOU* 456 ff.; but Untermann 1993: 94). – U. *marone*, SP. **maro(n)úm** (a functionary title) ← Etr. *maru* (*WOU* 454 f.). On O.-U. **uhtur** (a functionary title), see *WOU* 788 f.

3.3. For Greek loanwords in Latin, Weise ([1882] 1964) and Saalfeld ([1884] 1964) are still indispensable. O.-U. influences on Latin are treated in Ernout ([1909] 1928) and Hofmann and Szantyr (1965: 36*f.), while Etruscan elements in Latin are discussed in Morani (2000: 316 ff.). Lazzeroni (1972: 1 ff.) and Magni (1993: 85 ff.) deal with Sabellic, especially Oscan borrowings from Greek, while Prosdocimi (1978b: 1029 ff.) and Giacomelli (1983) present a thorough discussion of Greek elements in Italic and bilingualism in ancient Italy. For Latin loanwords and interferences in Oscan, Porzio Gernia (1970: 94 ff.) and Campanile (1976: 109 ff.) should be consulted. Camporeale (1956: 33 ff.) deals with magisterial titles and terms. A complete list of the Latin, Greek, and Etruscan loanwords in Sabellic can be found in Heidermanns (1999: 436 ff.).

4. Word-formation

The augmentation of the lexicon is mainly achieved by the formation of new complex items (derivations, compounds) from basic words. Italic word-formation does not much differ, either in whole or in detail, from the principles of IE derivation and composition. Surprisingly, though, in the Italic languages compounding is rather limited and achieved a degree of productivity only in literary and poetic Latin via the influence of Greek (cf. Lindner 1996, 2002; Heidermanns 2002). It would be pointless to attempt to delineate

both Latin and Sabellic word-formation within the restricted space of this article. Reference works dealing with these matters include, for Latin, Leumann ([1928] 1977: esp. 273–403) and for Sabellic, Heidermanns (1999). A detailed treatment of Oscan-Umbrian word-formation can be found in von Planta (1897: 1–275) with shorter discussions in Buck ([1904] 1928: 182–194) and Bottiglioni (1954: 95–103). For further references, see Heidermanns (2005: 565 ff. for Latin and 626 f. for Sabellic).

5. References

Bottiglioni, Gino

1954 *Manuale dei dialetti italici (Osco, Umbro e dialetti minori). Grammatica, Testi, Glossario con note etimologiche*. Bologna: Società Tipografica Editrice Bolognese.

Buck, Carl Darling

1928 [1904] *A Grammar of Oscan and Umbrian*. 2nd edn. Boston: Ginn.

Campanile, Enrico

1967 Note sulla stratificazione del lessico italico. *Studi e saggi linguistici* 7: 106–141.

Campanile, Enrico

1976 La latinizzazione dell'osco. In: Giacomo Devoto, Antonio Pagliaro, and Vittore Pisani (eds.), *Scritti in onore di Giuliano Bonfante*, Vol. 1. Brescia: Paideia Editrice, 109–120.

Camporeale, Giovannangelo

1956 La terminologia magistratuale nelle lingue osco-umbre. In: *Atti e Memorie dell'Accademia Toscana di Scienze e Lettere "La Colombaria"* 21 (n.s. 7). Florence: Olschki, 33–108.

Cousin, Jean

1951 *Bibliographie de la langue latine, 1880–1948*. Paris: Belles Lettres.

Cupaiuolo, Fabio

1993 *Bibliografia della lingua latina (1949–1991)*. Naples: Loffredo.

Eichler, Ernst, Gerold Hilty, Heinrich Löffler, Hugo Steger, and Ladislav Zgusta (eds.)

1995 *Namenforschung. Ein internationales Handbuch zur Onomastik*. 1. Teilband. (Handbücher der Sprach- und Kommunikationswissenschaft 11.1). Berlin: De Gruyter.

Ernout, Alfred

1928 [1909] *Les éléments dialectaux du vocabulaire latin*. 2nd edn. Paris: Champion.

Giacomelli, Gabriella

1963 *La lingua falisca*. Florence: Olschki.

Giacomelli, Roberto

1983 *Graeca Italica. Studi sul bilinguismo-diglossia nell'Italia antica*. (Studi grammaticali e linguistici 15). Brescia: Paideia Editrice.

Haas, Otto

1960 *Das frühitalische Element. Versuch über die Sprache der ersten Indogermanen Italiens*. Vienna: Notring.

Heidermanns, Frank

1999 *Sabellische Nominalbildung. Die Wortbildung der Substantive und Adjektive im Oskisch-Umbrischen*. University of Cologne Habilitation. (To appear as Handbuch der italischen Dialekte, Vol. 4. Heidelberg: Winter.)

Heidermanns, Frank

2002 Nominal Composition in Sabellic and Proto-Italic. *Transactions of the Philological Society* 100: 185–202.

- Heidermanns, Frank
 2005 *Bibliographie zur indogermanischen Wortforschung. Wortbildung, Etymologie, Onomasiologie und Lehnwortschichten der alten und modernen indogermanischen Sprachen in systematischen Publikationen ab 1800*, Vol. 1. Tübingen: Niemeyer.
- Hofmann, J. B. and Anton Szantyr
 1965 *Lateinische Syntax und Stilistik. Allgemeiner Teil*, Vol. 2 of the *Lateinische Grammatik* by Manu Leumann, J. B. Hofmann, and Anton Szantyr. (Reprinted with revisions 1972). Munich: Beck.
- Lazzeroni, Romano
 1972 Contatti di lingue e culture nell'Italia antica: elementi greci nei dialetti italici. *Studi e saggi linguistici* 12: 1–24.
- Leumann, Manu
 1977 [1928] *Lateinische Laut- und Formenlehre*. Munich: Beck. [Göttingen: Vandenhoeck & Ruprecht.]
- Lindner, Thomas
 1996 *Lateinische Komposita. Ein Glossar, vornehmlich zum Wortschatz der Dichtersprache*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Lindner, Thomas
 2002 *Lateinische Komposita. Morphologische, historische und lexikalische Studien*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Magni, Elisabetta
 1993 Ve 192: Episodi di interferenza fra greco e osco. *Studi e saggi linguistici* 33: 85–104.
- Meillet, Antoine
 1934 [1903] *Introduction a l'étude comparative des langues indo-européennes*. 7th edn. Paris: Hachette.
- Molinari, Maria Vittoria
 1965 Relazioni tra il lessico germanico ed i lessici latino ed oscumbro. In: *Memorie dell'Istituto Lombardo – Accademia di scienze e lettere*, Classe di lettere, scienze morali e storiche, XXVIII.3. Milan: L'Istituto Lombardo di Scienze e Lettere, 335–402.
- Morani, Moreno
 2000 *Introduzione alla linguistica latina*. Munich: LINCOM.
- Muller Jzn, Frederik
 1926 *Altitalisches Wörterbuch*. Göttingen: Vandenhoeck & Ruprecht.
- von Planta, Robert
 1897 *Grammatik der oskisch-umbrischen Dialekte*, Vol. 2. Strassburg: Trübner.
- Porzig, Walter
 1954 *Die Gliederung des indogermanischen Sprachgebiets*. Heidelberg: Winter.
- Porzio Gernia, Maria Luisa
 1961 Rapporti tra il lessico sacrale osco e latino. *Archivio Glottologico Italiano* 46: 97–138.
- Porzio Gernia, Maria Luisa
 1970 Aspetti dell'influsso latino sul lessico e sulla sintassi osca. *Archivio Glottologico Italiano* 55: 94–144.
- Prodocimi, Aldo Luigi (ed.)
 1978a *Lingue e dialetti dell'Italia antica*. (Popoli e civiltà dell'Italia antica 6). Rome: Biblioteca di Storia Patria.
- Prodocimi, Aldo Luigi
 1978b Contatti e conflitti di lingue nell'Italia antica: l'elemento greco. In: Aldo Luigi Prodocimi (ed.). *Lingue e dialetti dell'Italia antica*. (Popoli e civiltà dell'Italia antica 6). Rome: Biblioteca di Storia Patria, 1029–1088.
- Rix, Helmut
 1994 *Die Termini der Unfreiheit in den Sprachen Alt-Italiens*. (Forschungen zur antiken Sklaverei 25). Stuttgart: Steiner.

- Rix, Helmut
1995a Etruskische Personennamen. In: Eichler et al. (eds.), 719–724.
- Rix, Helmut
1995b Römische Personennamen. In: Eichler et al. (eds.), 724–732.
- Saalfeld, Günther Alexander Ernst Adolph
1964 [1884] *Tensaurus italograecus. Ausführliches historisch-kritisches Wörterbuch der griechischen Lehn- und Fremdwörter in Lateinischen*. Hildesheim: Olms. [Vienna: Gerold.]
- Untermann, Jürgen
1993 ‘Urverwandtschaft’ und historische Nachbarschaft im Wortschatz der italischen Sprachen. *Incontri linguistici* 16: 93–101.
- Untermann, Jürgen
1995 Die vorrömischen Personennamen der Randzonen des alten Italien. In: Eichler et al. (eds.), 732–738.
- Untermann, Jürgen
2000 *Wörterbuch des Oskisch-Umbrischen*. (Handbuch der italischen Dialekte 3). Heidelberg: Winter.
- Vetter, Emil
1953 *Handbuch der italischen Dialekte*. Vol. 1. *Texte mit Erklärung, Glossen, Wörterverzeichnis*. Heidelberg: Winter.
- Wallace, Rex E.
2007 *The Sabellic Languages of Ancient Italy*. Munich: LINCOM.EUROPA.
- Weise, Friedrich Oscar
1964 [1882] *Die griechischen Wörter im Latein*. (Preisschriften der Jablonowskischen Gesellschaft XV. 23). Leipzig: Zentral-Antiquariat der Deutschen Demokratischen Republik. [Leipzig: Hirzel.]
- WOU = Untermann 2000.

Thomas Lindner, Salzburg (Austria)

51. The dialectology of Italic

1. The status of Proto-Italic within Indo-European
2. Internal subgrouping of Sabellic
3. Internal subgrouping of Latino-Faliscan
4. References

1. The status of Proto-Italic within Indo-European

The differences between Latino-Faliscan and Sabellic are not trivial and have led some to view Italic as a pseudo-branch or *Sprachbund* that arose through convergence of geographically contiguous but phylogenetically not closely connected dialects, rather than as a node on the Indo-European Stammbaum. This view goes back to the 1910s and 1920s and originated in Germany, though it came to be identified especially with Italian scholars. See the overview of scholarship in Jones (1950: 62–63). Beeler (1966) takes a compromise position. The methodological issues surrounding this dispute cannot

- Rix, Helmut
1995a Etruskische Personennamen. In: Eichler et al. (eds.), 719–724.
- Rix, Helmut
1995b Römische Personennamen. In: Eichler et al. (eds.), 724–732.
- Saalfeld, Günther Alexander Ernst Adolph
1964 [1884] *Tensaurus italograecus. Ausführliches historisch-kritisches Wörterbuch der griechischen Lehn- und Fremdwörter in Lateinischen*. Hildesheim: Olms. [Vienna: Gerold.]
- Untermann, Jürgen
1993 ‘Urverwandtschaft’ und historische Nachbarschaft im Wortschatz der italischen Sprachen. *Incontri linguistici* 16: 93–101.
- Untermann, Jürgen
1995 Die vorrömischen Personennamen der Randzonen des alten Italien. In: Eichler et al. (eds.), 732–738.
- Untermann, Jürgen
2000 *Wörterbuch des Oskisch-Umbrischen*. (Handbuch der italischen Dialekte 3). Heidelberg: Winter.
- Vetter, Emil
1953 *Handbuch der italischen Dialekte*. Vol. 1. *Texte mit Erklärung, Glossen, Wörterverzeichnis*. Heidelberg: Winter.
- Wallace, Rex E.
2007 *The Sabellic Languages of Ancient Italy*. Munich: LINCOM.EUROPA.
- Weise, Friedrich Oscar
1964 [1882] *Die griechischen Wörter im Latein*. (Preisschriften der Jablonowskischen Gesellschaft XV. 23). Leipzig: Zentral-Antiquariat der Deutschen Demokratischen Republik. [Leipzig: Hirzel.]
- WOU = Untermann 2000.

Thomas Lindner, Salzburg (Austria)

51. The dialectology of Italic

1. The status of Proto-Italic within Indo-European
2. Internal subgrouping of Sabellic
3. Internal subgrouping of Latino-Faliscan
4. References

1. The status of Proto-Italic within Indo-European

The differences between Latino-Faliscan and Sabellic are not trivial and have led some to view Italic as a pseudo-branch or *Sprachbund* that arose through convergence of geographically contiguous but phylogenetically not closely connected dialects, rather than as a node on the Indo-European Stammbaum. This view goes back to the 1910s and 1920s and originated in Germany, though it came to be identified especially with Italian scholars. See the overview of scholarship in Jones (1950: 62–63). Beeler (1966) takes a compromise position. The methodological issues surrounding this dispute cannot

be adequately addressed here. In my view, a traditional Proto-Italic stage most easily accounts for the available evidence and involves the fewest additional assumptions. We will examine the evidence from phonology, morphology, lexicon, and syntax in 1.1–1.4 below and discuss Latin–Sabellic divergences in 1.5.

1.1. As Weiss (2009: 467 ff.) notes, the phonological isoglosses traditionally used as diagnostic of an Italic subgroup of Indo-European are a somewhat mixed bag of trivial and non-trivial innovations. The following discussion incorporates much of his list, augmented with some additional material; I have tried to include some items not usually discussed in this context. Space does not allow inclusion of every conceivable Proto-Italic innovation.

1.1.1. Four of Weiss’s “most probative” changes are shared with Venetic (1.1.1.1–1.1.1.3 and 1.1.1.5) and at least one is probably shared with Celtic (1.1.1.6). Given the uniqueness of these changes within Indo-European, they are unlikely to have arisen independently in these branches. This means that if one wants to use the first four of these as evidence of an Italo-Venetic subgroup (or however one chooses to label it), one must also use the fifth as evidence of an Italo-Celtic (or Italo-Veneto-Celtic) one.

1.1.1.1. Labialization of PIE **dh* and **g^wh* to *f* word-initially and *f* (often showing up as *v* or *β*) word-internally. At least the first of these is found in Venetic (*vha.g.s.to* ‘fecit’). Weiss adds **gh-/*ġh- > h-*. This is more complicated because of the difference between Latin and Faliscan in the development of **ghu-/*ġhu-*. See further below, 3.2.2.

1.1.1.2. PIE voiced aspirates > voiceless fricatives word-initially and voiced fricatives word-internally. Fricatives are an uncommon outcome of the voiced aspirates in branches where that series is kept distinct from the other stop series, and hardly ever met with word-initially (cf. Kümmel 2007: 66–67). Only in Greek (starting in the Hellenistic period, sporadically earlier) do the aspirates ultimately become fricatives across the board. Germanic turns them into fricatives word-internally. Sanskrit (orthographic) *h* from **ġh* and (before front vowel) **g^wh* seems to have been a voiced fricative originally (Wackernagel 1896: 243–244), and in Middle Indic, debuccalization of aspirated stops, producing *h*, is common intervocally. In Armenian, fricatives can result from original voiced aspirates in some word-internal contexts. Thus the fact that all the Italic languages share the development to *f/h* word-initially and *v/γ* word-internally is significant (that is, assuming that the Latin stop outcomes are hardenings of earlier fricatives still preserved in Sabellic; this is not universally agreed upon).

1.1.1.3. Unstressed **o* in the sequence **ou(H)V > *a*, yielding **au(H)V* (Thurneysen-Havet’s Law). This has recently been shown to be an Italic rather than a Latin-specific change and to have operated before the shift to word-initial stress (Vine 2004, 2006), cf. U. *sauitu* imper. ‘cut, slash’ < **ksou-ēje-*, Lat. *cauus* ‘hollow’ < **kouH-ó-*. Apparently also in Ven. *ho.s.tihavo.s*.

1.1.1.4. **j, ĭ > or, ol*. This is diagnostic of Italic according to various authors. Also in Venetic (*voltigenei*, perhaps *murtuvoi*). Weiss notes that several Greek dialects independently show *o*-vocalism here too, but it is otherwise rare.

1.1.1.5. $*\eta mV > *omV$ and secondary $*\eta > om$, as in *homō* and OLat. *esom* (with clitic reduction from $*esmi$, via $*es\eta$). Venetic has *dekomei* ‘tenth’ (loc.) $< *dek\eta mei$.

1.1.1.6. Syncope of $*-Cje-$ to $-Ci-$ in an open syllable after single light syllable. This may be shared with Celtic. Syncope also seems to have affected the sequence $*i\bar{V}$ - produced from $*iV-$ after a heavy syllable.

1.1.2. I would consider the following innovations to be as probative as the ones Weiss deems “most probative”:

1.1.2.1. $*-eje- > *-ije-$ when the first *e* is unstressed (Vine 2012), whence *o*-stem denominatives in $*-e-je-$ $> -\bar{i}-$ (Lat. *seruire*, U. **seritu**). Similarity to Celtic denominative \bar{i} -verbs from $*-e-je-$ is superficial only, since there $*-eje- > *-\bar{e}- > -\bar{i}-$ (McCone 1996: 49). While I find Vine’s account for the Italic \bar{i} -denominatives compelling, I am not certain that acrostatic *i*-stem nom. pl. ($*-ejes >$) $*'ijes$ would have further developed to $-\bar{is}$ as he claims (2012: 565–567), as, pace Vine, Weiss (2009: 145 n. 21) may be right that the syncope did not happen in closed syllables, cf. 4th-conjugation participles in *-ient-*, gerund(ive)s in *-iend/iund-*, 3pl. *-iunt*.

1.1.2.2. Stressed $CRHC > *CaRaC$, parallel to Greek but without the different coloring effects for the different laryngeals. This is strongly doubted by Schrijver (1991: 193–197), and I agree with him that most of the alleged examples are not very good. But at least *caluus* $< *kalawos$, probably the base of the Oscan gentilicium **Kalaviis**, and *palma* $< *palamā$ alongside Skt. *-kūlva-* and Gk. *παλάμη*, respectively, are difficult to explain otherwise.

1.1.2.3. The counting of $*(C)VR$ and sequences of two syllables as a heavy sequence for the purposes of “Sievers’ Law”, e.g. $*or-je- > *or-i\bar{je}- >$ Lat. *or-ī-* ‘rise’, $*her-je- > *her-i\bar{je}- >$ U. *her-ī-* ‘wish’, $*sepel-je- > *sepel-i\bar{je}- >$ *sepel-ī-* ‘bury’. I place quotations around this phenomenon because it is not clear whether we are dealing in Italic with Sievers’ Law proper or a similar, independent phenomenon; cf. Weiss (2009: 40) and now Byrd (2015: 188 with n. 21). The occurrence of Sievers’ Law following a sequence of two syllables is also found in Germanic.

1.1.2.4. Apocope of $*-i$. The limits of this change for Italic are disputed (see Weiss’s discussion, 2009: 468 with n. 17). The personal endings $*-si$, $*-ti$, $*-nti$ undergo apocope in some other branches also, including Insular Celtic. Contra Weiss (2009: 468) and with Schrijver (2006: 49) and McCone (1996: 100–102), the Celtic change is specific to Insular Celtic; and it is not limited there to verbs, cf. McCone (1996: 100–102). Hock (2007: 71–72), followed by Weiss, thinks this loss is attributable to phrase-final apocope, since verbs in an SOV language would have been at the end of a phonological phrase. But Insular Celtic is VSO, and that word-order was established before the loss of $*-i$ (see McCone 2006: 65 for remarks on the early date of verb-initial syntax in Insular Celtic). At any rate, apocope of $*-i$ in Italic is not limited to verbs, and Meiser’s suggestion (1998: 73–74) that $*-i$ was lost when unstressed (at a time before stress-retraction to initial syllables) may work better, whence e.g. $*éti > et$ but loc. sg. $*pedí > pede$.

1.1.2.5. **-ūiV- > *-īiV-* (the “*pius* rule”) has often been considered Italo-Celtic, cf. OIr. consuetudinal present *bíid* etc. < **bhūie-* like Lat. *fiō*. But Zair (2009) has shown that the other putative Celtic examples are not probative, and adduces some counterexamples, leading him to suggest a different explanation for *bíid*. Even if *bíid* should turn out to show the rule (I for one do not find Zair’s account of this form entirely convincing), the rule’s effects were clearly more limited in that branch – perhaps restricted to after labials. If the Italic and Celtic material is to be combined, the more restricted application of the rule in Celtic could mean it originated in pre-Italic and spread limitedly to pre-Celtic.

1.1.2.6. The dissimilation of **l...l > *l...r*, as in Lat. *mīlit-āris*, Vest. *Flusare* ‘Floralis’ (month name).

1.1.2.7. **-gi- > -ġi-*. Weiss (2009: 469) points out that Sabellic examples of this are lacking, but he notes that it must predate 1.1.1.6 above (whence *ait* < **aijet* < **ag-iet[i]*), which seems good evidence to me for its being Proto-Italic in date.

1.1.2.8. Development of thorn clusters to **ks* (> *s* in Latin) is widely regarded as diagnostic of Italic (thus e.g. Meiser 1986: 38). Though it is true there is no clear Sabellic evidence (Weiss 2009: 469), it is very difficult to imagine that the merger of **tġ*, **g^whdh*, and **ks* as (ultimately) *(k)s* in Latin (*ursus*, *sitis*, *situs* ‘mold’, *texō*) only happened at the pre-Latino-Faliscan level. The reduction of **dhġh* to **ġh* in ‘yesterday’, though found in Albanian and Germanic, is not found in the more closely related Celtic, where **dh* won out, suggesting the cluster reduction was *einzel sprachlich*, though surely before **dh* would have become *f* (it is inconceivable that a ***fh-* or ***fġ-* would have become *h-*).

Kloekhorst (2014: 46–49, 62–63) disputes Schindler’s (1977: 31–32) widely followed claim that the reduction of **dhġh(ŋ)m-* to **ġh(ŋ)m-* in ‘earth’ was *grund sprachlich*. He seems to be right that it is at least post-Anatolian, but the agreement among Gk. *χαμαί*, Lat. *humī*, *homō* (there probably was no **ġemōn*, contra Kloekhorst 2014: 49; see Livingston 2004: 34), Goth. *guma*, Lith. *žmuō*, and maybe the Ved. forms in *jm-* (if they replace expected **hm-*) cannot be argued away, and his alternative explanations seem far-fetched to me. The late PIE speakers who migrated into Europe as well as their pre-Indo-Iranian kin probably had undergone the reduction. (Kloekhorst is right, though, that in prevocalic position the attested reductions are *einzel sprachlich*.) Thus the agreement among Lat. *homō*, *humus*, SPic. **homanah** is not diagnostic of a specific Italic simplification in this word.

1.1.3. The following two changes appear to me less probative than Weiss allows:

1.1.3.1. Unstressed **-ōy- > *-āy-*. Only attested in *octāuus*, Osc. **Ūhtavis** ‘Octāvius’. The isolated example and the well-known irregularities in the development of the ordinal numerals across the family make it uncertain whether this is a regular sound change or a sporadic lexical change.

1.1.3.2. **-sy- > *-ry-* was proposed by Rix (1981) as a Proto-Italic change on the basis of *Minerua* < **menes-ūā* ‘the mindful one’ and *caterua* ‘throng’, Umb. **kateramu** ‘organize themselves into groups’ < **kates-ūā(-)*, cf. *catēna* ‘chain’ < **kates-nā*, and several other Latin etyma. But Meiser (1986: 184) points out that U. nom. sg. **mersus** ‘correct’

< pre-Umbrian **medes-uos* not §*mederuos* indicates the change was post-Italic. The reason for Rix's (1981: 112) dating of the change to Proto-Italic is Etruscan *Men(e)rva* and variants, with *-r-* already in the 6th century, too early for rhotacism in Italic. But, contrary to the prevailing view, Minerva's name does not seem to me securely etymologizable; certainly there is little if any guarantee that it has anything to do with 'mind'. In Etruria, where she was worshiped long before coming to Rome, *Menrva* was a goddess of lightning and war, and had chthonic associations as well as connections to health and the nursing of babies (see e.g. Jannot 2005: 148; de Grummond 2006: 71–78). In Rome she was primarily a patroness of craftsmen and tradespeople (Wissowa 1912: 253); associations with wisdom postdate the equation with Athena. The real figure of a divinized mind or sensibility would appear rather to be *Mens* (*Bona*), whose worship began in 217 BCE, when, following Rome's disastrous military defeat at Trasimene during the Second Punic War, she was introduced after consultation of the Sibylline Books. *Mens* is one of several deities that were apparently imported from Greece in that year (Wissowa 1912: 313–314). The putative sound change is somewhat suspect for a Proto-Italic date anyway, given that nowhere else did **s* rhotacize fully to **r* that early, and even the precursor of rhotacism, intervocalic **-s-* > **-z-*, is *einzel sprachlich* (1.1.4.1. below).

The parallel of Lat. *mergō* < **mezgō* cited by Rix (1981: 118) need not be of Proto-Italic date, either; Latin is the only witness. The comparison of **menes-uā* to Ved. *manas-vin-* (Rix 1981: 117) is of no value; *manas-vin-* first occurs in the Brāhmaṇas and belongs to a very productive adjective type (°*as-vin-*) of the post-Rigvedic period (Debrunner 1954: 917). In sum, Umbrian and Latin could have both inherited a Proto-Italic **-su-* (and **-zg-*, for that matter) with later independent change of **s* > **z* > *r*. Osc. **aisīvu** tPo 17, reasonably interpreted as 'of the gods' in *Imagines*: 2 793–795, was suggested by Fortson and Weiss (2013) to be from **aisuom* by anaptyxis, from the *u*-stem **aisu-*. But if **-su-* clusters behaved like *-CR-* clusters generally in Oscan, one would really expect §**aisīvu**. The idea in *Imagines* that **aisīvu** has taken on the last half of **deiv-** remains quite dubious, but the matter requires further investigation.

1.1.4. A few of the numerous items that are clearly not probative (see Weiss 2009: 468–470 for some others) deserve some comment:

1.1.4.1. Weiss (2009: 470) is probably right that intervocalic **-s-* > **-z-* is later than Proto-Italic. He cites *maximus* < **magisomos* before voicing, though he allows that **mag(i)zomos* cannot be excluded (Weiss 2009: 81 n. 8). Another example might be **esom*. If this had become **ezom* in Proto-Italic, we might expect Lat. §*rum* instead of *sum*: aphaeresis happened due to clisis, and an unstressed cliticized **zom* arguably would not have escaped rhotacism.

The initial *z-* of Fal. 3pl. *zot* 'sunt' cannot be separated from the initial *z-* of *zextos*, *zenatuo*, etc.; whatever their explanation, this problem is not relevant here (see Bakkum 2009: 85–86 for a recent assessment). Although it is moot for our present purposes, I am not sure Weiss is right to dismiss intervocalic **-s-* > **-z-* as a trivial change. It is, of course, commonly seen cross-linguistically, but it did not occur prehistorically in any other branch of IE except Germanic, where it was part of a larger set of changes (Verner's Law), and in Eretrian Greek, where it is a local (if early) development.

1.1.4.2. Weiss deems **mī* > *nī* trivial because of its also being found in Greek. For Proto-Italic, I am not convinced this even happened, trivial or not. The evidence is

essentially confined to *ueniō*, which, however, is very ambiguous. (The Greek evidence is just as uncertain; see e.g. Brugmann and Thumb 1913: 90; Schwyzer 1939: 309.) *Quoniam* < **k^wom-iam* has also been forwarded as an example (Leumann 1977: 126; Weiss 2009: 160), but I have little confidence that this conjunction is ancient. Furthermore, **m_i* > *n_i* is potentially contradicted by the word-initial change **m_i-* > *m-*, evidenced e.g. by *mouēre* < **m_ieuH-ē-* (on this verb see now in detail Vine 2006: 217 ff.). On all these issues, see Fortson to appear.

1.1.4.3. Kümmel (2008: 6) suggests **k^we(n)k^w-* > **k^wo(n)k^w-* as a possible Proto-Italic change, assuming pre-Sabellic **k^wonk^we* is not an innovation of that branch (Lat. *quīnque* is worthless because it owes its vocalism to *quīn[c]tus*). Indeed, Lat. *popīna*, a Sabellic loan, does show that **k^wek^w-* became **k^wok^w-* in both these groups. Such rounding, however, seems fairly unremarkable and probably not a solid isogloss. The rest of the family is of little help since **k^we(n)k^w-* sequences are limited to Italic and Celtic. The rounding in OIr. *cóic* < pre-Irish **k^wink^we* is a late change (McCone 1996: 118) and was not conditioned by a following labiovelar anyway.

1.1.5. The particular assemblage of the secure innovations shared among the Italic languages is unique and, even leaving aside the doubtful cases, they are most easily attributed to a stage of common development. Note that there are no sound changes specific to either Latino-Faliscan or Sabellic that must predate the innovations listed above. We might expect that there would be some if the convergence theory were correct, as it presupposes a space of time in which pre-Sabellic dialectal IE and pre-Latino-Faliscan dialectal IE were undergoing their own developments prior to the convergence period. This is the source of the awkward position, tellingly highlighted by Tikkanen (2009: 254), in which advocates of the convergence theory wind up: they “often find themselves assuming a continued close connection” between pre-Latino-Faliscan and pre-Sabellic after they split from Western IE, with the result that “[w]hat is proposed by the *Sprachbund* theory is thus a split that is not really a proper separation ... If so, one cannot help but wonder what kind of separation is actually meant.” (The detailed discussion of the Italic problem found in Tikkanen’s 2009 thesis, which I was only able to consult through the kindness of the author, is not included in its published version, Tikkanen 2011.)

1.2. Even more significant are the morphological innovations binding the Italic family together.

1.2.1. As is well-known, principal among these are innovative features of the Italic verb.

1.2.1.1. The development of the four conjugations with stem-vowels in *-ā-* *-ē-* *-e-/i-* and *-ī-* from the same inherited sources in each case. In large measure the development of these conjugations is due to sound change, especially the early loss of intervocalic yod and vowel-contractions, but since it is likely that some of the latter were analogical (e.g. **āo* > *ā*, **ēo* > *-ē-* in the 1pl. and 3pl. of 1st- and 2nd-conjugation verbs) it is methodologically simpler to ascribe the relevant analogies to a (pre-)Proto-Italic phase than to independent developments. Given that 1sg. *-āō* is still uncontracted (or restored) in Sabellic, 3pl. pan-Italic *-ānt* (Lat. *-ant*), *-ēnt* (Lat. *-ent*, SPic. **-īnt**) is noteworthy. It is entirely thinkable, however, that the phonology of the *ā*-denominatives of Italic, Venetic

(*dona.s.to*), and Celtic (Ir. *móraid*) reflects an inherited feature going much farther back than Italic.

1.2.1.2. Meiser (1993: 171) points out that cognate Latin and Sabellic verbs have the same present stems nearly three-quarters of the time, by contrast with the situation in the perfect (see below, 1.5.4.). This strongly suggests common development.

1.2.1.3. The gerund and gerundive morpheme, which becomes *-nd-* in Latin and *nm-* in Sabellic. The form, function, and unusual syntax of the gerundive, attested in both branches of Italic, form a cluster of non-trivial morphosyntactic innovations. For the most recent account of its origin, see Jasanoff (2006) (**-ntino-*).

1.2.1.4. The loss of the inherited imperfect indicative and its replacement with the formant **-βā-*. Although it is not assured that Osc. *fufans* is an imperfect rather than a pluperfect, the morpheme is the same in either case and had the same preteritizing function. (Pisani's [1963] attempt to derive it from **bheudh-* is unsuccessful, and his remark that *es-* rather than *fu-* is elsewhere the only imperfective stem does not seem accurate, given imperf. subj. *fusíd* = Lat. *foret* and the imperatives *futu*, etc.) It has now also been suggested by Dupraz (2010: 320–321) that the broken form *profafā[* on a North Oscan (Vestinian) inscription from Navelli (Mattiocco 1986: 92–95; *Imagines. Incerulae* 3, not in *ST*) is also an imperfect in *-fā-*. (The editors are agnostic about the final letter before the break in *profafā[*; Mattiocco's (1986) reading is based on autopsy and should probably be followed. The inscription has apparently since been lost.)

The similar loss of the imperfect in Insular Celtic is commonly believed to have been precipitated by the loss of **-i*, said to have erased the distinction between many primary and secondary endings (see e.g. Schrijver 1992: 189 ff.). This works well in the 2sg. and 1pl. But in the 3rd person, this account is complicated by the fact that final stops were realized as voiced already in late PIE, as clearly reflected in Celtic by the Celtiberian ablatives and 3rd-person imperatives in *-z* < **-d*. Thus what we write as **-(n)t* was really **-(n)d*, which would have been distinct from newly created **-(n)t* by apocope from **-(n)ti*. The same considerations apply to Italic. The loss of the imperfect in both these branches, then, may well be more than just an automatic byproduct of the apocope of **-i*.

Whether the future morpheme **-βe/o-* is also of Proto-Italic date, and was totally lost in Sabellic, is unclear. It is also unclear to me whether **-βe/o-* and Insular Celtic subjunctive **be/o-* necessarily share a common source (on **be/o-* see Schumacher 2004: 247–248). A specifically Latino-Faliscan analogy of the type **esā-* : **ese/o-* : **-βā-* : X, X = **-βe/o-* cannot be ruled out.

1.2.1.5. The imperfect subjunctive in **-sē-*. (see 1.5.1)

1.2.1.6. The 2nd-plural mediopassive endings in *-m^o* (Sabellic **-mō[r]*, Lat. *-minī*, *-minō*). Though the details of the preform are unclear, the base alone is a sufficiently non-trivial innovation.

1.2.1.7. It has been suggested by Meiser (2003: 57), Harðarson (2011), and most extensively Fortson (2012) that Latin mediopassive inf. *-rier* is cognate with Sabellic *-fē(r)*,

from PIE **-dhjeh₁* vel sim., with contamination of the dental (pre-Latin **-δ-*) with the **-z-* of the active infinitive morpheme. Fortson (2012: 89 ff.) further suggested that the mediopassive marker *-r* was added already by Proto-Italic times (Umbrian *-ff[e]i* could easily result from **-ff[e]ir*, contra the erroneous account in Rix 1976: 328).

1.2.1.8. The loss of all inherited participles besides the present active and perfect passive. Lexicalized traces of present mediopassive (*alumnus*) and probably aorist (*cliēns*) participles survive, but the death of these as productive categories probably preceded Proto-Italic. A few words have been etymologized as old perfect active participles (e.g. recently SPic. **vepses** by Meiser 2003: 48–49 and Martzloff 2007; my money is on **vepses** simply being the genitive of a *lāpsus*-type perfect passive participle: ‘... of Titus Alius buried [vel sim.] in this tomb’). I have not found these proposals convincing, but they do not change the picture in any case.

1.2.1.9. The generalization of the optative to an all-purpose modal or “subjunctive” category, with precise agreements in morphology across the branch: in the present, the full-grade optative suffix **-jē-* was added to the stem-vowel *-ā-* in the first conjugation, while the other conjugations used the morpheme **-ā-*.

1.2.1.10. The reduction of 1sg. **esmi* to **esm̄* > **esom*. For arguments that this is an Italic innovation, see Joseph and Wallace (1987).

1.2.2. We may add the following from non-verbal categories:

1.2.2.1. Loss of the instrumental and its replacement by the ablative. This may be shared with Celtiberian and the rest of Celtic, but only Italic is certain to have remade inherited *o*-stem instrumental **-eh₁* as **-ē-d* and to have used this as an adverb formant (Old Latin *facillumed*, Osc. *amprufid* ‘improperly’ Lu 1, SPic. **kupírih** AP 2 ‘well’ < **kuprēd*). The spread of ablative **-d* outside the *o*-stems is shared with Celtic (Celtiberian). In Young Avestan, which also spread the dental outside the *o*-stems, the new ablative formation was based on the genitive (replacement of gen. *-s* with abl. *-t*).

1.2.2.2. Loss of the dual. Not characteristic of any other ancient branch except, independently, Anatolian (if it had the dual in its prehistory) and Armenian.

1.2.2.3. The addition of a particle *-i* to demonstratives and the nom. sg. animate relatives: masc. **k^wo-i* fem. **k^wa-i* > Lat. *quī quae*, Osc. **pui paí**. A parallel development is found in Insular Celtic **k^wei* (> OIr. *cía*, MW *pwy*), but with different ablaut.

1.2.2.4. The creation/addition of an adverb **k^womk^we* to the relative/indefinite to form a more intensive indefinite ‘whosoever’: U. **piśi pumpe**, Lat. *quicumque*.

1.2.2.5. The suffixation of the particle **ke* to deictic pronouns: Lat. *hi-c*, Pael. *ecuc*, Osc. *izik*. This could of course be an areal feature, as is probably the case with the abstraction of a suffix of identity having the shape *-dem* in Latin and **-dom* in Sabellic (Lat. *idem*, Osc. m. nom. sg. **isídum**) from neut. nom.-acc. sg. **-d* plus a particle **em* or **om*. *-dom* did not spread quite as far as *-dem*, cf. Osc. m. nom. pl. **ius-um** ‘idem’.

1.2.2.6. The remaking of the 2sg. dat. personal pronoun **tebhi* as **tebhei* (Lat. *tibī*, Osc. **tfei**). Other branches have performed similar remodelings but with different dative endings; independently OPr. *tebbei* has used the same ending as Italic.

1.2.2.7. In derivational morphology, mention may be made of the innovatory adjective and noun suffixes **-āno-*, **-āri-/āli-*, **-āsijō-*, **-āto-* (see Hajnal 1993), **-dhli-*, **-idho-*, **-ijōn-*, **-mōnijom*; abstract nouns in **-itiā*; the generalization of the *o*-grade agent-noun suffix **-tōr-*; the feminine agent-noun suffix **-trīk-*; the diminutive suff. **-kelo-*; and the repetitive verbal suffix **-tā-* (in Sabellic in U. **etatu** etc. < **ei-tā-*; on the type, see Weiss 2009: 401–402 with inclusion of an analysis by A. Nussbaum from an oral presentation).

1.3. Although lexical correspondences are widely considered the least reliable material for purposes of subgrouping (so e.g. Tikkanen 2009: 47), it is clear from studies like Untermann (1993) and Rix (2005) that the lexical impact of either Latin on Sabellic or the other way around was modest at best, and semantically circumscribed. If two matching lexemes are innovative vis-à-vis PIE with respect to morphology and/or semantics, and have each undergone subbranch-specific sound changes such that later borrowing is unlikely or impossible, then a strong case can be made for their having common patrimony. There are close to 120 morphologically and/or semantically innovative exact lexical correspondences (or implied correspondences, as when one branch attests a denominative verb to a lost base preserved in the other branch) between Latin and Sabellic. This number is impressive given the limited surviving corpus. Unlike the probable or provable loanwords of later times, these correspondences span all semantic fields. Dismissal of them based simply on automatic fall-back to borrowability is in my opinion too cavalier.

1.3.1. It may be useful to list these items in one place. The list is primarily culled from Untermann's dictionary (2000); see also Untermann (1993) for discussion of specific semantic groups and individual lexemes. I have kept the list conservative (quite a few other, less certain correspondences could be added) and included only one Latin (or Faliscan) and one Sabellic form per entry. I write Proto-Italic fricatives as voiceless throughout for convenience: **ād-ro-* 'black': Lat. *āter*, U. *adro*; **ad-ser-e/o-* 'declare (a captured slave) free': Lat. *asserere*, O. *aserum*; **ajēs-* 'bronze': Lat. *aes*, U. adj. **ahesnes** (but tentatively suggested by Cowgill 1973: 294 n. 45 to be a borrowing from Latin); **aiyi-tāt-* 'lifetime': Lat. *ae(ui)tāt-*, O. **aítateis**; **ali-tero-* 'other (of two)': Lat. *alter*, O. **altram**; **arūo-* 'field': Lat. *aruum*, U. **arvam**^o; **atk^we* 'as' (vel sim.) > Lat. *atque* 'as, and', U. **ape** 'when'; **-dām* in time adverbs: Lat. *quon-dam* 'at one time', U. *nersa* 'until'; **deiuīno-* 'divine': Lat. *dīuīnus*, O. **deivinais**; **dek-ē-* 'be proper': Lat. *decēre*, U. **tiçit**; generalized stem **d(i)jē-* from the acc. sg. **d(i)jēm* 'day': Lat. *diē-*, Osc. *zicolom* < **djē-kelolo-* (also Venetic [loc. *diei*]); generalized oblique stem **dijou-* 'Jove': Lat. *Iou-*, Mars. *iou-*; **du-plo-* 'double, two each': Lat. *duplus*, U. *dupla*; **emelo-* 'buy': Lat. *emere*, U. **ematur**; **elis-to-* 'this': Lat. *iste*, U. **estu** (Celtib. *íste* 'and[?], or[?]' is possibly built of the same material, but remains uncertain); pres. **fak-je/o-* 'make, do': Lat. *faciō*, U. **façia**; **fameli(i)ā* 'household': Lat. *familia*, U. **fameñias**; **famelo-* 'slave': Lat. *famulus*, P. *famel*; **fatē-* 'speak, make solemnly known': Lat. *fatēri*, O. **fatium**; ? **fēl-* 'suckling' vel. sim.: Lat. *fēlā-* 'suck', U. **feliuf** 'suckling' (?); **fēsiai* 'religious holiday': Lat. *fēriæ*, O. **fiisiais**; **flōs-* 'flower': Lat. *flōs*, Vest. (month-name) *flusare*; **Flōsā* 'Flora': Lat. *Flōra*, O. **fluusaí**; **fraud-* 'wrongful act, flaw': Lat. *fraud-*, U. *frosotom* if

fraud-t-*; **gnārā-* ‘relate, recite’: Lat. *narrāre*, U. **naratu; **gnāti(ō)n-* ‘stock, group of related people’ vel sim.: Lat. *nātiōn-* ‘stock, race’, U. **natine** ‘patrician family’; **g^wrāt-* ‘favor, grace’: Lat. *grātēs*, O. *brateis*; **k^wanto-* ‘how much/great’: Lat. *quantus*, O. *pantes*; **hospot-* ‘stranger’: Lat. *hospit-*, P. *hospus*; **ifei* ‘there’: Lat. *ibi*, U. **ife**; **joko-* ‘word, speech’: Lat. *iocus* ‘joke’, U. **iuka** ‘prayers’; **iūuenkā-* ‘heifer’: Lat. *iūuēnca*, U. *iūengar*; **karō*, **karn-* ‘piece of meat’: Lat. *carō*, *carn-* ‘meat’, U. **karu**, **karn-** ‘part of sacrificial animal’; **kasē-* ‘lack’: Fal. **carefo**, O. **kasit**; **katelo-* ‘small animal (used in sacrifices), puppy’: Lat. *catulus* ‘puppy’, U. **katel** ‘puppy (?)’; **kates-* ‘chain’: Lat. *catēna* < **kates-nā*, U. **kazi**; **katesuā-* ‘group, throng’: Lat. *caterua*, U. **kateramu** ‘group together’; **ke-* ‘here, hither’ as a terminative preverb: Lat. *ce-dō* ‘hand over!’, Osc. *ce-bnust* ‘shall have come (to)’; **Ker-es-* ‘Ceres’: Lat. *Cerēs*, O. **kerrī**; **kersnā* (**kert-es-nā*) ‘portion, meal’: Lat. *cēna*, O. **kerssnaīs**; **kleitrā-* ‘container for transport’: Lat. dimin. *clitellae* ‘pack-saddle’, U. **kletram**; **klu-ē-* ‘be called’: Lat. *cluēre*, SPi. **kduíu**; **koisā-* ‘take care’: Lat. *cūrāre*, P. *coisatens*; **kontrād-* ‘against’: Lat. *contrā*, O. *contrud*; **kubā-* ‘lie’: Lat. *cubāre*, SPic. **qapat**; **k^wām-* ‘than, as’: Lat. *quam*, O. *pam*; **k^wām-dō-* ‘when’: Lat. *quandō*, U. *panu-peí* ‘whenever’; **k^wō-* ‘whither’: Lat. *quō*, U. **pu-e**; **k^wosjo-* ‘whose’: Lat. *cuius*, O. **púiuu**; **k^wufeí-* ‘where’: Lat. (-*c*)*ubi*, U. *pufe*; **lēg-* ‘law’: Lat. *lēg-*, Marr. *lixs*; **likē-* ‘be allowed’: Lat. *licēre*, O. **líkitud**; **louko-* ‘sacred grove’: Lat. *lūcus*, O. **lúvkeí**; **manu-þ-* ‘commit, hand over’: Lat. *mandā-*, O. **manafum**; **manu-* ‘hand’: Lat. *manus*, U. **manuve**; **mēnssā* (flat?) baked good’: Lat. *mēnsa*, U. **mefa**; **nei* ‘not’: Lat. *nī*, O. *nei* ‘except’; **nōmen-* ‘name; (political) state’: Lat. *nōmen*, U. *nome*; **ḡ-tag-ro-* ‘untouched, unused’: Lat. *integer*, U. **antakres**; **oit-* ‘use’: Lat. *ūtī*, P. *oisa* (perf. part.); **oll-* ‘that’: Lat. *olle*, O. **úlleis**; **oltḡmo-* ‘last’: Lat. *ultimus*, O. **últiumam**; **op* ‘up to, over against’: Lat. *ob*, O. **úp**; **op-sito-* ‘covered, buried’: Lat. *opsitus*, O. **úpstúst** (*Imagines* 2: 1239–1240); **pāk-* ‘peace’: Lat. *pāc-*, U. *pas-* (†*paš-*?); **parasā* ‘type of bird’: Lat. *parra*, U. **parfam**; **patnelo-* ‘open (tr.)’: Lat. *pandere*, O. **patensíns**; **pelp-men-* ‘meat’: Lat. *pulmentum*, U. *pelmner*; **pīūā-* ‘make expiation’: Lat. *piāre*, U. *pihatu*; **pīūāklom* ‘expiatory offering’: Lat. *piāculum*, U. *pihaclu*; **pīūo-* ‘dutiful’: Lat. *pius*, O. **pīhiúí**; **portā-* ‘carry’: Lat. *portāre*, U. *portatu*; **postḡmo-* ‘last’: Lat. *postumus*, O. **pustmas**; **potē-* ‘be able’: Lat. *potēns*, O. **pútiad**; **prai* ‘in front of’: Lat. *prae*, O. **prai**; **prai-tero-* adj. ‘in front’: Lat. *praeter* ‘before, beyond’, U. *pretra* ‘the first ones’; **prai-stat-ā* ‘protectress’: Lat. *Praestita*, U. **prestate**; **prismo-* ‘first’: Lat. *prīmus*, P. *prismu*; **profā-* ‘approve’: Lat. *probāre*, O. **prúfatted**; **profo-* ‘correct’: Lat. *probus*, U. **prufe**; **poplo-* ‘people’: Lat. *populus*, U. *poplom*; **rē-* ‘thing, matter, property’: Lat. *rēs*, U. **ri**; **re-ueid-s-e/o-* ‘check over, inspect’: Lat. *reūsere*, U. **revestu**; **sakrā-* ‘sanctify’: Lat. *sacrāre*, O. **sakrannas** (not **sak-ro-* itself given W. *hagr* ‘ugly’, see Maier 1987); **salayo-* ‘whole, safe’: Lat. *saluus*, O. **salavs**; **sankto-* ‘sacred’: Lat. *sānctus*, O. **sahtúm**; **sed-ē-* ‘sit’: Lat. *sedēre*, U. *sersitu*; **sedí-/sēdī-* ‘seat’: Lat. *sēdēs*, U. *sersi*; **sei* ‘if’: Lat. *sī*, U. *se-pis*; **sekā-* ‘cut’: Lat. *secāre*, U. **prusekatu**; **sekno-* ‘image’: Lat. *signum*, O. **segúnú** ‘figurine’; **serḡo-* ‘observing’: Lat. *seru-āre*, O. **serevkiđ** ‘responsibility’; **seupo-* ‘supine’: Lat. *suppus*, U. *sopa* (see now Weiss 2010: 358–383); **skapelā-* ‘shoulder(blade)’: Lat. *scapulae*, U. *scapla*; **sollo-* ‘every’: Lat. *soll-*, O. **sullus**; **stīpelā-* ‘agree upon contractually’: Lat. *stīpulāre*, U. *stīplatu*; **sup(e)rād* ‘above’: Lat. *suprā*, U. **subra**; **superno-* ‘upper’: Lat. adv. *superne*, U. *superne*; **supero-* ‘upper’: Lat. *superus*, O. **supruis**; **supmo-* ‘highest’: Lat. *summus*, U. **sume**; **taf(e)lā* ‘board, table’: Lat. *tabula*, U. **tafle**; **tanto-* ‘so much/great’: Lat. *tantus*, O. *etanto*; **ten-ē-* ‘hold’: Lat. *tenēre*, U. *tenitu*; **termen-* ‘boundary’: Lat. *termen*,

O. **teremníss**; **termino-* ‘boundary stone’: Lat. *terminus*, U. *termnom*; **tersā* ‘ground, earth’: Lat. *terra*, O. **teras**; **togā* ‘covering, toga’: Lat. *toga*, SPic. **tokam** (Vine’s 1993: 232 n. 44 alternate interpretation of **tokam** as from **tōkkā* < **toutikā-* ‘public’ [U. *toce*] requires a monophthongization already in the 6th century BCE; but cf. unmonophthongized **toútaih**); **transuorssō-* ‘transverse’: Lat. *trānsuersus*, U. *trahuorfi*; **trifu-* (or **tribu-*, Weiss 2010: 193 ff.) ‘tribe’: Lat. *tribus*, U. **trifu**; **tri-podā-* ‘dance the three-step’: Lat. *tripodāre*, U. *ahatripursatu*; **uapemōniom* ‘surety’: Lat. *uadimōnium*, O. **vaa-munim** (see most recently Fortson and Weiss 2013: 665–667); **ualē-* ‘fare/be well’: Lat. *ualēre*, O. $\varphi\lambda\epsilon$; **leifrā* ‘unit of weight’: Lat. *libra*, U. (abbrev.) *uef*; **u(e)ijā* ‘road’: Lat. *uia*, O. **viú**; **uinkelo-* ‘fetter’: Lat. *uinculum*, ? U. *preuišlatu* ‘bind fast (in advance)”; **uokā-* ‘call’: L. *uocāre*, U. *subocau*; **uostero-* ‘your (pl.)’: Lat. *uoster*, U. *uestra* (with secondary *e*-vocalism in Umbrian).

1.3.2. Note also the following quasi-exact correspondences (same morphemes in different order, or slightly different morphemes, different ablaut grades, etc.): Lat. *quoad* ~ O. **adpúd**; Lat. *extrā* ‘outside’ ~ O. **ehtrad** (**ek-* instead of **eks-*); Lat. *fānum* ‘temple’ (**fas-no-* < **dhh₁s-no-*) ~ O. **fiisnú** (**fēs-nā* < **dheh₁s-no-*); Lat. *cornīx* ‘crow’ ~ U. *curnaco*; **Mārt-/Mawort-/Mamert-* ‘Mars’; Lat. *mōnstrum* ~ SPic. **múfqlúm** < **mons-t/klo-* (I follow Vine 1993: 127–130 against the usual reconstruction **mones-t/klo-*; in addition to his argument that this is not an environment where we would expect syncope in either South Picene or Latin, secondary *ns* syncopated from **nVs* does not yield *f* elsewhere in Sabellic); Lat. *pīcus* ‘woodpecker’ (vocalism after *pīca* ‘magpie’; see Meiser 1986: 47–48 for the morphological analysis) ~ U. *peico/a* (*-ī-); Lat. *plautus* ‘flat-footed’ ~ O. **plavtad** ‘sole (of foot or shoe)’, U. *preplotatu* ‘trample (in advance)”; Lat. (*porcī*) *sācrēs* ‘sacrificial (pigs)’ ~ O. **sakrim** ‘sacrificial animal’; Lat. *sollemnis* ‘yearly’ ~ O. **súllemnais** (reading of *Imagines* 1: 418).

1.4. Tikkanen (2009: 180–181, 242–243) has argued that the Latin and Sabellic accusative with infinitive and the gerundive constructions (cf. 1.2.1.3) reflect syntactic innovations of Proto-Italic.

1.5. If the divergences between Sabellic and Latino-Faliscan indicated that the most recent common ancestor of the two branches can only have been PIE itself, then we would not be justified in setting up an intermediate Proto-Italic node. But even the most striking and pervasive differences – above all in the formation of the perfect – are part of a system that had to have arisen considerably later than PIE, a system that included the innovations enumerated above plus some others shared also with one or another branch (principally the specialization of unreduplicated *s*-desideratives as futures and the functional merger of aorist and perfect).

1.5.1. Some Sabellic innovations postdate innovations common to both Sabellic and Latin and therefore presuppose an earlier Proto-Italic stage. The change of word-internal **-nss-* to *-f-* in Umbrian (probably Oscan as well, but no examples have yet been recovered; word-final **-nss* < **-nts* becomes *-f* in both languages) must postdate the Italic change of inherited dental-plus-dental clusters to **-ss-*. The rounding of **k^wenk^we* ‘five’ to **k^wonk^we* followed the Italic or pre-Italic change of **p ... k^w* to **k^w ... k^w*. The Sabellic assimilation of **nd* > *nn* is found in the gerundive, an Italic innovation. If Jasanoff’s

(2006) analysis of the gerundive suffix as continuing **-ntno-* < **-ntino-* is correct, this statement might need to be modified, depending on how **-ntno-* was treated in pre-Sabellic (simplification to **nn-* without an intermediate **-nd-* stage?).

1.5.2. The difference between Latin nom. sg. *-iō* ~ oblique *-iōn-* vis-à-vis Sabellic **-iō* ~ **-īn-* represents paradigm leveling on the Latin side (Sabellic **-īn-* < **-i(i)en-* after heavy syllable, then generalized), and the suffix is a post-PIE conglomeration in any case.

1.5.3. The differences in declension are relatively trivial, and both branches share the changes enumerated above in 1.2.2.1.

1.5.4. The most significant differences are in conjugation, but it is not clear that they require an especially remote common ancestor. They would appear to indicate several centuries of independent development. Proto-Italic had not yet supplied all secondary presents with corresponding perfects, leaving its descendants to cobble together independent solutions to this problem. The task was still in progress in the historical period; witness U. **purtius** alongside morphologically younger **purtiñus** in Tables I and II but only *purdins(i)-* in the younger tables. Rarely do Sabellic and Latin agree on the formation of a given perfect; but it should be noted that such disagreements are found between Latin and Faliscan as well. In fact, of the four securely attested Faliscan perfect stems, three differ from their Latin counterparts: Fal. *poreded* with dereduplicated *-ded* vs. Lat. *-didit*; Fal. *fifiked*, *ff[i]iqod* vs. Lat. *fīnixit*; Fal. *facēd*, *facēt* vs. *fēcīt* or Praenestine *vhevhaked*. Only Fal. *pepara[i]* and Lat. *peperī* agree, and the Faliscan pattern *faciō* : **facī*, familiar also from Sabellic, is not even attested in Latin outside of composition. The generalization in Sabellic of (originally athematic?) *e*-grade 3pl. **-ent(i)/*-en(d)* where Latin and Faliscan have **-ont(i)/*-on(d)* is due to trivial leveling: a form like Osc. **sent**, U. *sent* is an archaism while Lat. *sunt*, Fal. *zot* has taken over thematic **-ont*; contrariwise, Osc. **fiēt** has generalized **-ent* at the expense of **-ont* (Lat. *fiunt*). The generalization of aorist endings in Sabellic over against perfect endings in Latin simply means that the old aorist and perfect were still formally distinct in Proto-Italic, much as in Old Irish (see in detail Meiser 2003). Faliscan continues an additional aorist ending (3pl. **-ont*) not found in Latin, unless *-ērunt* is a cross of *-ēre* and aorist **-ont*, but that would go against the inscriptional evidence, which lacks any trace of aorist **-ont* in early Latin. The fleshing-out of the perfect system was independently carried out in each branch but is most easily understood under the assumption of an already shared category merger (Meiser 2003: 84–85 and passim thinks the category merger did not happen until late in Proto-Italic). Thus in primary verbs Latin was more prone to generalize *s*-aorists, and Sabellic, root aorists; both preserve a certain number of reduplicated perfects and long-vowel aorists; and both have innovative formations for creating perfects to secondary presents.

1.5.5. The *s*-future houses an interesting formal difference between Latin and Sabellic. Besides the well-known fact that Latin *s*-futures are thematic while in Sabellic they are athematic, outside of the 1st conjugation they are almost always deradical in Latin (*ad-empsit* < **-em-s-*, *rupsit*, *faxit*, etc.) while in Sabellic they have generalized a formant *-es-* that is presumably depreential **-e-s-* in origin (O. *pertemest*; deradical only *fust*

and perhaps *e[est]* ‘will go’). The thematic inflection of the Latin *s*-future is probably secondary (cf. Jasanoff 1988: 233 n. 15); taking the formation as a subjunctive of a desiderative (or, for that matter, of the *s*-aorist) needlessly multiplies entities at the pre-Latino-Faliscan and/or Italic level. The Sabellic sequence *-es-* was generalized to the 2nd and 4th conjugations, whence U. *heriest*, *habiest* < **hab-ē-es-t* (but U. **staheren** ‘they will stand’ could be just **stā-īē-sent*); Bantian *herest* is ambiguous between /herrest/ < **her-īe-st*, the traditional interpretation, and athematic *her-* plus *-es-* (on the inheritance of old athematic *her-* into Oscan see Nussbaum 1976: 252). Note interestingly that in these same two conjugations in Latin, the formation is not found (except for *prohibēs-si[n]t* a few times in Cicero), and it is rather rare in the 3rd (except for 3rd-*iō* verbs). There are no such gaps in the corresponding subjunctives in *-sī-*. Importantly, in both subbranches, the purely future function of these formations is innovative vis-à-vis PIE. (For a very different analysis of the Sabellic future, see Meiser 1993: 176.)

1.5.6. Traditionally the two branches have been seen as having completely different infinitives. This is true in the active (Lat. **-si*, Sab. **-om*), but, as noted above, the Sabellic passive infinitive is likely cognate with Lat. *-rier*, both from **-diē(r)* (see 1.2.1.7). This appears to have been a true infinitive suffix all the way back (Rix 1976; Fortson 2012), but the other formations are nominals that underwent different generalizations in each branch, as is typical of IE infinitives.

1.5.7. The pronominal systems are also markedly different, but this is not terribly surprising either, given how frequently such systems are renewed cross-linguistically. Most of the elements making up the various stems are found throughout Italic; the arrangements and combinations of these elements are particular to each subgroup. Comparable situations are found in Celtic and Balto-Slavic, for example; and even two quite closely related languages like Old English and Old High German have some striking differences in the pronouns. Indeed, the diversity of forms and usages of personal and demonstrative pronouns simply across modern regional varieties of English is considerable.

1.5.8. The many lexical differences between Latino-Faliscan and Sabellic should not surprise: if one compares the Gothic and the various Old English and Old High German versions of the Lord’s Prayer, almost every lexical item is different in at least one, and sometimes all, versions (identical or near-identical across all the versions are only ‘our’, ‘thou’, ‘heaven’, ‘thy’, ‘name’, ‘[be]come’, ‘earth’, ‘give’, ‘us’, ‘we’, ‘not’, and ‘evil’).

2. Internal subgrouping of Sabellic

2.1. According to a widespread view, Sabellic, not unlike Gaul, is divided into three parts: Oscan (consisting of Oscan proper plus the “North Oscan” varieties Paelignian, Marrucian, Vestinian, and Hernican), Umbrian (consisting of Umbrian proper plus Volscian and Marsian; the sole supposed Aequian inscription [*ST VM 8*] is a forgery, see *Imagines* 1: 16 and 59 [item 3]), and a third group containing South Picene and Pre-Samnite.

2.1.1. Wallace (1985: 99–100 n. 16, 2004: 814) has objected to a strict internal division of Sabellic, preferring to view “Oscan” and “Umbrian” as situated at the ends of a dialect continuum. This view has now been given detailed and supportive treatment by Clackson (2015), who concludes that almost all the subgrouping criteria that have been proposed are problematic and that subgrouping is thus currently impossible.

2.1.2. Dialect continua in themselves are not incompatible with subgrouping if enough evidence is available documenting the relative chronology and direction of diffusion of features (for a clear and thorough demonstration of this, see Toulmin 2009). For Sabellic, such information is generally lacking. However, I believe some of the evidence dismissed by Clackson (2015) is more revealing of subgroups than he claims, as will be seen below. I agree with Wallace on general principles that labels like “Oscan” and “Umbrian” do not cleanly correlate with the messier situation on the ground; this is of course true of all language labels. But the evidence Wallace (1985) adduced in support of a dialect continuum is not persuasive. Marrucinian, he claimed, shares features with both Oscan (RC-epenthesis) and Umbrian (*-ns > -f). Marrucinian actually does not exhibit RC-epenthesis; his example *salaus* is not from **salwos*, but from **salawos* with syncope (cf. Weiss 2009: 162). As for *-ns > -f, most people now follow Rix’s (1986) claim that this occurred in the prehistory of Oscan, too. Wallace also claims that Paelignian, normally considered an Oscan dialect, has several Umbrian features, including -rf- < syncoated *-rVs-, **ū* > *ī*, palatalization of **k* before *i/y*, and *-ns > -f (99–100 n. 16). But -rf- only occurs in one form in one inscription (*cerfum* Pg 9), against plenty of examples of -rr- as in Oscan; **ū* > *ī* is highly doubtful for Paelignian (Meiser 1986: 53; Untermann 2000: s. v. *clisuist*); and palatalization in Paelignian is not attested with velars, as in Umbrian, but with dentals, as in Oscan.

2.2. Clackson considers a much broader range of features traditionally thought to be unique to one or another Sabellic language, as well as a host of putative isoglosses that indicate an Umbrian–South Picene/Pre-Samnite subgroup. He is surely correct that most of those isoglosses are weak or flat-out wrong (these points are also already laid out in Adiego Lajara 1990: 78–79): initial **l*- > *w*- (not surely attested in Pre-Samnite, and **l*- sometimes becomes *y*- rather than *w*- in South Picene), monophthongization of **ei* and **ou* (usually preserved in South Picene and Pre-Samnite), lenition of final **-d* (preserved in Pre-Samnite and on the Tortora stele), **ū* > *ī* in final syllables or more generally (South Picene evidence unclear), and creation of a demonstrative stem **esto*- (also in Latin, so older than Sabellic). In so doing, he strengthens the case for a taxonomic separation of South Picene from Umbrian.

2.2.1. To my mind, Clackson does not sufficiently consider the isoglosses that may unite Oscan and Umbrian, and the general cloak of doubt that he casts on the material is sometimes needlessly opaque. In his estimation, only the innovated 3pl. secondary ending -ns shared by Osco-Umbrian to the exclusion of South Picene holds up to scrutiny as a possible subgrouping criterion within Sabellic (Clackson 2015: 30, 33; SPic. **údiíns** is not likely to be a 3pl. contra *Imagines* 1: 190–191, see Fortson and Weiss 2013). This change is usually understood as **-nd* > **-n(n)* → **-n* plus plural **-es*, but he disputes the supposed weakening of **-nd* because of preserved -*d* (perhaps spelling -*nd*) in *fuḃfoð* etc. in the Pre-Samnite Tortora inscription (Clackson 2015: 18–19). But this does not mean that **-nd* could not have assimilated to -*nn* in the prehistory of Oscan and Umbrian,

and it definitely did intervocalically in U. *ponne* < **k^wonde*, pace Clackson (2015: 19 n. 58; correct his reference to p. 606 of Untermann 2000; the **k^wom-ne* there is an outdated suggestion of Thurneysen's that is not otherwise lent any credence by Untermann). That *-ns* is only attested in the material from after 400 BCE (Clackson 2015: 30) is an accident.

2.2.2. Clackson presents the rounding-cum-raising of word-final **-ā* to *-ō* as one of the changes that “appear to have taken place in Oscan and in Umbrian, but are questionable for the minor varieties lying between them” (2015: 23); although his treatment is a little unclear to me, he seems to count it among a small series of changes that in his estimation affected the general Sabellic area only after 400 BCE (Clackson 2015: 28). In the case of **-ā* > *-ō*, he bases this on the absence of direct evidence for the change in Umbrian prior to the later period of that language. But I see no account taken of the traditional view that there was early rounding of **-ā* to **-ā̃*, which was not reflected in the script, followed later by raising of **-ā̃* to *-ō*, which was. The first stage has to have happened prior to the loss of **-d* in Umbrian, since the abl. sg. in *-ā* < **-ād* never shows rounding, and we know that loss of **-d* predates the sixth century BCE (**face** Um 4, mid-500's). The simplest scenario is the traditional one, to assume **-ā* > **-ā̃* in the ancestor of Oscan and Umbrian, followed later (but still prehistorically) by **-ā̃* > **-ō* in central and southern Oscan (Paelignian evidences it in one inscription (Pg 9), against many others without. Marrucian does not show it; Vestinian is silent on the matter). Raising in Umbrian is reflected in the script only starting in the 2nd c. BCE and could in theory have spread there from Oscan; unfortunately we do not have relevant material from the large area between Samnite Oscan and Umbria, and it might be sociolinguistically a bit peculiar if the raising spread from Samnium all the way up through Umbroid Marsian and Sabine while having but limited effect on the Oscans' close Paelignian and Marrucian kin to the east. (I have no idea how best to account for the handful of word-internal cases, mostly before (*t*)s or *t*, of rounding/raising in later Umbrian: *Prestote* vs. older **Prestate**, *anderuacose/antervakaze*, *Casilos/Kaselate*, *pihos/pihaz*, *subotu/subator/subahtu*, and (before *k*) *Tesenocir/Tasenakes*.)

2.2.3. As per Fortson (2012: 90), I do not agree with the view, reflected in Clackson (2015: 25), that Oscan has innovated the addition of mediopassive *-r* to the imperative *centsamur* and the passive infinitive **sakrafir** by contrast with U. **persnimu** and *pihaf(e)i*. The views and data of Rix (1976: 328 with n. 66, 1986: 328), in part cited by Clackson (2015), with respect to the writing of final *-r* in Umbrian are wrong, as noted above (1.2.1.7.). Although I only discuss the passive infinitive in this context in Fortson (2012), I would extend those deliberations to the mediopassive imperative and reconstruct **-mōr* already at least as far back as Proto-Osco-Umbrian.

2.2.4. Oscan and Umbrian also share the innovation of the demonstrative stem **ekso-*. In Oscan, this stem is confined to the oblique cases, with the others furnished by **eko-*; in Umbrian, **ekso-* is the stem of the whole paradigm. Insofar as evidence is available, all varieties of Oscan agree on the distribution.

2.3. In some cases, Clackson and others have attempted to cast doubt on the validity of certain traditionally proposed Proto-Sabellic changes by claiming they happened later and diffused but not all the way to South Picene. I think such doubts are not warranted

for the first of these changes below; the others are unclear, but if Clackson should be right, then they only add potentially to the list of shared Osco-Umbrian features.

2.3.1. SPic. **opesa[?]úom** Sp RI 1 is generally taken as an infinitive of *opsā-* ‘make’ without syncope of the second syllable, leading to doubts about the validity of medial syncope as a pan-Sabellic sound change. But it is very uncertain what this form is (doubts voiced at Weiss 2010: 326 n. 266; and see now Clackson 2015: 10 with n. 28). There is a lot of damage, with the **p** barely visible and perhaps a missing character between the **a** and **ú** (see *Imagines* 1: 167–168, and cf. Marinetti 1985: 249). Furthermore, the Sabellic infinitive ending was **-om*, a sequence almost always written **-úm** in South Picene, not **-om**. Clackson (2015: 10 n. 28) seems implicitly to endorse the view that the perfects **opsút** and **[o]psúq** are built to a different stem or root and therefore are no evidence of a syncopated **op(e)s-*. The stem in question, **ops-* or **ōps-*, is taken by Untermann (2000: s. v. **úpsannúm**) as the *s*-aorist of the root **h₃ep-*. But given that e.g. U. *portā-* ‘carry’ (**puratatu**, etc.) had a perfect stem *port-* (*portust* VIIIb 3), I see no reason why *opsā-* could not simply have produced a perfect *ops-* in the same way. For the type, see Rix (1992: 239) and Dupraz (2010: 363); *ops-* would thus be a pseudo-root aorist, broadly comparable to Osc. **kúmbened**, Proto-Umbrian **face** (usually regarded as dederuplicated **fefak-*), U. *habe*, and **purtiuus**, where the present differed from the perfect only in having a present-stem-forming suffix. Osc. **ekķelled** Cp 42 (reading from *Imagines* 1: 440), if to the same **kellā-* as pf. **kellaked** Sa 10, 12, may be a further example. An analogically lengthened **ōps-*, evidenced by spellings like **upsed**, **uupsens**, *ουρκενς* from Alfedena, Pompeii, and Messina, respectively, is an Oscan innovation from Samnium and points south. To my knowledge there are no secure examples of SPic. <o> representing inherited **ō* except in final syllables before *-m* or *-h*, and in the otherwise deviant spelling **peteronis** AP 5 (reading of *Imagines* 1: 187; expected **petrún-is** AP 4). Vestinian has *osens* MV2, probably with short *o*, cf. *fadatruni[es]* in the same inscription with *-u-* < **ō-*. Presumably the vocalism of the perfect passive participle **upsatuh**, attested multiple times from Teanum (Si 4–6 and 20–1), was influenced by the perfect stem. U. *opset[a]* Um 6, *oseto* Um 7 are ambiguous. The refashioned U. perfects **usaçe**, **usaie** Ila 44, Ib 45 surely had the short *o-* of the present; on these forms see now Willi (2010: 2). Compare also SPic. pres. **praistaít** : perf. **adstaecoms** (I take this as a perfect, not present, following Rix 1992: 337–338), **adstaíuh**, where the present and perfect stems may not in fact have been distinguished except in the endings. In short, nothing stands in the way of assuming that SPic. **ops-** evidences syncope.

Even if **opesa[?]úom** is an unsyncopated form, it does not vitiate attributing syncope to Proto-Sabellic, any more than, say, the unexpectedly unrhacized *s* of **asa** ‘altar’ disproves the validity of positing rhotacism for Umbrian. Note anyway SPic. **amgenas** AP 3, whose *-mg-* can hardly have arisen except through syncope; similarly the consonant cluster in **úflfú[h?]** CH 2 (reading of *Imagines* 1: 2011: 251). Admittedly, forms like **iokipedu**, **haligatú**, **rakineví** give the impression of being unsyncopated, but with little idea as to what they mean – and the possibility that <i> represents *ī* – we cannot draw conclusions from them. It is most unlikely that **matereih** **patereih** AP 2 are unsyncopated by contrast with Osc. **maatreis**, U. **matres**, that is, preserve an original internal *-e-* (Meiser 1986: 131, followed by Tikkanen 2011: 15): the kinship terms should have inherited oblique stems in *-tr-*. Here we have anaptyxis of that cluster; anaptyxis is characteristic of this inscription. Alternatively, we could be dealing with generalized full-grade **-ter-* from the nominative, as perhaps also in Venetic **vhraterei**.

2.3.2. Oscan and Umbrian agree on spreading the *o*-stem acc. sg. to consonant stems. Whether South Picene shared in this development is not clear. It is possible that preservation of older consonant-stem **-em* < **-m̥* is seen in **dl̥i]kdeintem**, which certainly looks much more like the accusative of an *nt*-stem than a 1sg. verb form like **knúskem** or maybe **pdufem**. On the other hand, **áudaqum**, normally taken as equivalent to Lat. *audācem*, shows **-om*. But, as pointed out by Weiss (2010: 63–64), Fortson and Weiss (2013:190–191), and now Clackson (2015: 10), the connection of SPic. **áudaqum** with Lat. *audāx* is difficult because of the traditional derivation of *audāx/audēre* from *auidus*, whose *-d-* derives from **-dh-*; we would thus expect **šáufaqum** if this derivation is correct and the Latin and Picene forms were cognate. For more on these words, see now Fortson (2016), also with some remarks on *-em* vs. *-um*.

2.3.3. Clackson (2015: 28–29) disputes the traditional claim that stops were spirantized before **t* in Proto-Sabellic, citing possible counterexamples from South Picene and the Tortora inscription. But he notes that the two instances of apparent *kt* in the latter – *fri] }qto[d]* and *takiosqtod* – are probably 3sg. imperatives: the second has been argued to be syncopated from **-ske-tōd*, rendering it irrelevant, and of course that could be true of *fri] }qto[d]* as well, mutatis mutandis. The South Picene material is more difficult, consisting of the form **deiktam** CH 1 and the reasonably secure restoration **molk[t]a[h]**. Phonological or morphological restoration of the stop is always possible, as happened also in e.g. Osc. **fruktatiúf** ‘utilization’ < **frūktātiōns*; but it is not easy, especially in **molk[t]a[h]** (for **deiktam**, Weiss 2002: 356 with n. 27 notes that a syncopated **deikVtam* is at least theoretically possible). On the other hand, positive evidence for the spirantization rule in South Picene may be found in **oftorim**, also CH 1; it is difficult not to connect this with the Paelignian gentilicium *ofuries* Pg 48 (Meiser 2013: 36 with n. 7), which is probably based on either an **optōr-* or an **ok^wtōr*. (The only difficulty that I see is that we might expect a spelling ***oftúrim**, but sure examples of **ō* before *r* are lacking. **múreis** CH 1 is of disputed meaning and etymology, and **rtúr** TE 7 could reflect either **-tor* or **-tōr*.) But as that inscription was discovered a mere 40 km. away from the South Picene one, we may be dealing with a regional term that diffused from outside South Picene (and that stayed in the area for many centuries).

2.4. In sum, we are left with possibly a half-dozen innovations that support an Oscan-Umbrian subgroup.

3. Internal subgrouping of Latino-Faliscan

3.1. There is no question that Faliscan is the closest relative of Latin; at issue has been whether to treat Faliscan as a Latin dialect or a separate language, and how to evaluate the features Faliscan shares with Sabellic against Latin. For an exhaustive recent treatment see Bakkum (2009). The language/dialect question is not interesting or evaluable from a purely structural point of view (and the informal yardstick of mutual comprehensibility is not available to us). On the sociolinguistic status of Faliscan and on Faliscan self-identification, see Bakkum (2009: 341–342). The differences between Faliscan and Latin are mostly minor but not quite minimal, reflecting maybe a couple of centuries of prehistoric independent development. This would jibe with Rix’s (2005: 563) statement

that the Latino-Faliscan speech area was cloven in twain by the Etruscans moving into Veii in the 9th/8th c. BCE.

3.2.1. The only phonological deviation from Latin is in the word-internal treatment of the voiced aspirates, where Faliscan preserves the inherited Proto-Italic fricatives against the Latin stop outcomes. It is striking that **-δ-* (from **-dh-* and in some cases **-s-*) became labialized to *-β-* in both Sabellic and Faliscan across the board, but only some of the time in Latin. We could be dealing with a trait that diffused from Sabellic territory into Faliscan and (incompletely) Latin territory. This makes it all the more likely that Fal. *efil-* ‘aedile’ is not an early borrowing from Latin. For additional arguments see Bakkum (2009: 179).

Discrepancies on the side of Faliscan over against the rest of Italic in word-initial *f-* for expected *h-* (e.g. *foied* ‘today’ alongside Lat. *hodie*) are best explained as due to a general weakening of *f-* to *h-* in the 4th century with retention of *f-* in spelling; this resulted in an orthographic reanalysis whereby *f* could be interpreted as spelling [h], whence hypercorrect spellings like *foied*. See Wallace and Joseph (1991) and Bakkum (2009: 79–81).

3.2.2. Possibly more complicated is the history of **ghu-*, which becomes *fu-* in Latin (*fundō*) but *hu-* in (6th c. BCE) Fal. *huti[c]ilom* ‘futicillum, little vessel’ from **ghu-ti-* or **ghū-ti-*. The interpretation of this form is not assured, but the *h-* can only continue an aspirate and that aspirate should have become *f-* if the outcomes were as in Latin. Stuart-Smith’s (2004: 206) suggestion that *hu-* here is weakened from earlier **fu-* is disputed by Bakkum (2009: 72) because of the later date of general *f- > h-*. But perhaps the weakening happened earlier before *u* than before other vowels, by a kind of rounding dissimilation; a parallel to this can be found in dialectal Albanian (Kümmel 2007: 104). If the labialization of **χu-* to *fu-* is a Latin-specific change, as usually assumed, note that means that the weakening of **χ/*γ > h* was independent in Latin and Faliscan (**χ > h* was probably not Proto-Italic given U. *-veitu* < **uek-tōd* < **uegh-e-tōd*, not **ueχ-tōd* which would have produced §-*vehtu* or §-*veetu*; Buck 1928: 98).

3.2.3. The <c> of Fal. *lecet* ‘lies’ from **legħ-* has been phonetically interpreted by some as [ɣ], by others as [g]. It must be the latter; nowhere else to my knowledge does <c> stand for [ɣ]. But that would appear to mean that **gh > g*, uniquely in Italic (so most recently Stuart-Smith 2004: 58, 62–64). Bakkum (2009: 75) is surely right to doubt this. *lecet* should be considered together with the Italic ‘law’ word (Lat. *lēg-*, Osc. *lig-*), whose traditional derivation from **leg-* ‘choose, gather’ has never been semantically comfortable. The occasionally proposed alternative derivation from **legħ-* ‘lay down’ (see the references in Walde and Hofmann 1938–1953: s. v. *lēx*) is much more attractive and easy to parallel, cf. ON *lōg*, OE *dōm*, Gk. *θέμις*, OHG *gisezzida*, Lat. *statūtum*, Latv. *likums*, etc. Thus *lecet* and *lēg-* support each other and suggest that, for whatever reason, the root as inherited into Italic was **leg-* and not **legħ-*. The absence of Lachmann’s-Law lengthening in *lēctus* ‘bed’ is unremarkable in a non-paradigmatic form.

3.3. There is no morphology (in our limited corpus) that Faliscan shares with Sabellic to the exclusion of Latin. The Faliscan 3pl. perfect is of aorist rather than perfect origin, as in Sabellic, but with different vocalism. Latin has traces of the aorist endings in the singular (*feced*), so Faliscan has simply held on to one aorist ending longer than Latin

did. (I do not think it likely that Italic or pre-Latin inherited both an **-e* and an **-ei* in the 3sg. perfect, contra Leumann 1977: 606–607. The latter is assured [$\rightarrow -ei-t > -\bar{t}$ etc.], and the survival of the aorist is assured, so parsimony dictates taking *-ed* from the aorist, later losing out in Latin to *-eit* [and probably also in Faliscan, whence *-et*]. The curiosity in all this is why *-ei[t]* does not appear in early inscriptions if it had been in competition with *-ed* the whole time. One possibility, of course, is that *-ed* still retained a vestige of its aoristic sense in the early historical period: *fifiked* ‘fashioned’ rather than ‘has fashioned’, etc. Note the interesting development in Lucanian Oscan, which replaced 3sg. *-ed* with *-et* and perhaps even, as in Latino-Faliscan, with *-eit* [Bakkum 2009: 160 with n. 85], e.g. $\delta\epsilon\delta\epsilon\tau$, $\alpha\nu\alpha\phi\alpha\kappa\epsilon\tau$, and perhaps $\lambda\iota\omicron\kappa\alpha\kappa\epsilon\iota\tau$.) As long as Fal. 2pl. *ues* and U. *uestra* are unexplained, they cannot be used for subgrouping. (Possibly *ues* represents a replacement of **wōs* with the cons.-stem nom. pl. This ending might also have spread to *salues*; Bakkum’s [2009: 196] hesitant lemmatization of this under an *i*-stem *saluis* is, as he himself says, odd since the adjective is everywhere an *o*-stem in Italic [and cf. p. 413, where he suggests *-es* was imported from another declension]. The only assured *o*-stem nom. pl. is *lete* ‘beds’, which, purely theoretically, could have the same ending with omitted *-s*. Differently on *ues* Vine 1993: 179 n. 11.) The occasional agreement in stem-formation between Faliscan and Sabellic, principally in the perfect stem *fifik-* vs. Lat. *fīnx-*, point to different generalizations of perfect vs. aorist stems, if they point anywhere at all; there were several productive processes throughout Italic for forming perfect stems, and parallel independent developments (as is quite possibly the case in Fal. perf. stem *fac-*) cannot be ruled out.

3.4. None of the material surveyed suggests closer kinship of Faliscan with Sabellic; the two subbranches had clearly been distinct for many centuries before documentation begins.

4. References

- Adiego Lajara, Ignacio-Javier
 1990 Der Archaismus des Südpikenischen. *Historische Sprachforschung* 103: 69–79.
- Agostiniani, Luciano, Alberto Calderini, and Riccardo Massarelli (eds.)
 2011 *Screhto est: Lingua e scrittura degli antichi Umbri. Catalogo della mostra*. Perugia: Università degli Studi di Perugia.
- Bakkum, Gabriël Cornelis Leonides Maria
 2009 *The Latin Dialect of the Ager Faliscus: 150 Years of Scholarship*. Ph.D. dissertation. Amsterdam: Amsterdam University Press.
- Beeler, Madison
 1966 The interrelationships within Italic. In: Henrik Birnbaum and Jaan Puhvel (eds.), *Ancient Indo-European Dialects: Proceedings of the Conference on Indo-European Linguistics Held at the University of California, Los Angeles April 25–27, 1963*. Berkeley: University of California Press, 51–58.
- Brugmann, Karl and Albert Thumb
 1913 *Griechische Grammatik: Lautlehre, Stammbildungs- und Flexionslehre, Syntax*. 4th edn. Munich: Beck.
- Buck, Carl Darling
 1928 [1904] *A Grammar of Oscan and Umbrian*. 2nd edn. Boston: Ginn.

- Byrd, Andrew Miles
2015 *The Indo-European Syllable*. Leiden: Brill.
- Clackson, James
2015 Subgrouping in the Sabellian branch of Indo-European. *Transactions of the Philological Society* 113: 4–37.
- Cowgill, Warren
1973 Latin *stāre*. *The Journal of Indo-European Studies* 1: 271–303.
- Debrunner, Albert
1954 *Altindische Grammatik*. Vol. 2, part 2: *Die Nominalsuffixe*. Göttingen: Vandenhoeck & Ruprecht.
- Dupraz, Emmanuel
2010 *Les Vestins à l'époque tardo-républicaine: Du nord-osque au latin*. Mont-Saint-Aignan Cedex: Publications des universités de Rouen et du Havre.
- Dupraz, Emmanuel
2012 *Sabellian Demonstratives: Forms and Functions*. Leiden: Brill.
- Fortson, Benjamin W. IV
2012 Latin *-rier* and its Indo-Iranian congeners. *Indogermanische Forschungen* 117: 75–118.
- Fortson, Benjamin W. IV
2016 Latin *audēre* and related forms. In: H. Craig Melchert and Brent Vine (ed.), *Proceedings of the 26th UCLA Indo-European Conference*. Bremen: Hempen, 13–26.
- Fortson, Benjamin W. IV
To appear A cover-up: Latin *amiciō*. To appear in a festschrift.
- Fortson, Benjamin W. IV and Michael Weiss
2013 Review of *Imagines*. *Bryn Mawr Classical Reviews* 2013.06.17.
- de Grummond, Nancy Thomson
2006 *Etruscan Myth, Sacred History, and Legend*. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
- Hajnal, Ivo
1993 Das oskische Cognomen *Pukalatúí*. In: Helmut Rix (ed.), 125–142.
- Harðarson, Jón Axel
2011 The 2nd line of the Duenos inscription. In: Giovanna Rocca (ed.), *Atti del Convegno Internazionale Le lingue dell'Italia antica: Iscrizioni, testi, grammatica / Die Sprachen Altitaliens: Inschriften, Texte, Grammatik. In memoriam Helmut Rix (1926–2004) 7–8 marzo 2011, Libera Università di Lingue e Comunicazione, IULM, Milano (= *Ἀλεξάνδρεια / Alessandria* 5), 153–163.*
- Hock, Hans Henrich
2007 Morphology and *i*-apocope in Slavic and Baltic. In: Karlene Jones-Bley, Martin Huld, Angela Della Volpe, and Miriam Robbins Dexter (eds.), *Proceedings of the 18th UCLA Indo-European Conference, Los Angeles, Nov. 3–4, 2006*. Washington, DC: Institute for the Study of Man, 65–76.
- Imagines* = Michael H. Crawford (ed.)
2011 *Imagines Italicae: A Corpus of Italic Inscriptions*. 3 vols. In collaboration with William M. Broadhead, James Clackson, Federico Santangelo, S. Thompson, Margaret Watmough with computing by Errietta M. A. Bissa and Gabriel Boddard. London: Institute of Classical Studies, University of London.
- Jannot, Jean-René
2005 *Religion in Ancient Etruria*. Transl. by Jane Whitehead. Madison: University of Wisconsin Press.
- Jasanoff, Jay
1988 PIE **ǵnē-* 'recognize, know'. In: Alfred Bammesberger (ed.), *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter, 227–239.

- Jasanoff, Jay
 2006 The origin of the Latin gerund and gerundive: A new proposal. In: Harvey Goldblatt and Nancy Shields Kollman (eds.), *Rus' Writ Large: Languages, Histories, Cultures. Essays Presented in Honor of Michael S. Flier on his Sixty-Fifth Birthday* (= *Harvard Ukrainian Studies* 28 [2010]), 195–208.
- Jones, D. M.
 1950 The relation between Latin and Osco-Umbrian. *Transactions of the Philological Society* 49: 60–87.
- Joseph, Brian D. and Rex Wallace
 1987 Latin *sum* / Oscan *súm*, *sim*, *esum*. *American Journal of Philology* 108: 675–693.
- Kloekhorst, Alwin
 2014 The Proto-Indo-European “thorn”-clusters. *Historische Sprachforschung* 127: 43–67.
- Kümmel, Martin Joachim
 2007 *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.
- Kümmel, Martin Joachim
 2008 Die sabellischen Sprachen. Class handout, University of Freiburg. <http://www.indogermanistik.uni-freiburg.de/seminar/pers/kuemmel/umat/die-sabellischen-sprachen>
- Leumann, Manu
 1977 *Lateinische Laut- und Formenlehre*. Munich: Beck.
- Livingston, Ivy
 2004 *A Linguistic Commentary on Livius Andronicus*. New York: Routledge.
- Maier, Bernhard
 1987 Latin *sacer* ‘sacred’ and Welsh *hagr* ‘ugly’. *Celtica* 19: 96–97.
- Marinetti, Anna
 1985 *Le iscrizioni sud-picene*. Vol. 1: *Testi*. Florence: Olschki.
- Martzloff, Vincent
 2007 Latin *pollinctor*, grec *λίπ(α)*, picénien VEPSES: Phraséologie et élaboration poétique. In: Alain Blanc and Emmanuel Dupraz (eds.), *Procédés synchroniques de la langue poétique en grec et en latin*. Brussels: Safran, 171–189.
- Mattiocco, Ezio
 1986 *Centri fortificati vestini*. Sulmona: Museo civico di Sulmona.
- McCone, Kim
 1996 *Toward a Relative Chronology of Ancient and Medieval Celtic Sound Change*. (Maynooth Studies in Celtic Linguistics 1). Maynooth: Department of Old Irish, St. Patrick’s College.
- McCone, Kim
 2006 *The Origins and Development of the Insular Celtic Verbal Complex*. (Maynooth Studies in Celtic Linguistics 6). Maynooth: Department of Old Irish, National University of Ireland.
- Meiser, Gerhard
 1986 *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Meiser, Gerhard
 1993 Uritalische Modussyntax: Zur Genese des Konjunktiv Imperfekt. In: Helmut Rix (ed.), 167–195.
- Meiser, Gerhard
 1998 *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Meiser, Gerhard
 2003 *Veni vidi vici: Die Vorgeschichte des lateinischen Perfektsystems*. Munich: Beck.

- Meiser, Gerhard
2013 Review of Weiss 2010. *Kratylos* 58: 33–46.
- Nussbaum, Alan J.
1976 Umbrian *pisher*. *Glotta* 54: 241–253.
- Pisani, Vittore
1963 Oskisch *fufans*. *Zeitschrift für vergleichende Sprachforschung* 78: 101–103.
- Rix, Helmut
1976 Die umbrischen Infinitive auf *-fi* und die uridg. Infinitivendung **-d^hiōi*. In: Wolfgang Meid and Anna Morpurgo Davies (eds.), *Studies in Greek, Indic, and Indo-European Linguistics Offered to Leonard R. Palmer on the Occasion of His Seventieth Birthday*. Innsbruck: Institut für Sprachwissenschaft der Universität, 319–331.
- Rix, Helmut
1981 Rapporti onomastici fra il panteon etrusco e quello romano. In: Giovanni Colonna (ed.), *Gli etruschi e Roma: Incontro di studio in onore di Massimo Pallottino Roma, 11–13 dicembre 1979*. Rome: Bretschneider, 104–126.
- Rix, Helmut
1986 Die Endung des Akkusativ Plural commune im Oskischen. In: Annemarie Etter (ed.), *o-o-pe-ro-si: Festschrift für Ernst Risch zum 75. Geburtstag*. Berlin: De Gruyter, 583–597.
- Rix, Helmut
1992 Zur Entstehung des lateinischen Perfektparadigmas. In: Oswald Panagl and Thomas Krisch (eds.), *Latein und Indogermanisch: Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg, 23.–26. September 1986*. Innsbruck: Institut für Sprachwissenschaft der Universität, 221–240.
- Rix, Helmut (ed.)
1993 *Oskisch-Umbrisch: Texte und Grammatik. Arbeitstagung der Indogermanischen Gesellschaft und der Società Italiana di Glottologia vom 25. bis 28. September 1991 in Freiburg*. Wiesbaden: Reichert.
- Rix, Helmut
2005 Lehnbeziehungen zwischen den Sprachen Altitaliens. In: Gerhard Meiser and Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17–23. September 2000, Halle an der Saale*. Wiesbaden: Reichert, 559–572.
- Schindler, Jochem
1977 A thorny problem. *Die Sprache* 23: 25–35.
- Schrijver, Peter
1991 *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam: Rodopi.
- Schrijver, Peter
1992 The chronology of the loss of post-posttonic vowels between identical consonants and the origin of the Celtic first person singular imperfect. *Münchener Studien zur Sprachwissenschaft* 53: 179–196.
- Schrijver, Peter
2006 Review of Meiser 2003. *Kratylos* 51: 46–64.
- Schumacher, Stefan
2004 *Die keltischen Primärverben: Ein vergleichendes, etymologisches und morphologisches Lexikon*. Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Schwyzler, Eduard
1939 *Griechische Grammatik*. Munich: Beck.
- ST = Rix, Helmut
2002 *Sabellische Texte: Die Texte des Oskischen, Umbrischen und Süd-pikenischen*. Heidelberg: Winter.

- Stuart-Smith, Jane
2004 *Phonetics and Philology: Sound Change in Italic*. Oxford: Oxford University Press.
- Tikkanen, Karin
2009 *A Comparative Grammar of Latin and the Sabellian Languages: The System of Case Syntax*. Ph.D. thesis, Uppsala University.
- Tikkanen, Karin
2011 *A Sabellian Case Grammar*. Heidelberg: Winter.
- Toulmin, Matthew
2009 *From Linguistic to Sociolinguistic Reconstruction: The Kamta Historical Subgroup of Indo-Aryan*. Canberra: Centre for Research on Language Change.
- Untermann, Jürgen
1993 'Urverwandtschaft' und historische Nachbarschaft im Wortschatz der italischen Sprachen. *Incontri linguistici* 16: 93–101.
- Untermann, Jürgen
2000 *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- Vine, Brent
1993 *Studies in Archaic Latin Inscriptions*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Vine, Brent
2004 New thoughts on an old curse. In: Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson, and Thomas Olander (eds.), *Per aspera ad asteriscos: Studia Indogermanica in honorem Jens Elmegård Rasmussen sexagenarii Idibus Martii anno MMIV*. Innsbruck: Institut für Sprachen und Literaturen der Universität, 615–626.
- Vine, Brent
2006 On 'Thurneysen-Havet's Law' in Latin and Italic. *Historische Sprachforschung* 119: 211–249.
- Vine, Brent
2012 PIE mobile accent in Italic: Further evidence. In: Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen, and Jens Elmegård Rasmussen (eds.), *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics*. Copenhagen: Museum Tusulanum, 545–575.
- Wackernagel, Jacob
1896 *Altindische Grammatik*. Vol. 1: *Lautlehre*. Göttingen: Vandenhoeck & Ruprecht.
- Walde, Alois and J. B. Hofmann
1938–1953 *Lateinisches etymologisches Wörterbuch*. 3 vols. Heidelberg: Winter.
- Wallace, Rex E.
1985 Volscian *sistiatiens*. *Glotta* 63: 93–101.
- Wallace, Rex E.
2004 Sabellian languages. In: Roger D. Woodard (ed.), *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge: Cambridge University Press, 812–839.
- Wallace, Rex E. and Brian D. Joseph
1991 Is Faliscan a local Latin patois? *Diachronica* 8: 159–186.
- Weiss, Michael
1998 On some problems of final syllables in South Picene. In: Jay Jasanoff, H. Craig Melchert, and Lisi Oliver (eds.), *Mír Curad: Studies in Honor of Calvert Watkins*. Innsbruck: Institut für Sprachwissenschaft der Universität, 703–715.
- Weiss, Michael
2002 Observations on the South Picene inscription TE. 1 (S. Omero). In: Mark R. V. Southern (ed.), *Indo-European Perspectives*. Washington, DC: Institute for the Study of Man, 351–366.
- Weiss, Michael
2009 *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave.

- Weiss, Michael
 2010 *Language and Ritual in Sabellic Italy: The Ritual Complex of the Third and Fourth Tabulae Iguvinae*. Leiden: Brill.
- Willi, Andreas
 2010 The Umbrian perfect in **-nç-/-nš-**. *Transactions of the Philological Society* 108: 1–14.
- Wissowa, Georg
 1912 *Religion und Kultus der Römer*. 2nd edn. Munich: Beck.
- Zair, Nicholas
 2009 OIr. *biid* < **b^huH-ye/o-* and ‘hiatus’ verbs. In: Stephanie W. Jamison, H. Craig Melchert, and Brent Vine (eds.), *Proceedings of the 20th UCLA Indo-European Conference, Los Angeles 2008*. Bremen: Hempen, 213–220.

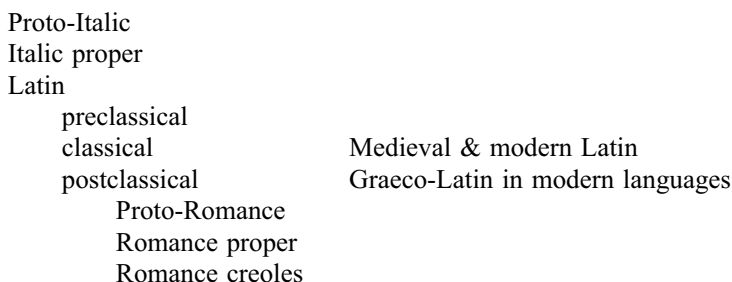
Benjamin W. Fortson IV, Ann Arbor, MI (USA)

52. The evolution of Italic

1. Overview
2. Italic and pre-Roman Italy
3. Latin
4. The Romance languages
5. Romance-based creoles
6. Outlook
7. References

1. Overview

The external and internal history of Italic can be divided into three macro-periods, each of which can be subdivided into three evolutionary steps. Latin, the most prominent member of the family, survived in a twofold manner: as a spoken language, orally transmitted from generation to generation resulting in the different Romance languages; and as a medium of written culture, transmitted by formal training in grammar schools over two millennia. It also survived in the international lexicon of science and technology. The overall evolution of the Italic branch of Indo-European can be summarized in the following diagram:



- Weiss, Michael
 2010 *Language and Ritual in Sabellic Italy: The Ritual Complex of the Third and Fourth Tabulae Iguvinae*. Leiden: Brill.
- Willi, Andreas
 2010 The Umbrian perfect in **-nç-/-nš-**. *Transactions of the Philological Society* 108: 1–14.
- Wissowa, Georg
 1912 *Religion und Kultus der Römer*. 2nd edn. Munich: Beck.
- Zair, Nicholas
 2009 OIr. *biid* < **b^huH-ye/o-* and ‘hiatus’ verbs. In: Stephanie W. Jamison, H. Craig Melchert, and Brent Vine (eds.), *Proceedings of the 20th UCLA Indo-European Conference, Los Angeles 2008*. Bremen: Hempen, 213–220.

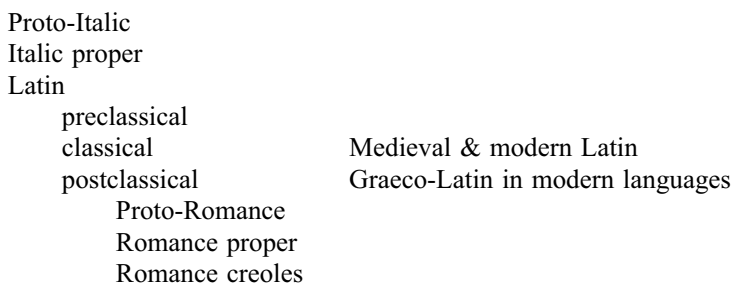
Benjamin W. Fortson IV, Ann Arbor, MI (USA)

52. The evolution of Italic

1. Overview
2. Italic and pre-Roman Italy
3. Latin
4. The Romance languages
5. Romance-based creoles
6. Outlook
7. References

1. Overview

The external and internal history of Italic can be divided into three macro-periods, each of which can be subdivided into three evolutionary steps. Latin, the most prominent member of the family, survived in a twofold manner: as a spoken language, orally transmitted from generation to generation resulting in the different Romance languages; and as a medium of written culture, transmitted by formal training in grammar schools over two millennia. It also survived in the international lexicon of science and technology. The overall evolution of the Italic branch of Indo-European can be summarized in the following diagram:



2. Italic and pre-Roman Italy

2.1. Little is known about Proto-Italic. This language family is derived from Proto-Indo-European and came into Europe from the East, wherever the original home of Indo-European might have been situated. What seems clear is that Proto-Italic was spoken to the north of the Alps before the Italic tribes settled in the Appenine Peninsula. While we do not know the exact location of their prehistoric habitat, there is universally recognized linguistic evidence that they lived in close contact with Celtic tribes and also with speakers of Proto-Germanic. If we accept Vennemann's hypothesis (2003: XV) that Europe north of the Alps was formerly inhabited by speakers of Vasconic, it must be assumed that prehistoric languages related to present-day Basque have also left some traces in Proto-Italic. An example of such substratal influence is possibly the word for 'year' which is *annus* in Latin and *akno-* in Oscan. This Italic word occurs otherwise only in Germanic (Gothic *apnam* [dat. pl.]) but is unattested in other Indo-European languages and has no clear Indo-European etymology (*pace* Ernout and Meillet [1932] 1985: 35; Walde and Hofmann 1938–1956: 51). As Vennemann (2003: 232) has convincingly shown, it can be derived from Basque *adin* 'age'. A substantial part of the Latin lexicon cannot be related to the Indo-European proto-language; this etymologically enigmatic vocabulary can partially be traced to Vasconic. Proto-Italic first evolved in a contact zone in Central Europe where it coexisted with dialects of Proto-Celtic, Proto-Germanic, and Vasconic (and surely still other language families).

In the second half of the 2nd millennium BCE groups of speakers of Indo-European dialects began to infiltrate into what was to become Italy. They are usually related to the Terramare and the Villanovan cultures (Pulgram 1958: 98 ff.), although admittedly archaeological and linguistic evidence in prehistoric times can seldom be equated with certainty.

2.2. The distribution of languages in pre-Roman Italy is known by inscriptions dating from the first millennium BCE. The picture emerging from this evidence is complex and puzzling. The Indo-European character of some languages is still a matter of dispute, as well as the internal subgrouping of languages of undoubtedly Indo-European origin (Woodard 2004). In the following paragraph I give a short overview, generally following the majority of specialists (being well aware that there are divergent opinions for almost every detail). The order is roughly geographic (North to South).

- Ligurian, spoken in Liguria, may have been an Indo-European language, perhaps belonging to the Celtic branch. There are arguments, however, to consider it as an isolate.
- Lepontic, spoken to the north of Liguria in the Western Alpine region, was clearly a Celtic language, closely related to Gaulish.
- Rhaetic, spoken in the Central Alps, was possibly related to Etruscan, but this is far from certain, since lexical and onomastic similarities can also be attributed to contact. It was certainly not Indo-European.
- Venetic, spoken in Venetia, was undoubtedly Indo-European. It is safe to assume that it formed an independent branch by itself, rather than a subgroup of Italic.
- Etruscan, spoken in Tuscany, parts of Umbria, and Northern Latium, is an isolate. Rix (1998) postulated a genetic relationship between Etruscan, Rhaetic and Lemnic (spo-

- ken on the Greek island Lemnos), grouping all three languages in a family named “Tyrsenian”; this whole group may be related to Nuragic.
- Nuragic, the language of the indigenous inhabitants of Sardinia (and possibly also Corsica), is clearly non-Indo-European. This language is not known from inscriptions, but only from toponyms and the substratal lexicon it left in Sardinian.
 - North-Picene, spoken in Eastern Umbria (Pesaro), is scantily documented. There is general agreement, however, that it was not Indo-European. As long as no more archaeological evidence is unearthed, it must be considered as an isolate.
 - South-Picene, spoken along the Adriatic coast in Southern Umbria, has been clearly identified as belonging to the Sabellic branch of Italic, being somewhat closer to Umbrian than to Oscan.
 - Umbrian, spoken in most of central Italy, is well attested by the famous Iguvine tablets. Oscan, spoken in central and southern Italy, mainly in Campania, is well-known from numerous inscriptions. Both languages, together with minor languages such as South-Picene and Volscian, form the Sabellic branch of Italic, often called Osco-Umbrian.
 - Faliscan, spoken in Velletri (Falerii) north of Rome, was the closest relative of Latin. Latin originated on the hills of Rome. Together these languages form the Falisco-Latin (or Latino-Faliscan) branch of Italic.
 - Messapic, spoken in Apulia, was an Indo-European language. It seems to be independent from Italic, possibly related to Illyrian, the group of Indo-European dialects spoken in the Western Balkan region on the opposite coast of the Adriatic Sea.
 - Sicel (or Siculian), spoken in Sicily, is poorly documented. Whether or not it was Indo-European remains an open question.

Let us now summarize. Leaving aside clearly non-Indo-European languages and languages of doubtful genetic affiliation, we can distinguish five major branches of Indo-European in the Apennine Peninsula: Celtic (represented by Lepontic), Venetic, Sabellic, Latino-Faliscan, and Messapic. Of these, Celtic unquestionably does not belong to Italic. It is safe to exclude also Venetic and Messapic from Italic, the former by itself constituting an independent branch of Indo-European, the latter belonging to the Illyrian family. So we are left with two main branches of Italic proper, Sabellic (Osco-Umbrian), and Latino-Faliscan. These two can confidently be assumed to descend from a common ancestor, appropriately termed Proto-Italic.

2.3. The commonalities between Sabellic and Latino-Faliscan are well-known (e.g. **dh* → *f*, as in Oscan *fakiiad* = Latin *faciat*; imperfect indicative in **-fā-*, as in Oscan *fu-fans* ‘they were’, Latin *portā-bant*); they permit us to postulate a valid genetic node “Italic”. On the other hand, the differences between these two branches are also important; compare lexical divergences such as the following, reflecting old dialectal variation in Indo-European: Latin *ignis* (= Sanskrit *agnī-*) ≠ Umbrian *pīr* (= Greek *pūr*); Latin *aqua* (= Gothic *ahwa*) ≠ Umbrian *utur* (= Greek *húdōr*). The overall morphosyntactic type was alike for all known Italic languages and dialects and stems directly from the Indo-European proto-language: use of a fully developed system of case declension for the expression of grammatical relations; a rich verbal system inflecting for person, tense, and mood; SOV as the dominant word order; use of prepositions – although in predominantly SOV languages postpositions would be *a priori* more likely to occur.

In prehistoric times, Latino-Faliscan formed a small island in the middle of a linguistic landscape dominated by Etruscan in the Northwest and by Sabellic languages in the North, East, and South. While Faliscan died out without leaving a trace, Latin was destined to become one of the most widespread and influential languages in human history.

3. Latin

3.1. The Falisco-Latin branch of Italic turned out to be one of the world's most successful language families. This is evidently due not to inherent qualities of Latin, but to the military and political superiority of the Romans (*senatus populusque romanus*). In the course of the Roman conquests, Latin spread successively over great parts of the known inhabited world, from the hills of Rome to an empire comprising 6 million km². The main stages of this expansion are the following:

- The conquest of Central Italy (Latin and Samnite Wars, 343–282 BCE) led to the latinization of formerly Etruscan and Sabellic territories.
- When Southern Italy was finally subdued (Pyrrhic War, 280–275 BCE), Messapic vanished completely, and the Greek of Magna Graecia was also doomed to disappear, except for two enclaves which have survived until today in Apulia and Calabria.
- As a result of the First Punic War (264–241 BCE), the great islands of the Tyrrhenian Sea, Sicily, Sardinia, and Corsica, became the first Roman colonies; Greek and Siculian disappeared in Sicily, and Nuragic in Sardinia and Corsica.
- In 222 BCE the Romans conquered Mediolanum (Milano) and established their rule in *Gallia cisalpina*; as a consequence, the languages of northern Italy, Celtic and others, faced extinction.
- The Second Punic War (218–201 BCE) was decisive in the mortal struggle between Rome and Carthage. As a result of the final Roman victory, the Iberian Peninsula came under Roman control, which led to the slow disappearance of most of its indigenous languages: Iberian, Celtiberian, Tartessian, Lusitanian, and several more whose names we do not even know. Iberian was still spoken in the 1st century CE. Basque has survived until today, thus becoming the only extant witness of the variegated world of pre-Indo-European languages not only in the Iberian Peninsula, but in Western Europe in general.
- The Third Punic War (149–146 BCE) ended with the complete destruction of Carthage and its power. As a result, parts of North Africa were latinized, at least in the urban centers, while the rural areas have to this day preserved their Berber languages.
- 146 BCE was also the date of the destruction of Corinth; Macedonia ceased to exist as a political power, and Greece came under Roman rule. But in contrast to the Western parts of its growing empire, the Latin language never had a chance to impose itself over Greek in the East. While Dalmatia was latinized, the Greek-speaking world (including Asia Minor and the Middle East) preserved its language because of the superiority of Greek culture. Latin was confined to administrative and military purposes.
- 122 BCE the Romans established their rule in *Gallia transalpina*, founding the cities of Aquae Sextiae (Aix-en-Provence) and Narbona (Narbonne), capital of the *Provincia Narbonensis*, the future *Provence*.

- Gaius Julius Caesar conquered the rest of Gaul in his famous *bellum gallicum* (58–51 BCE). The urban population rapidly adopted Latin, but Gaulish was still said to be heard in rural environments by Gregory of Tours (538–594 CE). Caesar also made an excursion to Britannia, which was conquered in the following century but never latinized in depth.
- The last Roman conquest was that of Dacia, a region north of the lower Danube. The emperor Trajan vanquished king Decebalus in the Dacian wars (101–106 CE), whereupon the Roman Empire reached its greatest extension. The local population subsequently gave up their Thrako-Dacian languages and adopted Latin.

3.2. Some of the latinized territories were lost in later centuries:

- The spread of Islam in North Africa (foundation of Kairouan in 670) led to its arabization. Henceforth, Arabic coexisted with the original Berber languages. However, Latin was deeply rooted. The famous Arabic geographer Muḥammad al-Idrīsī (1100 Ceuta–1166 Sicily) reports that the Tunisian city of Gafsa (Latin *Capsa*, Arabic *Qafṣa*) had just been “berberized” and that most of its inhabitants speak ‘the African Latin language’ (*al-lisân al-laṭîni al-ifrîqî*, Edrīsî 1866: 104 f., 124). This is a remarkable testimony of the persistence of “Latin” (which certainly means a Romance language derived from Latin) in Africa as late as the 12th century.
- The Iberian Peninsula also underwent partial arabization after the Islamic conquest (battle of Guadalete 712). In Andalusia and the Levant, Arabic coexisted with Mozarabic, the local offspring of Latin, until the 12th or even the 13th century. During the Nasrid reign Granada had become monolingual in Arabic. The Hispanic Reconquista (1085 Toledo, 1248 Sevilla and the Algarve, 1492 Granada) definitely ousted Arabic from Iberian soil, replacing it by Portuguese, Castilian, and Catalan. Thus, de-romanization was followed by re-romanization in the southern and eastern parts of the Iberian Peninsula.
- In the Roman province of Dalmatia, the indigenous Indo-European language Illyrian survived for many centuries in the countryside, while in urban centers latinization was complete. From the 6th century CE on, Avar and later Slavic invasions pushed Latin back to the coastal cities where it developed into Dalmatian. This Romance language, especially the dialects of Ragusa (Dubrovnik) and Vegliot (Krk) played an important role in Mediterranean trade until the 15th century. It survived until 1898 when the last speaker of Vegliot died. Dalmatian was completely replaced by Croatian.
- In Central Europe, the romanized (although probably not completely latinized) regions south of the *limes* were gradually invaded by speakers of Germanic languages during the first millennium. What is today Bavaria, South-West Germany, the Rhinelands, Alsatia, Austria, and Alemannic Switzerland passed from Latin to a form of German at different moments in late Antiquity. This part of the formerly Latin speaking world is often referred to as *Romania submersa*.

3.3. Latin belongs to the classical languages of humanity, together with Greek, Sanskrit, Hebrew, Arabic, and Chinese. This means that it has been fixed once and for all in a given historical moment and transmitted in an unaltered written form over millennia. It also means that it was venerated as a model of linguistic perfection, that its prominent

authors were considered as true “authorities”, and that it had profound and lasting influence on its civilization. Western Europe, the civilization of the Occident, bears the mark of Latin and its linguistic culture to this day, even after the gradual disappearance of Latin itself.

Latin was established as a classical language in the 1st century BCE and the first decades of the 1st century CE, the so-called “golden age” (*latinitas aurea*); the later decades up to the beginning of the 2nd century still belong to the era of classical perfection, they are referred to as the “silver age”. The history of Latin gravitates around this epoch; it can be subdivided into a pre-classical, a classical, and a post-classical period.

3.4. The earliest attestations of Latin appear around 600 BCE. The authenticity of the famous *fibula praenestina* is still a matter of debate, but here I follow Pulgram (1978: 161): “I think it is genuine.” The second-oldest inscription is the no less famous Quirinal vase with the DUENOS inscription, easy to read but difficult to interpret. Other inscriptions from the same period are classified as Faliscan, the status of which relative to Latin (dialect or closely related language) is unclear (see Fortson, this handbook, 3). From that time on, we can trace the history of pre-Classical Latin by the epigraphic evidence from Rome and various parts of its growing dominion. Up to the middle of the 2nd century BCE (conquest of Carthage and Greece) the language was uniform; no differences between “higher” and “lower” styles can be detected. The only observable changes are in the domain of phonetics and morphology: diphthongs are monophthongized, final consonants vanish (*Iovei* → *Iovī*, *iouidicatod* → *iūdicātō*), the old Indo-European ablative merges with the instrumental and the locative (*Gnaivōd* → *Gnaeō*), verbal forms change or are lost (old root subjunctives like *advenat* ‘may arrive’ die out), etc.

From the typological point of view, the language of the old inscriptions is similar to its Italic neighbours. Grammatical relations are expressed by a rich array of nominal and verbal inflections. Word order is flexible, but SOV is predominant; this can be seen in the following sequence of three sentences, two of which are OV and one VO: [*C. L. Scipio Barbatus*] *Taurasia Cisauna Samnio cepit, subigit omne Loucanam opsidesque abdoucit* ‘he took Taurasia, Cisauna, and Samnium, subjects all Lucania and leads away hostages’ (Pulgram 1978: 183). Interestingly, this text, written shortly after 290 BCE, shows clear vestiges of the spoken language on the phonetic level: final *-m* disappears before consonant (*Taurasia*, *Cisauna*, *Samnio*), but is retained before a vowel (*Loucanam*). Such phenomena can be observed elsewhere; in a text dated ca. 200 BCE we read: *Manlio cosol pro poplo* (Pulgram 1978: 1984; Iliescu and Slusanski 1991: 25). The final *-s* of the nominative singular in *Manlius* is dropped, *-n-* before *-s-* in *consul* disappears, and the middle vowel in *populo* is syncopated. Phonetically, these early texts show features already close to the future Romance languages.

3.5. Literature starts in the second half of the 3rd century BCE. For historical linguistics, Plautus (254–184) is the most interesting author; in his comedies we find many features which were obscured under the varnish of the classical language and which surface again many centuries later in Romance. Here are just three textual examples (from Iliescu and Slusanski 1991: 53, 55):

- *nimis bella es*: colloquial *bellus* ‘beautiful’ has survived in all Romance languages; in classical Latin *pulcher* was used instead.

- *desiste percontarier*: colloquial *percontari* ‘to question, to ask’ was not used in classical Latin, but reappears in Spanish (*preguntar*) and Portuguese (*perguntar*).
- *hoc quod fabular*: in classical Latin, ‘to speak’ is expressed by *loqui*; this verb did not survive in any Romance language; instead, in some languages forms derived from *fabulari* are used (Spanish *hablar*, Friulian *fevelâ* etc.).

The military conquest of Greece led to the hellenization of Roman culture. Inevitably, the Horatian formula comes to mind: *Graecia capta ferum victorem cepit et artes / intulit agresti Latio* ‘Captured Greece captivated her fierce conqueror and brought arts into rustic Latium’ (Ep. II, 1, 156 f). Under Greek influence, literary registers were created. The indigenous Saturnian verse, based on stress accent (Livius Andronicus, d. 207 BCE), was replaced by quantitative meters, such as the hexameter (Quintus Ennius, 238–169 BCE). The classical language took shape, becoming phonetically and morphologically regularized. In accordance with Greek models, Cicero (106–43 BCE) created the scientific and philosophical vocabulary which was destined to become the linguistic foundation of Western civilization ever since. Virgil (70–19 BCE), Horace (65–8 BCE), and Ovid (43 BCE–17 CE) set the standards for lyric and epic poetry for millennia to come. Caesar (100–44 BCE) and Tacitus (55–116 CE) were everlasting models for historiographic prose.

3.5. The spoken vernacular became almost invisible for centuries. Only rarely can we catch a glimpse of what was really going on in the spoken language. The catastrophic eruption of Mount Vesuvius in 79 CE had the welcome side-effect that many graffiti on the walls of Pompeii were conserved under thick layers of ashes. One particularly felicitous finding is a distichon (Pulgram 1978: 215 f.) which was preserved in two versions, a classical and a post-classical form. Here is the classical version:

*quisquis amat valeat pereat qui nescit amare
bis tanto pereat quisquis amare vetat*

‘May whoever loves prosper; may he perish who does not know [how] to love;
may he perish twice over whoever forbids [one] to love.’

And this is the colloquial version which shows the degree to which Latin had changed towards the end of the 1st century CE:

*quisquis ama valia peria qui nosci amare
bis [t]anti peria quisquis amare vota*

The second version, despite its errors and vulgarisms, is just as much “Latin” as the first one; neither *quisquis* nor *peria* are “Proto-Romance”. These lines simply reflect the fact that spoken and written registers had definitely become differentiated by that time. A gap between “classical” and “vulgar” Latin had opened. The written literary language became an immutable standard, while the spoken vernacular began to drift away from it, as spoken language can never be stopped in its course. But at the beginning of the millennium, the basic morphosyntactic typology was not yet affected.

What was affected was the phonetic type; vowel quantity, on which the classical meters are based, had disappeared. The “vulgar” form of this distichon cannot be read anymore according to the rules of classical prosody, long and short syllables have be-

come indistinguishable. On the phonetic level, colloquial Latin as reflected in the Pompeian graffiti was definitely on its way towards Romance.

3.6. In the following centuries, a profound change in syntactic typology took place. While preserving its complex nominal inflection and most other typological characteristics mentioned above, Latin evolved from a predominantly SOV to a predominantly VSO language, at least in its spoken form. This evolution must be discussed a bit more in detail.

Classical Latin is notorious for its “free word order”; in verses like the famous Virgilian

*Tityre, tu patulae recubans sub tegmine fagi
silvestrem tenui Musam meditaris avena* (Ecl. 1, 1–2)
‘Tityrus, lying under the cover of a spreading beech you
ponder the forest Song with a thin pipe’

patulae refers to *fagi*, *silvestrem* to *Musam*, and *tenui* to *avena* – adjectives and nouns have, so to speak, been completely scrambled for meter’s sake. Even if it must be assumed that such a use of language had an artificial flavor, it was undoubtedly not only possible, but even highly appreciated. And after all, even in these verses the basic order is SOV (*tu Musam meditaris*)! Of course, prose yields more reliable data than poetry, as far as word order is concerned. Indeed, statistically speaking (S)OV was the most frequently represented word order in all authors analyzed by Linde (1923), from a maximum of 90% in Caesar dropping to approximately 50% in Cicero. This means that individual style played an important role, but not to the point that the basic rules were overturned. Classical Latin, like its Italic ancestors and relatives, was basically a verb-final language.

This picture changed dramatically during the period of the Roman Empire. If we compare documents of “vulgar” Latin with classical texts, the frequency of the basic word-order types is reversed. From the 4th century on, V(S)O turns out to be the most frequent and stylistically unmarked order, at least in texts which are relatively free of the classical norm. A comparison between the respective beginnings of Tacitus’ *Annals* and the Christmas Story in Hieronymus’ *Vulgata* makes this change unmistakably clear:

urbem Romam a principio reges habuere. libertatem et consulatum L. Brutus instituit. dictaturae ad tempus sumebantur. neque decemviralis potestas ultra biennium neque tribunorum militum consulare ius diu valuit. non Cinnae, non Sullae longa dominatio. et Pompei Crassique potentia cito in Caesarem, Lepidi atque Antonii arma in Augustum cessere, qui cuncta discordiis civilibus fessa nomine principis sub imperium accepit. (Tacitus, Ann. I,1)
‘At the beginning, kings ruled the city Rome. L. Brutus established freedom and the consulate. Dictatorships were adopted on occasion. Neither the decemviral power held sway beyond two years nor (was) the consular law of the military tribunes long in effect. Neither the despotism of Cinna nor of Sulla were long. And the power of Pompey and Crassus quickly yielded to Caesar, and the weapons of Lepidus and Antony to Augustus, who, when the world was wearied by civil strife, by the name of ‘Prince’ accepted (it) under (his) authority.’

Factum est autem in diebus illis exiit edictum a Caesare Augusto ut describeretur universus orbis. haec descriptio prima facta est a praeside Syriae Cyrino. et ibant omnes ut

profiterentur *singuli in suam civitatem. ascendit autem et Joseph a Galilaea de civitate Nazareth in Judaeam in civitatem David, quae vocatur Bethlehem, eo quod esset de domo et familia David, ut profiteretur cum Maria desponsata sibi uxore praegnante. factum est autem, cum essent ibi, impleti sunt dies ut pareret. et peperit filium suum primogenitum, et pannis eum involvit, et reclinavit eum in praesepio; quia non erat eis locus in diversorio.* (Luc. 2, 1–7)

‘It happened, however, that in those days an edict went out from Augustus Caesar that the entire world should undergo a census. This census was first enacted by Cyrenus, governor of Syria. And everyone went so that they could be registered, each into his own city. And Joseph also went up from Galilee, from the city of Nazareth in Judea into the city of David, which is called Bethlehem, because he was of the house and family of David, so that he might be registered with Mary who was betrothed to be his wife and pregnant. It happened, however, while they were there, that the days were fulfilled that she would give birth, and she bore her first-born son, and wrapped him in swaddling clothes, and laid him in a manger; because there was not a place for them in the inn.’

Tacitus (55–116 CE) places the verb in final position in 6 out of 6 sentences; Hieronymus (347–419 CE) puts it in initial position in 15 out of 17 sentences (omitting the relative clause as non-indicative of basic word order); only *pannis eum involvit* has the verb at the end. These figures are confirmed if we compare Caesar’s *Bellum Gallicum* (ca. 50 BCE) with Egeria’s *Itinerarium* (ca. 384 CE) or any other classical and overtly post-classical Latin texts. Between the 1st and the 4th century a profound syntactic change has taken place. It must be stressed that the later authors, even writing in an overtly “vulgar” style, still conform strictly with the morphological rules of the classical language; neither the nominal nor the verbal inflection show any indication of decline. As far as morphology is concerned all this is still regular Latin. But the syntactic type is turned upside down. This major change in syntactic typology took place inside Latin itself, before its disintegration into separate languages. VSO was the common starting point for the emerging Romance vernaculars.

4. The Romance languages

4.1. Diastratic and diaphasic variation is normal in a language, all the more so in Latin where, being used as a means of everyday communication by at least 20, perhaps 45 million people, a certain degree of divergence is inevitable. Nevertheless, during the entire period of the Roman Empire (until its division in 495 CE) Latin does not show signs of disruption into major geographical varieties. Latin remained diatopically uniform over a period of centuries. Löfstedt (1959: 48) even posulates a kind of “Received Standard Imperial” which predominated even among the most “vulgar” writers (except, of course, with respect to word order, which was not sufficiently taken into account by earlier scholars). It is only in the second half of the 1st millennium CE that local varieties were disrupted, so as to form new independent languages. As a result of the breakdown of the central power of Rome and of the system of roads (*viae*), regular commercial contact and population exchange declined. The formerly unified Empire disintegrated. Gradually, new languages made their appearance. Latin evolved into Romance.

In the “real world” of spoken language, there is no break in the evolution of Latin. In a certain sense, Latin has never ceased to be spoken, since present day French, Span-

ish, Italian, etc. are modern forms of Latin, transmitted from generation to generation in an uninterrupted chain since Antiquity. But classical Latin in its written form had a life of its own, being preserved – more or less purely – in its original form for two millennia. Inevitably, a moment came when the distance between written Latin and its spoken offshoots was sufficiently great that they were felt to be different languages (Elcock [1960] 1975). There is no sharp boundary in real evolution, but there is a sudden change in the consciousness of speakers. The moment when they become aware of speaking something clearly different from Latin can be considered as the birth of the respective Romance languages.

4.2. This moment came in 842 for French. As Nithard reports in his Latin chronicle, Louis the German swore allegiance in *romana lingua*, whereas his brother Charles the Bald did so in *teudisca lingua*, in the presence of their respective armies. From a strictly linguistic point of view, the Strasbourg Oaths are just an instantaneous snapshot in the long evolution from Latin to French, but their fundamental importance lies in the fact that here a Romance text is explicitly opposed to a surrounding text formulated in Latin. Romance is clearly presented as something different from Latin.

The same “sudden awareness” is documented more than a century later in Italy. The *Placiti Cassinesi* preserve the testimonies of peasants in the surroundings of Monte Cassino in their original, i.e. proto-Italian form, because these peasants evidently had no knowledge of Latin and their depositions had juridical value only if they were written down in their exact wording.

In Spain, the *Glosas Emilianenses* and *Silenses* are added as notes to ecclesiastical Latin texts in order to explain expressions which had fallen into disuse and thus become incomprehensible. The *Glosas Emilianenses* include, in addition to 365 annotations in proto-Spanish (a mixture of Castilian and Aragonese), also two sentences in Basque; the early Romance language of the Rioja region was felt to be an entity almost as different from Latin as Basque! The dating of these glosses is a matter of dispute; traditionally they are thought to have been written ca. 960–980, but new palaeographical analyses seem to push that date forward (ca. 1075).

In the Iberian Peninsula, Romance as a separate entity arose in yet another context, namely in Hebrew and Arabic strophic poetry. Some final verses, or *kharjas*, are in Mozarabic, the Ibero-Romance language of southern and eastern Spain. In such a context there is evidently no reference to Latin whatsoever. The oldest datable *kharja* concludes a Hebrew poem by Yoseph ibn Caprel between 1039 and 1042, that is, in the middle of the two competing datings for the *Glosas* (Bossong 2010: 191–227).

Other languages followed: Occitan ca. 1000 (*Lo poema de Boecis*), Rhaeto-Romance ca. 1100 (*Einsiedler Interlinearversion*), Sardinian 1102 (*Carta arborense*), Portuguese 1192 (*Auto de partilhas*), Catalan 1204 (*Homilies d'Organyà*). Rumanian, spoken far away from the *Romania continua* of Western Europe, is a late-comer; its first known document is a letter of Neacșu of Câmpulung, written in the Cyrillic alphabet and dated in 1521 (Bossong 2008: 270).

4.3. From a typological perspective, the emerging Romance vernaculars were strikingly different from Latin (for this section, see further Bossong 2008). During the “dark centuries” between the end of the Western Roman Empire (476) and the appearances of Romance, profound changes had taken place; these changes are scantily documented and

must be reconstructed mainly from the final results. As we have seen, on the level of macro-syntax the basic change from verb-final to verb-first ([S]OV → V[S]O) was already accomplished in spoken Latin by the end of the Empire. In the subsequent period, the complex system of synthetic nominal declension broke down, giving way to a new analytic type where grammatical relations were expressed by the order of elements and by prepositions. In the West, only in Old French and in Old Occitan (vestigially in Old Rhaeto-Romance) did a much reduced system of two cases survive until the 13th century. The *casus rectus* (derived from the Latin nominative) was used for the subject, the *casus obliquus* (from the accusative) for the direct object and other oblique relations. The two-case system of Gallo-Romance disappeared because of its internal weaknesses: odd syncretisms combined with a counterintuitive distribution of endings (marked *murs* rect.sg. = obl.pl., opposed to unmarked *mur* obl.sg = rect.pl), morphological irregularity, and its absence in most feminine nouns. As a result of this early loss, all modern Romance languages of Western Europe lack case declension.

In Eastern Europe, Rumanian shows case inflection to this day. It has conserved elements of Latin, but the system was restructured, resulting in a two-case system different from Old French and Old Occitan: the first case is used for the nominative and accusative functions, the second one for the dative and genitive functions. Contact with Eastern European languages showing richly developed case systems (Church Slavonic, Greek, Hungarian, etc.) might have favored the preservation of nominal declension in Rumanian.

In many parts of the Romance-speaking world, a system of “Differential Object Marking” (DOM, Bossong 1991, 1998a) arose, creating a distinction between subjects and objects. In Romance languages, such a system consists of a preposition which must accompany an animate noun but is lacking with an inanimate one and is therefore called “differential” (of course, rules are more complex than this simple statement). The preposition used is mostly *a* (from Latin *ad*), but also *onde* in Peruvian Spanish, *da* in the Ligurian dialect of Nicosia in Sicily, and *pre* → *pe* (from Latin *per*) in Rumanian. Such systems frequently evolve in languages which, as a consequence of the decline of nominal inflection, cannot distinguish any longer the main grammatical functions subject and object; e.g. the Semitic languages offer instructive parallels. DOM exists in the three Ibero-Romance languages, in south-western Occitan, in Lower Engadinian, in all Southern Italian dialects, including Corsican, Albanian, and Sicilian, in Sardinian, and in Rumanian.

4.4. Comparing the diachronic dynamics of Indo-European languages, we may note that in most branches noun declension was substantially reduced or disappeared completely (Bossong 2004). Only Lithuanian has preserved the original Indo-European case system almost intact; its closest relative Latvian has reduced it somewhat. Most Slavic languages still express grammatical relations essentially by case endings; only Bulgarian and Macedonian have given up the old case system at a relatively recent date (19th century, Feuillet 1999); today these two varieties of Balkan-Slavic are as analytical as the modern Romance languages. In the Germanic family, only Icelandic and Faeroese still preserve the old case declension system, while in German and Dutch it is in decline; in continental Scandinavian and English it has disappeared. Modern (Insular) Celtic has lost case endings completely. Modern Greek still relies on the remnants of the case system of the ancient language, but in a substantially reduced form. In most Iranian languages, the

Indo-European case system has collapsed, giving rise to new DOM systems (Bossong 1985). Armenian still has four cases, but the most notable exception to this general tendency of simplification is Ossetian, the only Indo-European language with a richer case system than proto-Indo-European or Sanskrit; this Iranian language is surrounded by languages with richly developed case systems (Chechen, Georgian, Daghestanian languages), so that it is not surprising that it has reversed the tendency of case loss which predominates in most other subfamilies of Indo-European, reconstituting a case system which is extremely rich according to Indo-European standards (although only moderately so when compared to Daghestanian languages).

A comparison of Romance with Indo-Aryan is particularly instructive. The system of Sanskrit with its eight cases and three numbers has been reduced drastically in the development of Neo-Indic languages. Modern Hindi has a two-case system comparable to that of Old French. In addition to the distinction between a *casus rectus* and *casus obliquus*, it shows Differential Object Marking, like most other Indo-Aryan languages. Grammatical relations are expressed, not by prepositions but by postpositions, in accordance with the (S)OV syntax predominating in Indo-Aryan (which, in its turn, was favored by contact with Dravidian). This postposition is added to the oblique case endings; it has the form *-ko* in Hindi-Urdu, *-ke* in Bengali, etc. Resuming the foregoing discussion, we can conclude that Romance languages follow a tendency towards simplification of the case system widespread in modern Indo-European.

4.5. In comparison with nominal inflection, verbal inflection is much more conservative. The categories of person, tense, and mood have been preserved in their original form to a much higher degree than nominal case. Only voice behaved differently; the Latin passive was lost completely, and was replaced by analytical constructions of much rarer use than the synthetic passive of Latin. In view of the high frequency of the passive voice in classical Latin, and comparing Latin to Greek, where the passive voice has been preserved to this day, one might wonder why this is so. For the moment, I have no answer, and apparently nobody else has even asked the question, although it is a striking difference between classical Latin and its modern offspring.

In the tense system, the imperfect was the most stable category; it was retained unaltered in all languages, and even served as a basis for the construction of new tense morphemes in Romance Creoles. In contrast, the Latin perfect was lost, or is presently being lost, in numerous Romance languages (Rhaeto-Romance, Northern Italian, spoken French, Sardinian, Rumanian). All Romance languages make use of a new analytic perfect, formed with *esse* and *habere*, or later on with *habere* alone, + past perfect (originally passive) participle. The roots of this periphrastic construction can be detected in classical Latin; it developed into a fully fledged verbal tense in the first half of the 1st millennium CE, thus forming part of the verbal system of common proto-Romance. In Catalan, a new analytical construction, using the auxiliary *anar* 'to go' + infinitive, has replaced the old synthetic perfect of Latin.

The synthetic future of Latin was replaced by analytical constructions in proto-Romance, which in their turn evolved into fused, i.e. synthetic forms again (Fleischman 1982; Schwegler 1990). Since Latin future forms stem generally from older analytical constructions, and since the synthetic future of Romance tends to be replaced by new analytical forms, we can observe in this domain several subsequent cycles from analytic to synthetic, back to analytic, and so forth; this can be exemplified by Spanish:

Italic **cantāfō* → Latin *cantābō* → proto-Romance *cantare habeo* → Mozarabic *cantar-aiō*
 → Standard Spanish *cantaré* → spoken Spanish *voy a cantar* → modern dialects *v'a-cantá*

As this example convincingly shows, the development from synthetic to analytic is not a one-way street. The evolution from Latin to Romance goes from synthetic to analytic in many domains, but surely not in all. At any moment in history, analytical constructions can evolve into synthetic ones.

4.6. The expression of person requires a commentary of its own. While in most Romance languages grammatical person relies uniquely on the verbal endings inherited from Latin (and ultimately from Indo-European), in Northern Italian, Rhaeto-Romance, and also partly in modern spoken French, cliticized subject pronouns have taken this role (Bossong 1998b). These are used either instead of the old endings (as in French *ž-šāt/ty-šāt/i(l)-šāt* instead of classical Latin *cant-o/cant-as/canta-t*) or as a supplement to these (as in Lombardian *a-cant-i/te-cant-et/el-cant-a*). In an extended zone in the center of the Romania, the exclusively suffixing character of verbal conjugation has given place to a partially prefixing conjugation which may, in the long run, change profoundly the typological make-up of these languages.

4.7. As we have seen above, classical Latin SOV had changed to VSO already in the first half of the 1st millennium CE (Dardel and Haadsma 1983; Bossong 2006). VSO is a relatively unstable type which tends universally towards SVO. All Romance languages have undergone the evolution from VSO to SVO during the 2nd millennium CE, but at different speeds. French was first, having completed the grammaticalization of SV at the turn of the 17th century. Portuguese and Italian followed in the 19th century, whereas Spanish still preserves remarkable vestiges of the older VS type up to the present day (Bossong 1998c).

5. Romance-based creoles

5.1. Romance-speaking peoples were the pioneers of the age of discoveries. Portuguese and Spanish seafarers made their way around Africa to South and East Asia, as well as across the Atlantic to the Americas. They were followed by Dutch, French, and English merchants and colonizers – and slave traders. Indigenous people, mainly from West Africa, were deported by force to serve as workers in plantations around the globe. Contacts between masters and slaves required pidgins which later developed into full-fledged creoles. As a consequence, a series of languages was formed which can legitimately be considered as the third generation of the Italic branch of Indo-European: the Romance creoles.

5.2. Whether or not creole languages should be classified as belonging to the families of their respective lexifiers is a thorny question. An alternative is to group them together in a branch of their own; in such a special “family” labelled “Creoles” we find Portuguese, Dutch, or Arabic-based creoles side by side – an unsatisfactory solution from the standpoint of historical-comparative linguistics, despite the undeniable typological

resemblances between creoles of different origin. It seems more appropriate to classify creoles under the heading of their respective lexifier languages. According to this view, the Romance language family includes not only Portuguese, but also Papia Kristang; the Germanic family not only Dutch, but also Berbice; and the Semitic family not only Arabic, but also KiNubi. The genetic relationship of a creole and its lexifier is a special one: language change was dramatic, not gradual; the grammatical system of the lexifier was drastically reduced and then subsequently restructured according to certain recurrent patterns. We cannot delve here into the question of the origin of these recurrent patterns: are they due to a common historic origin, such as a proto-Creole derived from Portuguese? Or are they caused by a common West African substrate? Or are they the effect of universals of human language? Whatever the answer to this question may be, there remains an undeniable fact, namely that historically all creoles derive immediately from their respective lexifiers, although in a special way. This lineal descent from a direct ancestor can best be accounted for by integrating all creoles in the families of their lexifiers.

5.3. Adopting this point of view, we can affirm that Romance-based creoles belong to Romance and, consequently, to the Italic branch of Indo-European, where they represent the latest generation of languages, formed between the 16th and 18th century. There are three Romance lexifier languages: Portuguese, Spanish, and French.

- Portuguese creoles are the most numerous of the three. In Africa they are found in two regions: the Atlantic coast of northwestern Africa (Cape Verde and Guinea Bis-sau), and the Gulf of Guinea (São Tomé & Príncipe, four different languages). In Asia, we can distinguish three zones: India with Sri Lanka (numerous languages and dialects); Malaysia (Papia Kristang); and Macau.
- French creoles are the second largest group. They are found in two regions: in and around the Caribbean, from Louisiana through Haiti and the Lesser Antilles (Martinique, Dominica, Guadeloupe) to French Guiana; and in the Indian Ocean (Seychelles, Réunion, Mauritius).
- Spanish creoles are just three in number: Palenquero, spoken in Colombia; Papiamentto, spoken in Aruba, Bonaire, and Curaçao; and Chabacano, spoken in the Philippines.

5.4. Typologically, all creoles have a number of features in common. In the nominal system, gender has disappeared; the expression of number was simplified. Grammatical relations are expressed by position alone. Differential Object Marking is only found in Papia Kristang, where *kung* (from Portuguese *com*) is used for marking animate objects. The verbal system of Romance was completely restructured. Tense, aspect, and mood (TAM) are expressed by sets of morphologically transparent prefixes which can often be traced to Romance verbs (*te-* PAST from French *était*; *ta-* HABIT from Portuguese *está*). Since endings are lost, person is expressed by cliticized pronouns preceding the verb with its TAM marker in subject function and following it in object function (Cape Verde Portuguese *nn-sera-kónta-ñózi* ‘I will tell you’, Karifuna French *li-kupé-l* ‘he cut it’ (Holm 1988–1989, 2: 274, 381). The syntactic type of creole languages is invariably SVO.

Some creole languages have remained relatively close to their lexifier languages, e.g. Réunionnais with respect to French. Others have drifted far away, e.g. Principense from

Portuguese. A certain degree of creolization has been detected in spoken Brazilian Portuguese, but this is not sufficient to classify it as a full-fledged creole. The mere existence of such transitional forms between ‘dialect of X’ and ‘creole derived from X’ is a strong argument in favor of the solution adopted here, namely to integrate creoles into the families of their respective lexifiers and to overcome the still widespread practice (e.g. www.ethnologue.com) to classify them as a separate class of “creoles”. This argument is further corroborated by the fact that in many creole-speaking communities a “post-creole continuum” has developed, i.e. a whole array of transitional forms between the basilectal “deep creole” and the acrolectal lexifier as a target language.

6. Outlook

The Italic branch of Indo-European has turned out to be one of the most widespread and most influential language families in human history (Janson 2006; Stroh 2007). Language has truly proved to be a “companion of the Empire” (compare Nebrija’s famous dictum: *siempre la lengua fue compañera del imperio*, [1492] 1984: 97). Latin followed the gradual expansion of Roman rule (340 BCE to 106 CE). Portuguese, Spanish, and French followed the European powers in their worldwide expansion (15th to 19th century CE). Latin generated the Romance languages; some Romance languages had creole offspring. But beyond military and political power, Latin and Romance had a tremendous impact on Western civilization and human culture in general. Written Latin became a classical language for two millennia. French left an indelible mark on English, a language of Germanic origin whose vocabulary was overwhelmingly romanized. In Europe and worldwide, Latin (and Greek integrated into Latin) has become the lexical base for the ever-growing technical and scientific vocabulary of the contemporary world; Graeco-Latin is fundamental in all modern languages of civilization. Even if Latin is no longer used as a living language (except in the Vatican), it survives transformed into a means of international communication which has deeply marked even non-Indo-European languages all over the world.

7. References

- Baldi, Philip
 1999 *The foundations of Latin*. Berlin: Mouton de Gruyter.
- Bossong, Georg
 1985 *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuromanischen Sprachen*. Tübingen: Narr.
- Bossong, Georg
 1991 Differential Object Marking in Romance and Beyond. In: Douglas Kibbee and Dieter Wanner (eds.), *New Analyses in Romance Linguistics*. Amsterdam: Benjamins, 143–170.
- Bossong, Georg
 1998a Le marquage différentiel de l’objet dans les langues d’Europe. In: Jack Feuillet (ed.), *Actance et valence*. Berlin: Mouton de Gruyter, 193–258.

- Bossong, Georg
1998b Vers une typologie des indices actanciels. Les clitiques romans dans une perspective comparative. In: Paolo Ramat and Elisa Roma (eds.), *Atti del XXX Congresso Internazionale della Società di Linguistica Italiana*. Roma: Bulzoni, 9–43.
- Bossong, Georg
1998c La inversión narrativa y la tipología del español. In: José Perona, Herminia Provencio Garrigós, Estanislao Ramón Trives, and Agustín Vera Luján (eds.), *Estudios de lingüística textual. Homenaje al Profesor Muñoz Cortés*. Murcia: Universidad de Murcia, 79–88.
- Bossong, Georg
2004 Analytizität und Synthetizität. Kasus und Adpositionen im typologischen Vergleich. In: Uwe Hinrichs (ed.), *Die europäischen Sprachen auf dem Wege zum analytischen Sprachtyp*. Wiesbaden: Harrassowitz, 431–452.
- Bossong, Georg
2006 La sintaxis de las Glosas Emilianenses en una perspectiva tipológica. In: José Jesús Bustos Tovar and José Luis Girón Alconchel (eds.), *Actas del VI Congreso Internacional de Historia de la Lengua Española*. Madrid: Arco, 529–543.
- Bossong, Georg
2008 *Die romanischen Sprachen. Eine vergleichende Einführung*. Hamburg: Buske.
- Bossong, Georg
2010 *Poesía en convivencia. Estudios sobre la lírica árabe, hebrea y romance en la España de las tres religiones*. Gijón: Trea.
- Dardel, Robert de and Rinze Haadisma
1983 Déclaratives romanes à verbe initial. *Vox Romanica* 42: 24–39.
- Edrisi
1866 *Description de l'Afrique et de l'Espagne*. Edited and translated by Reinhart Dozy and Michael Jan De Goeje. Leiden: Brill.
- Elcock, William Denis
1975 [1960] *The Romance languages*. 2nd edn. London: Faber and Faber.
- Ernout, Alfred and Antoine Meillet
1985 [1932] *Dictionnaire étymologique de la langue latine*. 4th edn. Paris: Klincksieck.
- Feuillet, Jack
1999 *Grammaire historique du bulgare*. Paris: Institut d'Études Slaves.
- Fleischman, Suzanne
1982 *The future in thought and language. Diachronic evidence from Romance*. Cambridge: Cambridge University Press.
- Holm, John
1988–1989 *Pidgins and creoles*. 2 vols. Cambridge: Cambridge University Press.
- Holtus, Günter, Michael Metzeltin, and Christian Schmitt
1996 *Lexikon der Romanistischen Linguistik (LRL)*. II, 1. *Latein und Romanisch. Historisch-vergleichende Grammatik der romanischen Sprachen*. Tübingen: Niemeyer.
- Iliescu, Maria and Dan Slusanski
1991 *Du latin aux langues romanes. Choix de textes traduits et commentés (du I^e siècle avant J.C. jusqu'au X^e siècle après J.C.)*. Wilhelmsfeld: Egert.
- Janson, Tore
2006 *Latein. Die Erfolgsgeschichte einer Sprache*. Hamburg: Buske.
- Linde, Paul
1923 Die Stellung des Verbs in der lateinischen Prosa. *Glotta* 12: 153–187.
- Löfstedt, Einar
1959 *Late Latin*. Oslo: Aschehoug.

- Nebrija, Antonio de
1984 [1492] *Gramática de la lengua castellana*. Antonio Quilis (ed.). Madrid: Editora Nacional.
- Palmer, Leonard R.
1954 *The Latin language*. London: Faber and Faber.
- Pulgram, Ernst
1958 *The tongues of Italy. Prehistory and history*. Cambridge, MA: Harvard University Press.
- Pulgram, Ernst
1978 *Italic, Latin, Italian. 600 B.C. to A.D. 1260. Texts and commentaries*. Heidelberg: Winter.
- Rix, Helmut
1998 *Rätisch und Etruskisch*. Innsbruck: Institut für Sprachwissenschaft, der Universität.
- Schwegler, Armin
1990 *Analyticity and Syntheticity. A Diachronic Perspective with Special Reference to Romance Languages*. Berlin: Mouton de Gruyter.
- Stroh, Wilfried
2007 *Latein ist tot, es lebe Latein! Kleine Geschichte einer großen Sprache*. Berlin: List.
- Vennemann, Theo
2003 *Europa Vasconica – Europa Semitica*. Berlin: Mouton de Gruyter.
- Walde, Alois and J. B. Hofmann
1938–1956 *Lateinisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Woodard, Roger D. (ed.)
2004 *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge: Cambridge University Press.

Georg Bossong, Zurich (Switzerland)

IX. Germanic

53. The documentation of Germanic

1. Proto-Germanic
2. North Germanic
3. East Germanic
4. West Germanic
5. References

1. Proto-Germanic (PGmc.)

Proto-Germanic (*Urgermanisch*) is the ancestral language of the attested old Germanic (Gmc.) languages. As the comparative method of historical linguistics requires, PGmc. is reconstructed dialect-free (*thesauri*: Falk and Torp 1909; Orel 2003; Kroonen 2013) but it is almost certain that it never was a uniform proto-language.

It is unclear when and where PGmc. evolved as a “condensation” of an individualized culture or of a quasi-individualized *ethnos*. A likely approximate date for the *Ausgliederung* of Gmc. is 500 BCE. Germanization can hardly be linked to the expansion of the Jastorf culture alone (Müller 1999: 316; Steuer 1999: 326).

Factual documentation from the earliest times is provided in only a few external sources (*Nebenüberlieferung*) of Latin and Greek provenience. In a *periplus*, Pytheas of Massilia (ca. 380–310 BCE) describes his journey to ancient North Sea territories reaching the island of Θούλη; unfortunately, the text has not survived – only one quotation and a few paraphrases remain – so that the names mentioned in the extant Pytheas excerpts are of no significance. By the medium of epigraphic sources we reach the late 3rd c. BCE; the oldest source is the so-called Protogenes inscription from Olbia on the northern coast of the Black Sea in which the (East) Gmc. tribe of the Σκίροι (‘the pure ones, purebreds’; cf. Goth. *skeirs** ‘clear’ etc.) is mentioned (Σκιροῦς acc. pl., Syll. 495; LaN I: 592). A few decades later, the antonymic tribe name of the Βαστάρνοι (‘the bastards, underbreds’; cf. NHG *Bastard* etc.) is recorded several times in Greek inscriptions as well as in literature (LaN I: 117 ff.). The first Gmc. *gentes* that came into contact with Rome were the *Cimbri* and the *Teutones*, who challenged the Roman empire at the end of the 2nd c. BCE (LaN I: 218 ff., 656 ff.; substitutes for Gmc. **Himbrōz* and **Deudanōz*, cf. *Himber-sysæl* and *Thythæ-sysæl*, two areas in northern Jutland known from 1231 CE).

The most important record from pre-Christian times is the inscription on helmet B of Ženjak-Negau (Slovenia), presumably from the 3rd or 2nd c. BCE. Incised in north Italic (Venetic) letters, the inscription *Harigasti Teiwā* (cf. PGmc. **harja-* ‘army’ < PIE **korǵo-*, **gasti-* ‘stranger, guest’ < **g^hosti-*, **teiwa-* ‘god’ < **deǵuó-*) is best interpreted as a possessor’s inscription (Nedoma 1995); it seems that the helmet belonged to a Gmc. soldier who was involved in combat in pre-Roman northern Italy. Another early Gmc. anthroponym is attested in east Celtic coinage: a set of Boiiian silver tetradrachms, coined in the Bratislava region in the mid-1st c. BCE, exhibits a legend *Fariarix* (LaN I: 267;

cf. PGmc. **farjan-* ‘ferryman’, **rīk-* ‘ruler’; see Birkhan 1971: 28 ff.), obviously a Gmc. name of a Celtic sovereign.

To a large extent, we have to rely on the data provided by Roman and Greek authors; most of them had never been in direct contact with Gmc. people. However, sets of Gmc. names are recorded first by Caesar (ca. 52–50 BCE); our major sources of the first two centuries CE are the geographical works of Strabon and Ptolemy, the *Naturalis historia* by Pliny the Elder and, particularly, the *Germania* written by Tacitus. There are several hundred Gmc. names from the Pre-Migration period (collected in LaN I–II; cf. furthermore Neumann 1953: 53 ff. [ad Scand. Θρουσκάνος, dated to the 1st c. CE]; Polomé 1994: 3 ff.; Tiefenbach 1995), sometimes mutated by means of sound substitution. The same occurs with Gmc. appellatives in early Latin and Greek texts (cf., e.g., Neumann 1994: 95 ff.; Nedoma 2008: 55 ff. [ad **álhiz* ‘elk’]). The earliest external sources – whether epigraphical or literary – provide no secure evidence for the *Ausgliederung* of the Gmc. dialects.

As for Gmc. loanwords in contact languages, the most important borrowings appear in the Baltic-Finnic and Sami branches of Finno-Ugric (dictionary: Kylstra et al. 1991–2012). The oldest loan layers – they can only be dated relatively – possibly trace back to PGmc. times; however, it is doubtful that there are any Pre-Gmc. borrowings (Ritter 1993; on the contrary, *inter al.* Koivulehtu 2002: 586 ff.). Because of the partially conservative phonetic character both of Baltic-Finnic and Sami, though, some of the borrowings reflect a rather archaic *Lautstand* of the Gmc. base form; thus, Finn. *rengas* ‘ring’ (← **hrengaz* masc.) and *kenno* ‘cell’ (← **hennōn* fem. > OIcel. *hinna* ‘membrane’) do not yet show the Gmc. raising of *e* to *i* before the cluster *NC*.

2. North Germanic (NGmc.)

2.1. Ancient Norse (AN)

During the first two centuries CE the Gmc. dialect continuum covered roughly the territory between the Rhine in the west, the Vistula in the east, and the Danube in the south, including Denmark and southern Scandinavia in the north (cf., e.g., Seebold 1998: 297 f.). The earliest internal documentation within this “core *Germania*” is of Scandinavian provenance, where ca. 350 runic inscriptions have survived, written in the so-called older *fuþark* consisting of 24 letters.

The origin of the runes is controversial (cf. Düwel 2008: 175 ff.). However, there are more formal resemblances to letters of pre-Christian North Italic alphabets (in particular of the Val Camonica type; cf. Schumacher 2007: 336) than to Latin letters; admittedly, there is a chronological gap, since the earliest reliable runic inscriptions – the Vimose comb (RäF 26) and the Øvre Stabu lancehead (RäF 31) – are archaeologically dated to the 2nd half of the 2nd c. CE. The language of the early Scandinavian runic inscriptions, in some respects close to PGmc. (cf. Krause 1971: 23 ff.; Nielsen 2000: esp. 271 ff.; 2002: 615 f.), is called Ancient Nordic (or Proto-Nordic, NHG *Urnordisch*, Scand. *urnordisk*). Recently, Nielsen (2000: 77 ff., 89, 294 f.) has pointed out that unstressed PGmc.

Tab. 53.1: The older *fubark* (normalized forms) – The division in three groups (OÍcel. *ættir* literally ‘families’) of eight runes is long-established. The items in the 4th row are for presentational purposes only: as far as can be seen, the runes had no numerical values. Rune no. 4 𐌱 **a** in the early Scandinavian inscriptions seems to represent unstressed *ǣ*, too (Nedoma 2005). For the PGmc. rune names (**fehuⁿ* ‘cattle, goods’, **ūruz* ‘aurochs’, ..., **ōþalaⁿ* ‘[inherited] property’), see Nedoma (2003: 558 ff. [with lit.])

rune	ƿ	ᚱ	ᚲ	ᚳ	ᚴ	ᚵ	ᚶ	ᚷ
translit.	f	u	þ	a	r	k	g	w
phoneme	/f/	/u(:)/	/þ/	/a(:)/	/r/	/k/	/g/	/w/
number	1	2	3	4	5	6	7	8
name	<i>*fehuⁿ</i>	<i>*ūruz</i>	<i>*þurisaz</i>	<i>*ansuz</i>	<i>*raidō</i>	<i>*kaunaⁿ?</i>	<i>*gebō</i>	<i>*wunjō?</i>

rune	ᚨ	ᚩ	ᚪ	ᚫ	ᚬ	ᚭ	ᚮ	ᚯ
translit.	h	n	i	j	ī	p	z (r)	s
phoneme	/h/	/n/	/i(:)/	/j/	/i(:)/	/p/	/z/ (/r ₍₂₎ /)	/s/
number	9	10	11	12	13	14	15	16
name	<i>*haglaz</i>	<i>*naudiz</i>	<i>*īsaⁿ</i>	<i>*jēraⁿ</i>	<i>*ei^h/_waz</i>	<i>*perþō??</i>	<i>*algiz??</i>	<i>*sōwulō</i>

rune	ᚹ	ᚺ	ᚻ	ᚼ	ᚽ	ᚾ	ᚿ	ᚰ
translit.	t	b	e	m	l	ŋ	d	o
phoneme	/t/	/b/	/e(:)/	/m/	/l/	/ŋg/	/d/	/o(:)/
number	17	18	19	20	21	22	23	24
name	<i>*Teiwaz</i>	<i>*berkanaⁿ</i>	<i>*ehwaz</i>	<i>*mannaz</i>	<i>*laguz</i>	<i>*Ingwaz</i>	<i>*dagaz</i>	<i>*ōþalaⁿ</i>

-ōⁿ* on the one hand and **-ō*, **-au^o* on the other hand merged into AN *-ō(°)* – cf. **run-o acc. sg. fem. *ō*-stem (Einang, 4th c.; RāF 63), **arbijan-o** gen. pl. masc. *n*-stem (Tune, ca. 400; RāF 72), **mag-o-z** gen. sg. masc. *u*-stem (Vettelund, 4th c.; RāF 60) > OÍcel. *-a*, *-a*, *-ar* – so that the language of the early Scandinavian inscriptions cannot represent the predecessor of (Ingvaenic) WGmc. (PGmc. **-ōⁿ* > OE OFris. *-æ* > *-e* vs. PGmc. **-ō*, **-au^o* > OE OFris. *-a*). If there was something like a Northwest-Gmc. unity (as was claimed, among others, by Antonsen 2002: 31 ff.), the split occurred not later than the late 3rd c.; this is proved by the new found runic inscription **kab-a** = WGmc. *ka(m)b-a* ‘comb’ on the Frienstedt comb (Nedoma and Düwel 2012: 136 ff.) that shows WGmc. loss of final *z*.

“Classical” AN is a dialect-free *Trümmersprache* that remained practically unchanged for several centuries, viz. until the beginning of a transitional period in the late 5th c. (NHG *Spätturnordisch*, Scand. *yngre urnordisk*). Most of the AN runic inscriptions (the standard edition is RāF) are merely short texts found on portable objects like jewelry (esp. fibulae, bracteates) and weapons as well as on non-portable objects like stones. A considerable part of the runic texts is of a profane kind

(manufacturer's, possessor's and carver's inscriptions, memorial inscriptions) although there are a few religious and magic inscriptions as well. The runic epigraphic tradition assumes prime importance because it represents intentional, original, and contemporary documentation (within an oral culture of remembrance). One of the most famous AN inscriptions is engraved on the Gallehus gold horn B (ca. 400; RāF 43): *ek, Hlewagastiz Holtijaz, horna tawiðō* 'I, Hlewagastiz Holtijaz (a second name, a patronymic or even a denomination of provenance: 'wood-dweller'), made the horn' is the earliest example of a regular Gmc. alliterative verse.

2.2. Old Norse (ON)

Around 700 there is a caesura in the history of NGmc. epigraphical documentation, insofar as the 24-letter older *fupark* was replaced by the 18-letter younger *fupark* (cf. Schulte 2006 [with lit.]) that is attested in two main variants, long-branch and short-twig runes, as they are called. There are only a few extant runic inscriptions from the 8th c., but as of the beginning of the Viking Age (ca. 800) the source material increases significantly over the next centuries up to a total of ca. 6,000–6,100 inscriptions in the younger *fupark* known today, including ca. 3,500 inscriptions from Sweden alone (dictionary of Viking-Age runic Swedish: Peterson 1994). It is a very rare case in the history of script that a reduced grapheme inventory is used for an extended phoneme inventory (referring to ON, following umlaut and breaking processes during the transitional period; cf. Andersson 2002: 297 ff. [with lit.]). Because of the plurivalent phoneme-grapheme correlations it is difficult to examine the sound change processes that yield dialect divisions within ON (mainly, between East and West Norse varieties). The runic epigraphical tradition, however, lasts throughout the Middle Ages, with runes being used for "functional" texts of various kinds (memorial inscriptions on stones, situational private messages on wooden sticks, etc.).

Old Icelandic (OIcel.), the most conservative ON language and thus sometimes – inaccurately – called ON (*per se*), is definitely a *Großkorpussprache* (OIcel. texts are enregistered by Simek/Pálsson 2007; standard dictionaries: Fritzner [1886–1896] 1972; Egilsson and Jónsson 1931; ONP 1989 ff.; *etymologica*: de Vries 1962; Blöndal Magnússon 1989). The OIcel. vernacular literary tradition is extensive, and many of the texts reproduce ancient ("Gmc.") *fabulae*, plots, and motives. The earliest extant manuscripts can be dated to the mid-12th c.; the oldest written sources are, as expected, non-fiction texts on Christian, legal, and historical matters. One of the most important pieces is the *Íslendingabók* ('Book of Icelanders') written by Ari Þorgilsson, describing the early history of Iceland. Furthermore, several genres of prose literature – in particular Kings' Sagas (*konungasögur*), Family Sagas (*Íslendingasögur*), and Legendary Sagas (*fornaldarsögur*) – originate in the period from the late 12th to the early 13th c. As to date of origin, the oldest texts are poetic: several Eddic and Skaldic poems go way back even to the late 9th c.; probably the earliest text is the *Ragnarsdrápa* ('[laudatory] Poem addressed to Ragnarr') by Bragi Boddason. From the time of their composition, Eddic and Skaldic poems had been memorized and passed down orally for several centuries (cf. Jónsson 1921: 236 ff.), until they were fixed in 13th c. manuscripts (such as the

Codex regius of the famous *Poetic Edda*, containing a number of mythological and heroic poems, written ca. 1270).

Compared to OIcel., the other ON languages are documented to a lesser extent. The earliest of the few manuscripts of Old Norwegian – that was still close to OIcel. in the early 13th c. – are from ca. 1150–1200; the vocabulary is included in the “ON” dictionaries (see above). Similar to Iceland, the oldest written sources in Norway are non-fiction texts of Christian and legal kinds. During the reign of king Hákon Hákonarson (Hákon IV, 1217–1263), courtly literature flourished, especially prose translations of Arthurian romances and *chansons de geste* (Chivalric Sagas, *riddarasögur*).

The literary attestation of the East Norse languages is even more sparse. Most of the Old Swedish and Old Danish texts (recorded from ca. 1250) are on legal matters; fictional literature stays on the sideline (cp. Nedoma 2010: 157 ff., 166 ff. [with lit.]). Old Gutnish (on the isle of Gotland; recorded from 1350) is a *Trümmersprache*.

3. East Germanic (EGmc.)

3.1. Gothic (Goth.)

Migrations of the EGmc. *gentes* from the Baltic Sea coast in a southeastern direction during the (2nd and) 3rd c. CE caused EGmc. to be separated from the common Gmc. dialect continuum. There is only a small number of EGmc. runic epigraphic texts (less than 10, most of them probably Goth.) spread over eastern and central Europe; the earliest inscriptions are from the 1st half of the 3rd c. CE. One of the most interesting of these is found on the golden neck-ring of Pietroassa in Romania (1st half of the 5th c.; *Gutanī ō(pal) wī(h) hailag* ‘property of the Goths, sacred and holy [or: sacrosanct]’ RÄF 41), where **ai** (in **hailag**) definitely renders a diphthong /ai/.

Goth. is the earliest Gmc. language preserved in a longer text; unfortunately, its documentation is rather limited in regard to both quantity and quality (Goth. texts are enregistered by Braune/Heidermanns 2007: 6 ff.; edition/glossary: Streitberg 2000; lemmatized concordance: Snædal 1998; standard *etymologicon*: Feist 1939 [or Lehmann 1986]). Except for another few short texts, there is one (main) source by one author, viz. the Bible translation of the Visigothic bishop Wulfila (Οὐλφίλας); thus we are familiar with only one variety of Gothic.

Wulfila’s Bible (ca. 350–380) is based on a Greek source; it has survived in one epigraphical testimony (a folded lead tablet from Hács Béndekpuszta in Hungary, late 5th c.; cf. Schwab 2005: 101 ff.) and several fragmentary manuscripts (5th/6th c.; the most famous is the Uppsala Codex argenteus) that cover the greater part of the New Testament. Wulfila invented his own alphabet, deriving most of the letters from Greek, whereas some seem to be borrowed from Latin (e.g., <h>, <r>) and the older *fupark* (e.g., <u>, <o>).

Late (i.e. post-biblical Goth.) developments can be observed in names, e.g., *ai* > *ē* in *Gesila* (6th c., LaN I: 353; cf. Wagner 2002: 266) vs. *Radagaisus* (died a. 406, LaN I: 546 f.). Gothic became extinct after the collapse of the *regna* of the Ostrogoths in Italy (mid-6th c.) and the Visigoths in Spain (early 8th c.).

Tab. 53.2: Wulfila's Gothic alphabet (S-type or type II) – /i:/ is rendered by <ei>, /e/ [ɛ] by <ai> *ai* and /o/ [ɔ] by <au> *ai*; runic and external evidence (**hailag**, s. above; *Radagaisus*, s. below) suggest that <ai> and <au> represent /ai/ and /au/, too. ↑, the sign for '900', is attested only in the so-called Salzburg-Vienna Alcuin codex (ÖNB Wien, MS 795, fol. 20^v; entry ca. 800 or early 9th c.). The letter names are preserved *ibidem*, most of them exhibit somewhat curious forms (e.g., *eyz* ~ Goth. **aihus* 'horse?', *noicz* = °*ts* ~ Goth. *naups* 'need'). Most (but not all) of these "crypto-Goth." terms are identical with the rune names of the older *futhorc* (see above, Tab. 53.1)

letter	ⱦ	Ⱨ	ⱨ	Ⱪ	ⱪ	Ⱬ	ⱬ	Ɑ	Ɱ
translit.	a	b	g	d	e	q	z	h	þ
phonem	/a(:)/	/b/	/g/	/d/	/e:/	/k ^w /	/z/	/h/	/p/
num. value	1	2	3	4	5	6	7	8	9
name	<i>aza</i>	<i>bercna</i>	<i>geuua</i>	<i>daaz</i>	<i>eyz</i>	<i>quertra</i>	<i>ezec</i>	<i>haal</i>	<i>thyth</i>

letter	Ɐ	Ɒ	ⱱ	Ⱳ	ⱳ	ⱴ	Ⱶ	ⱶ	ⱷ
translit.	i (i)	k	l	m	n	j	u	p	–
phonem	/i/	/k/	/l/	/m/	/n/	/j/	/u(:)/	/p/	–
num. value	10	20	30	40	50	60	70	80	90
name	<i>iiiz</i>	<i>chozma</i>	<i>laaz</i>	<i>manna</i>	<i>noicz</i>	<i>gaar</i>	<i>uraz</i>	<i>pertra</i>	–

letter	ⱸ	ⱹ	ⱺ	ⱻ	ⱼ	ⱽ	Ȿ	Ɀ	↑
translit.	r	s	t	w	f	x	h	o	–
phonem	/r/	/s/	/t/	/w/	/f/	[x]	/h ^w /	/o:/	–
num. value	100	200	300	400	500	600	700	800	900
name	<i>reda</i>	<i>sugil</i>	<i>tyz</i>	<i>uiinne</i>	<i>fe</i>	<i>enguz</i>	<i>uuaer</i>	<i>utal</i>	–

3.2. Minor EGmc. languages

The documentation of the other old EGmc. languages – according to Procopius (*de Bello Gothico* III,2,5), they were close to Gothic – is very poor (cf. Francovich Onesti 2002; Tischler 2003: 340 ff.). Most of them have only survived in a few names. There are two Vandalic syntagmata (viz. *froia arme* [cf. Tiefenbach 1991: 251 ff.] and *eils ... scapia matzia ia drincan* [Anth. Lat. I, no. 285; the "Gothic" epigram]) and one (probably) Burgundian runic inscription (on the Charnay fibula, 2nd half of the 6th c.; R&F 6).

The latest attested EGmc. language is called Crimean Gothic. It was not until 1560–1562 that Ogier Ghislain de Busbecq, a Habsburg diplomat in Constantinople, handed down a list of several dozen vocables and three lines of a *cantilena* in a Gmc. dialect spoken in the Crimea (cf. Stearns 1989 [with lit.]); most of the 101 recorded forms (e.g., *ada* 'ovum' ~ Biblical Goth. **addi*, showing *Verschärfung*) suggest a late EGmc. dialect.

4. West Germanic (WGmc.)

4.1. Early West Germanic

The N/WGmc. “residual” dialect continuum was broken after Angles, Jutes, and (parts of the) Saxons left their homelands to settle in Britain: thus, a language border developed between NGmc. and WGmc. by the 6th c CE. The most important early sources from the WGmc. area are Latin dedication inscriptions addressed to *matronae* (or *matres*) in the Rhineland, dated ca. 160–250/260 CE. The majority of the venerated mother goddesses bear Gmc. (by-)names (cf. Neumann 1987), and a few of the theonyms exhibit post-PGmc. loss of unstressed short vowels in disyllabic endings *^oǃmǃz > *^oǃmz (e.g., *Aflims* CIL XIII 8157, *Vatvims* CIL XIII 7892 < PGmc. *^omiz dat. pl.). The Fienstedt comb runic inscription *ka(m)ba* ‘comb’ dating to the late 3rd c. (s. above) is of particular interest, since it proves that apocope of *z* occurred before loss of thematic vowel *a* (thus, PGmc. **-az* > WGmc. *-a* > OS OE OFris. OHG *-Ø*; cf. Nedoma/Düwel 2012: 141 ff.).

4.2. Old Saxon (OS; also Old Low German)

We know of only a few runic texts from Saxony. As for their linguistic significance, two 5th c. inscriptions indicate WGmc. gemination (pre-OS **kunni** < PGmc. **kunja*- ‘kin, kind’, Weser bones) and weakening of the linking vowel after a heavy syllable in compounds (**algu-skapi** = (*a*)*lgǝ*- < **algiskapi* ‘stag damage’, Wremen footstool; cf. Nedoma 2008: 57 f.).

Like Goth., OS is a *Kleinkorpussprache*, and like in Goth., the documentation focuses on biblical works (OS texts are enregistered by Krogh 1996: 111 ff.; dictionary: Tiefenbach 2010). The major OS literary source is the *Heliand* (‘Savior’), a gospel harmony that was probably composed between 830–840; this voluminous work with its 5,983 alliterative verses is preserved in two large codices, one of them nearly complete, and four fragments (including that from Leipzig; Schmid 2006). Apart from the *Heliand*, only a limited number of shorter (mostly Christian) texts of different kinds have survived. The OS onomastic evidence is meagre.

The descendant of OS is called Middle Low German (MLG; ca. 1200–1600), the language of the Hanseatic merchants.

4.3. Old English (OE)

The English runic tradition starts in the late 5th c. Some 80–90 inscriptions have survived, written in an extended Anglo-Frisian runic alphabet called *futhorc* (better: *fuporc*) consisting of 26–33 characters: the inclusion of additional runes was caused by sound changes (no. 26 ǃ **a**₁ renders /æ(:)/, no. 25 ǃ **a**₂ /a(:)/, no. 4 ǃ **a**₃ /o(:)/, no. 24 ǃ **o** /ø(:)/, etc.) that had probably occurred already on the continent. Familiarity with and usage of runes outlasted the time of the Anglo-Saxon Christianization: the latest inscriptions date from the 10th c.

It is worth mentioning that in England from the very beginnings of the Anglo-Saxon settlement both Latin and the vernacular language were employed as media for writing (most of the OE texts are enregistered by Wenisch 1979: 19 ff.; [outdated] dictionary: Bosworth, Toller, and Campbell 1898–1972, cf. furthermore DOE; etymological concordance: Holthausen 1963). The earliest known OE text is king Æþelberht's code of Kentish laws that was written around 600 CE but has survived only in a 12th c. manuscript; the earliest extant texts are glossaries and charters of the (early) 8th c. The great bulk of OE poetry is preserved in four manuscripts, all of them written in the decades around 1000 CE in (late) West Saxon; the original texts, most of them of Anglian provenience, date back a few centuries earlier. The most famous piece is *Beowulf*, a heroic epic poem in 3,182 alliterative verses that is set in Scandinavia; the text, an artificial meshwork of intertextuality, represents a kind of *summa litterarum* (Harris 1985: 260 ff.) but also contains quite a few oral-formulaic patterns. Aside from further heroic lays in ancient (“Gmc.”) tradition, a number of OE poems deal with religious heroes. As for prose, many Latin texts were translated into the vernacular in the time of Alfred the Great; comprehensive medical texts and annalistic literature (in particular the important *Anglo-Saxon Chronicle*) date to the 9th c., too. Representative for late OE (or, to be more exact, late West Saxon) are the numerous works of abbot Ælfric, named *Grammaticus*.

The descendant of OE is Middle English (ME; ca. 1100–1400/1500), the most prominent text of which is Geoffrey Chaucer's *Canterbury Tales*.

4.4. Old Frisian (OFris.)

Approximately 20 runic inscriptions are associated with early medieval Frisia, written in (a variant of) the Anglo-Frisian *futhorc*. The pre-OFris. inscriptions, dating from the 6th–9th c., preserve an archaic feature: PGmc. nom. sg. *-az > WGmc. -a is retained as -u = -ǫ (e.g., in **a₂dug₂islu** = -gīslǫ < *-gīslaz on the Westeremden weaving slay, ca. 800; cp. Nedoma 2014) as against the other WGmc. languages, where it has been lost.

The literary tradition starts with psalter glosses from ca. 1200. OFris. is a *Kleinkorpusssprache*: apart from the late charters, there are somewhat more than 20 extant manuscripts that stem from a period ranging from the late 13th c. to ca. 1600. Most of the manuscripts were written in the area east of the Lauwers (enregistered by Johnston 2001: 571 ff.; cf. furthermore Bremmer 2004; dictionary: Hofmann and Popkema 2008; *etymologicon*: Boutkan/Siebinga 2005 [merely deals with one codex, R₁]). Most of the OFris. texts concern laws, treatises, privileges, and statutes (some of them probably originating in the 11th c.). Though the extent of the Old Frisian legal tradition is unique within the old Gmc. languages, it reflects only a particular segment of medieval Frisian culture.

4.5. Old High German (OHG)

Some 80–90 runic inscriptions have been discovered in the (later) High German area, all on portable objects. The majority of the inscriptions are found in southwestern Germany and date from the 6th c.; soon after 600 CE the runic tradition ends abruptly due

to the growth of Christianization and the change of mortuary practice. The attested forms exhibit clearly pre-OHG features (Nedoma 2006: 129 ff.) with the sole exception of **dorih** = *Dōr(r)īχ* on the Wurmlingen lancehead (RäF 162), the first attested instance of the second sound shift (< *-rīk). Vernacular terms in the *leges barbarorum* (6th–8/9th c.), in particular the so-called Malberg glosses of the *Lex Salica*, and a mass of OHG names (as of ca. 730 CE) provide further early evidence.

During the Carolingian period, canonized Latin texts were glossed, translated, and annotated in ecclesiastical scriptoria, almost each of them using different dialects or dialect mixtures (OHG texts are enregistered by Schützeichel 2006: 9 ff.; dictionaries: Karg-Gasterstädt 1968 ff. and Schützeichel 2006, cf. furthermore Seebold et al. 2001; *etymologicon*: Lloyd et al. 1988 ff.). It deserves mention that approximately two-thirds of the OHG vocabulary is preserved by means of glosses: there are ca. 250,000 entries (representing ca. 27,000 lexemes) in manuscripts covering the whole of the Middle Ages. The earliest OHG glosses are found in the *Maihinger Evangeliar* from Echternach (dry-point; early 8th c., Middle Franconian); the famous *Abrogans* glossary (ca. 750–770, Bavarian) was written a few decades later. The first literary texts, such as the *Isidor* and the Monsee-Vienna fragments, date to the late 8th c.; the only OHG example of ancient (“Gmc.”) heroic poetry, the *Hildebrandslied* (a fragment of 68 [64] alliterative verses), may be even older. Unique pre-Christian remnants are the two *Merseburg charms*, recorded in a 10th c. manuscript; their actual age remains uncertain. Most of the longer OHG texts were written in the 9th c., e.g., the *Tatian* (East Franconian) and Otfrid’s Gospel harmony (South Rhenish Franconian); the most important author of the late OHG period was the St. Gall monk Notker, named *Labeo* (died 1022).

The successor of OHG is Middle High German (MHG; ca. 1050–1350); its most prominent works, such as the *Nibelungenlied* and the Arthurian romances, were written in the classical period of MHG (ca. 1170–1250).

4.6. Minor WGmc. languages

Langobardic is a *Trümmersprache*: except for three 6th c. runic inscriptions, the only extant material consists of vernacular appellatives and names in *Nebenüberlieferung* (6th–10th c.; cf. Francovich Onesti 1999). However its linguistic status is to be defined, Langobardic in Italy shows the second sound shift (e.g., *zāva* ‘[seditious] uniting’: Goth. *tēwa** ‘order’).

Old Dutch is a sparsely attested Franconian variety that did not undergo the second sound shift (called also Old Low Franconian, cf. Krogh 1997). Its most important source is the Limburgish psalter fragments of the now lost Wachtendonck codex (10th c.). Old Dutch is succeeded by Middle Dutch (ca. 1150–1500).

5. References

- Andersson, Thorsten
 2001 Nordgermanische Sprachen. In: Beck et al. (eds.), vol. 21, 289–306.
 Antonsen, Elmer H.
 2002 *Runes and Germanic Linguistics*. Berlin: Mouton de Gruyter.

- Bandle, Oskar, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman (eds.)
 2002 *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*, Vol. 1. Berlin: De Gruyter, 703–719.
- Beck, Heinrich, Dieter Geuenich, Hans Kuhn, Kurt Ranke, Heiko Steuer, and Richard Wenskus (eds.)
 1968–2007 *Reallexikon der germanischen Altertumskunde, begründet von Johannes Hoops*. 2nd edition. 35 Vols. Berlin: De Gruyter.
- Birkhan, Helmut
 1971 Die “keltischen” Personennamen des boiischen Großsilbers. *Die Sprache* 17: 23–33.
- Blöndal Magnússon, Ásgeir
 1989 *Íslensk orðsifjabók* [Icelandic etymological dictionary]. Reykjavík: Orðabók Háskólans.
- Bosworth, Joseph, T. Northcote Toller, and Alistair Campbell
 1898–1972 *An Anglo-Saxon Dictionary*. 3 vols. Oxford: Oxford University Press.
- Boutkan, Dirk and Sjoerd Michiel Siebinga
 2005 *Old Frisian Etymological Dictionary*. Leiden: Brill.
- Braune, Wilhelm and Frank Heidermanns
 2004 *Gotische Grammatik*. 20th edn. Tübingen: Niemeyer.
- Bremmer, Rolf H. Jr.
 2004 *Hir is eskriven: Lezen en schrijven in de Friese landen rond 1300* [Here is written: Reading and writing in the Frisian lands around 1300]. Leeuwarden: Hilversum.
- CIL I–XVI [+ no.] = *Corpus inscriptionum Latinarum*
 1862 ff. Berlin: Reimer.
- Dittenberger, Wilhelm
 1915–1924 *Sylloge inscriptionum Graecarum*. 4 vols. 3rd edn. Leipzig: Hirzel.
- DOE see di Paolo Healey, 2005.
- Düwel, Klaus
 2008 *Runenkunde*. 4th edn. Stuttgart: Metzler.
- Egilsson, Sveinbjörn and Finnur Jónsson
 1931 *Lexicon poeticum antiquae linguae septentrionalis*. (= *Ordbog over det norske-islandske Skjaldesprog* [Dictionary of the Norse-Icelandic poetic language]). 2nd edn. Copenhagen: Møller.
- Falk, Hjalmar and Alf Torp
 1909 *Wortschatz der germanischen Spracheinheit*. Unter Mitwirkung von Hjalmar Falk gänzlich umgearbeitet von Alf Torp. (*Vergleichendes Wörterbuch der indogermanischen Sprachen* von August Fick. 4. Auflage.). Göttingen: Vandenhoeck & Ruprecht.
- Feist, Sigmund
 1939 *Vergleichendes Wörterbuch der gotischen Sprache: mit Einschluß des Krimgotischen und sonstiger zerstreuter Überreste des Gotischen*. 3. neubearbeitete und vermehrte Auflage. Leiden: Brill.
- Francovich Onesti, Nicoletta
 1999 *Vestigia longobarde in Italia (568–774). Lessico e antroponomia*. Rome: Artemide.
- Francovich Onesti, Nicoletta
 2002 *I Vandali. Lingua e storia*. Rome: Carocci.
- Fritzner, Johan
 1972 [1886–1896] *Ordbog over Det gamle norske Sprog* [Dictionary of the Old Norse language]. 4 vols. Oslo: Universitetsforlaget. [Kristiania: Norske forlagsforening.]
- Harris, Joseph
 1985 Die altenglische Heldendichtung. In: Klaus von See (ed.), *Europäisches Frühmittelalter* (Neues Handbuch der Literaturwissenschaft 6). Wiesbaden: Aula, 237–276.
- Hofmann, Dietrich and Anne Tjerk Popkema
 2008 *Altfriesisches Handwörterbuch*. Heidelberg: Winter.

Holthausen, Ferdinand

1962 *Altenglisches etymologisches Wörterbuch*. 2nd edn. Heidelberg: Winter.

Johnston, Thomas S. B.

2001 The Old Frisian Law Manuscripts and Law Texts. In: Horst Haider Munske, Nils Århammar, Volkert Faltings, Jarich Freark Hoekstra, Oebele Vries, Alastair Walker, and Ommo Wilts (eds.), *Handbuch des Friesischen*. Tübingen: Niemeyer, 571–587.

Jónsson, Finnur

1921 *Norsk-islandske Kultur- og Sprogforhold i 9. og 10. århundrede* [Norwegian-Icelandic cultural and linguistic relations in the 9th and 10th centuries]. Copenhagen: Høst.

Karg-Gasterstädt, Elisabeth, Theodor Frings (Vol. 1), Rudolf Grosse (Vol. 2–4), Gotthard Lerchner, and Hans Ulrich Schmid (Vol. 5–6)

1968 ff. *Althochdeutsches Wörterbuch*. 6 vols. published (6: M–N). Berlin: Akademie-Verlag.

Koivulehtu, Jorma

2002 Contact with non-Germanic languages II: Relations to the East. In: Bandle et al. (eds.), 583–594.

Krause, Wolfgang

1971 *Die Sprache der urnordischen Runeninschriften*. Heidelberg: Winter.

Krause, Wolfgang and Herbert Jankuhn

1966 *Die Runeninschriften im älteren Futhark*. 2 vols. Göttingen: Vandenhoeck & Ruprecht.

Krogh, Steffen

1996 *Die Stellung des Altsächsischen im Rahmen der germanischen Sprachen*. (Studien zum Althochdeutschen 29). Göttingen: Vandenhoeck & Ruprecht.

Krogh, Steffen

1997 Zur Entstehung des Niederländischen. In: Elvira Glaser, Michael Schlaefler, and Ludwig Rübekeil (eds.), *Grammatica ianua artium. Festschrift Rolf Bergmann*. Heidelberg: Winter, 21–31.

Kroonen, Guus

2013 *Etymological dictionary of Proto-Germanic*. Leiden: Brill.

Kylstra, Andries Dirk, Sirkka-Liisa Hahmo, Tette Hofstra, and Osmo Nikkilä

1991–2012 *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen*. 3 vols. Amsterdam: Rodopi.

LaN I–II: see Reichert 1987–1990.

Lehmann, Winfred P.

1986 *A Gothic Etymological Dictionary. Based on the third edition of Vergleichendes Wörterbuch der Gotischen Sprache by Sigmund Feist*. Leiden: Brill.

Lloyd, Albert L., Otto Springer, and Rosemarie Lühr, with the assistance of Karen Purdy et al.

1988 ff. *Etymologisches Wörterbuch des Althochdeutschen*. 5 vols. published (5: iba–luzzilo). Göttingen: Vandenhoeck & Ruprecht.

Müller, Rosemarie

1999 Germanen, Germania, Germanische Altertumskunde: Die Bewertung von Realien für die Gesch[ichte] der Germ[anen]. In: Beck et al. (eds.), vol. 11, 309–317.

Nedoma, Robert

1995 *Die Inschrift auf dem Helm B von Negau. Möglichkeiten und Grenzen der Deutung norditalischer epigraphischer Denkmäler*. Vienna: Fassbaender.

Nedoma, Robert

2003 Runennamen. In: Beck et al. (eds.), vol. 25, 556–562.

Nedoma, Robert

2004 *Personennamen in südgermanischen Runeninschriften*. Heidelberg: Winter.

Nedoma, Robert

2005 Urnordisch **-a** im Nominativ Singularis der maskulinen *n*-Stamme. *Northwestern European Language Evolution* 46/47: 155–191. (= *Papers on Scandinavian and Germanic*

Language and Culture. Published in Honour of Michael Barnes on his Sixty-Fifth Birthday 28 June 2005).

- Nedoma, Robert
 2006 Schrift und Sprache in den sudgermanischen Runeninschriften. In: Alfred Bammesberger and Gaby Waxenberger (eds.), *Das fuþark und seine einzelsprachlichen Weiterentwicklungen* (Ergänzungsbände zum Reall. 51). Berlin: De Gruyter, 109–156.
- Nedoma, Robert
 2008 Voraltsächsisch **lguskabi** auf dem Schemel von Wremen. In: Barbara Theune-Großkopf and Robert Nedoma. Ein Holzstuhl mit Runeninschrift aus dem frühmittelalterlichen Gräberfeld von Trossingen, Lkr. Tuttlingen (Baden-Württemberg). *Die Sprache* 46 (2006): 52–64.
- Nedoma, Robert
 2010 *Kleine Grammatik des Altisländischen*. 3rd edn. Heidelberg: Winter.
- Nedoma, Robert
 2014 Vorfriesisch **-u** im Nominativ und Akkusativ Sg. der maskulinen *a*-Stämme. *Amsterdamer Beiträge zur älteren Germanistik* 73: 343–368. (= Rolf H. Bremmer Jr., Stephen Laker, and Oebele Vries [eds.], *Directions for Old Frisian Philology*.)
- Nedoma, Robert and Klaus Düwel
 2012 Die Runeninschrift. In: Christoph G. Schmidt, Robert Nedoma, and Klaus Düwel. Die Runeninschrift auf dem Kamm von Friestedt, Stadt Erfurt. *Die Sprache* 49 (2010/2011): 123–186.
- Neumann, Günter
 1953 Thruskanos. *Beiträge zur Namenforschung* 4: 53–55.
- Neumann, Günter
 1987 Die germanischen Matronen-Beinamen. In: Gerhard Bauchhenß and Günter Neumann (eds.), *Matronen und verwandte Gottheiten*. Cologne: Rheinland-Verlag, 103–132.
- Neumann, Günter
 1994 Früheste Zeugnisse unserer Sprache. *Jahrbuch der Akademie der Wissenschaften in Göttingen*, 93–110.
- Nielsen, Hans Frede
 2000 *The Early Runic Language of Scandinavia. Studies in Germanic Dialect Geography*. Heidelberg: Winter.
- Nielsen, Hans Frede
 2002 Delimitation of Ancient Nordic from Common Germanic and Old Nordic. In: Bandle et al. (eds.), 615–619.
- ONP = *Ordbog over det norrøne prosasprog, A Dictionary of Old Norse Prose*
 1989 ff. Register and 3 vols. published (3: de–em). Copenhagen: The Arnamagnæan Commission. Available online: URL <http://onp.ku.dk> [Last accessed 9 April 2016].
- Orel, Vladimir
 2003 *A Handbook of Germanic Etymology*. Leiden: Brill.
- di Paolo Healey, Antonette with John Price Wilkin and Xin Xiang
 2005 *The Dictionary of Old English Web Corpus* Online: URL <http://www.doe.utoronto.ca> [Last accessed 9 April 2016].
- Peterson, Lena
 1994 *Svenskt runordsregister* [Swedish rune word register]. (Runrön 2). 2nd edn. Uppsala: Inst. för Nordiska Språk, Uppsala Univ.
- Polomé, Edgar C.
 1994 Proto-Germanic and the Reconstruction of Proto-Indo-European. *Northwestern European Language Evolution* 23: 3–40.
- RäF [+ no.] see Krause and Jankuhn, 1966.

- Reichert, Hermann
1987–1990 *Lexikon der altgermanischen Namen*. 2 vols. Vienna: Österreichische Akademie der Wissenschaften.
- Ritter, Ralf-Peter
1993 *Studien zu den ältesten germanischen Entlehnungen im Ostseefinnischen*. Frankfurt am Main: Lang.
- Rübekeil, Ludwig
1992 *Suebica. Völkernamen und Ethnos*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Schmid, Hans Ulrich
2006 Ein neues ‘Heliand’-Fragment aus der Universitätsbibliothek Leipzig. *Zeitschrift für deutsches Altertum* 135: 309–323.
- Schulte, Michael
2006 The transformation of the older fuþark. *Arkiv för Nordisk Filologi* 121: 41–74.
- Schumacher, Stefan
2007 Val Camonica, Inschriften. In: Beck et al. (eds.), vol. 35, 334–337.
- Schützeichel, Rudolf
2006 *Althochdeutsches Wörterbuch*. 6th edn. Tübingen: Niemeyer.
- Schwab, Ute
2005 Bemerkungen zur siebenten Auflage des ‘Streitberg’ durch Piergiuseppe Scardigli mit Blicken auf das ‘Gotische Elementarbuch’ von Wolfgang Binning. *Die Sprache* 44 (2004): 91–116.
- Seebold, Elmar
1998 Germanen, Germania, Germanische Altertumskunde: Sprache und Schrift. In: Beck et al. (eds.), vol. 11, 275–304.
- Seebold, Elmar, Brigitte Bulitta, Elke Krotz, Judith Stieglbauer-Schwarz, and Christiane Wanzeck
2001 *Chronologisches Wörterbuch des deutschen Wortschatzes. Der Wortschatz des 8. Jahrhunderts (und früherer Quellen)*. Berlin: De Gruyter.
- Simek, Rudolf and Hermann Pálsson
2007 *Lexikon der altnordischen Literatur*. 2nd edn. Stuttgart: Kröner.
- Snædal, Magnús
1998 *A Concordance to Biblical Gothic*. 2 vols. Reykjavík: University of Iceland Press.
- Stearns, MacDonald Jr.
1989 Das Krimgotische. In: Heinrich Beck (ed.), *Germanische Rest- und Trümmersprachen* (Ergänzungsbände zum Reall. 3). Berlin: De Gruyter, 175–194.
- Steuer, Heiko
1999 Germanen, Germania, Germanische Altertumskunde: Die Germ[anen] und der Norden. In: Beck et al. (eds.), vol. 11, 318–327.
- Streitberg, Wilhelm (ed.)
2000 *Die gotische Bibel*. 2 vols. 7th/6th edn. (by Piergiuseppe Scardigli). Heidelberg: Winter.
- Syll. [+ no.] see Dittenberger, 1915–1924.
- Tiefenbach, Heinrich
1991 Das vandalische *Domine miserere*. *Historische Sprachforschung* 104: 251–268.
- Tiefenbach, Heinrich
1995 Älteste germanische Namen der Völkerwanderungszeit in lateinischen und griechischen Quellen. In: Ernst Eichler, Gerold Hilty, Heinrich Löffler, Hugo Steger, and Ladislav Zgusta (eds.), *Namenforschung. Ein internationales Handbuch zur Onomastik*. Vol. 1. Berlin: De Gruyter, 774–778.
- Tiefenbach, Heinrich
2010 *Altsächsisches Handwörterbuch*. Berlin: De Gruyter.
- Tischler, Johann
2003 Ostgermanen: Sprachliches. In: Beck et al. (eds.), vol. 22, 338–344.

- de Vries, Jan
 1962 *Altnordisches etymologisches Wörterbuch*. 2nd edn. Leiden: Brill.
- Wagner, Norbert
 2002 *Gaisericus und Gesiric. Zu ai und au im späteren Ostgermanischen und bei Wulfila. Beiträge zur Namenforschung*, N.F. 37: 259–270.
- Wenisch, Franz
 1979 *Spezifisch anglisches Wortgut in den nordhumbrischen Interlinearglossierungen des Lukasevangeliums*. Heidelberg: Winter.

Robert Nedoma, Vienna (Austria)

54. The phonology of Germanic

- | | |
|-----------------|-------------------------|
| 1. Introductory | 4. Post-tonic Syllables |
| 2. Consonants | 5. References |
| 3. Vowels | |

1. Introductory

1.1. Preliminaries

This account will concentrate on Proto-Germanic (PGmc), with side-glances at important developments in the daughter languages. No systematic notice will be taken of fragmentarily attested idioms, such as Langobardic or “Crimean Gothic”. The emphasis will be on correspondences rather than change processes and on data rather than phonological “theories”. Reconstructions of the parent language will incline to “classical” late (1) Proto-Indo-European (PIE) – even though this is a convenient fiction.

Conventions: **word* represents a (posited or reconstructed) lexeme that is not attested; *word* * a particular spelling or form (typically nom. sg. or infinitive as a headword) that is not found in the texts although the lexeme itself is attested (this is common for Gothic forms, given the limited nature of the corpus); ***word* indicates hypothesized pre-forms that presumably existed but have not left a (direct) reflex; †*word* designates an unattested form that might be expected to result from a given reconstruction but shows a different reflex. Laryngeals are depicted by x_1 , x_2 , and x_3 for *e*-, *a*-, and *o*-colouring respectively, *X* for when the quality is unknown or unspecified; resonants, by *R*, semi-vowels by *W*; *C* = any consonant; *V* = any vowel. A small circle under a resonant indicates syllabicity. Analogical forms are enclosed in []. For reasons of space, references will be kept to a minimum. I apologize that not all scholars will receive due acknowledgment. For a recent treatment, see Ringe (2006).

<https://doi.org/10.1515/9783110523874-009>

- de Vries, Jan
 1962 *Altnordisches etymologisches Wörterbuch*. 2nd edn. Leiden: Brill.
- Wagner, Norbert
 2002 *Gaisericus und Gesiric. Zu ai und au im späteren Ostgermanischen und bei Wulfila. Beiträge zur Namenforschung*, N.F. 37: 259–270.
- Wenisch, Franz
 1979 *Spezifisch anglisches Wortgut in den nordhumbrischen Interlinearglossierungen des Lukasevangeliums*. Heidelberg: Winter.

Robert Nedoma, Vienna (Austria)

54. The phonology of Germanic

- | | |
|-----------------|-------------------------|
| 1. Introductory | 4. Post-tonic Syllables |
| 2. Consonants | 5. References |
| 3. Vowels | |

1. Introductory

1.1. Preliminaries

This account will concentrate on Proto-Germanic (PGmc), with side-glances at important developments in the daughter languages. No systematic notice will be taken of fragmentarily attested idioms, such as Langobardic or “Crimean Gothic”. The emphasis will be on correspondences rather than change processes and on data rather than phonological “theories”. Reconstructions of the parent language will incline to “classical” late (1) Proto-Indo-European (PIE) – even though this is a convenient fiction.

Conventions: **word* represents a (posited or reconstructed) lexeme that is not attested; *word* * a particular spelling or form (typically nom. sg. or infinitive as a headword) that is not found in the texts although the lexeme itself is attested (this is common for Gothic forms, given the limited nature of the corpus); ***word* indicates hypothesized pre-forms that presumably existed but have not left a (direct) reflex; †*word* designates an unattested form that might be expected to result from a given reconstruction but shows a different reflex. Laryngeals are depicted by x_1 , x_2 , and x_3 for *e*-, *a*-, and *o*-colouring respectively, *X* for when the quality is unknown or unspecified; resonants, by *R*, semi-vowels by *W*; *C* = any consonant; *V* = any vowel. A small circle under a resonant indicates syllabicity. Analogical forms are enclosed in []. For reasons of space, references will be kept to a minimum. I apologize that not all scholars will receive due acknowledgment. For a recent treatment, see Ringe (2006).

<https://doi.org/10.1515/9783110523874-009>

1.2. Characterization

Germanic is a “centum-dialect” of Indo-European, thus merging the PIE palatal and plain dorsal stops at a pre-Germanic (pGmc) stage. Dialectally, it belonged to two groupings: North IE (together with Baltic and Slavic) and West IE (together with Italic and Celtic). It is archaic in keeping the proto-language’s three orders of stops separate, although their phonetic realization was transformed by Grimm’s Law.

Phonologically, PIE had three classes of sounds that can loosely be defined according to how they function in roots: always syllabic (vowels); never syllabic (obstruents); sometimes syllabic (liquids, nasals, semi-vowels, and laryngeals – de Saussure’s “*coefficients sonantiques*” 1879: 8, 135). Note that this definition of the sounds that are sometimes syllabic could encompass *s, which patterns to some extent like the laryngeals. However, unlike the laryngeals, PIE *s is never replaced by vowel reflexes in the non-Anatolian languages, and the sound is here treated with the obstruents.

The main changes that transformed the PIE sound-system and gave Gmc its phonological character were:

1. Consonant shifts, including Grimm’s and Verner’s Laws (GL and VL), 2.2.1.
2. Vowel changes, principally merger of *a* and *o* qualities and **ey* > PGmc **ī*.
3. The demise of the system of “*coefficients sonantiques*” as a separate class of sounds, 2.3.1.
4. The accent shift. This took place later than Verner’s Law, which was conditioned by the place of the IE (tonic) accent. The fixation of a stress accent on the initial syllable of the word weakened following syllables with far-reaching consequences, leading to a long drawn-out series of changes that produced different sound developments and inventories in accented and unaccented syllables. Most final consonants and some short final vowels were lost in early (e) PGmc (4.1, 4.3.1). Later vowel shortenings and losses took place in the daughter languages, some of them triggering variousumlaut phenomena in the North and West Germanic (NWGmc) dialects (cf. 3.3.5).

2. Consonants

2.1. PGmc Inventory

**p*, **t*, **k*, **k*^w
 b* ~ **b̥*, **d* ~ **d̥*, **z*, **g* ~ **ɣ*, ?g*^w ~ **ɣ*^w
 **f*, **p*, **s*, **h* ~ **χ*, **h*^w ~ **χ*^w
 **m*, **n*, **r*, **l*; **w*, **j*

It is not clear whether, or for how long, a voiced labiovelar existed as a unit phoneme (in medial position only) in PGmc (2.2.3.1). All the consonants except **z* could occur as geminates intervocalically, but only after short vowels in initial syllables.

2.2. Obstruents

2.2.1. Basic correspondences with PIE (Grimm's and Verner's Laws)

2.2.1.1. Grimm's Law

The IE stops changed their manner of articulation in Gmc. The most complex development is undergone by the voiceless series (see below and 2.2.1.2). The basic correspondences between the stops of (centum) PIE and Gmc were formulated (provisionally) by the Dane Rasmus Rask (1818: 169–71) and then more fully and accurately by the German Jacob Grimm and bear the latter's name (first published by Grimm 1822: 584–92). Also known as the First Consonant Shift, this states that:

1. Voiceless stops, when not preceded by an obstruent or *s*, became voiceless fricatives (probably via a stage with aspiration): **p, *t, *k, *kʷ* > **f, *þ, *χ, *χʷ*.
2. Voiced aspirated stops became voiced fricatives (although the reflexes became voiced stops in several positions by the end of PGmc, 2.2.1.3): **bʰ, *dʰ, *gʰ, *gʷʰ* > **b, *d, *ɣ, #*b-* (2.2.3.2) ~ **-ɣʷ-* (2.2.3.1).
3. Voiced plain stops became voiceless stops: **b, *d, *g, *gʷ* > **p, *t, *k, *kʷ*.

Examples:

1. Gk *pós* OE *fōt* 'foot'; OInd. *trín* Lat. *trīs* Go. *þrins* '3' M acc.; Lat. *centum* Go. *hunda* pl. '100'; OInd. *katará-* Gk *póteros* Go. *hvaþar* 'which (of two)?'.

Note that this shift does not affect the second of two contiguous obstruents: Lat. *captus* 'captive' Go. *hafts** 'held to, tied to (something)'; Lat. *stāre* Go. *standan* 'to stand', OFris. *strot-* beside *throt-* 'throat' to PIE $\sqrt{*(s)treud-}$; Lat. *piscis* Go. *fisks** 'fish'.

2. OInd. *bhrátar-* 'brother' Gk (Attic) *phrátēr* 'member of a clan' Go. *brōþar* 'brother'; OInd. *dádadhāmi* 'I place' Gk *títthēmi* OE *dō* 'I do'; OInd. *hamśás* Gk (Doric) *khán* OE *zōs* OHG *gans* 'goose'; PIE $\sqrt{*g^{wh}ed^{h}-}$: Av. *ǰaiðiieti* 'asks, prays' OI *guidid* 'asks for, prays' Go. *bidjan* 'to pray', PGmc **banan-* 'death, killer' (to PIE **g^{wh}en-* 'strike, kill') ~ Gk *omphé* 'divine voice' Go. *saggs** OHG *sang* 'song'.
3. OI *ubull* W *afal* Lith. *óbuolas, obuolys* OS *appul(-)* 'apple'; Lat. *decem*, Lith. *dėšimt* Go. *taihun* '10'; Gk *ágō* 'I drive' ON *aka* 'to travel'; Gk *guné* OI *ben* Go. *qinō* OE *cwene* 'woman'.

Word-finally, stops seem to have been neutralized to voiced unaspirated in PIE (cf. Byrd, this handbook, 5.1). Therefore, Gmc reflexes show up as voiceless in this position: e.g., OS *that* 'that' nom. sg. N of the demonstrative pronoun (cf. OInd. *tád*); OS *up* 'up' (cf. Lat. Osc. *s-ub*).

2.2.1.2. Verner's Law

1. A series of apparent exceptions to Grimm's Law, where voiced fricatives appear non-initially as reflexes of the PIE voiceless stops rather than the expected voiceless ones, were accounted for by the Dane Karl Verner (1877), who explained them by linking the

reflexes with the accent of Old Indic cognate forms, thereby also reconstructing the place of the PIE accent. The Gmc voiceless fricative reflexes remained: word-initially, in voiceless clusters, and where the accent immediately preceded; in other positions they were voiced. This brought a partial merger of the reflexes of PIE voiceless and voiced aspirated stops.

The seemingly arbitrary outcomes among the relation nouns of PGmc **mōdēr* ‘mother’ beside **brōpēr* ‘brother’ received their explanation, as the IPIE pre-forms were respectively **māter-* (cf. OInd. *mātār-*) and **b^hrāter-* (cf. OInd. *bhrātar-* Gk *phrātēr* ‘member of a clan’). The effects are also seen in the principal parts of strong verbs: PGmc **keusan-* ‘to test, choose’ pret. sg. **kaus* pl. **kuzun* p.p.p. **kuzana-*, OE *cēosan* *cēas curon coren* (with **z > r*, 2.2.2.1), to PIE $\sqrt{*gheus-}$, cf. OInd. pf. *jujōṣa* ‘enjoys’, as well as isolated vocabulary items, such as Go. *haidus** ‘way, manner’ OE *hād* ‘condition, state’: OInd. *ketú-* ‘form’.

Further instances of voicing: There seem to be instances of voicing word-initially in forms that were never stressed, e.g., the prefix *ga-* < PIE particle **ko*. Word-final fricatives (in effect *-s#*) seem to have been voiced under all circumstances; the only potential problem appears to be the *o*-stem genitive ending IPGmc **-as*, but such cognates as OInd. *-asya*, Gk (Homer) *-oio* indicate that the *s* was unlikely to have been word-final.

2. The voiced alternants were extensively eliminated in Gothic: compare such archaisms as *fulgins** ‘hidden’ adj. beside *fulhans** ‘hidden’ p.p.p.; *aihvatundi** ‘thorn bush’ (literally ‘horse-tooth’) beside *tunþus* ‘tooth’. Voiced and voiceless spirants subsequently merged medially in NGmc and, apart from **đ*, in the Ingvæonic dialect area of WGmc: OE, OFris., and parts of OS. VL is best preserved in OHG (in altered form because of the second consonant shift, 2.2.2.2), except that **-b̥-* > *-f-* immediately before *l* and *r* suffixes (von Bahder’s Law, 1903), with some apparent exceptions that are the result of anaptyctic vowels in the nom. sg. (thus: OHG *sweval* and *swebal* beside Go. *swibls** ‘sulphur’). Compare the principal parts of the verb ‘to journey’:

Go.	<i>-leiþan</i>	<i>-laiþ</i>	<i>-liþun</i>	<i>-liþans*</i>
ON	<i>liða</i>	<i>leið</i>	<i>liðu</i>	<i>liðinn</i>
OE	<i>liðan</i>	<i>lāð</i>	<i>lidon</i>	<i>liden</i>
OHG	<i>liðan</i>	<i>leid</i>	<i>litun</i>	<i>gilitan</i>

2.2.1.3. Further PGmc changes

The voiced fricatives **b̥*, **đ*, **g̥* as they existed after VL became the stops **b*, **d*, **g* in the following positions: word-initially (except for **g̥-*) and after nasals (cf. also 2.2.3.1).

2.2.2. Main developments in the dialects

2.2.2.1.

1. In Gothic the voiced fricatives also became stops after any consonant. Examples are, for **đ*: acc. sg. *gard* ‘house(hold)’, Nt. nom. sg. *skuld* ‘lawful, fitting’, nom. sg. *gazds*

‘sting’, acc. sg. *gahugd* ‘mind’. Remaining spirants were devoiced word-finally and before *s*: *bidjan* ‘to ask, pray’ pret. sg. *baþ*, nom sg. *hlaiþs* ‘bread’ beside gen. sg. *hlaiþis* (but *lamb* ‘lamb’). In unaccented syllables, an interchange is found such that voiceless spirants appear after voiced segments and vice versa (Thurneysen’s Law, 1898): e.g., dat. sg. of former *s*-stems *agisa* ‘fear’ but *riqiza* ‘darkness’.

In Gothic, the voiceless fricative **χ* early became **h* (a breathing or weakly articulated fricative) word-initially before vowels, as indicated by spellings: initial <h> is sometimes left out or wrongly inserted, suggesting the sound it depicted was unstable. That the sound written by <h> was not the voiceless equivalent of the fricative written <g> in Gothic explains why the latter sign is used for its devoiced variant, as in *magt* 2sg. ‘thou canst’. (However, this was not a PGmc development, as French loans from WGmc indicate that **χ* retained its articulation: F *flanc* ‘flank’ ← Frankish **χlank*, F *freux* ‘rook’ ← Frankish **χrōk*.)

2. In N and WGmc, **z*, insofar as it was not lost or assimilated (e.g., PGmc **izwiz* ‘you’ dat. pl. > WGmc **iwwiz*), eventually became *r*; in NGmc via a sound traditionally transcribed *r*, which has its own runic symbol (𐀚). Examples: ON *meiri*, OE *māra*, OHG *mēro* beside Go. *maiza* ‘more’. For final **-z*, see 4.1.

3. In NGmc, **ð* also became a stop after *l* (*halda* ‘to hold’), while intervocalic **f* and **þ* merged with **b* and **d*, largely eradicating the effects of VL (2.2.1.2). However, **f* and **þ* were assimilated after nasals (and in the case of *þ* after *l*): *fimm* : Go. *fimf* ‘4’, *annar* : Go. *anþar* ‘other’; *gull* : Go. *gulþ** ‘gold’.

4. In WGmc, consonants were geminated before resonants, especially **j*. (*r* of whatever origin was unaffected, except in parts of High German, cf. *erren*, beside *erien*, OE *erian* ‘to plough’.); Go. *bidjan* : OS *biddian* ‘ask, pray’; gen. sg. Go. *kunjis* : OS *kunnias* gen. sg. ‘race, lineage’ (beside combining-form *kuni-*); Go. *akrs* ‘field, acre’ : OS *akkar*. (NGmc has a similar but independent change, which only affects *k* and *g* before *j* and *w*: Go. *lagjan* ‘to lay’ : ON *leggja*; ON *rōkkva* ‘to darken’, to *rōkr* ‘darkness’ [on the vocalism, see 3.3.5, end], Go. *riqis*.)

5. In WGmc, **ð* was occluded to **d* in all positions: e.g., OE *bēodan* vs. ON *bjóða* ‘to bid, offer’.

2.2.2.2. High German Consonant Shift

In most of HG, the other two voiced fricatives (i.e., **þ* and **γ*) were also occluded earlier than the Second or High German Consonant Shift, which changed the manner of articulation of stops again. This started in the south of the speech-area (Upper German) and spread north with diminishing force. Implementation thus varies from dialect to dialect and according to date. Reflexes also depend on the position in the word. In broad terms, voiceless stops became affricates (word-initially and medially after consonants or when geminated) or geminate spirants (word-finally and medially between vowels, simplifying after long vowels); voiced stops were devoiced without regard to their position in the word, but this was mainly an Upper German phenomenon. The dictum holds:

dentals move first and furthest, velars last and least. As with GL, voiceless stops preceded by spirants were unshifted: e.g., ‘stone’ OS *stēn*: OHG *stein*. In addition, the sequence WGmc **tr-* remained: ‘to tread’ OE *tredan* : OHG *tretan*, OS, OHG *bittar* ‘bitter’ beside ON *bitr*. The general situation can be illustrated by the following examples from the East Frankish of the Tatian translation (the basis of grammarians’ standard OHG) and – where there is a difference – Bavarian, as a representative of Upper German, with OS as a comparison:

	OS	EFrk.	Bav.
‘10’	<i>tehan</i>	<i>zehan</i>	
‘heart’	<i>herta</i>	<i>herza</i>	
‘to eat’	<i>etan</i>	<i>ezzan</i>	
‘foot’	<i>fōt</i>	<i>fuoz</i>	
‘pepper’	<i>peper</i> (MLG)	<i>pfeffer</i>	
‘to throw’	<i>werpan</i>	<i>werphan</i>	<i>werfan</i>
‘to sleep’	<i>slāpan</i>	<i>slāf(f)an</i>	
‘ship’	<i>skip</i>	<i>scif</i>	
‘corn’	<i>korn</i>	<i>korn</i>	<i>chorn</i>
‘to drink’	<i>drinkan</i>	<i>trinkan</i>	<i>trinchan</i>
‘to seek’	<i>sōkian</i>	<i>suochen</i>	
‘sick’	<i>siok</i>	<i>sioch</i>	
‘day’	<i>dag</i>	<i>tag</i>	<i>tac</i>
‘to bind’	<i>bindan</i>	<i>bindan</i>	<i>pintan</i>
‘mountain’	<i>berg</i>	<i>berg</i>	<i>perc</i>
‘to give’	<i>geban</i>	<i>geban</i>	<i>kepan</i>

2.2.2.3. “Continental WGmc” (those dialects left after the departure of the English) underwent a change whereby WGmc **þ* was voiced and later became the stop *d*. The change began probably in Bavarian in the mid-8th century and spread northwards, affecting Low German and Dutch from about 1200, and Frisian dialects (but not word-initially) apart from Weser East Frisian and North Frisian. Thus, forms of the weak verb ‘to invite’: Go. *laþōn*; eOHG *kilathote* p.p.p. beside OHG *ladōn*, OS *lathian* beside *ladoian*, OFris. *lathia* beside OFris. *ladia*, NEFris. (Saterland = Ems) *lede*, (Wangerooze = Weser) *laðii*, NNFris. (Mooring) *lââsie* (-s- < **þ-*); word-initial Go. *þaurnus* ‘thorn’; eOHG *thorn* G *Dorn*, OLFrk *thorn* Du. *doorn*; O(E)Fris. *thornen** adj. ‘made of thorns’ NWFris. *toarn*, NEFris. (Saterland = Ems) *Touden*, (Wangerooze = Weser) *thoon*, NNFris. (Mooring) *tom*.

2.2.3. Labiovelars

Unit labiovelar phonemes existed in Proto-Germanic and Gothic, both continuing those inherited from PIE and augmented by secondary examples. However, their number was depleted by delabialization and full labialization. In attested North and West Germanic, the surviving labiovelars became sequences of velar plus *w* and usually lost the *w*, except word-initially.

2.2.3.1.

1. Unit labiovelar phonemes are to be reconstructed for PGmc and are attested in Gothic in the case of the two voiceless ones: *q* (< PIE **g^w*) and *hw* (< PIE **k^w*). By contrast, PIE **g^wh* (which became *b-* word-initially, 2.2.1.1(2), 2.2.3.2), appears in Gothic medial position after nasals as the sequence *-gw-* with a stop (cf. 2.2.1.3): Go. *siggwan* ‘sing’, cf. ON *syngva*; cf. OFris. *siunga* (with labial breaking caused by the lost *-w-*) beside OHG *singan*. In other positions, ePGmc root-final **ɣ^w* has been simplified, usually to *-w-* (cf. also 2.2.4.2 end), sometimes to *-g-*, under unclear conditions and with the reflexes disturbed by analogy. For *-w-*: Go. *saiws* ‘lake; sea’ (< PGmc **saig^w-iz*, to **sei^whan* ‘to trickle’), Go. *snaiws* ‘snow’ (< PGmc **snaig^w-az* < PIE **snoy^wh-os*), OHG *nioro* ‘kidney’ (< PWGmc **neuran-*, PGmc **ney^wran-* from PIE **neg^wh-ró-*, cf. Gk *nephros*), Go. *siuns* ‘sight’ (< PGmc **seg^w-niz*, to PGmc **seh^wan* ‘to see’). For *-g-*: ON *leiga* ‘hire, loan’, to Go. *leih^wan* etc. ‘to lend’. Note OE WS *sāwon* vs. Angl. *sēgon* pret. pl ‘saw’, which must be the result of different levellings.

2. The labiovelars inherited from PIE were augmented in PGmc by: a) sequences of velar plus *-w-*, which coalesced as labiovelars: *ai^wha-* ‘horse’ (< PIE **[x₁]ek^wo-*: OInd. *ásvas*, cf. OLith. *ešva* ‘mare’); and b) secondary labiovelars that developed mainly from root-final velars before *-w-* in motivated feminines of *u*-stem formations ending in a velar (if the root vocalism was not *-u-*, see 2.2.3.3[2]; see 2.3.2.3[2] for some examples involving other sources). That this was a PGmc change is shown by forms such as: Go. *mawī* etc. ‘girl’, F of *magus* ‘boy, son’ (< PIE **mag^hus*), with a labiovelar reflex (pre-Gmc **mag^hwī* > ePGmc **may^wī* > IPGmc **mawī*); Old (West) Norse neo-*wa*-stem adjective acc. sg. M *qngvan* ‘narrow’ (Go. *aggwus**) – as OInd. *amhú-* shows, this was originally a *u*-stem adjective, a category that otherwise became *a*-stem adjectives in Old (West) Norse. The labiovelar developed in the feminine forms in *-wī* and was generalized.

3. Pre-consonantly or word-finally, labiovelars were only retained in Gothic if they were part of a morphological pattern: Go. *leihts** ‘light(weight)’ (cf. Gk *elaphros* ‘light; nimble’, 3.2), *jah* ‘and’ < PIE *(*x₁*)*yo-k^we* vs. *nēhw* adv. ‘nearby’ beside formations in *nēhw-*; pret. 2sg. *sah^t* ‘sawst’, 1/3sg. *sah^w* ‘saw’ to *sai^wan* ‘to see’, *sag^q* ‘sank’ to *sig^qan** ‘to sink’.

4. In attested North and West Germanic, the voiceless labiovelars were resolved into sequences of velar plus labial. These have either been retained (primarily in word-initial

position: ON *hvat*, OE *hwæt* ‘what’, ON *kvað*, OE *cwæð* pret. sg. ‘said’), or have lost the labial element (OS *sehan*). In the WGmc languages, the loss is relatively late, as shown by OS *ehu-* ‘horse’ < **ehwa-* (first element of compounds) and OFris. forms such as *siunga* (above [1]) and *thiukke** ‘thick’ (< PGmc **þekkw-*, 3.3.1.3) with labial breaking, a change not found in the other WGmc dialects. Note also OE (Nh.) *ginēhwade** ‘approached, drew nigh’, with the sequence apparently retained medially.

2.2.3.2. In a number of positions, PIE labiovelars became plain labials. Thus, the regular outcome of word-initial **#g^wh-* was Gmc *b-* (2.2.1.1[2]).

In the case of the other two labiovelars, occasional development to the corresponding plain labial when there is another labial in the word has been held to account for examples such as Go. *fidwōr* (< PIE **k^wetwōr*) and *wulfs* ‘wolf’ (< PIE **w^lk^wos*), beside ON *ylgr* with delabialization, 2.2.3.3), *finf* ‘five’ (< PIE **pénk^we*), *wairpan* ‘throw’ (< PIE **werg^w-* ‘twist, throw’). But in other examples that must be instances of the same phenomenon, there is no other labial, e.g., Go. *ainlif** ‘11’, beside Lithuanian *vienúo-lika*. Moreover, some words that do contain a labial keep their labiovelar, e.g., Go. *qiman*, OHG *queman* ‘to come’. It seems likely that the development is a spontaneous tendency of the labiovelars. The change probably preceded the Gmc stage, not least because **k^w* (which in PGmc merged with **k^w*, 2.2.3.1) is not known to have become *f*. Also, *#b-* as a reflex of PIE **#g^wh-* is more easily obtained phonetically from pre-GL **b^h-* than post-GL **ɣ^w-*.

2.2.3.3. Delabialization

1. Delabialization took place within PGmc word-initially before the rounded vowels *u* (continuing a PIE rule, but extended to *u* developed from **ŕ*, 2.3.3.1) and – in at least some instances – *o*, before merger with **a* (the change presumably only “went to completion” when the labiovelar was followed invariably by **o*, so one finds such forms as Go. M nom. sg. *huas* beside gen. sg. *huis*, inst. sg. *hvē*, etc.).

Examples include:

- pre-Gmc **k^wntó-* ‘dog’ (to PIE **k^h[u]won-*: OInd *ś[u]vā*; Gk *kúōn*; Lith. *šuo*) > PGmc **hundaz* > Go. *hunds**.
- IPIE **g^wrx₂ús* ‘heavy’ (OInd. *gurús*; Gk *barús*) > PGmc **k^wuruz* Go. > *kaurus**.
- pre-Gmc *g^wh^htis* (to PIE *√*g^wh^hen-* ‘to strike, kill’) > PGmc **g^wunþijō* > OHG *gund-* ‘battle’.
- PIE **k^wól(X)-so-* ‘*turner, neck’ (with the “de Saussure Effect” [see Byrd, this handbook, 3.3]; to *√*k^welX-* ‘to turn’, cf. Lat. *collum* ‘neck’) > ePGmc **χ^wolsa-* > **χolsa-* > **halsa-* > Go. *hals* ‘neck’.
- PIE **g^wolb^hos* ‘(fruit of) the womb’ > ePGmc **k^wolbaz* > **kolbaz* > **kalbaz* > OIc. *kálfr* ‘calf’.
- Compare PIE *√*k^wex₂s-* ‘to cough’ (cf. MI *casachtach* ‘a cough’) ePGmc. **χ^wās-tan-* ‘a cough’ > **χ^wōs-tan-* > OE *hwōsta*.

2. Labiovelars were also delabialized – or, in the case of potential secondary ones, did not develop – when preceded by *u* (with a resonant allowed to intervene). A sequence

*uRC^w- becomes uRC-, as in Go. *tuggōn*- ‘tongue’ < ePGmc *tung^wō-n*- < PIE *d_hǵ^hwéx₂ (cf. OLat. *lingua*); Go. du. *ugk(-is)* ‘us two’ etc. < ePGmc **unk^w*- < PIE **ux₃-wé* (2.3.2.3[2]); also Go. *augō* ‘eye’ < **aug^w-an-*, from PIE **x₃ek^w-*, via contamination with ‘ear’ PGmc **auz-an-*).

Note also PGmc **murguz* ‘short’ (PIE **myǵ^hu-*), Go. *tulguz* ‘firm’ (PIE **d_lǵ^hu-*), forms that, with a different stem-vowel, would be expected to have a secondary labio-velar, cf. Go. *aggwus** (2.2.3.1).

The fact that we have Go. *wulfs* not †*wulhs* supports the pGmc date of the passage of **k^w* to **p* suggested above (2.2.3.2). Note, though, that Gmc **nagl-az* m. ‘nail’ from PIE **x₃nog^wh-(l-*) (cf. Lat. *ungula* ‘hoof’) unaccountably has delabialization (one might expect †*naula-*, cf. 2.2.3.1). A zero-grade allomorph **x₃ng^wh-* could be posited, becoming PGmc ***ung^wla-* (or analogical ***nug^wla-*, 2.3.3.1) and delabializing to ***ungla-* (***nugla-*), before being eliminated after -*g-* had been generalized throughout the paradigm.

3. Delabialization also took place before *j*: **sag^(w)jaz* ‘member of a retinue, man’ < pre-Gmc **sok^wyós* (cf. Lat. *socius*) > OIc. *seggr*, OE *secg*, OS *segg*; ON *ylgr* ‘she-wolf’ from an earlier shape **w(u)lg^(w)j-* F to PIE **w_lk^wos* (unless it was delabialized later by the preceding secondary *u* by [2]), the other Gmc languages reflecting a form rebuilt on the M (although still before VL) e.g., OHG *wulpa*, MHG *wülpe* < **wulb_ijō-*, (later rebuilt again as *Wölfin*).

2.2.4. Geminate stops

2.2.4.1. The only PIE “surface” geminate, /t/ [t^st], developed to -*ss-* in Gmc, as in Go. (*un-*)*wiss** ‘(un-)certain’ (cf. OInd. *vittás* ‘known’ < PIE morphological **wid-tos*).

2.2.4.2. In contrast to PIE, Gmc shows a large number of geminate consonants, especially voiceless geminate stops. These latter have been derived by a sound-law, known as Kluge’s Law (KL – the *locus classicus* in Kluge 1884; the most detailed recent treatment is Lühr 1988). This extends to obstruents the idea that resonants were geminated (via assimilation) before *n* (cf. 2.3.3.2). Before *n* immediately followed by the IE word-accent, fricatives (which in this environment had been voiced by VL) were occluded and geminated. They merged with the PIE voiced stops, which were also geminated by following *n*, and were devoiced along with them. (After long vowels, geminates do not occur, either because phonotactics prevented their creation or led to their simplification, cf. 2.1, 2.3.4.1) The change must be placed later than VL but earlier than the devoicing phase of GL (cf. 2.2.1.1).

Schematically for the dentals:

PIE	Fricativization	VL	Occlusion	Gemination	Devoicing
*-tn-’	> *-pn-’	} > *-dn-’	} *-dn-’	> *-dd-’	> *-tt-
*-d ^h n-’	> *-ḍn-’				
*-dn-’	>				

Many examples are found among *n*-stem nouns or belong to roots that formed nasal-infix presents in PIE; it can be assumed that many of these were rebuilt as nasal-suffix presents in pGmc. The change led (via analogies) to consonant alternations in these categories. In fact, the morphological distribution of geminates within Gmc is a strong internal argument for KL.

KL can also account for instances of PGmc root-final *-p* without having to posit the rare PIE **b*.

Examples:

- PGmc **deup(p)a-* Go. *diup-* ‘deep’, to PIE $\sqrt{*d^heub^h-}$ (for the root shape cf. Gk *Typháōn* a monster, and OE *dūfan* ‘to dive’ without KL), cf. **d^hub^h-nó-* ‘deep’ (Gaul. *Dubno-rix*, CSL. *duno* ‘ground’, Lett. *dibens*, Lith. *dugnas* with secondary velar); cl II weak verb **duppōjan-* Nw. *duppa* ‘dip, float’, cf. OE *doppettan* ‘immerse’ < pre-Gmc **d^hub^hnā-*.
- PGmc F *n*-stem nom. sg. **lápōn* ~ gen. sg. **lađnés* > **lattes* ‘plank’ OHG nom. sg. *lada* (on the *d*, see 2.2.2.3) and rebuilt OS *latta*.
- PGmc **hwitt-* ~ **hwūt(t)a-* ‘white’ Du. *witt*, Go. *hveits**, to PIE $\sqrt{*k^hweyt-}$ ‘to shine’ (IOInd. 3sg. pres. mid. *śvindate*, with secondary *d*) OInd. *śvitna-* (for expected †*śvitná-*) beside *śvetá-* ‘bright, white’.
- PGmc **likkōjan-* ‘to lick’ OS *likkon* < pre-Gmc **liġ^hnā-* (cf. Lat. *lingere*).
- Go. *n*-stem nom. pl. *smakkans* ‘figs’, probably to PGmc **smakkōjan-* ‘to savour’ and family, further related to OHG *gismag* ‘tasty, pleasant’ and perhaps Lith. *smaguriaĩ* ‘delicacies’, suggesting pre-Gmc **smog^h-*.
- Cf. the above with the following with preceding accent: Go. *aþn** ‘year’ < pre-Gmc **átnos* (cf. Lat. *annus*), ON *svefn*, OE *swefn* ‘sleep’ < pre-Gmc **swépnos* (cf. OInd. *svápnas*).

It is necessary to assume that *-ɣ^wn-*, which never undergoes gemination, became *-wn-* (2.2.3.1) between VL and Occlusion and that the change of **-mn-* to **-bn-* (2.3.3.3) was later.

Apparent counter-examples can be taken as having been formed after the sound-law ceased to operate or having incorrect pre-forms or etymologies. Still, it should be noted that Kluge’s Law is rejected by many, who instead invoke the nebulous concept of expressive gemination. However, expressivity cannot explain why the geminated outcome is so often voiceless. Nor can it explain the proliferation of geminates in specific morphological categories, such as *n*-stem nouns; nor the varying outcomes of split paradigms where one variant is not appreciably more expressive than another (e.g., G *Knabe* vs. *Knappe*).

2.3. Laryngeals, liquids and nasals, and semi-vowels

2.3.1. Demise of the system of “coefficients sonantiques”

The IE class is identified on the basis of its phonological patterning: less sonorous than vowels but able to form a syllable-peak (and excluding PIE **s*, 1.2).

1. Laryngeals ceased to exist. In the overwhelming majority of cases they were simply lost, although between a short vowel and a consonant they caused lengthening of the vowel (3.1.1). Otherwise, they merged with vowels (*a* and perhaps *u*), possibly other consonants (reflex Gmc **k*, 2.3.2.3), or left indirect reflexes in accented (2.3.4.1) and unaccented syllables (4.4.2).
2. Resonants: their syllabic versions developed into sequences *uR*.
3. Semi-vowels split into vocalic and consonantal phonemes (including the second element of diphthongs).

2.3.2. Laryngeals

2.3.2.1. Reflexes of syllabic laryngeals

In initial syllables, syllabic laryngeals became *a*: PIE **px₂tér-* > **faðar-* ‘father’, PIE **stx₂tís* (OInd. *sthiti-* ‘standing’) > **stadiz* ‘place’ Go. *staps*, OE *stede* (cf. 3.1.1, 3.2).

In final syllables, a development to *-u-* has been suggested, for which the best candidate is PGmc **anuð* OHG *anut*, ON *ǫnd* (with labial mutation, see end of 3.3.5) ‘duck’ nom. sg. (on the assumption of an IE *t*-stem nom. sg. **x₂enX-t-s*, cf. Lat. *anas* gen. sg. *anatis*).

2.3.2.2. Reflexes of non-syllabic laryngeals

Non-syllabic laryngeals were lost word-initially, before consonants (PIE **x₁dónt-* > PGmc **tanþ-* ‘tooth’) as well as vowels, where they left only their indirect coloring effects. For medial position: PIE **dugx₂ter-* (OInd. *duhitár-*, Gk *thugátēr*) > **duxtar-* ‘daughter’; PIE **x₂erx₃trom* (Gk *árotron*, OI *arathar*) Oic. *arðr* ‘plough’ (see also the examples in 2.3.3.1–2., 2.3.4.1).

2.3.2.3. Supposed consonantal reflexes (“laryngeal hardening”)

1. In some instances, Gmc appears to show **k* as a reflex of a PIE consonantal laryngeal. As there are numerous examples of secondary *-k-* in Gmc, the conditions under which this may be taken to be a laryngeal reflex need careful constraining. (Particularly, there is a tendency for the semi-vowel *-w-* to develop into *-g-* when not adjacent to an obstruent, and one source of secondary **k* in the Gmc languages is probably **g* of this origin via the Gmc consonant shift. After the consonant shift, the tendency produced examples of exrescent (spirantal) *-γ-*. Examples with **k* of uncertain origin: OHG *hackōn* ‘to hack’, which might have some connection with OHG *hauwan* ‘to hew’; ON *nokkvi*, OHG *nacho*, *nacko* ‘small boat’, which looks as though it might be connected with PIE **néx₂-us* ‘boat’; examples with **g*: OHG *jugund* ‘youth’ beside Go. *junda* dat. sg., OE *nigon* ‘9’ < **niwun* beside OHG *niun*; OE, OS *sugu* ‘sow (pig)’ from the *suw-* of oblique cases (cf. 3.1.1).

Laryngeal hardening is proposed for instances where the laryngeal was pre-tonic and immediately followed by **w*. To be certain that supposed laryngeal instances of **k* do not reflect **w*, one needs to be confident the **w* itself was retained and it is the laryngeal, not this, that is the source.

2. The strongest candidates all feature the third laryngeal preceded by a semi-vowel (pure vowels in this position underwent lengthening, 3.1.1):

- PGmc **unk^w*- ‘us two’ > Go. *ugkis*; ON *okkr*; OS *unk* from PIE **ux₃-wé* 1du. personal pronoun oblique (cf. OInd. *āvām* < **āva-* + *-ām*; Gk *nó*, *nóe* < **nōwe*).
- PGmc **ink^w*- ‘you two’ > Go. *igqis*; ON *ykk*; OS *ink*, remodelled from PIE **ux₃-wé* 2du. personal pronoun oblique (cf. OInd. *yuvām* for ***ūvām* < **ūva-* + *-ām* after nom. *yuvām*), with the Gmc initial vocalism adopted from the presumed nom. sg. Go. **jut*, supported by the 2pl. form as in Go. *izwis*, and the nasal contaminated from the 1st person form.
- PGmc **k^wik^waz* ‘alive’ (ON *kvikr*, acc. sg. M *kwikkvan*; OE *cwic*, OHG *queh*, *quec*) < **g^wix₃-wós* (OInd. *jīvās*, Gk *zōós*).

In the first example, the PIE starting-point indicates the former presence of a *w*, which may be presumed to have led to a secondary labiovelar by 2.2.3.1(2), subsequently delabialized according to 2.2.3.3(2).

The remodelling of the second example is to be seen in the context of the parallelism between the dual and plural oblique forms in Gmc, where in the 1st person they both begin with *un-* (cf. Go. pl. *uns[-is]*, reflecting PGmc **us-wé*, cf. 2.2.3.3[2]) and in the 2nd with *i-* (cf. Go. pl. *izwis*, from **us-wé* with analogical initial vowel from nom. *jus*). But once a remodelling is accepted, the form is less probative as an example of hardening.

Concerning the third example, which also developed a secondary labiovelar, we may note that the N-Gmc accusative form shows a reflex of the **-w-* (2.2.3.1[4]), which has geminated the preceding **k* (2.2.2.1[4]); gemination also in OHG *quec*. Go. *qius* lacks the **k*, however. The situation shows similarities with that regarding Dybo’s Rule (cf. 3.1.1), in that the laryngeal seems sometimes simply to have been lost without producing the expected reflex. (Cf. also OI *béo* ‘living’, as if from **g^wiwós*.)

Because the third laryngeal is considered to have been distinctively voiced (Byrd, this handbook, 3.3), it must have undergone the consonant shift to produce the voiceless reflex. Thus its hardening (i.e. occlusion) precedes this change.

2.3.3. Liquids and nasals

2.3.3.1. **ŕ* > **uR*, unless in a paradigmatic relationship with a full-grade *RV* form, in which case the outcome was often analogically remodelled as **Ru* (laryngeals were presumably already lost, otherwise **RuX* would have become †**rū*). Thus, the PIE negative prefix **ŕ-* > Gmc *un-*, PIE **k^wntóm* ‘100’ in Go. *hundafaps* ‘centurion’, PIE **k^wldos* (Gk *kládos* ‘branch’) > OSw. *hult* ‘wood’, pre-Gmc **g^wXnó-* ‘crushed, ground’ (Lat. *grānum* ‘grain’, Lith. *žirnis* ‘pea’) > ORu. **-kurne** dat. sg. ‘corn’. See also the examples in 2.3.3.2.

Analogically metathesized examples of **Ru* are: PGmc **nus-* after **nas-* ‘nose’, Go. *lustus** ‘desire’ Ø-grade formation to √**las-* ‘be avaricious, unrestrained’; but note OE *bred* beside *bord* (with *a*-umlaut, 3.3.1.2) ‘board’. In contrast, **snuzō-* ‘daughter-in-law’ (PIE **snusó-* F) provides an instance of Gmc *u* next to a resonant not developed from a syllabic resonant but of IE ancestry (cf. 3.2).

2.3.3.2. Gemination of at least *l* occurred before *-n-* when the accent followed: Go. *fulls* ‘full’ < **p_lx₁-no-* (OInd. *pūrṇás*, Lith. *pilnas*), OE *hyll* ‘hill’ < **hulli-* < **k_lXnis* (Lat. *collis*; cf. Lith. *kálnas* ‘hill, mountain’), Go. *wulla** ‘wool’ < **Xw_lx₁nex₂* (Hitt. *hulana-*; OInd. *úrṇā*; Lat. *lāna*; Lith. *vilna*).

2.3.3.3. The sequence **-mn-* became **-ḅn-* (after *Cn*-gemination, 2.2.4.2, as the outcome is not subject to it): Go. *ibns** ON *jafn*; OE *efen*, OHG *eban* ‘even’ < PIE **emnós* (OInd. *amnás* ‘just now; immediately’); also OIc. nom. sg. *nafn* ‘name’ from PGmc **naman-*, backformed to case-forms where **namn-* occurred. The change is subject to analogical disturbance and by-forms abound: ON *jamn*, OE *emn* ‘even’; Go. *stibna*, OE *stefn* ‘voice’ beside *stemma*, OHG *stimna*.

2.3.4. Semi-vowels

2.3.4.1. Intervocalic semi-vowels were geminated after a short vowel in initial syllables in PGmc (mostly resolving a hiatus after the loss of a following laryngeal: *-VWXV- > -VWWV-). This phenomenon is known as Holtzmann’s Law (first outlined 1835: 862–63); also called *Verschärfung* (intensification). WGmc reflects simple geminates, whereas Gothic and NGmc shifted *-Vww- to -Vggw- (PGmc **glawwa-*: OWN *gløgg_r* ‘clear-sighted, clever’, Go. *glaggwō* adv. ‘precisely’ vs. OE *glēaw*, OS *glauuu_a* M nom. pl. ‘prudent’) and *-Vjj- to -Vggj- in Norse and to -Vddj- in Gothic (PIE genitive **d[u]woy-Xow*; PGmc **twajjō*”: ON *tveggja*, Go. *twaddjē* vs. OHG Isidor *zuueiio* ‘2’, with regularized ending). Cf. Jasanoff (1978).

Some geminates are of “non-laryngeal” origin, for example in the word for ‘egg’ ON *egg*, “Crim.Go.” *ada* (implying Gothic **addja*) vs. OS, OHG *ei* gen. pl. *eiiero*, from IPGmc. **ajja*”, shortened by Osthoff’s Law (Byrd, this handbook 5.4) from **ōjjo-*, itself probably assimilated from something like pre-Gmc. **ōwyom*; PGmc. **dajjan* (Go. *dad-djan** OSw. *dæggja*) ‘suckle’ probably represents a secondary causative **d^hoy-éye/o-* (presumably from the same secondary stem that underlies Sl. *dojō* ‘I suckle’), with ePGmc. **daijjan* simplifying to **dajjan*.

2.3.4.2. Where not geminated in initial syllables (2.3.4.1), intervocalic *-j- was lost everywhere in Gmc, except after **i* (but *-iji- yielded *-ī-, as in PIE **tréyes*, OInd. *tráyas*, via PGmc. ‘virtual’ **prijiz* > OIc. *prír* ‘3’ M nom., beside Lat. *trēs*): Go. *aiz* ‘money’ (< PIE **áyes*, cf. OInd. *áyas-* ‘metal’); Go. *frijōn* ‘to love’ (< PIE **prix₂-ex₂-ye/o-*). Cf. Thórhallsdóttir (1993). For examples in unaccented syllables, see 4.2.2.

In North Germanic, **j-* was lost word-initially: *ár* ‘year’ (Go. *jēr*; on the stem vowel, see 3.3.2.1), *ostr* ‘cheese’ (cf. Finnish loanword *juusto*); while word-initial **w-* was lost before *u* (*ull* ‘wool’ Go. *wulla**, *ulfr* ‘wolf’ Go. *wulfs*) and *o* (*ormr* ‘snake’ Go. *waurms*;

orþ ‘word’ Go. *waurd*) and before liquids (*litr* ‘figure, appearance’ Go. *wlits*; *rata* ‘to wander’ Go. *wratōn**).

2.3.4.3. The sequence **-nw-* was assimilated to **-nn-* in PGmc, as in: Go. *kinnus** ‘chin’ < **kinw-* < PIE **ĝenw-* (Gk *génus*, Toch A *śanweṃ* ‘jaw’, Lat. *genuīnus* adj ‘belonging to the jaw’); OIc. *þunnr* ‘thin’ (*u*-stem adj, cf. OInd. *tanús*, Lat. *tenuis*, and 2.2.3.1(2) ad Go. *aggwus**).

3. Vowels

3.1. PGmc inventory

3.1.1. Compared with the consonants, the developments affecting vowels from late IE to PGmc were straightforward. The post-IE inventory of five short and long vowels (the latter including those developed from **VX*-sequences, see examples in 3.2) was reduced to four by the merger of *a* and *o* qualities (for evidence that the qualities remained distinct for a time in Gmc, see above 2.2.3.3). The final outcomes were short *a* and long *ō*, although an initial merger under one value (*ā*) is probable. PGmc *u* was augmented by the development of syllabic **ʀ* to **uR* (2.3.3.1). PIE syllabic laryngeals yielded short *a* in initial syllables and possibly short *u* in final syllables (2.3.2.1). The PIE sequence **ey* became long *ī* (= *ij*) everywhere.

PIE sequences of (semi-)vowel plus laryngeal in initial syllables, rather than becoming long vowels, are sometimes reflected by short vowels in Germanic and other West-Indo-European languages, with loss of the laryngeal when pre-tonic before a non-syllabic resonant (so-called Dybo’s Rule, 1961). The clearest examples are: PGmc **sunuz* ‘son’ (beside OInd. *sūnú-*, Lith. *sūnūs* < PIE **suX-nú-z*, to root $\sqrt{*sewX-}$ ‘to give birth’); PGmc **wiraz* ‘man’ (cf. Lat. *vir*, OI *fer*; beside OInd. *vīrá-*, Lith. *vyras* < PIE **wiX-ró-s*, probably to the root $\sqrt{*weiX-}$ ‘exhibit vitality’). Compare also Go. *qius* < PIE **g^wix₃wós*, with apparent loss pre-tonically before a non-syllabic semi-vowel, where North and West Germanic have forms with a consonant (see 2.3.2.3[1]). Note, however, monosyllabic PGmc, **sūz* ‘sow (pig)’ (< PIE **súX-s*) with no shortening.

In sequences of the type short vowel plus (velar) nasal plus χ , a phonetic long nasalized vowel developed within PGmc. These were subject to denasalization in all the daughter languages (3.3.3).

3.1.2. According to the standard reconstruction of late PGmc, the vowel system consisted of 14 items; four short vowels, four long vowels, three nasalized long vowels, and three diphthongs.

Short	Long	(Long) Nasalized	Diphthongs
i u	ī ū	ĩ ũ	Ø eu
e a	ē ō	ã	ai au

On long **ē₂*, see 3.3.2.2. On developments of vowels in unaccented syllables, see 4.

3.2. Basic correspondences with “IPIE”

- PGmc *i (< PIE *i) Go. *fisks** ‘fish’ (cf. Lat. *piscis*), Go. *widuwō*. OS *widowa* ‘widow’ (cf. OInd. *vidhāvā*).
- PGmc *e (< PIE *e) OE, OS *etan* ‘to eat: (PIE $\sqrt{*x_1ed-}$, as in Hitt. *ēdmi*, OLith. *ėmi*; Lith. *ėdu*, Lat. *edō* ‘I eat’); NGmc (ORu.) **swestar** voc. ‘sister’ (OInd *svāsā*, Lith. *sesuō*).
- PGmc *a (< IPIE *a) Go. *akrs* ‘field’ (cf. Gk *agrós*, Lat. *ager*). (< IPIE *o) Go. *ahtau* (cf. Gk *oktṓ*, Lat. *octō* ‘8’). (< PIE *X) in Go. *fadar* voc., OHG *fater* nom. ‘father’ (cf. OInd. *pitā́*, Gk *patér*, Lat. *pater*) < PIE * $px_2tér$.
- PGmc *u (< PIE *u) Go. *juk** ‘yoke’ (cf. OInd. *yugám*, Gk *zugón*), **snuzō*- F OHG *snur* ‘daughter-in-law’ (PIE **snusó*- F, cf. Gk *nuós*, OInd. *snuśā́*). Also from syllabic \tilde{r} (2.3.3.1).
- PGmc *ī (< IPIE *ī) Go. *swein** ‘swine’ < **swīno*- (cf. Lat. *suīnus* ‘porcine; pork’), 2sg. Go. *wileis* ‘thou wilt’ < **welīs* (cf. Lat. *velīs*). (< IPIE *ey) Go. *steigan** OHG *stīgan* ‘to climb’ (cf. Gk *steikhō* ‘I go, step’).
- PGmc *ē (< IPIE *ē) Go. *-dēps* < PGmc **dēdiz*, to PIE root **d^hex₁-* ‘to put, place’.
- PGmc *ō (< IPIE *ā) **brōpēr* ‘brother’ (cf. Gk *phrātēr* ‘member of a clan’). (< IPIE *ō) **dōmaz* ‘judgment’ < PIE **d^hóx₂mos* ‘thing put, set down’ (cf. Gk *thōmós* ‘heap’).
- PGmc *ū (< IPIE *ū) OE, OHG *mūs* ‘mouse’ (cf. Gk *mūs*, Lat. *mūs*).
- IPGmc *ī, as Go. *leiht** ‘light(weight)’ < **lixtaz* < PIE **x₁leng^{wh}-t-* (cf. 2.2.3.1[3]), beside OE *lungre* adv. ‘quickly’ (cf. Gk *elaphrós* ‘light; nimble’ < PIE **x₁lŋg^{wh}-*’).
- IPGmc *ā, as Go. *hahan** ‘to hang’ to **kenk-* (cf. OInd. *sāṅkate* ‘hesitates’); Go. *brahta* ‘brought’ 3sg pret. to *briggan* ‘to bring’.
- IPGmc *ū, as Go. *puhta* ‘seemed’ 3sg pret. to *puhkjan** ‘to seem’.
- PGmc *eu (< IPIE *ew) OE *cēosan* ‘to choose’ (3.3.4) PIE $\sqrt{*g^hews-}$ (cf. Gk *s*-aorist *geúsasthai* ‘to taste’).
- PGmc *ai (< IPIE *ay) Go. *gaits* ‘goat’ IPIE **g^haid-* (cf. Lat. *haedus*). (< PIE *oy) Go. *wait*, OE *wāt*, OHG *weiz* ‘I know’ < PIE **wóidx₂e* (Gk *oída*).
- PGmc *au (< IPIE *aw) Go. *aukan** ‘to increase’ (cf. Lat. *augēre*). (< IPIE *ow) Go. *ausō* ‘ear’ IPIE **ows-os-* (cf. Gk [Attic] *oús*).

3.3. Main developments in the dialects

Most of the early Gmc dialects developed complex vowel systems, largely as a result of the assimilatory effects on stem-vowels of following consonants or the vowels in following syllables. NGmc and OE were the most subject to such changes. Only some of the major ones will be mentioned here.

3.3.1. Short vowels

3.3.1.1. In Gothic, *e* became *i* including in the diphthong **eu*: Gothic *itan* vs. OS *etan*; *kiusan* beside OE *cēosan*. However, *i* and *u* were lowered before *h*, *hv* and *r* and spelled <ai> and <au>, respectively: *taihun* ‘10’, *saihvan* ‘to see’, *waurd* ‘word’.

3.3.1.2. “A-umlaut”

The North-West-Germanic languages created a new short /o/ from PGmc. **u* as part of a complex of changes, conveniently labelled “a-umlaut”, which involved lowering of short [u] and lowering and raising of short [i] and [e]. There were three relevant environments: those that conditioned a high vowel (before *i*, *ī*, [iu] and *j* in the following syllable, and when the intervening consonantism began with a nasal); a low vowel (before short *a*, long *ō*); or were neutral (before *u*, *ū* and *ē* or *Ø*), preserving the etymological vowel.

Examples: OHG *gibit* ‘gives’ beside *geban* ‘to give’, *ring* ‘ring’ (< **hrengaz*, cf. Finnish loanword *rengas*), *metu** ‘mead’, *sehs* ‘6’; *horn* ‘horn’, *gold* ‘gold’ (beside *gulđin* ‘golden’ and Go. *-gulþ* ‘gold’), *hunt* ‘dog’.

The outcomes are much disturbed by analogy and double-forms abound. It may also be noted that so-called Crimean Gothic perhaps is to be classed as a North-West-Germanic dialect, as it attests a-umlaut; note especially *Schuos* [for *Schnos**]: *Sponsa* ‘fiancée’ in the list of words not recognized by Busbecq (‘cum nostra lingua non satis congruentia [not sufficiently in agreement with our language]’).

3.3.1.3. In North Germanic, PGmc short **e* underwent Breaking before heterosyllabic *a* and *u*: **eĥnaz* ‘even’ > *jafn*, **etunaz* ‘giant’ > *jōtunn*. (This change preceded Labial Mutation, see end of 3.3.5, because any form that could have either change always has breaking, thus *jór* ‘horse’ – not †*ór* – < PNGmc **ehwaz*, *rōkkr* ‘darkness’ < PNGmc **rekkwaz* – no breaking after *r* – beside *þjōkkr* ‘thick’ < NGmc **þek^wuz* ~ *þekk^waz*.)

3.3.1.4. In OE and – to a lesser extent – OFris., short front vowels are diphthongized by tautosyllabic grave consonants (Breaking). A few examples must suffice: Breaking affects only short *i* and *e* in Fris., but all front vowels, both short and long in OE: OE *feohtan*, OFris. *fiuchta*, beside OHG *fehtan* ‘to fight’; OE *meaht* beside OFris. *macht*, OHG *maht*; OE WS *nēah* beside Go. *nēhv*, OHG *nāh* ‘near’. For Frisian labial breaking, see examples in 2.2.3.1.

Back-mutation – OE only – is caused by *u* and *a* in the following syllable *meodu* beside OHG *metu** ‘mead’, Kentish *neofan* beside OHG *nefun* ‘nephews’.

3.3.2. Long vowels

3.3.2.1. Long **ē* was retained in Gothic, but became long **ā* in NWGmc accented syllables (while retaining its quality in unaccented syllables, cf. 3.3.2.2). In the Ingvæonic dialect area – OE, OFris., and parts of OS – this accented long **ā* split into positional variants that came to belong to different phonemes: before nasal consonants, it was nasalized and backed and subject to rounding, often appearing spelled <o>; elsewhere it was fronted.

Go. *ga-dēps** ‘deed’, *mēna* ‘moon’; ON *dáð*, *máni*; OHG *tāt* and *māno*, and (parts of) OS *dād*, *māno*.

OE, OFris. *mōna* ‘moon’, OS <monohtlic>, for *mōnohtlic**, ‘monthly’, beside non-Ingvæonic *mānuth* ‘month’; as opposed to OE (WS) *dæd* ‘deed’, (Anglian) *dēd*, OFris. *dēd(e)*, OS *gēbun* ‘gave (pl.)’, beside *gābun*.

3.3.2.2. Long $*\bar{e}_2$

1. The NWGmc languages have two long mid/low unrounded vowels in accented syllables: $*\bar{a}$, the continuation of Gmc (and PIE) long $*\bar{e}_{(1)}$ (3.3.2.1), and $*\bar{e}$ (> OHG *ea*, *ia*, *ie*, 3.3.2.3), known as $*\bar{e}_2$. NWGmc retained the quality of long $*\bar{e}_{(1)}$ in unaccented syllables, where it was augmented by monophthongization of $*-ai\#$ to $*-\bar{e}\#$ (4.5.3). $*\bar{e}_2$ is not obviously the continuant of any PIE sound and there is only one relevant lexeme that can be securely reconstructed for PGmc in which it appears (see 3). It is best represented in a limited set of categories where it is the result of secondary developments. These comprise (a) the stem-vowels of rebuilt class 7 strong verb preterits (see Harðarson, this handbook), (b) the product of various phonological developments within WGmc, and (c) loanwords, mainly from (late) Latin.
 - a) The NWGmc opposition between present and preterit stems in class 7, as in OE (WS) *slāpan slēp*, could only be implemented if ($*\bar{e}_1$ had become long $*\bar{a}$ and) \bar{e}_2 already existed.
 - b) On the one hand, such monosyllabic pronominal forms as OE *wē*, OS *wē* (also *wī*) ‘we’ beside Go. *weis*, ON *vér* (and OHG *wir*); IWGmc $*s\bar{e}$ M nom. sg. ‘the, that’ OE *sē* (< WGmc. creation $*siz \neq$ Go. *sa*, ON *sá*); OE *mē* ‘to me’, but OHG *mir*, beside Go. *mis*. On the other, OE (WS) *mēd* (but Angl. *meord*) OHG *miata* ‘meed, reward’, beside Go *mizdō*. And, outside Ingvæonic, $*-īr-$ > $*-\bar{e}r-$, as OHG *wiara* ‘ornament of precious metal’ beside OE *wīr* ‘wire’.
 - c) OHG *spiagal* ← Lat. *speculum* ‘mirror’; some of the loanwords are also found in Gothic: *Krēkōs* (with sound-substitution of the initial consonant, as pre-Go. $*\gamma$ - was still a fricative at the time of the loan, 2.2.1.3), OHG *Kriachi* (with the Second Consonant Shift) cf. Lat. *Graecus* ‘Greek’; *mēs*, OHG *mias* ← Lat. *mensa* ‘table’, but they can hardly be used to establish a PGmc phoneme.
2. Gothic, by contrast, cannot be demonstrated – or assumed – to have (reflexes of) more than one long \bar{e} vowel synchronically, if one analyzes the ending of *r*-stem nom. sg. *brōþar* etc., as containing a short vowel (see 4.4.3.1); and interprets outcomes of the sequence $*an$ as still containing a nasalized vowel (3.3.3).
3. The position taken here is that long $*\bar{e}_2$ is a post-Proto-Germanic creation. The place adverb of rest ‘here’ is the only inherited word common to all the Gmc languages that displays $*\bar{e}_2$. It is hard to see how a phoneme /e:₂/ existed in PGmc, as a “contrastive sound” can hardly be posited for a single lexeme.

The expected PGmc pre-form for ‘here’ is $**hir$ with a short vowel; the Go. imperativized derivatives *hiri*, *hirjats* ‘come hither!’ sg. and pl. with a short stem-vowel may be noted, also *hidrē* ‘hither’. One can account for $*\bar{e}_2$ by positing a lowering before *r* and a lengthened allophone [he:r] that developed under emphatic accent, only achieving phonemic status in the daughters. (However, Fris., OS and the OHG Tatian translation show reflexes of *i*-quality: $*hīr$.)

The only other word to contain long \bar{e}_2 that looks as though it could go back to Proto-Germanic is Go. *fēra* ‘side, region, limb of the body’, OHG *fiera* ‘side, direction’. It has no convincing etymology and is found in only two daughters, and the Old High German word could be one of this language’s loans from Gothic.

3.3.2.3. Long \bar{e}_2 and \bar{o} were diphthongized in OHG (and some OS texts): *hear*, *hiar*, *hier* ‘here’; *bruoder*, *bruader* ‘brother’.

3.3.3. Long nasalized vowels

In Gothic, analogical developments suggest that PGmc **ī̃* was no longer nasalized at this stage: e.g., PGmc class 3 verb **þinχan* > class 1 **þīχan* Go. *þeiħan* ‘to thrive’. By contrast, *ã* was presumably still nasalized; otherwise Gothic would have a long oral *ā*-vowel that occurred in only one context (cf. 3.3.2.2[2]).

In North Germanic, the First Grammarian indicates that such vowels were still nasalized in the 12th century (cf. Benediktsson 1972).

In Ingvæonic, PGmc nasalized *ã* (3.1.1) was joined by WGmc accented long **ā* before nasals (3.3.2.1). A category of nasalized vowels (also from additional sources) is retained to this day in the archaic dialect of Älvdalen in northern Dalarna in central Sweden.

3.3.4. Diphthongs

The passage of PIE **ey* to long *ī* (3.1.1) created a hole, making the system of diphthongs unbalanced (3.1.2). In most Gmc languages, *eu* develops in isolation and *ai* and *au* enter into a front–back opposition based on the nature of their second element (and they tend to monophthongize). In OE, the sole diphthong in final *-i* was eliminated by monophthongization to long *ā*, while *eu* and *au* became “horizontal diphthongs” *ēo* and *ēa* (*āea*), entering into a vertical relationship based on the degree of opening of their first element: *wāt* ‘I know, knows’ (Go. *wait*), *bēodan* (Go. *-biudan**) ‘to command’, pret. *bēad* (Go. *-baup*).

3.3.5. A note on *i*-mutation (and NGmc Labial Mutation)

The NWGmc languages all show a change called *i*-mutation (*i*-umlaut or, particularly in ON, front mutation), which involves the fronting and raising of low and back simple vowels and diphthongs by a process of feature transfer from high-front segments in succeeding syllables. It was independent in each of them, as the different implementations show, and is a result of the weakening of post-tonic syllables following the fixation of stress on initial syllables (1.2.[4]). It was characteristically caused by an **i* or **j* in the following syllable. In NGmc, *-R* could also cause front-mutation on its own (or in combination with *i*, see below): ON *gler* ‘glass’ < **glara*, *eyra* ‘ear’ < **aurō* (cf. Go. *ausō*).

Similarities and differences can be seen in the following set of examples.

IPGmc	Go.	ON	OE	OHG
<i>*stapiz</i> ‘place’	<i>staps</i>	[<i>staðr</i>]	<i>stede</i>	[<i>stat</i>]
<i>*gastiz</i> ‘guest’	<i>gasts</i>	<i>gestr</i>	<i>giest</i>	<i>gast</i>
<i>*gastiz̄</i> ‘guests’	<i>gasteis</i>	<i>gestir</i>	<i>giestas</i>	<i>gesti</i>
<i>*hauzijan</i> ‘to hear’	<i>hausjan</i>	<i>heyra</i>	<i>hieran</i>	<i>hōren</i>

In ON, vowels in light disyllables did not undergo the change before *i* (unless the next syllable contained *r*, thus **batir* > *betr* ‘better’ adv.). The lack of *i*-mutation in nom. sg. *staðr* is to be explained by levelling from the other case-forms.

In OHG, *i*-umlaut took place later than the loss of short *-i* after heavy syllables (hence *gast* beside the OE and ON singulars); the reflexes were not noted in the spelling apart from the product of short **a* (cf. G *hören*). Certain intermediate consonants hindered *i*-mutation, thus *mahti* pl. to *maht* ‘power’, beside *ensti* to *anst* ‘grace’. OHG *stat* has analogical lack of both final *-i* and *i*-mutation.

NGmc shows a parallel set of changes, known as Labial Mutation, involving the rounding of vowels and caused by labial elements in the following syllable: *land* pl *lōnd* < **landu*, *sökkva* ‘to sink’ < **sekkwa*, corresponding to both Go. *sigqan* ‘to sink, go down (intransitive)’ and *sagqjan* ‘to cause to sink, bring down’, **blīwa* > *blý* ‘lead (the metal)’.

4. Post-tonic syllables

As a result of the adoption of a stress accent on initial syllables and subsequent loss of the autonomy of the syllable (cf. 1.2[4]), the development of sounds in tonic and post-tonic syllables diverged in Gmc.

4.1. Final consonants

PIE final stops (in effect dentals) were lost except in monosyllables as a single change that took place before any vowel losses and shortenings.

PIE final **-m#* became **-n#* in Gmc and then disappeared (except in light monosyllables), causing nasalization of the preceding vowel (denoted *-Vⁿ*). The nasalization was subsequently lost.

Examples: Go. *wili* (from **welīt*, cf. Lat. *velit*), 3sg. of the verb ‘will’, has a shortened vowel, while the 2sg. Go. *wileis* (from **welīs* – cf. Lat. *velīs* – and thus formally identical apart from the final consonant) does not.

PGmc **in* < PIE **(x₁)en*, Go. acc. sg. M *þana* < PGmc **þan* < PIE **tóm* (cf. OInd. *tám*, Lat. *is-tum*) + particle **-ōⁿ*; Go. acc. sg. F *þō* < IPIE **tām* (cf. OInd. *tām*, Lat. *is-tam*).

s and *r* are the only IE final consonants to survive the PGmc period in polysyllables. Final **-z* from IE **s* by VL (cf. 2.2.1.2) was devoiced in Gothic (M nom. sg. *arms* ‘arm’, *fōtus* ‘foot’, 2.2.2.1), became *-r* in NGmc (*armr* ‘arm’, *fōtr* ‘foot’), and was lost in WGmc polysyllables (OE *earm* [with breaking, 3.3.1.4], *fōt*). However, in unaccented light monosyllables, **-z* was retained and rhotacized in certain OHG pronominal forms, as: *er* ‘he’, *ir* ‘ye’, *wir* ‘we’.

4.2. Semi-vowels

4.2.1. Sievers’ Law (SL). Whatever the situation in PIE (see Byrd, this handbook, 5.3), it seems that Gmc had a version of the regulation of the semi-vowels, such that pre-

vocalic post-consonantal *j was replaced by *ij after a heavy syllable. Germanic shows no reflexes of initial SL (Lindeman's Law variants, see Byrd, this handbook, 5.3) in monosyllables: PIE *d(u)wo- etc. : PGmc *twa- 'two'.

SL was alive in ePGmc, as shown by adjustments consequent on changes in syllable weight caused by loss of medial laryngeals (which made some heavy syllables light) and the resolution of syllabic resonants (which made some light syllables heavy). Thus Go. *ga-tamjan*, ON *tamja* 'to tame' < PIE *domx₂-éye- (OInd. pres. indic. 3sg. *damáyati*) acquired a light stem-syllable when the laryngeal was lost; whereas Go. pres. indic. 3sg. *waurkeiþ* 'works' < *wurkīþ < IPGmc *wurkijþ < PIE *wŕǵ-ye-ti had a heavy first syllable only after vocalic *-j- became *-ur- in PGmc. Note also causative with suffix *-éye-: *log^h-éye-ti > ePGmc *lagijīþ → IPGmc *lagijīþ > OHG *legit* '(s)he lays'. In each Gmc branch, the automatic regulation was probably ended by loss of *-a, which destroyed many -(i)jV- contexts.

4.2.2. *-j- was lost before *i, and both *-j- and *-w- were lost between vowels in non-initial syllables when preceded by a back vowel. Accordingly, *-ijV- was the only sequence where intervocalic yod was retained (unless V was i, 2.3.4.2).

Examples: PIE 3sg *léghyeti* '(s)he is lying down' > PGmc. *ley(j)idi > OHG *ligit*, beside OHG 3pl *liggent* (with WGmc gemination, 2.2.2.1) < *ligjandi < PIE 3pl. *léghyonti*; present inflection of weak class II verbs, as in 1sg. *salbô < pre-Gmc *salbājō; 2sg. impv. *salbô < *salbōe < pre-Gmc *salbāje (cf. also 4.4.3.1).

Go. 1st du. pres. *bidjōs* < PGmc *bedjōz < PIE *g^{wh}ed^h-yowelos (but pre-Gmc nom. pl. *sunewes > PGmc *suniwiz > Go. *sunjus* 'sons').

IE *-eye- yielded *-ī- (via *-ije-, 3.1.1, 4.5.1, and virtual *-iji-, 2.3.4.2): *g^hósteyes 'strangers' > IPGmc *gastīz Go. *gasteis*, ON *gestir*, OHG *gesti* 'guests' (beside Lat. *hostēs*).

4.3. Short vowels

4.3.1. Absolute-final mid-vowels were lost in PGmc: Go. 1/3sg. pret. pres. *wait* (= Gk *oída*, -e), Go. 2pl. pres. *bairiþ* (= Gk *phérete*). The examples show that final stops were lost earlier (4.1).

Nasalized short vowels (4.1) were retained: cf. ORu. *a*-stem acc. sgs. from *-aⁿ# **horna** (Gallehus), **staina** (Tune). Similarly in *i*- and *u*-stem acc. sgs.

Unaccented *e merged with *i, except in the sequence -er > -ar (although *-eri yielded *-iri). Thus Go. *ufar*, OHG *obar*, eOE *ofær* (the latter two with *a*-umlaut, 3.3.1.2) beside ON *yfir*, OHG *ubiri*; eOE (Auzon Casket) nom. pl. 'gibroþær' (< PIE *-eres via PGmc *-ariz).

And, of course, *a and *o merged (3.1.1).

These changes yielded an inventory by the end of PGmc of *i*, *a*, *u*, all of which had nasalized pendants.

4.3.2. Main post-Proto-Germanic developments

Short vowels in third or higher syllables were lost (although sometimes restored analogically). See examples in preceding section.

In NWGmc, **-am#* > **-um#*: dat. pl. ON *doggum* (on the stem-vowel, cf. 3.3.5 end) OHG *tagum* versus Go. *dagam* ‘days’; 1pl. pres. indic. ON *kveþum* OHG *quëdum-ēs* (with extended ending) versus Go. *qīþam* ‘we say’. NGmc and Ingvæonic have gone further and merged all final **-Vm#* sequences under *-um*.

ON shortened all unstressed long vowels, re-establishing a system of *i, a, u*.

OHG augmented the short vowel inventory by shortening of final *-ē* and *-ô* (including from **-ai* and **-au*, 4.5.3) to *-e* and *-o*.

Final vowels in the Ingvæonic dialects were characterized by the fronting of WGmc **-a* to *-æ* (spelled <-a> or <-e> in Old Saxon; weakened to *-e* in later Old English and in Old Frisian), while WGmc **-o* was liable to unrounding to *-a*. Shortened *-ē* merged with *-æ*, producing an early four-member system, preserved in various instantiations in OS and in eOE; OE merged *-i* and *-æ* under *-e*, to give a system: *e, a, (o) u*.

4.4. Long vowels

4.4.1.1. The data require that a distinction between two kinds of long vowels be reconstructed for syllables which were – or became – final in PGmc. The evidence for this distinction comes both from within Gmc and from comparison of Gmc with Indo-European; the contrast is not preserved as such in any of the Old Germanic dialects.

4.4.1.2. Internal Gmc evidence is provided by Go. correspondents of Older Runic long vowels in absolute final position. In some instances these have retained their length in Go., but elsewhere they have undergone shortening – without there being any apparent conditioning environments to account for this dual treatment. For example, there is a final long *-ō#* in the genitive plural endings ORu. *arbijanō* ‘of heirs’ and Go. *qinōnō* ‘of women’ (both *n*-stems); but corresponding to ORu. weak preterit 1sg. *tawidō* ‘I made’, Gothic has *tawida* with a short vowel. Similarly, ORu. 3sg. *tawidē* ‘made’ also corresponds to Go. *tawida*. Beside this latter, the Go. gen. pl. *wulfē* ‘of wolves’ shows a long vowel.

Those PGmc long vowels that shorten in Go. final syllables are here called “bimoric” (symbolized as in **-ō*); those that do not are called “trimoric” (symbolized as in **-ô*). (Alternative designations are “long” and “overlong”; or “acute” and “circumflex”, the latter terms reflecting the assumption that the Germanic distinction was tonal as in Lithuanian.) The greater resistance to shortening evidenced by PGmc “trimoric” long vowels is to be attributed to some extra feature of length or intonation.

4.4.1.3. The same dual treatment of earlier long vowels emerges from comparison of Gmc with IE. The ending of the genitive plural, Go. *qinōnō*, continues post-Indo-European **-ōm#* (cf. Greek *n*-stem *akmōnōn* ‘of anvils’). But in the *ā*-stem accusative singular ending, post-IE **-ām#* (cf. Gk *khōrān* ‘place’), Go. shows a short vowel: *giba* ‘gift’. The Gothic reflexes in these two categories are not to be correlated with the quality of

the Indo-European vowels. Late Indo-European long **ā* and **ō* underwent qualitative merger in Germanic final syllables, as elsewhere. (That positing a Proto-Germanic distinction between **ā* and **ō* in final syllables is not the answer can be surmised from the inner-Germanic correspondence ORu. *tawidō* : Gothic *tawida*.) And the distinction between two kinds of long *ē*-vowel can hardly be linked with a supposed earlier qualitative difference. (See examples in 4.4.1.2)

OHG corresponds with *n*-stem gen. pl. *zungōno*, *ā*-stem acc. sg. *geba*, both with short final vowels. These represent a later stage of development, in which the correspondent of Go. *-ō* has also shortened and the reflex of the “bimoric-trimoric” contrast resides primarily in a qualitative distinction, cf. 4.4.3.4.

4.4.2. Gmc trimoric vowels are mainly the products of contraction of vowels across lost laryngeals (see Jasanoff 2004). However, some “trimoric” vowels arose as the result of Gmc (post-laryngeal) contractions, as in the present inflection of the weak class II verbs (4.2.2).

The position taken here requires that the PIE nom. pl. ending be reconstructed as **-x₁es* to account for the reflex **-ōz* in M thematic nouns such as OS *dagos* (4.4.3.4[3]). I reject the view that the WGmc ending reflects a hypercharacterized PIE **-ōs-es*, otherwise found only in Indo-Iranian.

4.4.3. Main developments in the dialects

4.4.3.1. In Gothic, any PGmc long vowel before a word-final consonant retained its length; in absolute-final position, “bimoric” long vowels shortened, while “trimoric” long vowels retained their length. Examples: *fidwōr* ‘four’ (< PIE Nt. pl. **k^wetwōr*), *galeikōm* ‘we compare’ (< IPIE **-āyome#*); *baira* ‘I carry’ (cf. Gk *phērō* ‘I bear’), *giba* nom. sg. ‘gift’ (< IPIE **-ā*, cf. Gk *khōrā* ‘place’); *galeikō* ‘I compare’ (< IPIE **-āyō#*). This analysis entails that the Go. *er*-stem nom. sg., as in *brōþar*, is etymologically the vocative form (< PIE **b^hrēx₂ter*) and does not reflect a long vowel (cf. Stiles 1988: §3).

4.4.3.2. The North-West-Germanic languages share the change of absolute-final bimoric **-ō#* to **-u#*. Thus ORu. strong verb 1sg. pres. indic. *gibu* ‘I give’ OHG *biru* ‘I carry’, nom. sg., ON *gjǫf* (cf. 3.3.1.3) and Old English *giefu* ‘gift’ (< **geþu*) correspond to the Gothic categories cited above.

4.4.3.3. North Germanic lost the distinction between “bimoric” and “trimoric” long vowels some time after the NWGmc change of **-ō#* to **-u#*. All North-West-Germanic long vowels were retained in Older Runic; all vowels in final syllables are short in Old Norse. Thus nasalized “bimoric” **-ō̃* was treated the same as “trimoric” **-ō̃n* and **-ō̃#*, with all of them yielding Old Norse *-a*. Examples: strong adj. fem. acc. sg. *blindā* ‘blind’ (Go. *blindā*) from PGmc **-ō̃n*; gen. pl. *kvinnā* ‘of the women’ (cf. Go. *qinōnō*) from PGmc **-ō̃n*; 2sg. ipv. *laða* ‘invite!’ (Go. *laþō**) from PGmc **-ō̃#*. For covered long vowels, compare *brōðir* (< **brōþēr*) and nom.-acc. pl. *giafar* (< both nom. **geþōz* and acc. **geþōz*).

4.4.3.4. 1. In the West Germanic languages, the distinction is observable in the quality of the reflexes. This presumably mirrors differences in the chronology of the shortening

of bimoric and trimoric vowels. This branch differs from both Go. and NGmc in offering three reflexes of the merger product of IE long **ā* and **ō* in syllables which were – or became – final in PGmc. Go. attests the dual outcome *-ō* and *-a*, NGmc offers the twin reflexes *-a* and *-u*, but WGmc, as represented by OHG, shows *-o*, *-a* and *-u*.

Old High German provides the most important evidence. In absolute final position it distinguishes between five (short) vowels, which essentially preserve the WGmc qualities (these have been altered to a greater or lesser extent in the Ingvaenic languages: OE, OFris., OS, cf. 4.3.2). Moreover, the Alemannic dialect, alone among WGmc, retains long vowels in medial and covered final syllables (most notably in the early ninth century *Benediktinerregel*, where length is indicated by vowel doubling, and in certain manuscripts of the works of the monk Notker III of Sankt Gallen, died 1022, which place a circumflex above long vowels). Compare dat. pl. *blintēm* ‘blind’ (Go. *blindaim*) beside uncovered masc. nom. pl. *blinte* (Go. *blindai*); wk. vb. cl. II 2pl. ipv. *ladōt* ‘invite ye!’ (Go. *lapōb**) beside 2sg. ipv. *lado* ‘invite!’ (Go. *lapō**; cf. 2.2.2.3). As these two sources, separated from one another by some 200 years, agree in great detail as to which vowels they mark as long, their testimony can be regarded as reliable.

2. Compare the following correspondences:

	OHG	OE	NGmc	Go.	IPGmc
gen. pl.	<i>zungōno</i>	<i>tungena</i>	<i>tungna</i>	<i>qinōnō</i>	*-ō ⁿ
adv. ‘like’	<i>galthho</i>	–	<i>glīka</i>	<i>galeikō*</i>	*-ō#
<i>ā</i> -stem acc. sg.	<i>geba</i>	<i>giefe</i>	<i>spaka</i> ¹	<i>giba</i>	*-ō ⁿ
1 st sg. pres. indic.	<i>quidu</i>	<i>cweoðu</i> ²	<i>w^aritu</i> ³	<i>qīþa</i>	*-ō#

Note: 1. ON *spaka* ‘wise’ is an adjectival form, as the noun has replaced the acc. form with the nom.; other examples ‘gift’. 2. West Mercian. 3. ORu. ‘I inscribe’; other examples ‘I say’.

3. In contrast to Go. and NGmc, WGmc also shows differentiated **-a* and **-o* reflexes of the merger product of IIE long **ā* and **ō* in instances where it was followed by a consonant in PGmc.

Late Proto-Germanic	*-ōC#	*-ōC#;	*-ēC#
Old High German	<i>ladōt</i>	<i>wazzar</i> ;	<i>bruoder</i>
Old Saxon	<i>dagos/-as</i>	<i>uatar/-er</i> ;	<i>fiuuar</i>
Old English	<i>dagas</i>	<i>wæter</i> ;	<i>feower</i>
Old High German	<i>kebo</i> < nom. pl. *-ōz	<i>geba</i> < acc. pl. *-ōz	

In OHG, before a final consonant, the primarily qualitative distinction between the reflexes is reinforced by a distinction in vowel length (cf. 4.4.3.4[1]).

4.5. Diphthongs

4.5.1. **ey* early became long *ī* (*ij*), as everywhere (3.1.1). **eu* became *iu*, except perhaps when tautosyllabic, when the outcome may have been *au* (depending on the interpretation of certain case-endings of *i*- and *u*-stems, see Harðarson, this handbook).

4.5.2. There is no need to reconstruct a distinction between long and short diphthongs in PGmc final syllables. Only short diphthongs are required, with **-ai* shortening to **-a* in Gothic third syllables at least, e.g., *haitada* (cf. OInd. *-te*, Gk *-toi*). Diphthongs were retained in monosyllables and apparently in disyllables, cf. Go. *ā*-stem F dat. sg. *gibai* ‘gift’. Such *ā*-stem F dat. sgs. as *gibai* ‘gift’ could be derived from a PGmc. pre-form containing a long diphthong, which is more costly phonologically, as it requires a more complex system. Note that Go. M nom. pl. *blindai* etc. could be influenced by monosyllabic *pai*, however; compare gen. sg. *dagis* for †*dagas*, after *bis*. Elsewhere, difficulties are presented by uncertain pre-forms or the possibility of levelling within paradigms.

4.5.3. The NWGmc monophthongization of PGmc **-au* yielded “trimoric” **-ô*: cf. OHG *ahto* (Go. *ahtau*), with the same final vowel as the genitive plural, cf. 4.4.1.3, 4.4.3.4 PGmc **-ai* monophthongized to long **-ē* in NWGmc and was lost medially in NGmc and OE: cf. Go. *allaizō* – ON *allra*, OE *eal(l)ra* – OHG *allero* adj. ‘all’ strong gen. pl.

Acknowledgement

I am responsible for errors and shortcomings in this chapter, but would like to thank Anna Helene Feulner, Sergio Neri, Roland Schuhmann, and Paolo Vaciago for minimizing them by suggesting improvements to earlier versions.

5. References

- von Bahder, Karl
1903 Zur hochdeutschen Lautlehre (hd. *f* = wgerm. *b*). *Indogermanische Forschungen* 14: 258–265.
- Benediktsson, Hreinn
1967 The Proto-Germanic Vowel System. In: *To Honor Roman Jakobson. Essays on the Occasion of his Seventieth Birthday*. The Hague: Mouton, 174–196.
- Benediktsson, Hreinn
1972 *The First Grammatical Treatise*. (University of Iceland Publications in Linguistics 1). Reykjavik: Institute of Nordic Linguistics.
- Dybo, Vladimir A.
1961 Sokrašćenie dolgot v kel’to-italijskix jazykax i ego znaćenie dlja balto-slavjanskoj i indoevropskoj akcentologii [The shortening of long vowels in Italo-Celtic and its significance for Balto-Slavic and Indo-European Accentology]. *Voprosy Slavjanskogo Jazykoznanija* 5: 9–34.
- Grimm, Jacob
1822 *Deutsche Grammatik*. Erster Teil. Zweite Ausgabe. Göttingen: Dieterich.
- Hollifield, Patrick Henry
1980 The Phonological Development of Final Syllables in Germanic. *Die Sprache* 26: 19–53 (Part 1), 145–178 (Part 2).
- Holtzmann, Adolph
1835 Review of H.F Massmann, *Skeireins aivaggeljons thairh Johannem*. *Heidelberger Jahrbücher der Literatur* 28: 854–863.

- Jasanoff, Jay
1978 Observations on the Germanic Verschärfung. *Münchener Studien zur Sprachwissenschaft* 37: 77–90.
- Jasanoff, Jay
2004 Acute vs. circumflex: Some notes on PIE and post-PIE prosodic phonology. In: Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson, and Thomas Olander (eds.), *Per aspera ad asteriscos. Studia Indogermanica in honorem Jens Elmegård Rasmussen sexagenarii Idibus Martii anno MMIV*. (Innsbrucker Beiträge zur Sprachwissenschaft 112). Innsbruck: Institut für Sprachwissenschaft der Universität, 247–255.
- Kluge, Friedrich
1884 Die germanische Konsonantendehnung. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 9: 149–186.
- Kuryłowicz, Jerzy
1952 The Germanic vowel system. *Biuletyn Polskiego Towarzystwa Językoznawczego* 11: 50–54.
- Lühr, Rosemarie
1988 *Expressivität und Lautgesetz im Germanischen*. (Monographien zur Sprachwissenschaft 15). Heidelberg: Winter.
- Moulton, William G.
1954 The stops and spirants of early Germanic. *Language* 30: 1–42.
- Pedersen, Holger
1933 *Études Lituanienes*. Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser 29, 3. Copenhagen: Levin & Munksgaard.
- Rask, Rasmus
1818 *Undersøgelse om det gamle Nordiske eller Islandske Sprogs Oprindelse* [An investigation of the ancient Scandinavian or Icelandic language]. Copenhagen: Gyldendal.
- Rasmussen, Jens Elmegård
1983 Two phonological issues in Germanic. *Acta Linguistica Hafniensia* 18: 201–219.
- Ringe, Don
2006 *From Proto-Indo-European to Proto-Germanic*. (A Linguistic History of English 1). Oxford: Oxford University Press.
- Rooth, Erik
1974 *Das Vernersche Gesetz in Forschung und Lehre*. Lund: Gleerup.
- de Saussure, Ferdinand
1922[1879] *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. In: *Recueil des Publications scientifiques de Ferdinand de Saussure*. Geneva: Société anonyme des éditions Sonor. [Leipzig: B. G. Teubner.]
- Stiles, Patrick V.
1988 Gothic nominative singular *brōþar* ‘brother’ and the reflexes of Indo-European long vowels in the final syllables of Germanic polysyllables. *Transactions of the Philological Society* 86: 115–143.
- Thórhallsdóttir, Guðrún
1993 The development of intervocalic *j in Proto-Germanic. Ph.D. dissertation, Cornell University.
- Thurneysen, Rudolf
1898 Spirantenwechsel im Gotischen. *Indogermanische Forschungen* 8: 208–214.
- Verner, Karl
1877 Eine ausnahme der ersten lautverschiebung. *Zeitschrift für vergleichende Sprachforschung* 23: 97–130.

Patrick V. Stiles, London (United Kingdom)

55. The morphology of Germanic

- | | |
|-----------------|---------------|
| 0. Introduction | 4. Pronouns |
| 1. Substantives | 5. Adverbs |
| 2. Adjectives | 6. Verbs |
| 3. Numerals | 7. References |

0. Introduction

In Gmc, the rich morphological system of PIE has been heavily simplified. This applies to the inflection of both nouns and verbs. The eight cases of the parent language, all of which are preserved in Sanskrit, have been reduced to five in the Old Gmc languages. A comparison of the verbal systems of Gmc and Ancient Greek shows how much simpler the former is. Gmc has only two tenses (present and preterite), while Greek has seven, according to the traditional terminology (present, imperfect, future, aorist, perfect, pluperfect, and future perfect). In PIE, nouns and verbs were inflected in three numbers, singular, dual and plural. In Gmc, the dual is preserved only in the 1st and 2nd persons. The loss of the 3rd person dual led to the extinction of the dual inflection of all nouns and most pronouns. The dual inflection, therefore, only exists in the pronouns of the 1st and 2nd persons, i.e. in the personal and possessive pronouns as well as in the 1st and 2nd persons of verbs. The forms of the numeral ‘2’ in the Gmc languages preserve some relics of the dual inflection (cf. 3.0).

The features that distinguish the morphological system of Gmc from those of the other IE languages are notably the following:

1. The extension of the *n*-declension of substantives, the so-called “weak” substantive declension.
2. The formation and distinction of a strong and weak adjective declension.
3. The systematic extension of the ablaut alternations, inherited from PIE, in the strong verbs.
4. The development of the weak preterite.

1. Substantives

In PIE, nouns formed with the same suffix could show different accent and ablaut alternations, and this distinction depended on both the derivational type and the semantic function of each formation. In Gmc, this system is no longer productive, but its remnants appear in the allomorphs of the various inflectional classes, the selection of which differs in the Old Gmc languages.

As already mentioned, the case system of PIE has been simplified in Gmc (case syncretism). Of the original eight cases (nom., voc., acc., gen., abl., dat., inst., and loc.) no Gmc language preserves more than five. Beside the four core cases in the nominal inflection, i.e. nom., acc., gen. and dat., Gothic has retained the vocative in some declensions. And in the *a*-declension of WGmc, a separate inst. sg. still exists. In Old High

German and Old Saxon, it is a direct descendant of the PIE instrumental, but in Old English, it continues the locative. Old Norse has, by contrast, only four cases. Here, the old inst. sg. of *a*-stems survives only in the dat. sg. neut. of strong adjectives (and certain pronouns).

Some scholars (cf. Griepentrog 1995: 14 and Schaffner 2001: 625 f., 636) assume that in the singular inflection of PGmc consonantal stems the locative ending **-i* and the dative ending **-ei* coexisted, and that the latter is continued at least in the dative singular of West Germanic. There is, however, no conclusive evidence for the existence of the dative singular ending **-ei* (> **-ī*) in Germanic (cf. Hollifield 1980: 33). In both North and West Germanic, the continuants of dat. **-ei* and loc. **-i* would most likely have coincided.

However, the continuants of PIE datives, instrumentals, and locatives in the Old Gmc languages (see Table 55.1) permit the assumption that early PGmc still preserved the distinction between these cases, at least in some declensions. Moreover, in the formation of adverbs, relics of the ablative of thematic adjectives still exist (Pre-PGmc **-ōd*, **-ēd*, cf. 5).

The dual inflection of nouns was lost in PGmc (cf. Introduction).

1.1. The inflectional classes of substantives

Substantives are traditionally divided into vocalic stems and consonantal stems. A further distinction in the morphology of substantives is that between thematic and athematic stems. Thematic stems are characterized in Indo-European terms by the vowel **e/o* (Gmc **i/a*) which appeared directly before the case desinences. All athematic stems share some important inflectional features which distinguish them from thematic stems. In what follows, we will adhere to the traditional classification of substantives.

1.1.1. Vocalic stems

1.1.1.1. *a*-stems

The *a*-stems, which continue PIE **o*-stems, fall into four subclasses: *a*-, *ja*-, *ija*-, and *wa*-stems. The so-called “pure *a*-stems” can, indeed, be divided into different classes depending on the form of the suffix, i.e. whether it consists of the pure thematic vowel or a combination of a (vowel plus) consonant, other than *j* and *w*, and *a* (*-a*-, *°la*-, *°ma*-, *°na*-, *°ra*-, etc.) (cf. Bammesberger 1990: 63 ff.). The same applies, *mutatis mutandis*, for the “pure *ō*-stems” (cf. Bammesberger 1990: 112 ff.). The *a*-stems comprise masculine and neuter nouns, cf. Go. masc. *wulfs* ‘wolf’, *harjis* ‘army’, *hairdeis* ‘herdsman’, *þius* ‘servant’, neut. *waurd* ‘word’, *kuni* ‘race’, *reiki* ‘rule, power’, *triu* ‘wood’. The distinction of *ja*- and *ija*-stems is due to the operation of Sievers’ Law, which regulated the distribution of suffix alternants of the type **-jV-* and **-ijV-* (on Sievers’ Law cf. Mayrhofer 1986: 164–166 w. lit. and Barrack 1998). These stems go back to PIE formations characterized by the suffixes **-jō-* (Sievers-variant **-ijō-*) and **-ih₂o-* (for a diachronic morphological analysis of these suffixes cf. Harðarson 2009: 20 f.).

After the change of **-ih₂o-* to **-iĵo-* this morpheme merged with the Sievers-variant **-iĵo-* of the suffix **-iĵo-*. In PGmc, the corresponding suffix **-iĵa-* was maintained when preceded by a long root-syllable or a polysyllabic stem, but was changed to *-ja-* if preceded by a short root- or stem-syllable (this is called the “converse of Siever’s Law”). This shows that a Gmc *ja-* or *iĵa-* stem can be assigned to either a PIE **iĵo-* or **ih₂o-* formation only after a thorough diachronic analysis has been carried out.

On the inflection of PGmc *a*-stems, see Table 55.1.

1.1.1.2. *ō*-stems

Similar to the *a*-stems, the *ō*-stems, which continue PIE **eh₂*-stems, are divided into four subclasses: *ō-*, *jō-*, *iĵō-*, and *wō-* stems (cf. comment on “pure *ō*-stems” in 1.1.1.1). These stems contain feminine nouns only, cf. Go. *giba* ‘gift’, *halja* ‘hell’, *bandi* ‘band, fetter’, *nidwa* ‘rust’. The distinction of *jō-* and *iĵō-* stems has the same motive as that of *ja-* and *iĵa-* stems (see section 1.1.1.1). The *jō-* and *iĵō-* stems derive from PIE formations with the suffixes **-iĵeh₂-* (Sievers-variant **-iĵeh₂-*) and **-ih₂eh₂-*. These formations embrace 1. **eh₂*-abstracts from adjectives in **-iĵo-* and **-ih₂o-* (cf. Go. *sibja* ‘relationship’, OIce. *Sif* (Thor’s spouse), pl. *sifjar* ‘kin, relationship’, OE *sib[b]* ‘relationship, etc.’ vs. Go. *un-sibjis* ‘lawless, impious’, OIce. *sifr* ‘close relative’, OE *sib[b]* ‘related, relative’, etc.), 2. deverbals with the ablauting suffix **-ih₂/iĵeh₂-* (cf. Gk φύζα ‘flight’, gen. φύζης) or non-ablauting **-iĵeh₂-* (cf. Ved. *vidyā-* ‘knowledge’), and 3. feminines derived from masculines by means of the suffix **-ih₂/iĵeh₂-* (*devī*-type). Note that in early PGmc, the feminine derivatives of the so-called *vrkī*-type adopted the inflection of the *devī*-type. A similar development can be observed within Sanskrit. Thus, OIce. *ylgr* ‘she-wolf’, gen. sg. *ylgiar* (< PN **wulgi+R*, **wulgijōR*), that originally belonged to the *vrkī*-type, is declined in the same way as all other *iĵō*-stems, cf., e.g., *heiþr* ‘heath’, gen. sg. *heiþar* (< PN **haiþi+R*, **haiþijōR*). It is a widespread misunderstanding that the nom. sg. OIce. *ylgr* corresponds to Ved. *vrkīh*. PGmc **-īz* would have yielded ON **-ir*, as already pointed out by Schmidt (1889: 72).

In PGmc, the inflections of stems with the ablauting suffix **-ih₂/iĵeh₂-* and those with non-ablauting **-iĵeh₂-* merged. The result was a new inflectional type with a generalized full-grade suffix, except in the nom. sg., where the two endings **-ī* (< **-ih₂*) and **-jō* (< **-iĵeh₂*) were preserved, but redistributed according to the syllabic structure of the root. Nouns with a short root had the ending **-jō* (< **-iĵeh₂*), and those with a long root the ending **-ī* (< **-ih₂*). The *devī*-type derivatives retained the ending **-ī* of the nom. sg., even when phonological changes led to the development of a short root syllable. They are represented by, e.g., Go. *mawi* ‘maiden’, gen. sg. *maujos*, OIce. *mér* ‘id.’, gen. sg. *meýiar* < PN **mawi+R*, **maujōR* (< PGmc **magwī*, **magwijōz*, fem. of **magu-* ‘son, boy, etc.’), and Go. *frijondi* ‘(female) friend’, gen. sg. *frijondjos* (fem. of *frijonds* ‘friend’).

The PGmc continuants of the suffixes **-ih₂eh₂-* and **-iĵeh₂-* (Sievers-variant of **-iĵeh₂-*) merged in the same way as those of the suffixes **-ih₂o-* and **-iĵo-* (cf. section 1.1.1.1). The further development was conditioned by the syllabic structure of the root. Following a heavy root, **-iĵō-* remained; after a light root, it changed to **-jō-*.

For the transformation of Gmc *ō*-stems to *n*-stems, cf. 1.1.2.2. On the inflection of PGmc *ō*-stems, see Table 55.1.

1.1.1.3. *i*-stems

This class contains all three genders. In PGmc, the masculines and feminines were inflected alike. They continue the proterokinetic accent and ablaut type of the PIE *i*-stems, cf. OHG nom. sg. *kuri* f. ‘choice’, *tāt* f. ‘deed’, gen. sg. *kuri*, *tāti* < PGmc **kuziz*, **kuzeiz*, **ḍāḍīz*, **ḍāḍēiz* < pre-PGmc **ǵus-i-s*, **ǵus-éj-s*, **d^heh₁-ti-s*, **d^heh₁-téj-s* ← PIE **ǵéus-i-s*, **ǵus-éj-s*, **d^heh₁-ti-s*, **d^heh₁-téj-s*. The original ending of the gen. sg. has undergone analogical changes in North and East Gmc. Some of the PGmc *i*-stems belonged originally to the acrostatic type, cf. **balg-i-/balg-ei* ‘skin, bag, bellows’ (Go. *balgs* ‘wine-skin’, OIce. *belgr* ‘skin, bellows, etc.’) ← PIE **b^hólǵ^h-i-/b^hélǵ^h-i*. Also some acrostatic *u*-stems show this assimilation to the proterokinetic inflection (cf. 1.1.1.4).

In PGmc, there are only a few vestiges of neuter *i*-stems, which already in PIE were scantily represented. Gothic and Old Norse have no examples, but some are attested in Old English, Old Saxon, and Old High German. An original neuter *i*-stem is OHG *meri* ‘sea, lake’ (with dat. sg. *meri* beside *mere* by analogy with the *a*-stems), cf. Go. *marisaihws* ‘lake’, OIr. *muir* n. ‘sea’, Lat. *mare* < **mori*, gen. sg. **mrej-s* (for older **meri-s*) (cf. Klingenschmitt 1992: 125 and Weiss 2009: 314; cf. also Schaffner 2010: 132), OCS *morje* n. ‘sea’ < **morjo-*. In Old Icelandic (*marr*) and Old English, (*mere*) it has become a masculine.

1.1.1.4. *u*-stems

u-stems comprise all three genders. They mainly reflect the proterokinetic accent and ablaut type of the PIE *u*-stems, cf. Go. sg. nom. *sunus* ‘son’, gen. *sunaus* (< PIE **-eu-s*; in unstressed position, tautosyllabic PIE **eu* became *au* in PGmc, cf. Rasmussen 1999a: 88 fn. 10, 95 and Neri 2003: 132), pl. nom. *sunjus* (< PIE **-eu-es*), gen. *sunuwe* (← PIE **-eu-ōm*), further PN dat. sg. **Kunimudiu** = *Kunimundiu* (Tjurkö), OHG *suniu* (< PIE loc. **-eu-i*). But remnants of the acrostatic (cf. Go. *kinnus* f. ‘chin, cheek’ with *nn* < *nw* in an early PGmc paradigm with nom. sg. **kenu-z*, gen. **kenw-ez* ← PIE **ǵén-u-s*, gen. **ǵén-u-os*, cf. Gk γένυς, -υος, f. ‘jaw’, Ved. *hánuh* f. ‘id.’, Toch. A *śanwem* ‘[two] jaws’ < **ǵén-u-*; on Ved. *hánu-* [with unexpected *h-* for *j-*], see Mayrhofer 1986: 139; on Toch. A *śanwem*, see Hilmarsson 1989: 104 f., 134 f. and Pinault 2008: 505) and holokinetic type (cf. the Go. variants sg. nom. *-aus*, gen. *-us*, acc. *-au*, and OE nom. pl. *sunā* < PIE **-ou-es*; for the details see Neri 2003: 122–124, 133 f., 178 w. lit.) also exist. Some Gmc *u*-stems that are inflected according to the proterokinetic type show *o*-grade of the root, cf. Go. *haidus* ‘manner’ (< **kojt-u-*). They continue a mixed type with proterokinetic accent and *o* : zero ablaut of the root. This can be labeled proterokinetic II. On this type, see Neri (2003: 23–25, 83–89), who prefers the label acrodynamic Ib (“acrodinamico Ib”), because it is, in fact, a transformation of the acrostatic type with *o* : *e* ablaut of the root (called “acrodinamico Ia” by Neri).

The inflection of the masculines and feminines is identical. Of the neuters only a few traces are extant in the Old Gmc languages, and no plural forms are attested. The inflection of the neuters differs from that of the masculines and feminines only in the nominative and accusative, cf. Go. nom.-acc. *faihu* ‘cattle’, gen. *faihaus*, dat. *faihau* (< PGmc **feh-u*, **feh-auz*, **feh-æu*).

1.1.2. Consonantal stems

Consonantal stems are divided into several subclasses (root stems, *r*-stems, *n*-stems, heteroclites, *z*-stems, *b/d̄*-stems, *nd*-stems). Due to space restrictions, the following discussion will be confined to root stems and *n*-stems.

1.1.2.1. Root stems

The Gmc root nouns are continuants of a very archaic morphological class of PIE. They represent a significant number of the lexicalized root nouns that can be posited for the parent language. In this respect, Gmc is rather conservative compared with other IE languages (cf. Griepentrog 1995: 12). In PIE, the overwhelming majority of root nouns were of feminine gender, but masculines and neuters also occurred. The inherited root nouns in Gmc are feminines, with the exception of the word for ‘foot’, which is masculine. None of the neuters has survived.

In PIE, this class of nouns had two accent and ablaut types, namely the acrostatic and amphikinetic ones (cf. Schindler 1972). The former had *o*-grade of the root in the strong cases and *e*-grade in the weak, cf. **pod-/ped-* m. ‘foot’ continued, e.g., in Gk πῶς (Att. πούς), gen. ποδός, Lat. *pēs*, gen. *pedis*, Gmc **fōt-* (with generalized *ō* from the nom. sg.). The latter had *e*-grade of the root in the strong cases and zero-grade in the weak, cf. **leuk-/luk-* f. ‘light, brightness’ in Lat. *lūx*, *lūcis* and Ved. *rūc-*. Beside the normal amphikinetic type, there existed a mixed type with *o* : zero ablaut (in roots of the structure **CeRT*), cf. **morġ-/myġ-* ‘border, demarcation’ in Gmc **mark-*, Gaul. *brig-* (Αλλό-βριγες, Νιτιό-βριγες) < **mrig-* < **myġ-* (cf. Schindler 1972: 34 f. and Delamarre 2003: 90 f.).

In PGmc, however, the paradigmatic ablaut alternations of the root nouns had been given up (due to paradigmatic leveling). One of the most prominent representatives of this class is the word for ‘cow’, which appears in the forms **kō-* (OHG, OS, OLF) and **kū-* (ON, OE, OFris.). The former stems from the PIE acc. sg. **gʷōm* and acc. pl. **gʷōs* (< **gʷou-m* and **gʷou-ms*, respectively, from **gʷou-/gʷeu-*). The latter is the result of secondary change of *ō* to *ū*, which probably took place under the influence of the word for ‘sow’, i.e. **sū-* (cf. Schindler 1973: 152 f. fn. 20).

Root nouns had the same set of endings as the consonantal stems ending in an obstruent.

1.1.2.2. *n*-stems

These stems contain all three genders. The masculines are divided into four subclasses, *an-*, *jan-*, *ijan-*, and *wan-* stems, the feminines into five, *ōn-*, *jōn-*, *ijōn-*, *wōn-*, and *īn-* stems, and the neuters into two, *an-* and *ijan-* stems. (To this last subclass belongs, in fact, only one word, i.e. ON *nýra* ‘kidney’ [< **neurijan-*] vs. OHG *nioro* m., MHG, MDu. *niere*, etc. [< **neuran-*].) While the masculine and neuter stems continue PIE formations, the feminine stems are PGmc innovations. The masculines reflect two PIE accent and ablaut types, the holokinetic and hysterokinetic ones, but the vast majority

of them were inflected according to the former type. The nom. sg. had the ending $*-\tilde{o}$ in PGmc, which is continued in WGmc (cf. OHG, OS *hano*, OE *hona* ‘cock’) (on the nominative singular ending of masculine *n*-stems in PIE and PGmc, see Harðarson 2005: 218–229). The gen. sg. ended in $*-n-ez$, which was later changed to $*-en-ez$ under the influence of the dat. ($*loc.$) sg. in $*-en-i$. In the plural paradigm, the acc., gen., and dat. had the zero-grade of the suffix, cf. acc. $*ar-n-unz$, $*ber-n-unz$, gen. $*ar-n-\tilde{o}^n$, $*ber-n-\tilde{o}^n$, dat. $*ar-n-umi/az$, $*ber-n-umi/az$ (from $*ar-an-$ ‘eagle’ and $*ber-an-$ ‘bear’); the forms of the acc. and dat. gave rise to the *u*-stem declension attested in ON *orn*, gen. *arnar*, and *biorn*, gen. *biarnar* (vs. ON *are*, ODan. *Biari*, OSw. *Biæri* [cf. ON *Biarne*], OHG *aro*, *bero*, etc.) (cf. Benediktsson 1968: 10, Harðarson 2001: 102 fn. 143). The PGmc acc. pl. in $*-n-unz$ developed by regular sound change ($*ar-n-unz$, $*ber-n-unz$ < pre-PGmc $*h_3or-n-ŋs$, $*b^her-n-ŋs$, cf. Ved. acc. pl. *rāj-ñ-aḥ* (from *rāj-an-* m. ‘king’) < $*-n-ŋs$). But the dat. and inst. pl. in $*-n-umaz$ and $*-n-umiz$ (instead of $*-un-ma/iz$ < $*-ŋ-mol/is$) were formed on analogy with those cases which had the suffix alternant $-n$. By adding the ending allomorph $*-ŋi/os$ used with obstruent stems, or its PGmc continuant $*-umi/az$, to the weak stem in $*-n-$, an unfavorable allomorphy with the zero-grade suffix alternant $*-un-$ beside $*-n-$ (as opposed to the full-grade alternants $*-en-$ and $*-an-$) was prevented. The *r*-stems show a parallel development in the dat. and inst. pl. ($*-r-umaz$ and $*-r-umiz$, instead of $*-ur-ma/iz$ < $*-r-mol/is$, after those forms of the paradigm that had the zero-grade suffix alternant $-r$). There are several comparable cases such as $*trud-$ in Go. *trudan* ‘tread’, OIce. *troða* (instead of $*turð-$) with *tr-* by analogy with $*trað-/træð-$ in the preterite and $*trað-$ in OIce. *treþia* ‘trample down’, OE *treddan*, cf. $*treð-$ in OE *tedan*, OHG *tretan*, etc.

Further reflexes of the original ablaut alternations within the paradigm are by-forms like OHG *chnodo* (and *knoto*) ‘knot’ and OE *cnotta*, OFris. *knotta*, MLG *knotte* ‘id.’ (with *tt* < *dn*; here the intermediate stages of this development are disregarded; on the sound change in question see Kluge’s Law in 6.3.2) < $*knub-an-$ (and $*knud-an-$)/*knud-n-* (cf. Schaffner 2001: 553, 2003: 210 f.); OE *grēofa* ‘pot’ and MDu. *groppe(n)* ‘iron pan’, late MHG *groppen* ‘iron pan, cauldron’ (from LG) (with *pp* < *bn*) < $*greub-an-/grub-n-$ (cf. Kroonen 2011: 161); OIce. *bole* ‘bull’, OE *bula* ‘young bull’ and MLG *bulle*, MDu. *bolle* ‘bull’ (with *ll* < *ln*) < $*bul-an-/bul-n-$ (cf. Schaffner 2001: 553 f.). But already in PGmc, the suffix alternant $*-an-$ was generalized in the plural paradigm. In some dialects, this development spread also to the gen. and dat. sg. (NGmc, OS, OE, and OFris.). In OS, the homonymy of the acc., gen., and dat. sg. seems to be the standard case, whereas the use of the gen. and dat. ending $-en$ in Heliand C can be ascribed to Old Dutch influence (cf. Krogh 1996: 306).

The hysterokinetic type has only left a few traces in Gmc, cf. OIce., OSw. nom. pl. *yxn* ‘oxen’ (from *uxe*) < PN $*uhsniR$ ← PGmc $*uhsenez$ (cf. MW *ychen* ‘id.’ < PC $*uhsenes$), Go. gen. pl. *auhsne* (from *auhsa*). The ending of the nom. sg. was $*-\tilde{e}^n$, which is represented in North and East Gmc (cf. OIce. *hane*, Go. *hana* ‘cock’).

It should be noted that the Gmc dialects, with the exception of PN, have selected and generalized one of the two nom. sg. endings of the masculines, $*-\tilde{o}$ and $*-\tilde{e}^n$, which, as we have seen, originally belonged to different accent and ablaut types. In PN, both endings were still extant, apparently in free distribution (cf. *nipijo*, *wagnijo* vs. *farawisa*, *swarta* with <a> for \tilde{e}^n), but later the continuant of $*-\tilde{e}^n$ was generalized (cf. Harðarson 2005: 225–228 and Nedoma 2005: 155–158, 161–165).

Tab. 55.1: Sample of early PGmc noun paradigms

	<i>a</i> -stems (*wulfaz ‘wolf’)		<i>ō</i> -stems (*gebō ‘gift’)	
	sg.	pl.	sg.	pl.
nom.	*wulfaz	*wulfōzēz, -ōsez	*gebō	*gebōz
voc.	*wulfe	= nom.	= nom.	= nom.
acc.	*wulfa ⁿ	*wulfanz	*gebō ⁿ	*gebōz
gen.	*wulfasa, -eza	*wulfō ⁿ	*gebōz, -ōz	*gebō ⁿ
dat.	*wulfōi	*wulfamaz	*gebōi	*gebōmaz
inst.	*wulfō	*wulfamiz	*gebō	*gebōmiz
loc.	*wulfai, -ei		*gebai	

Note: *a*-stem (*wurdaⁿ ‘word’): nom.-acc. sg. *wurdaⁿ, pl. *wurdō (in the rest of the cases like *wulfaz)

	<i>i</i> -stems (*weniz ‘friend’)		<i>u</i> -stems (*sunuz ‘son’)	
	sg.	pl.	sg.	pl.
nom.	*weniz	*wenejez	*sunuz	*sunewez
voc.	*weni	= nom.	*sunu, -au	= nom.
acc.	*weni ⁿ	*weninz	*suni ⁿ	*sununz
gen.	*weneiz	*wenejō ⁿ	*sunauz	*sunewō ⁿ
dat.	(*wenei?)	*wenimaz	(*sunewei?)	*sunumaz
inst.	*wenī	*wenimiz	*sunū	*sunumiz
loc.	*wenēi		*sunēu, *sunewi	

Note: *u*-stem (*fehu ‘property, cattle’): nom.-acc. sg. *fehu, caret plurali (in the other sg. cases like *sunuz)

	<i>r</i> -stems (*faðær ‘father’)		<i>n</i> -stems (*berō ‘bear’)	
	sg.	pl.	sg.	pl.
nom.	*faðær	*faðerez	*berō	*beranez
voc.	*faðer	= nom.	(*bera ⁿ ?)	= nom.
acc.	*faðeru ⁿ	*faðrunz	*beranu ⁿ	*bernunz
gen.	*faðrez, -urz	*faðrō ⁿ	*bernez	*bernō ⁿ
dat.	(*faðrei?)	*faðrumaz	(*bernei?)	*bernumaz
inst.		*faðrumiz		*bernumiz
loc.	*faðri		*bereni	

The neuter stems reflect the holokinetic type of the PIE collective. In the oblique cases of both the singular and plural, they have the same endings as the masculines. The PGmc ending **-ōn* of the nom.-acc. sg. continues the corresponding ending of the PIE collective, that was reinterpreted as singular in PGmc (cf. Gk nom.-acc. sg. ὕδωρ ‘water’ ← PIE coll. **μῆδ-ōr*). The endings of the nom.-acc. pl. are, by contrast, PGmc innovations: **-ōn-ō* (with the regular plural marker **-ō* of the neuters added to the singular form in **-ōn*, cf. Go. *hairtona* ‘hearts’, OSw. *ōghon* ‘eyes’), **-un-ō* (with the zero-grade suffix alternant *-un-* after a long root syllable, cf. OS *hertun* ‘hearts’, OHG *herzun*, OIce. *hiǫrto*), and **-n-ō* (with the zero-grade suffix alternant *-n-* after a short root syllable, cf. Go. *namna* ‘names’, OIce. *nǫfn*) (cf. Harðarson 2005: 230 f.). The neuters that formed a nom.-acc. pl. in **-n-ō* have adopted the plural inflection of the *a*-stems (in NGmc, the singular was also assimilated to the *a*-stems, cf. OIce. *nafn*, gen. sg. *nafns*).

The feminine *ōn*-stems originate in an *n*-extension of the *ō*-stems. This development started in the adjective declension, where the feminine gender imitated the “weak inflection” of the masculine gender (cf. 2.2). PGmc *ō*-stem nouns denoting female beings were as a rule transformed into *ōn*-stems, cf. PIE **gʷén-h₂-/gʷ(η)n-éh₂-* ‘womanhood, woman’ > PGmc **kʷenō-/kʷ(ʷ)unō-* → **kʷenōn-*, **kunōn-* (Go. *qino*, OHG *quena*, OE *cwene* vs. ON *kona*; in PGmc, the original ablauting paradigm was split into two [cf. Harðarson 1989: 86 f.]). But abstract nouns were also affected by this development, cf. PGmc **fullō-* ‘fullness’ (fem. abstract from **fulla-* ‘full’ < **fulna-*) → **fullōn-* (Go. *fullo* ‘id.’, OHG [theonym] *Folla* = ON *Fulla*, the personified ‘fullness’). The *n*-extension of *ō*-stems brought about the corresponding addition of *n* to many *ī/jō*-stems, cf. **lug-ī/jō-* ‘lie’ → **lug-īn-*. This goes parallel with the transition of *ī/jō*-feminines of *nd*-participles and comparatives to the *n*-inflection in North and East Gmc.

Unlike the masculine and neuter *n*-stems, that showed paradigmatic stem allomorphy, the feminine *n*-stems were unchangeable, i.e. a word such as **kʷenōn-* had the same stem form throughout the paradigm (sg. nom. **kʷenōn*, acc. **kʷenōn-u*, gen. **kʷenōn-ez*, loc. **kʷenōn-i*, etc.).

See Table 55.1 for a sample of early PGmc noun paradigms. Some points to be noted are that the gen. sg. ending **-eza* (beside **-asa*) of *a*-stems is posited on the basis of the Gothic evidence (cf. *anþariz-uh* ‘another’s’, see Bethge 1900: 353 and Krause 1968: 150). Moreover, for the inst. sg. of athematic stems in early PGmc, the ending **-mi/-umi* (< **-mī/-mī* for older **-bʰi*, cf. Gk ἰ-φι ‘by violence’) can be postulated. This ending, which corresponds to the inst. pl. ending **-miz/-umiz* (< **-mīs/-mīs* for older **-bʰis*) has left some traces in Germanic, cf. OHG *zi houbitun*, OE *æt hēafdum* ‘at the head’, *meolcum* ‘(with) milk’, OFris. *melokon* ‘(with) milk’, *bi dumhedum* ‘by stupidity’ (see Streitberg 1896: 228, Brugmann 1904: 386, and Bammesberger 2001, who, however, offers another diachronic analysis). It is, therefore, possible that the nouns **weniz*, **sunuz*, **faðēr* and **berō* once had the inst. sg. forms **wenimi*, **sunumi*, **faðrumi* and **bernumi*.

2. Adjectives

2.1. Stem formation and inflection

Like the substantives, the adjectives were divided into various classes according to their stem formation. In the course of Germanic language evolution, the *a/ō*-stems gradually

grew stronger at the expense of the other stems. Gothic distinguishes between *a/ō-*, *ja/jō-*, *ija/ijō-*, *i-*, and *u-*stems, but the *a/ō-*stems constitute by far the bulk of all adjectives. In North and West Gmc, the *i-* and *u-*stems went over into the *ja/jō-*declension (less frequently into the *a/ō-*declension).

The inflection of the adjectives was originally identical with that of the corresponding substantives. In the masc. and neut. genders *a*-stem adjectives such as **gōđā-* ‘good’ and **swarta-* ‘black’ were declined like the substantives **đaga-* m. ‘day’ and **wurđā-* n. ‘word’, respectively. Their feminine, characterized by the suffix **-ō-* (< PIE **-eh₂-*), followed the inflection of *ō*-stem substantives such as **gebō-* f. ‘gift’. The *i*-stems did not distinguish between masculine and feminine inflection (cf. Ved. nom. sg. *śúciḥ* m./f. ‘bright’). The feminine of the *u*-stems was either declined in the same way as the masculine or like an *ī/ijō-*stem (cf. Ved. nom. sg. *svādúḥ* m./f. and *svādvī* f. ‘sweet’). But already in PGmc, several endings of the demonstrative pronoun **sa*, f. **sō*, n. **pat(ōⁿ)* ‘that’ invaded the (strong) adjective inflection, cf. Table 55.2. Moreover, the feminine *ī/(i)jō-*inflection of *u*-stems and later also of *i*-stems caused the partial merger of these stems with the *ja/jō-* and *ija/ijō-*stems. The result of this development was a mixed inflection of both the *i-* and *u-*stems as still preserved in Gothic; cf. *þaurus* ‘dry’ (< **þurz-u-z*), nom. sg. fem. *þaurus*, acc. sg. fem. *þaurja*, acc. sg. m. *þaurjana*; *sels* ‘good, kind’ (< **sæl-i-z*), nom. sg. fem. *sels*, dat. sg. n. *seljamma*, nom. pl. m. *seljai*. In the North and West Gmc languages, on the other hand, the transition of the *i-* and *u-*stems to the other declensions was complete (see above), but they show remnants of the older situation; cf. by-forms such as ON *starkr* and *sterkr*, acc. sg. m. *sterkian* (beside *sterkan*) < **stark-ulija-* (based on the older opposition **stark-u-* : f. **stark-ijō-* < **stark-u-jō-*); *þurr* and *þyrr* ‘dry’ (cf. OE *þyrre*, OHG *durri*). On the reconstruction of a PGmc *u*-stem **stark-u-*, cf., e.g., Falk and Torp (1909: 486), Hellquist (1957: 1067), and Magnússon (1989: 956) (Heidermanns 1993: 546 f. refutes it too hastily). For the postulated change of **-u-jō-* to **-i-jō-* cf. section 6.3.2.

2.2. Distinction of strong and weak adjectives

One of the main features of PGmc was the distinction between indefinite or “strong” and definite or “weak” forms of the adjectives. The strong forms are inherited from PIE, the weak ones are a Germanic innovation. The weak adjective inflection developed in a linguistic situation where both single substantives and substantives qualified by adjectives could be used either with indefinite or definite meaning, cf. Lat. *arbor* ‘a tree’ and ‘the tree’, *arbor viridis* ‘a green tree’ and ‘the green tree’. The weak adjective originates in the substantivization of the strong adjective by means of an *n*-suffix (cf. Osthoff 1876). The underlying PIE formation had an individualizing or personifying function, cf. OAv., YAv. *marətan-/marəθn-* ‘a mortal being’ from *marəta-* ‘mortal’, Gk οὐρανίωτες ‘the heavenly gods’ from οὐράνιος ‘heavenly’, Lat. *Catō* lit. ‘the sharp one (intellectually)’ from *catus* ‘sharp’, and Go. *liuta* ‘hypocrite’ from *liuts* ‘hypocritical’. Since ablauting suffixes were added in their zero-grade form to thematic stems (cf., e.g., **-o-nt-* : **-ent-/nt-*, **-o-ih₁-* : **-iēh₁-/ih₁-*; see Klingenschmitt 1994: 240 and Schaffner 2003: 212) one would expect that *n*-stems derived from *o*-stems with the suffix **-(e/o)n-* should show an invariable stem in **-on-* (< **-o-n-*), but such derivatives have been assimilated either to the holokinetic inflec-

tion of primary and secondary individualizing formations from athematic base words or to the inflection of possessive formations in **-ōn-* (< **-o-Hn-*). (On the distinction of these two types of *n*-formations, cf. Hoffmann 1955: 35–40 and Harðarson 1989: 79 f. As examples of the possessive type, cf. Gk Γνάθων ‘the one who has great jaws’ vs. γνάθος ‘jaw’ and Lat. *Nāsō* ‘the one who has a great nose’ vs. *nāsus* ‘nose’.) The former holds true for Avestan and Gmc, the latter for Greek and Latin. Thus, in PGmc an individualizing formation such as **blend-an-* ‘the blind one, a blind man’ (cf. Go. *blinda* ‘id.’), derived from **blend-a-* ‘blind’, was inflected in the same way as a primary *n*-stem of the type **han-an-* ‘cock’ (← ‘singer’, cf. Lat. *canere* ‘sing’).

The definite meaning of the weak adjectives comes from the use of the individualizing *n*-stems as appositions to proper names or substantives denoting persons (cf. Delbrück 1909). The place of the apposition was following the principal word, cf. OIce. *Illuge svarte* ‘I. the Black’, *Sigurþr unge* ‘S. the Young’. It is worth mentioning that such phrases, in which the weak form of the adjective is used without the definite article, preserve the original semantic function of the *n*-stems in question.

In PGmc, the individualizing *n*-derivatives from *a*-stem adjectives were increasingly used as an expression of definiteness. By analogy, the *n*-formation was extended to the corresponding feminine form in **-ō-* and later also to all other classes of adjectives. The result was the grammaticalization of the definite or “weak” forms of the adjectives. After that the situation was as follows:

strong/indefinite form	weak/definite form
m. <i>*jung-a-</i> , f. <i>*jung-ō-</i>	<i>*jung-an-</i> , <i>*jung-ōn-</i>
m. <i>*sæl-i-(-ija-)</i> , f. <i>*sæl-i-(-ijō-)</i>	<i>*sæl-ijan-</i> , <i>*sæl-ijōn-</i>
m. <i>*nem-and-</i> , f. <i>*nem-and-ī/ijō-</i>	<i>*nem-and-an-</i> , <i>*nem-and-īn-</i> and <i>*nem-and-ijōn-</i>
f. (comp.) <i>*hard-iz-ī/jō-</i>	<i>*hard-iz-īn-</i>

Gothic and North Germanic selected **nem-and-īn-* as the feminine to *and*-stems, but West Germanic selected **nem-and-ijōn-*. The feminine comparative type **hard-iz-īn-* is continued in Gothic and North Germanic. In West Germanic, on the other hand, the weak feminine of the comparative was refashioned on the basis of the corresponding positive and superlative (masc./neut. pos. **jungan-*, sup. **jungistan-* : fem. **jungōn-*, **jungistōn-* = masc./neut. comp. **jungiran-* : fem. X [= **jungirōn-*]).

2.3. Comparison

There were two modes of forming the comparative and superlative in PGmc. The comparative had either the suffix **-izan-* or **-ōzan-*, the superlative **-ista-* or **-ōsta-*. These suffixes were added to the stem of the positive deprived of its final vowel, cf. Go. *hardus* ‘hard’ (< **hard-u-*), comp. *hardiza*, sup. *hardists*, *arms* ‘poor, pitiable’ (< **arm-a-z*), comp. *armoza*, sup. *armosts*.

In PIE, the comparative, which belonged to the holokinetic accent and ablaut type, was formed with the suffix **-jos-/-jes-/-is-*. In PGmc, the zero-grade of the suffix has

Tab. 55.2: The strong inflection of the PGmc adjective **leuþa-* ‘beloved, dear’ (the non-italicized forms have pronominal endings)

	singular		
	masc.	neut.	fem.
nom.	<i>*leuþaz</i>	<i>*leuþaⁿ</i> , -at(\bar{o}^n)	<i>*leuþō</i>
acc.	<i>*leuþanōⁿ</i>	<i>*leuþaⁿ</i> , -at(\bar{o}^n)	<i>*leuþōⁿ</i>
gen.	<i>*leuþasa</i> , -eza		<i>*leuþezōz</i>
dat.	<i>*leuþazmōi</i>		<i>*leuþōi</i> , -ezōi
inst.	<i>*leuþō</i>		<i>*leuþezō</i>
loc.	<i>*leuþei</i>		
abl.	<i>*leuþō(t)</i> , -ē(t)		

	plural		
	masc.	neut.	fem.
nom.	<i>*leuþai</i>	<i>*leuþō</i>	<i>*leuþōz</i>
acc.	<i>*leuþanz</i>	<i>*leuþō</i>	<i>*leuþōz</i>
gen.	<i>*leuþaizōⁿ</i>		
dat.	<i>*leuþaimaz</i>		
inst.	<i>*leuþaimiz</i>		

been extended by the individualizing suffix **-e/on-*, i.e. **-is-* → **-is-e/on-*. This extension has a complete match in Greek and can therefore be ascribed to late PIE.

The feminine of the comparative was of the *devī-*-type, i.e. the strong stem ended in **-iēs-ih₂-*, the weak in **-is-iéh₂-*. In pre-PGmc, this stem allomorphy was leveled to **-is-ih₂-/is-iéh₂-*; later, PGmc **-iz-ī-* developed a weak form in **-iz-īn-*, which eventually was generalized at the expense of the strong form.

The PIE superlative in **-is-t(h₂)o-* (beside **-is-ṛh₂o-* attested in Italic and Celtic) was derived from the weak stem of the comparative by means of the suffix **-t(h₂)o-*, which was also used in the formation of ordinal numbers. The suffix **-is-t(h₂)o-* is directly continued in PGmc **-ista-*.

Unlike the comparatives in **-izan-* and the superlatives in **-ista-* the forms in **-ōzan-* and **-ōsta-* are Gmc innovations. Their origin is controversial.

PGmc also had comparatives in *-(t)era-* and superlatives in *-(t)uma-* (sometimes extended by the suffix **-ista-* to produce *-(t)umista-*), which were derived from adverbs, cf. OE *æfter* ‘after’, *æfter(r)a* ‘hinder, next, second’, *æftemest* ‘last, aftermost’, OHG *after* ‘after, behind’, *aftaro (aftero)* ‘posterior’, Go. *afta* ‘behind’, *aftuma* ‘the following, next, hindmost, last’, *aftumists* ‘last, aftermost’. These forms continue PIE derivatives characterized by the suffixes *-(t)ero-* and *-(t)ṛh₂o-*, respectively, cf. Ved. *upāri*

‘above’, YAv. *upairi* ‘id.’, *upara-* ‘upper, higher’, Ved. *upamá-* ‘upmost, highest’, YAv. *upāma-*, *upama-* ‘id.’, Lat. *super*, *superus* (cf. *superior*), *summus* (< **supmo-*), etc.

3. Numerals

The Gmc numerals 1–10 and 100 continue more or less directly the corresponding PIE numerals. 1–4 and 100 were declinable. PGmc 1–10 can be reconstructed as follows: **ainaz* m., **ainō* f., **ainaⁿ/ainat(ōⁿ)* n., **twō/twōu* m. (cf. Ved. *d(u)vá*, *d(u)váu*), **twōz* f. (for **twai* = n., cf. Ved. *d(u)vé*, or, less likely, < **d_ueh₂ih₂*, cf. Lat. *duae*), **twai* n. (< **d_uoih₁*, cf. Ved. *d(u)vé*), **prejez* m./f., **prijō* n., **fed_warez* m./f., **fed_wōr* n., **semfe*, **sehs*, **sehun(t)* (with *t* from 9 and 10, < **sef(t)ún[t]*), **ahtōu*, **newun(t)* (with *t* from 10), **tehun(t)* (< **dék_md /dek_mt/*).

In PIE, the nom.-acc. m. of the numeral ‘two’ had the by-forms **d(u)uóh₁* and **duóh₁u/d(u)uóh₁u* (/_#V). The nom.-acc. du. m. of thematic nouns shows the same variation, i.e. **-oh₁* and **-oh₁u/-oh₁u* (Ved. *-ā* and *-au*). The longer forms are extended by the deictic particle **u* = Ved. *u*, etc.; cf. Eichner (1992: 85) where he discusses the Anatolian and PIE numeral 8, which was an old dual. The form **dék_md /dek_mt/* is a neuter *t*-stem noun derived from the cardinal **dék_m* (cf. Lat. *decem*, Ved. *dása* etc.), i.e. **dék_m* ‘10’ → **dék_m-t-* ‘group of ten; 10 (subst.)’ (cf. Rau 2009: 48 f.).

11 and 12 are compounds with *-lib-* (< **-lik^w-*) as second element: cf. Go. *ainlif* ‘11’ (dat. *ainlibim*), *twalif* ‘12’ (dat. *twalibim*) < **aina-lib-*, **twa-lib-*, which originally meant something like ‘(ten and) one left over’ and ‘(ten and) two left over’, respectively (cf. Lith. *vienuolika* ‘11’ and *dvylika* ‘12’). The change of **-lik^w-* to *-lib-* first took place in the numeral 12 owing to the preceding labial, cf. Gmc **wulfā-* ‘wolf’ < PIE **w₁lk^wo-* and **fimf* ‘5’ < **semfe* < PIE **pénk^we*.

13–19 are formed from the units 3 to 9 plus ‘ten’, cf. OHG *drīzehan*, Go. *fidwortaihun*, *fimftaihun*, etc.

20–60 are phrases composed of the units 2 to 6 and the plural of the masc. *u*-stem **tegu-* ‘decade’: **twō(u) tegewez*, **prejez tegewez*, etc. These phrases were inflected. The stem **tegu-* derives from an *u*-stem **dekú-/dek-éu-* ‘decade’ (← PIE **dék-u-/dék-éu-*), that also existed in Italic, cf. Umbr. nom. pl. or gen. sg. **tekvias** < **deku-ijā-* (whence Osc. **dekkviarím**), Lat. *decuria* ‘a body of ten men’ and Umbr. dat. abl. pl. **tekuries**, *dequrier* (for the reconstruction of the *u*-stem **dék-u-* for Italic and Germanic cf. Schulze 1904: 145 f. and Feist 1939: 150 w. lit.).

70–90 are transformations of PIE compounds formed from the units 7 to 9 and the neuter plural (collective) **-k_mt₂* ‘decades’, cf. Go. *sibuntehund*, *ahtautehund* (← **ah-tōtēhund*), *niuntehund* < PGmc **sebuntēhundā*, **ahtōtēhundā*, **newuntēhundā*. The *ē* in 70 and 90 comes from the once existing **semfēhundā* ‘50’ (< **penk^wēk_mt₂*, cf. Gk πεντήκοντα) and **sehsēhundā* (cf. Szemerényi 1960: 34 f.). After the numerals **sebunt* ‘7’ and **newunt* ‘9’ had lost their final *t*, a reanalysis of the decade compounds **sebuntēhundā* and **newuntēhundā* as **sebunt-tēhundā* and **newunt-tēhundā* took place, whence **ahtō-tēhundā* (for older **ahtō-hundā*). Ringe (2006a: 206) holds a view similar to the one presented here, but according to him the reanalysis which led to the creation of the new element **-tēhund-* was only based on the development of the numeral 7 and its relationship to 70. Some scholars prefer to interpret **-tēhund-* as a *v₂ddhi*-formation

from **tegunþ-* f. ‘decade’ (cf. Darms 1978: 34–48 w. lit.). But the assumption that this noun ever existed in Gmc is *ad hoc*.

100: PGmc **hundaⁿ* (cf. Go. *hund*, OE, OS *hund*, OHG *hunt*) < PIE **k̑ntóm* < **dk̑ntóm* ‘unity of ten (decades)’ (on the simplification of initial *TK* to *K* before syllabic nasal in PIE see Lipp 2009: II, 27, 87 f.). Beside this form the North Sea languages had the compound **hunda-raðaⁿ* (cf. ON *hundraþ*, OE *hundred*, OFris. *hundred*, *hunderd*, OS *hunderod*), the second element of which is related to Go. *raþjo* ‘number’. It subsequently spread to High German (cf. MHG *hundert*).

200–900 were expressed by single units and the plural of **hundaⁿ*, that was fully inflected.

1000: PGmc **þūsund-īlijō-* (cf. Go. *þūsundi*, gen. sg. *þūsundjos*) continues the feminine of a secondary adjective **tuh₂-s-ont/nt-* ‘great, swollen’ (from **teuh₂-s-* ‘swelling’). The same formation is attested in Balto-Slavic, cf. OCS *tysqštb/tysqštb* and Lith. *tūkstan-tis* (cf. Neri 2009: 8 and Rau 2009: 15 fn. 12).

For reasons of space, a discussion of the ordinals and other numeral expressions cannot be included here.

4. Pronouns

4.1. Personal pronouns

The pronouns of the 1st and 2nd persons are non-gendered pronouns, i.e. they show no difference of gender. Together with the corresponding possessive pronouns they are the

Tab. 55.3: The PGmc inflection of the pronouns of the 1st and 2nd persons

	1 st person		
	sg.	du.	pl.
nom.	<i>*ek, *ekaⁿ</i>	<i>*weta</i>	<i>*wejez</i> or <i>*weiz</i>
acc.	<i>*meke/a</i>	<i>*unke, unkiz</i>	<i>*unse, *unsiz</i>
gen.	<i>*meina-</i>	<i>*unkera-</i>	<i>*unsera-</i>
dat.	<i>*mez</i>	<i>*unkiz</i>	<i>*unsiz</i>

	2 nd person		
	sg.	du.	pl.
nom.	<i>*þū</i>	<i>*juta</i>	<i>*jūz</i>
acc.	<i>*þeke/a</i>	<i>*ink^we, *ink^wiz</i>	<i>*izwe, *izwiz</i>
gen.	<i>*þeina-</i>	<i>*ink^wera-</i>	<i>*izwera-</i>
dat.	<i>*þez</i>	<i>*ink^wiz</i>	<i>*izwiz</i>

only pronouns that preserve the number category dual in PGmc. A further peculiarity of their inflection is the use of different stems, even within each of their numbers. The PGmc inflection is given in Table 55.3 (cf. Ringe 2006: 208–211, 290 f. [w. lit.] and Neri 2009: 8 [w. lit.]). Instead of the old genitives the corresponding possessive pronouns, **meina-*, **unkera-*, **unsera-* etc., were used (see below), cf. Lat. *ego*, gen. *meī* from *meus*, etc.

As a pronoun of the 3rd person, the anaphoric pronoun **i-/ei-/e-* (< PIE **(h₁)i-/ (h₁)e_i-/ (h₁)e-*) was used. It distinguished all three genders, cf. nom. sg. m. **i-z*, f. ***ī* → **sī* (which originally belonged to the demonstrative pronoun **sa-/pa-* ‘that’, see 4.4), n. **i-t(ōⁿ)*, acc. sg. m. **i-nōⁿ*, f. **ej-ōⁿ*, gen. sg. m./n. **e-sa*, f. **e-zō-z*.

On the PIE paradigm of the anaphoric pronoun, cf. Weiss (2009: 340–342); on the PGmc paradigm, cf. Ringe (2006: 289). For the opposition of initial **-sī-* in gen. sg. f. Ved. *asyāh* and **-s-* in PGmc **ezōz*, see 4.4.

4.2. The reflexive pronoun

In PGmc, as in PIE and most of the daughter languages, the non-gendered reflexive pronoun had no nominative and was inflected alike in the singular and plural. The forms correspond to the acc., dat., and gen. sg. of the 1st and 2nd person pronouns. For PGmc they can be reconstructed as follows: acc. **seke/a*, gen. **seina-*, dat. **sez*.

4.3. Possessive pronouns

The possessive pronouns, which followed the inflection of strong adjectives, were derived from the 1st and 2nd person pronouns and the reflexive pronoun. In PGmc, the 1st and 2nd persons had one stem for each number, i.e. 1st pers. sg. **meina-*, du. **unkera-*, pl. **unsera-*, and 2nd pers. sg. **peina-*, du. **ink^wera-*, pl. **izwera-*. The possessive of the reflexive pronoun had only one stem, i.e. **seina-*. The possessives belonging to the singular were derived from the PIE locatives **mej*, **tej*, **sej* by means of the suffix **-no-* (**mej-no-*, **tej-no-*, **sej-no-*), the others were formed from the stems of the oblique cases with the suffix **-ero-*, cf. similar formations in Latin (*noster*, *vester*) and Greek (ἡμέτερος ‘our, ours’, ὑμέτερος ‘your, yours’).

4.4. Demonstrative pronouns

In addition to the anaphoric pronoun **i-/ei-/e-* (see 4.1), PGmc had several demonstrative pronouns, most of which were inherited from PIE. The following stems can be ascribed to PGmc:

**sa-/pa-*, f. **sō-/pō-* ‘that’ (< PIE **so-/to-*, f. **seh₂-/teh₂-*). In PIE, the feminine singular of this pronoun could apparently be formed in two ways, either according to the *eh₂-* or the *iĥ₂/iēh₂-* type. The former had the nom. **seh₂*, acc. **teh₂-m*, gen. **te-seh₂-(e)s*, dat.

Tab. 55.4: The PGmc inflection of the demonstrative pronoun **sa-/pa-* ‘that’

	singular			plural		
	masc.	neut.	fem.	masc.	neut.	fem.
nom.	<i>*sa</i>	<i>*pat(ōⁿ)</i>	<i>*sō</i>	<i>*pai</i>	<i>*pō</i>	<i>*pōz</i>
acc.	<i>*panōⁿ</i>	<i>*pat(ōⁿ)</i>	<i>*pōⁿ</i>	<i>*panz</i>	<i>*pō</i>	<i>*pōz</i>
gen.	<i>*pesal/pasa</i>		<i>*pezōz</i>	<i>*paizōⁿ</i>		
dat.	<i>*pazmōi</i>		<i>*pezōi</i>	<i>*paimaz</i>		
inst.	<i>*pæ, paimi</i>		<i>*pezō</i>	<i>*paimiz</i>		
loc.	<i>*pei</i>					

**te-seh₂-ei* (cf. Go. *so, þo, þizos, þizai*). The corresponding forms of the latter were: nom. **sih₂*, acc. **sih₂-m*, gen. **te/o-sjeh₂-s*, dat. **te/o-sjeh₂-ei* (cf. OIr. *sí*, Go., OHG *si*, OAv. *hī*, OAv., YAv. *hīm*, Ved. *sīm, tásyāh, tásyai*). (For alternative explanations, see Cowgill 2006.) The PGmc inflection is displayed in Table 55.4.

**pi-* ‘this’ (< PIE **ti-*, from which **tjo-/tijo-* was derived, see below s. v. **þija-*).

**sija-/sijō-* ‘that’ (< PIE **sjo-/sijo-*, f. **sjeh₂-/sijeh₂-*, cf. Ved. *syá-*, f. *syá-* ‘id.’).

**þija-/þijō-* ‘that’ (< PIE **tjo-/tijo-*, f. **tjeh₂-/tjeh₂-*, cf. Ved. *tyá-*, f. *tyá-* ‘id.’).

**ena-/enō-* ‘that’ (< PIE **eno-*, f. **eneh₂-*, cf. Gk *ἔνη-* ‘the day after tomorrow’).

**jena-/jenō-* ‘that’ and, with ablaut alternation, **jana-/janō-* ‘id.’. It is open to question whether Go. *jains* ‘id.’ continues a third PGmc stem **jaina-*, related to the two others, or descends from **jena-*. For an analysis of these stems as well as of **ena-*, cf. Klingenschmitt (1987: 174–176).

**hi-/hei-/he-/(hija-)* ‘this’ (< PIE **k_i-/ke_i-/ke-*, cf. Lith. *šis* ‘this’ [gen. sg. m. *šiō*], OCS *sb* ‘id.’ [nom.-acc. sg. n. *se*, gen. sg. m./n. *sego*], Gk *σήμερον* ‘today’ < **k_i-āmeron*, Hitt. *kā-* [ka/ke-/ki-] ‘this’ [nom. sg. c. *kāš*, inst. *kēdanta, kēt*, nom.-acc. sg. n. *kī*]).

**hina-/hinō-* ‘that’ (< PIE **k_i* ‘here’ and the pronoun **no-*, a variant form of **eno-*, see above s. v. **ena-*; the distal deictic function of the pronoun is determined by the second element, cf. Klingenschmitt 1987: 177 f.).

In North and West Gmc, a new proximal demonstrative ‘this’ was formed on the basis of **sa-/pa-* ‘that’ and **pi-*, primarily by adding the deictic particle **sai* ‘ecce’ (in Go. *sai*, OHG *sē* < PIE **so_i*, cf. Brugmann 1904: 28) or **si* (contained in **sija-* ‘that’ < PIE **sjo-/sijo-*, see above) to their inflectional forms or by reduplication; cf. *no_i si* (or *no_isi*) ‘if not’ in the 2nd line of the Duenos Inscription (on its interpretation, see Harðarson 2011: 156 ff.).

4.5. Relative pronouns

A relative pronoun proper did not exist in PGmc. In the individual Gmc languages, relative clauses were introduced in various ways. Gothic had a relative particle *-ei* (*/ī*) that was added to the forms of the personal pronouns or the demonstrative *sa, so, þata*, cf. *ikei, þuei* (dat. *þuzei*), *izei, sei* (< *si-ei*), *saei, soei, þatei* (gen. *þizei*, etc.). When used

alone, *ei* served as a conjunction of complement clauses ('that') or final clauses ('so that'). In Old Icelandic, relative clauses were most frequently expressed by the particles *es*, *er*, or *sem* (with or without the demonstrative). In Old Saxon and Old High German, they were generally introduced by the simple demonstrative, in Old English by the relative particle *þē* alone or in combination with the personal pronouns or the demonstrative.

4.6. Interrogative pronouns

The PIE stems **k^wi-/k^wēi-* and **k^wo/e-*, which were paradigmatically combined with each other, had two functions. Under the accent they served as the interrogative, but when enclitic as the indefinite, cf. Gk τίς 'who' (< **k^wi-s*) vs. τις 'someone'. The masculine and feminine were inflected alike. In the singular inflection, the nom., acc., and inst. of all genders were formed from the *i*-stem, the other cases from the *o*-stem, cf. nom. m./f. **k^wi-s*, nom.-acc. n. **k^wi-d*, gen. **k^wé-s(i)o*, dat. **k^wo-smōi*, and inst. **k^wi-h₁* (of all genders). All cases of the plural were formed from the *i*-stem, cf. nom. m./f. **k^wéi-es*, nom.-acc. n. **k^wi-h₂*. The *o*-stem could be extended to other cases, which were the domain of the *i*-stem, cf. Ved. *káh* 'who' (< **k^wó-s*; the nom. **k^wís* is reflected as a petrified form in *ná-kis* and *mā-kis* 'no one, nothing'), acc. m. *kám*, nom. f. *kā* (< **k^wé-h₂*), nom.-acc. n. *kád*. This was also the case in Gmc, where the *i*-stem was ousted. Moreover, the PGmc interrogative **h^wa-/h^we-* 'who, what' was only inflected in the singular. The inflection is best preserved in Gothic, where a special form of the feminine exists, cf. nom. m. *hwas*, nom.-acc. n. *hva* (< **k^wó-d*), gen. m./n. *hvis*, nom., acc. f. *hvo*, dat. *hwizai*, gen. **hwizos*.

A further interrogative that PGmc inherited from PIE is **h^waþera-* 'which of two' < **k^wó-tero-*, cf. gr. πότερος, Ved. *kataráh* (with shifted accent), Lith. *katràs*, OCS *kotorъ* 'id.'. It is continued in Go. *hwaþar*, ON *hvaþarr*, *hvárr* (with *hvár-* < **hvaðr-* before a vocalic ending), and OE *hwæðer*. By contrast, OHG (*h*)*wedar*, OS *hweðar* and OE *hweðer* apparently descend from an ablaut variant **k^wé-tero-*. For the by-forms *hwæðer* and *hweðer* in OE, see Brunner (1965: 40).

Other interrogatives in Gmc are innovations.

4.7. Indefinite pronouns

Many of the indefinite pronouns can also function as other parts of speech, depending on context. Thus, e.g., the numeral for 'one' is used as indefinite in several Gmc languages. The same goes for the substantive **man-an-/man-n-* 'human'.

As in other IE languages, the indefinites of Gmc are often formed by composition with the negative **ne* and/or the generalizing particle **k^we*, **k^we-ne*. (The proper meaning of the compound particle **k^we-ne* was 'and not, not even', but when used with another negative it could throw off its own negative sense. In this way, it developed the meaning 'also, even' (or the like) which could be extended to contexts without a preceding negative.) Such pronominals are, e.g., Go. *ni hwas-hun* and *ni ains-hun* 'no one, no, none, nothing' (with *-hun* < **-h^wene* < **-k^we-ne*), *hwaz-uh*, *hwarjiz-uh* 'each, every' and

hvaparuh ‘each of two’ (with *-uh* < **-k^we*, cf. Mottausch 2001), OIce. *neinn* (used after a negative) ‘any, anybody’, *enge* ‘none’, *hverge* ‘each, every’, *hvárge* ‘each of two’ and ‘neither of two’ (with *-ge* < **-g^(w)en(e)* < **-k^wene*), OE *nān*, OS *nēn* ‘none’ (= OIce. *neinn* < **ne-aina-*). Go. *-hun* and OIce. *-ge* (cf. WGmc *-gin* in the adverbs OE *hwerzen*, OS *hwergin*, *hwargin*, OHG *io wergin* ‘somewhere’) exhibit “grammatical change”, i.e. they are Verner’s variants.

Some of the Gmc indefinites are inherited from PIE, cf. **suma-* ‘some’, which is related to Ved. *sama-* ‘someone’, Gk *ἄμῶς* ‘somehow’ etc. (< **sm̥-ó-*), **anþera-* ‘other’ (also ‘second’) and **ne h^was-h^wene* ‘no one, no, none’. Others are innovations.

5. Adverbs

In Gmc, as in the other IE languages, most adverbs are original case forms that are no longer in declensional use. Some of them show endings that have become obsolete in Gmc. Although the adverbs are mainly inflectional in origin, some of them were derivational already in PGmc.

Especially productive was the formation of adverbs in **-ō̃(t)* or **-ē̃(t)* derived from thematic adjectives, cf. **gernō̃(t)* / **gernē̃(t)* ‘eagerly, willingly, gladly’ (ON *giarna*, OE *zeorne*, OS, OHG *gerno*) from **gerna-* ‘eager, desirous, willing’ (Go. *faihu-gairns* ‘avaricious’, ON *giarn* ‘eager, willing’, OE *zeorn* ‘eager, desirous, etc.’, OS *gern* ‘coveting, desiring’, OHG *gern* ‘eager’). They continue pre-PGmc ablatives in **-ōd* and **-ēd*, respectively (< PIE **-o-ad*, **-e-ad* or rather **-o-h₁ad*, **-e-h₁ad*, cf. Harðarson 2016: 157 f.).

The comparative and superlative degrees of adverbs had two sets of endings: 1. comp. in **-iz*, sup. in **-istaⁿ*, 2. comp. in **-ōz*, sup. in **-ōstaⁿ*. The latter set belonged to adverbs derived from adjectives that formed their comparatives and superlatives with the corresponding suffixes **-ōzan-* and **-ōsta-*, cf. **armō̃(t)/-ē̃(t)* ‘miserably, wretchedly’, **fastō̃(t)/-ē̃(t)* ‘fast, firmly’ → comp. **armōz*, **fastōz* (ON *fastar*, OE *earmor*, *fæstor*, OHG *fastor*; in the WGmc languages the final *r* of the adverbial comparative in **-ōr* was analogical after the corresponding adjectival comparative in **-ōran-*), sup. **armōstaⁿ*, **fastōstaⁿ* (ON *fastast*, OE *earmost*, *fæstost*, OHG *fastost*), cf. the adjectival comparatives **armōzan-*, **fastōzan-* and the superlatives **armōsta-*, **fastōsta-*. Other adverbs took the former set of endings, cf. **furi* ‘before’, **wel-ǣ* ‘well’, **langō̃(t)/-ē̃(t)* ‘for a long time’ → comp. **furiz*, **batiz*, **langiz* (ON *fyr*, *bet*, *leng*, OE *bet*, *lenz*, OS *bet*, *leng*), sup. **furistaⁿ*, **batistaⁿ*, **langistaⁿ* (ON *fyrst*, *bezt*, *lengst*, OE *betst*, *lenzest*, OS *best*), cf. the adjectival comparatives **furizan-*, **batizan-*, **langizan-* and the superlatives **furista-*, **batista-*, **langista-*.

The preverbs and prepositions of the IE languages generally have an adverbial origin (cf. Kuryłowicz 1964: 171). As for the relationship of adverbs and “particles”, a sharp distinction between the two cannot be made.

6. Verbs

6.1. Changes in the verbal system from Indo-European to Germanic

In the PIE verbal system, the opposition of present and aorist stems was fundamental. This opposition was not motivated temporally, but rather aspectually, i.e. it marked the imperfective and perfective aspect, respectively. The present stems either consisted of the bare root or of formations that were characterized by certain “present formatives” such as the thematic vowel *-e/o-*, *-je/o-*, *-ske/o-*, infix *-ne/n-*, etc. Originally the present formatives were elements of derivation serving as expressions of different *aktionsarten*. (The distinction here made between *aspect* and *aktionsart* corresponds to that between *grammatical aspect* and *lexical aspect*, especially in Anglophone linguistics.) The semantic emptying that affected some of the old *aktionsart*-formations led to the emergence of the semantically unmotivated “present stem” (cf. Rix 1986: 16). Beside the morphologically marked verbal stems, there were (amphikinetic) root stems whose assignment to either the aorist or the present depended solely on the inherent aspectual character of the root in question, i.e. whether it was perfective (telic) or imperfective (atelic). The perfect denoted an attained state (of the subject) as a consequence of completed action. In fact, it was a stative present, as is best shown by Homeric Greek and the so-called preterite-presents of the IE languages. On the semantic function of the PIE perfect, see di Giovine (1996: 176, 248 f., 259–261, 273–276 w. lit.) and Kümmel (2000: 65 w. lit.).

The distinction of tenses was only possible in the indicative mood, which represented the predicate as a reality. The present tense could be formed only from the present stem. The past tense of the present stem was imperfective and corresponded to the imperfect of the classical languages. On the other hand, the past tense of the aorist stem was perfective and equivalent to the Greek aorist. As PIE did not have any “future tense”, futurity was expressed by desiderative stem formations and the subjunctive mood.

Originally, the category tense was expressed exclusively by inflectional endings (primary or secondary endings), but later also by means of the augment. The category mood, on the other hand, was either expressed by inflectional endings (imperative) or by derivational suffixes (subjunctive and optative).

In Gmc, the following changes have taken place:

1. The augment vanished without a trace.
2. The subjunctive mood, that expressed volition and expectation, was lost as a formal category.
3. Beside its resultative function, the perfect developed a preterital meaning that merged with that of the aorist or the imperfect depending on the lexical aspect of the verb in question.
4. Due to the extensive dereduplication of the perfect its dual and plural indicative forms coalesced with the corresponding forms of the root aorist. The same applies to the optative forms.
5. The aorist was lost as a formal category.
6. The imperfect was lost as a formal category.
7. As a consequence of the loss of the aorist and the imperfect a preterite tense of the passive ceased to exist.

8. Ablaut became an important and characteristic feature of primary (strong) verbs, as their “tense categories” were reduced to those of the present and perfect. Thus, ablaut was an indicator of tense, beside the endings.
9. A new “weak” preterite developed from a periphrastic construction.

6.2. The inflectional categories of Germanic verbs

The inflectional categories of Gmc verbs are the following:

Tenses: present and preterite. A periphrastic perfect (and a pluperfect) developed secondarily in North and West Gmc.

Moods: indicative, optative (subjunctive), and imperative. The imperative is only used in the present tense. The 3rd person of the imperative (only preserved in Gothic) continues the so-called “future imperative” (cf. 6.5).

Voices: active and passive. A synthetic (simple) passive exists only in the present tense. Gothic is the only Gmc language that preserves it, cf. pass. *nimada* ‘is taken’ vs. act. *nimip* ‘takes’.

Persons: 1st, 2nd, and 3rd person.

Numbers: singular, dual, and plural. The dual of verbs survives only in Gothic, and there only in the 1st and 2nd persons. The absence of verbal dual forms in the 3rd person is caused by the general loss of the third person dual in PGmc.

The non-finite (or nominal) verb forms are: infinitive, present participle, and past participle.

6.3. The classification of Germanic verbs

Germanic verbs distinguish between a thematic and an athematic inflection (i.e. between an inflection with and without a stem vowel). Apart from a few relics of athematic presents the distribution is as follows: the present is thematic, the preterite athematic.

Further, Germanic verbs are divided into two great classes: strong and weak. The strong verbs have preterite forms which generally reflect the formation of the PIE perfect. The weak verbs, on the other hand, form their preterite with a dental suffix. This formation is a PGmc innovation.

Beside these two main classes, Gmc has the so-called “preterite-presents”, which have a mixed inflection: The present tense is inflected as a strong preterite, while the preterite is of the weak type.

6.3.1. Strong verbs

Strong verbs are divided into seven classes, the first six of which are characterized by ablaut alternations between the present and the preterite. To class VII belong those verbs which form their preterite by means of reduplication. Some of them also show ablaut alternations. The basic forms (principal parts) of strong verbs are the following: the

infinitive, the 3sg. and the 3pl. ind. pret., and the past participle, cf., e.g., the Class III verb PGmc **werþana* ‘become’, **warþe*, **wurðun(t)*, **wurðanaz*.

In the present tense, the strong verbs show different stem formations: athematic root presents (cf. **es-/s-* ‘be’, and **dō-* ‘do, make’); thematic presents, which fall into two types, the barytone with an *e*-grade root (cf. **þeite/a-* ‘bite’, **keuse/a-* ‘choose’) and the oxytone with a zero-grade root (cf. **wige/a-* ‘fight’, **sūbe/a-* ‘sleep’) (cf. also **lūke/a-* ‘close’; on this type, see Vine 1985 w. lit.); *j*-presents (cf. **þegje/a-* ‘receive, accept’, **haffe/a-* ‘heave, lift, raise’); and nasal presents (cf. **stande/a-* ‘stand’, **fregne/a-* ‘ask, question’) – in the majority of cases the nasal infix or suffix has become a part of the root, cf. pres. **skeine/a-* ‘shine’, pret. **skain-/skin-*, pres. **spurne/a-* ‘spurn, kick’, pret. **sparn-/spurn-* (cf. Lat. *spernō*, pf. *sprēvī*). Of these formations, the barytone thematic presents are by far the most productive.

The preterite of strong verbs derives, directly or indirectly, from the PIE perfect. In PIE, this formation had an important restriction. It could only be formed if the inherent meaning of the root allowed the subject of the sentence to be affected by the state which resulted from a completed action (cf. section 6.1). In other words, its meaning had to be subject-resultative. Therefore, a verbal root like **deh₃-* ‘give’ was unable to form a perfect, because the act of giving (normally) does not affect the giver/subject – thus, it is not accidental that Homeric Greek, which has maintained the original function of the PIE perfect more faithfully than any other language (cf. Harðarson 2001b: 38 w. lit.), does not have any perfect form of the verb δίδωμι ‘I give’. On the other hand, a perfect like PIE **se-smór-e* ‘has (got) his/her share’ (cf. Hom. Gk/Aeol. ἔμμορε ‘id.’) from the root **smer-* ‘get one’s share’ was quite in order. In those languages which preserve the PIE perfect, it has gradually developed an object-resultative meaning. The consequence of this change was the expansion of the perfect to practically all action verbs. Perfect forms which on semantic grounds cannot claim to be of PIE age are called “neo-perfects”. Most preterite forms of strong verbs in Germanic are of this type.

In PGmc, the reduplication of the perfect was lost, except when it played an important morphological role. Only those verbs which belong to class VII have preserved the reduplication. By virtue of Osthoff’s Law (according to which a long vowel was shortened in the position before a resonant or semi-vowel plus consonant [$\check{V}RC/\check{V}UC \rightarrow \check{V}RC/\check{V}UC$]) verbs like **haitana* ‘call, name’ or **halðana* ‘hold’ (with Gmc *a*-vocalism) could not form their preterites by lengthening the root vowel as the verbs of class VI did, cf. **alana* ‘nourish’, **hlahjana* ‘laugh’, pret. **ōle*, **hlōhe*. Therefore, they maintained the reduplication as a feature of the preterite, cf. **he-gaite*, **he-galðe* (Go. *haihait*, *haihald* with analogical root-initial *h*, <ai> = [ε] in the reduplication). Gothic is the only Germanic language that has systematically preserved the reduplication in class VII of strong verbs. In North and West Gmc, reduplication was abandoned, with the exception of a few relics. Instead, a new ablaut pattern was developed as shown in Table 55.5 (cf. Hilmarsson 1991: 38–40 and Jasanoff 2007: 281). But later, analogical transfer could lead to a slightly different outcome in the separate languages, cf. pret. OS *held*, OIce. *helt* (< **held-*), OE *hēold* (< **heuld-*), OHG *hialt* (< **hē,ld-*). For a detailed account of the development in North and West Germanic, see Jasanoff (2007). For a review of the different explanations of the strong preterite of class VII verbs in the North and West Germanic languages, see Hill (2009).

In classes I–III of strong verbs, the formation of the IE perfect is directly continued, except for the regular loss of the reduplication, cf. **þait-/þit-*, **kaus-/kuz-*, **warþ-/wurð-*

Tab. 55.5: The new ablaut pattern in strong verbs of class VII in North and West Gmc

Present	Preterite
<i>a</i> (* <i>hald-</i>)	<i>e</i> (* <i>held-</i>)
<i>au</i> (* <i>hlaup-</i>)	<i>eu</i> (* <i>hleup-</i>)
<i>ai</i> (* <i>hait-</i>)	<i>ē₂</i> (* <i>hē₂t-</i>)
<i>ā</i> (* <i>lāt-</i>)	<i>ē₂</i> (* <i>lē₂t-</i>)
<i>ō</i> (* <i>blōt-</i>)	<i>ē₂</i> (* <i>blē₂t-</i>)

(from **beitana* ‘bite’, **keusana* ‘choose’, **werþana* ‘become’) < pre-PGmc **b^he-b^hōid-/b^he-b^hid-*’, **ġe-ġóus-/ġe-ġus-*’, **ue-uórt-/ue-urt-*’.

The verbs of classes IV and V have introduced a lengthened *e*-grade root into the indicative plural forms, cf. (**nam-* :) **nāmun(t)*, (**gab-* :) **gābun(t)* (from **nemana* ‘take’, **geþana* ‘give’). That the form **nāmun(t)* is a substitute for expected **numun(t)* is suggested by the evidence of the preterite-presents, cf. Go. (*man* :) *munun* (from *munan* ‘think’). The form **gābun(t)* (with substitution of the ending **-un(t)* for **-ur* < **-r*) comes from pre-PGmc **g^hēb^h-r* which, in turn, has developed from a preform **g^he-g^hb^h-r* through simplification of the phonotactically difficult sequence **g^hb^h* and compensatory lengthening of the preceding *e*. In such cases or, more generally, in weak forms of the structure **T₁e-T₁T₂-* (*T* = obstruent), a new lengthened-grade stem of the perfect developed. A parallel development, which has been explained as dissimilatory loss with accompanying compensatory lengthening, is attested in other languages, cf. Ved. ptc. pf. act. *dās-vāms-* (< **dēk-uós-* < **de-dk-uós-*) from *dās* (**das*) ‘honour, offer, grant’ and *sāh-vāms-* (< **sēġ^h-uós-* < **se-sġ^h-uós-*) from *sah* ‘overcome, conquer, win’ (see Klingenschmitt 1982: 129 and Kümmel 2000: 244, 565), Lat. pf. *sēdī* (= Go. *set-* < **sēd-* < **se-sd-*, see Meiser 2003: 156), and Alb. aor. *dogja* (< **d^hēg^{wh}-* < neo-pf. **d^he-d^hg^{wh}-*) from *djeg* ‘I burn (tr.)’ (see Klingenschmitt 1982: 131 and Schumacher 2005: 617f., the latter also on the derivation of the preterites of class V strong verbs from reduplicated perfects). This new stem with lengthened vowel gradually spread to all roots of the type **CeC-*, **CRc-*, or **CeR-*; cf. weak perfect stems in Sanskrit like *pet-* and *ten-* (beside *papt-* and *tatn-* from *patⁱ* ‘fly, fall’ and *tan* ‘stretch’), which got their radical *e* from perfect stems like *sed-* (< **sazd-* from *sad* ‘sit down’, where **az* became *e* by regular sound change). This development benefited from the opposition of the present root form **et-* and the preterite **ēt-* of the verb **etana* ‘eat’, which continue PIE **h₁ed-* and **h₁e-h₁d-*, respectively (cf. Neri 2011: 181). However, in PGmc the lengthened-grade perfect could not spread to the strong verbs of classes I–III because their root structures would have made them subject to shortening by Osthoff’s Law, as was the case with *CaRC*-roots of class VII. (According to another theory, the lengthened-grade preterite forms in question derive from the imperfects of Narten presents [cf. Jasanoff 2012]. It is, however, difficult to see how a lengthened-grade imperfect could have become a part of the perfect morphology, i.e., how it could have come to be used as an equivalent of the perfect, and in such a way that it was limited to the “weak” [plural and dual] stem.)

In class VI the ablaut alternation *a* : *ō* originates in verbs such as **akana* ‘drive’ and **haffjana* ‘heave’, whose singular preterite forms **ōke* and **hōfe* go back to pre-

Tab. 55.6: The ablaut alternations in strong verbs of classes I–VI

I	(PGmc ablaut: <i>ei – ai – i – i</i>) Cf. Go. <i>beitan</i> (= /bītan/) ‘bite’, <i>bait</i> , <i>bitun</i> , <i>bitans</i> .
II	(PGmc ablaut: <i>eu/ū – au – u – u</i>) Cf. Go. <i>niutan</i> ‘enjoy’, <i>naut</i> , <i>nutun</i> , <i>nutans</i> , <i>ga-lūkan</i> ‘shut, lock’, <i>-lauk</i> , <i>-lukun</i> , <i>-lukans</i> .
III	(PGmc ablaut: <i>eRC – aRC – uRC – uRC</i> [<i>ReC – RaC – RuC – RuC</i>]) Cf. Go. <i>bindan</i> ‘bind’, <i>band</i> , <i>bundun</i> , <i>bundans</i> , OIce. <i>bresta</i> ‘burst’, <i>brast</i> , <i>brusto</i> , <i>brostenn</i> .
IV	(PGmc ablaut: <i>eR/uR – aR – ēR – uR</i> [<i>Ru – Ra – Rē – Ru / u – wa – wē – u</i>]) Cf. Go. <i>niman</i> ‘take’, <i>nam</i> , <i>nemun</i> , <i>numans</i> , OHG <i>neman</i> ‘id.’, <i>nam</i> , <i>nāmun</i> , <i>ginoman</i> , OS <i>kuman</i> ‘come’, <i>quam</i> , <i>quāmun</i> , <i>kuman</i> , OIce. <i>troþa</i> ‘tread’, <i>traþ</i> , <i>tróþo</i> , <i>troþenn</i> , <i>sofa</i> ‘sleep’, <i>svaf</i> , <i>svófo</i> (<i>sófo</i>), <i>sofenn</i> .
V	(PGmc ablaut: <i>eC – aC – ēC – eC</i>) Cf. Go. <i>giban</i> ‘give’, <i>gaf</i> , <i>gebun</i> , <i>gibans</i> , <i>bidjan</i> ‘ask, beg’, <i>baþ</i> , <i>bedun</i> , <i>bidans</i> , <i>fraihnan</i> ‘ask (questions)’, <i>frah</i> , <i>frehun</i> , <i>fraihans</i> , OIce. <i>fregna</i> ‘ask, hear, be informed’, <i>frá</i> , <i>frógo</i> , <i>fregenn</i> .
VI	(PGmc ablaut: <i>aC – ōC – ōC – aC</i>) Cf. Go. <i>faran</i> ‘go’, <i>for</i> , <i>forun</i> , <i>farans</i> , <i>wahsjan</i> ‘grow, increase’, <i>wohs</i> , <i>wohsun</i> , <i>wahsans</i> , <i>standan</i> ‘stand’, <i>stoþ</i> , <i>stoþun</i> (the past ptc. is missing).

PGmc **h₂a-h₂ôg-e* and **ke-kóh₂p-e*, respectively. The *ō*-vocalism of the plural is partly phonetically regular (**ōkun(t) < *h₂a-h₂ôg-*), partly analogical (**hōþun(t)* instead of **hāþun(t) ← *ke-kə₂p-*). Some verbs of class VI are iteratives which have turned into primary verbs in PGmc, cf. **daujanaⁿ* ‘die’, **swarjanaⁿ* ‘swear’ and **wahsijanaⁿ* ‘grow’. The original iterative present **wahsija-* ‘grow’ (cf. Go. *wahsjan*, 3sg. pres. *wah-seiþ*, pret. *wohs*) has been transformed into **wahsa-* in North and West Germanic by analogical back-formation based on the pret. **wōhs-*. Beside the iterative verb, PGmc had a homonymous causative verb, which formed a weak preterite (cf. ON *vexa* ‘increase’, 3sg. pret. *vexte*). The ablaut alternations in strong verbs of classes I–VI are displayed in Table 55.6.

Most verbs of class VII must originally have shown ablaut alternations in their reduplicated preterites. The “strong” stem of the singular had *o*-grade of the root, whereas the “weak” stem of the dual and plural had zero-grade, cf. **re-rō-/re-r-* (OIce. *rera*, *rerol/røro*), **se-zō-/se-z-* (OIce. *sera*, *serol/søro*), **le-lōt-/le-lt-* (Go. *lailot*, OE [Angl.] *leorton*, dissimilated from **leltun*), **re-rōþ-/re-rð-* (Go. *-rairoþ*, OE *reordon*) from **rō-je/a-* ‘row’, **sā-je/a-* ‘sow’, **lēt-e/a-* ‘let’ and **rād-e/a-* ‘advise’. But it is not quite clear whether those preterites that had *a*-vocalism in the root forms of the singular, as did the corresponding presents, still maintained the zero-grade of the root in the dual and plural in late PGmc or had generalized the vocalism of the singular. In other words: We do not know for sure whether PNGmc and PWGmc had the pattern **he-gait-* : **he-git-*, **he-*

gald : **he-guld-* or **he-gait-* : **he-gait-*, **he-gald-* : **he-gald-* before the reduplicated forms were compressed or replaced by new ablaut forms (cf. OE *heht* and *hēt* from *hātan* ‘call, command, promise’ and *leolc* < **lelk-* and *lēc* from *lācan* ‘swing, move, play’). The possibility of explaining some of the preterite forms with secondary ablaut as the regular outcome of reduplicated forms with a zero-grade root favors, however, the first alternative, cf., e.g., **e-uk-* (OIce. *iók*) from **auk-i/a-* ‘increase’, **be-but-* (> **beut-* or **befst-* in OE *bēot* and *beoftun*, *beafton*, cf. northern ME pret. and pp. *befst* in *Cursor Mundi*; for these forms cf. Hogg and Fulk 2011: 254; in the supposed preform **be-but-* [beβut] either the fricative *β* or the vowel *u* was ejected) from **baut-i/a-* ‘beat’, and **he-git-* > *[xe-jit-] (> **heht-* or **hej(i)t-* > **heit-* > **hē₂t-*) from **hait-i/a-* ‘call, command, promise’. Later, the stem forms of the plural would have been extended to the singular. Preterite forms like **hleup-* (OIce. *hlióp*, OE *hlēop*) from **hlaup-i/a-* ‘leap, run’ or **lē₂k-* (OIce. *lék*, OE *lēc*) from **laik-i/a-* ‘play’ are best explained as analogical formations. Jasanoff (2007: 265 ff.) explains plural stem forms like **heht-* and **held-* (for phonotactically impermissible ***hegld-*) as a result of the proportion **le-lōt-*, **re-rōp-* : **le-lt-*, **re-rā-* = **he-gait-*, **he-gald-* : pl. *X*.

6.3.2. Weak verbs

Weak verbs are divided into four classes, depending on their original stem formation. The vast majority of them are “derivative verbs”, either denominatives or deverbatives. In pre-PGmc, they were restricted to the present system and, thus, only formed the present and imperfect tenses. Later, they developed a new preterite formation at the expense of the old imperfect. This was the so-called “weak preterite”.

The classes of the weak verbs are as follows:

Class I: The verbs of this class continue four different types of formations: 1. iteratives and causatives in **-eje/a-* (e.g. **war-eje/a-* ‘hinder, defend, protect’, **flaut-eje/a-* ‘set afloat, launch’), 2. denominatives in **(-e/i/u)-je/a-* (cf. **dōme-je/a-* ‘judge’, **late-je/a-* ‘hold back, dissuade’ (lit. ‘make lazy’), **warme-je/a-* ‘warm’, **mekele-je/a-* ‘enlarge, praise’, **mati-je/a-* ‘take a meal, eat’, **naudī-je/a-* ‘force, compel’, *(*ga-*)*hardu-je/a-* ‘harden’, **namn-ije/a-* ‘name’), 3. iterative-intensives in **-at-ije/a-* (also **-et-ije/a-*, **-it-ije/a-*, **-ut-ije/a-*; the suffix clusters **-at-ije/a-* and **-it-ije/a-* have comparanda in Gk -άζω and -ίζω [< **-ad-iō* and **-id-iō*, respectively]; for a historical account of their derivational bases see Rau 2010: 162–170) (cf. Go. *swogatjan* ‘sigh’ : *gaswogjan* ‘id.’, OE *hlēapettan* ‘leap up’ : *hlēapan* ‘leap’, OIce. *skrækta* ‘screech, shriek’ : *skrækia* ‘id.’), 4. primary verbs in **-(i)je/a-* (cf. **bugja-* ‘buy’, **brūkija-* ‘use’, **purzija-* (**pursija-*?) ‘be thirsty’, and **wurkija-* ‘work, make’).

After the raising of *e* to *i* in unstressed syllables and after the fronting and unrounding of unstressed *u* before *j*, the stem endings **-eje/a-*, **-ije/a-*, and **-uje/a-* fell together, yielding **-iji/a-* (on the change of **-ujV-* to **-ijV-* in PGmc, see Hill 2012: 12 et passim). As Sievers’ Law, which regulated the distribution of the suffix alternants **-jV-* and **-ijV-*, was still operative in PGmc, verbs with a long root-syllable or a polysyllabic stem retained the suffix allomorphs **-iji-* and **-ija-*, while verbs with a short root- or stem-syllable changed them into **-ji-* and **-ja-* (by the “converse of Siever’s Law”, cf. 1.1.1.1). From this, it is obvious that the loss of intervocalic *j* in PGmc (except under the circumstance that it was preceded by *i* and followed by a vowel other than *i*, where

it remained as a glide) is younger than both the raising of unstressed *e* to *i* and the fronting and unrounding of unstressed *u* before *j*. Later, the loss of *j* before *i* (< *e*) led to the rise of two semi-thematic conjugations, of which one had the suffix allomorphs **-ī-* and **-ija-*, the other **-i-* and **-ja-*. As a result of this development, early PGmc stem forms like **war-eje/a-*, **flaut-eje/a-*, and **mekele-je/a-* were changed into **war-ijja-*, **flaut-ī/ija-*, and **mekil-ī/ija-*.

The preterite of weak verbs was based on the corresponding past participle. Of the different verbs belonging to class I only the types 1. and 4. were able to form a past participle in pre-PGmc. According to the evidence of Italic and Celtic, the (late) PIE intensives and causatives in **-éje/o-* formed a past participle in **-etó-* (cf. Meiser 2003:136). For early PGmc, we may therefore reconstruct forms like **wardēda-* and **flautēda-*. Descriptively, the past participle was formed from the present stem deprived of the suffix **-je/o-*. All indications are that the denominatives in **-ejé/ó-* and, subsequently, also the denominatives in **-ijé/ó-* and **-ujé/ó-* adopted this kind of participial formation. Thus, we can also posit forms like **lateda-*, **warmeda-*, **naudīda-*, **namnīda-*, and **hardūda-* for early PGmc. The change of the stem termination **-uje/a-* into **-ije/a-* in denominatives of the type **hardūje/a-* led to a corresponding transformation of the allied participle: **hardūda-* → **hardīda-*.

In pre-PGmc, the primary verbs in **-je/o-* formed the related past participle by adding the suffix **-tó-* to the bare root. PGmc has retained this formation, cf. **bugja-*, pp. **buhta-*, **brūkija-*, pp. **brūhta-*, **wurkija-* (= Av. *vərəziia-* ‘work, make’), pp. **wurhta-* (= Av. *varšta-* ‘made’ < **uršta-*). On the other hand, the past participle **purzīda-* (instead of the adjective-turned **pursta-*) ‘thirsty’ from **purzija-* (**pursija-*?) ‘be thirsty’ was formed on analogy with the secondary verbs.

Class II: The verbs of this class continue the PIE suffix cluster **-eh₂-je/o-* (> **-ah₂-je/o-*), which in PGmc first became **-ō-je/a-* (via **-ā-je/o-*) (for the view that the present inflection of Germanic weak verbs of class II can be accounted for wholly on the basis of pre-PGmc thematic stems in **-ā-je/o-*, see Cowgill 1959), but the later development was as follows: **-ō-je/a-* > **-ō-ji/a-* > **-ōi-/ō-* (loss of *j* before *i* and between vowels, contraction of **-ōa-* to **-ō-*) > **-ai-/ō-* (*-ōi-* > *-ai-* /_C : shortening by Osthoff’s Law). This awkward allomorphy was eliminated in favor of **-ō-*.

A corresponding substitution of **-ō-* for **-ai-* (< **-ōi-* < **-ōai-* < **-ōjai-*) took place in the paradigm of the optative. A verb such as **salb-ōje/a-* ‘anoint’ had the optative **salb-ōjai-* in early PGmc. This form should ultimately have become **salb-ai-*. Since this development would have led to the merger of the optative in question with that of strong verbs and weak verbs of class III (i.e. denominatives in *-ai/a-*), there was good reason to remodel it. The only way to link it formally with the corresponding present indicative, which had generalized the stem alternant **salb-ō-*, was to use the same stem form. Accordingly, the optative is distinguished from the indicative by separate endings only (cf. opt. 3sg. Go. *salbo*, OHG *salbo*, 3pl. Go. *salbona*, OHG *salbōn* vs. ind. 3sg. Go. *salbop*, OHG *salbōt*, 3pl. Go. *salbond*, OHG *salbōnt*).

Synchronically, the verbs of class II fall into two groups: 1. Deverbatives: these are often semantically opposed to their base verbs in showing iterative or intensive semantics, cf. Go. *hvarbon* ‘wander, go about, walk’, OIce. *hvarfa* ‘wander, stroll about, turn round’, OE *hwearfian* ‘wander, turn, roll or toss about’ (: Go. *hvarban* ‘walk’, OIce. *hverfa* ‘turn round, turn out of sight’, etc.), and 2. Denominatives, cf. Go. *karon* ‘care for, be concerned about’, OE *cearian* ‘take care, be anxious’ (: Go. *kara* f. ‘care, anxiety’,

OE *caru*, *cearu* ‘care, sorrow, grief’); OE *nīwian*, OHG *niuwōn* ‘renew, renovate’ (cf. OIce. *endrnýia* ‘renew, repeat’) beside Go. *ananiujan*, OS *niuwian* ‘id.’ (: Go. *niujis* ‘new’, etc.).

Class III: The verbs of this class overwhelmingly have stative meaning. They can be divided into deverbatives on the one hand and denominatives on the other. To the former group belong verbs like Go. *haban*, OIce. *hafā*, OE *habban* (with analogical *a* instead of *e*), OS *hebbian*, OHG *habēn* ‘have’ and OIce. *hanga* (2 and 3sg. pres. ind. *hanger*), OHG *hangēn* ‘hang (intr.)’, *pendēre*’ (: Go. *hāhan*, OE *hōn*, OS *hāhan* ‘hang (tr.)’, *suspendere*’); the second group includes verbs like Go. *fastan*, OHG *fastēn* ‘fast’ (← ‘be steadfast, firm’) (: OIce. *fastr*, OE *fæst*, OHG *fast* ‘firm’, etc. < PGmc **fasta-*), Go. (*ga*)*leikan* ‘please’, OHG (*gi*)*līhhēn* ‘resemble, please’ (← ‘be similar’) (: Go. *galeiks*, OE *zēlic*, OHG *gilīh* ‘like, similar’ < PGmc **ga-līka-*).

Beside these verbs, class III also comprises denominatives with a factitive meaning, cf. Go. (*ga*)*weihan* ‘consecrate’, OE *ful-wian* ‘baptize’ (secondarily a verb of class II; for the assignment of this verb to the original class III verbs, see Heidermanns 1993: 663) (: Go. *weihs* ‘holy’, OHG *wīh* ‘id.’ < PGmc **weiha-*) and Go. *ana-þiwan*, *ga-þiwan* ‘enslave’, OE *þeowian* ‘enslave, serve’ (with relic forms of class III), OHG *dewēn* ‘subjugate, humiliate’ (: **þewa-* ‘servant’ in Go. *þius*, etc.).

In PIE, the primary verbs and deverbatives had the suffix cluster *-*h₁-ǵé/ó-*, cf. **k₂p-h₁-ǵé/ó-* ‘have’ (from the root **keh₂p-/k₂p-* ‘seize, grasp’, cf. Gk κάπτω ‘I snatch, swallow’, Lat. *capiō* ‘I seize, grasp’), which yielded late PGmc **hab-ai-/hab-ja-* (via pre-PGmc **k₂p₂ǵé-/k₂p(h₁)ǵó-* (with loss of *h₁* according to Saussure’s Law, which states that a laryngeal was lost in the tautosyllabic sequences **oRH* and **HRo*) > early PGmc **hab₂je-/hab₂ja-* > **hab₂(j)i-/hab₂ja-*). (This explanation was suggested by Sergio Neri 2003: 264–265, 283. On the non-application of the so-called “Pinault’s rule” [*-*CH₂-* > *-*C₂-*] in PIE, see Lipp 2009: II, 449–458.) The suffix cluster *-*h₁-ǵé/ó-* originates in *ǵel/o*-derivations from primary aoristic (“fientive”) stems in *-*eh₁-* (cf. Harðarson 1998: 328–334). For the semantics, which correspond to that of the perfect (cf., e.g., ‘seize’ → pf. ‘have seized’ = ‘hold, have’), see Harðarson (1998: 334–337). The vocalization of the laryngeal was phonetically regular after a root ending in an obstruent (cf. Neri 2009: 8). The irregular (marked) suffix allomorphy *-*ai/ja-* in Gmc has been retained in some verbs of high frequency, cf. the verb for ‘have’ in the Ingvaenic dialects of West Gmc and the verbs for ‘say’ and ‘be silent’ in North Gmc.

Denominatives derived from *e/o*-stems by means of the suffix *-*h₁ǵé/ó-* had the stem ending *-*eh₁ǵé/ó-*, which ultimately developed into late PGmc *-*ai/a-* (via *-*ēǵé/ēǵó-* > early PGmc *-*āje/āja-* > *-*āji/āja-* > *-*āi/āa-* > *-*ai/ā-* > *-*ai/a-* (the shortening of the diphthong **āi* and the monophthong **ā* was caused by Osthoff’s Law, cf., e.g., 3sg. pres. ind. **āiþi* > **aiþi* and 3pl. **ānþi* > **anþi*). Here we apparently have the main source for the normal stem alternation *-*ai/a-* of class III weak verbs in Gothic and North Gmc (another explanation of this stem alternation is given by Harðarson 1998: 330 f. fn. 18).

As for the denominatives with a factitive meaning, they are best connected with the Greek type δουλόω ‘I make a slave of, enslave’ (: δούλος ‘slave’), ἐλευθερώω ‘I set free’ (: ἐλεύθερος ‘free’), whose stem termination reflects PIE *-*o-ǵe/o-*. The desubstantives are presumably older than the deadjectives, which have partially replaced the factitives in *-*eh₂(-ǵe/o)-*.

In PGmc, the stem ending **-o-je/o-* developed as follows: PIE **-o-je/o-* > PGmc **-aje/a-* > **-aji/a-* > **-ai/ā-* > **-ai/a-* (cf. above). Here is the second source for the stem allomorphy **-ai/a-* of class III weak verbs in Gothic and North Gmc (cf. Harðarson 1998: 331 and fn. 22).

Finally, there are continuants of reduplicated formations from roots of the shape **C(R)ejH-* which ended up in class III of weak verbs, cf. Go. *reiraiþ** ‘trembles’ (pret. *-reiraida*) < PIE intensive **h₃r_éj-h₃rojH-ti*, (cf. Ved. *alelet* ‘trembled’, see LIV² 307 f.) and OHG *bibēn* ‘tremble, quiver’, OIce. *bifask* ‘id.’ < **beḅai-/*biḅja-* ← PGmc **ḅeḅai-/*ḅeḅī-* < PIE pf. **b^he-b^hoiH-/b^he-b^hiH-* ‘be frightened’ (cf. Ved. *bibhāya* ‘is frightened’) (see Harðarson 2001a: 98 f. and LIV² 72 f.).

Class IV: In PGmc, verbs of this class had the suffix allomorphs **-nō-* and **-na-*, cf. **wak-nō/na-* ‘awake’ in Go. *ga-waknan*, OIce. *vakna*, OE *wæcnan* (strong verb of class VI, pret. *wōc*), *wæcnian* (weak verb of class II, pret. *wæcnode*). The singular of the present indicative had the stem alternant **wak-nō-*, the plural **wak-na-*. In Nordic, the full grade of the suffix has been generalized, while in Gothic, the present inflection has been remodeled on the analogy of thematic verbs (cf. *bairam* : *bairand* : *bairiþ* = *-wak-nam*, *-waknand* : X [= *-wakiþ*]; the preterite is, on the other hand, formed from the stem alternant in *-nō-* [*-wakiþ*]). In the West Gmc languages, both generalizations occur (beside other transformations).

The suffix **-nō/na-* has its origin in PIE *n*-infix presents derived from roots ending in **h₂* or **h₃*. An example of such a verb is PGmc **spurnō/na-* ‘kick, strike with the foot, spurn’ < PIE **sp^hr-né-H-/sp^hr-n-H-* (from the root **sp^herH-* ‘kick’), cf. OIce. *sporna* (pres. *spornar*, pret. *spornaþe* or *sparn*, *spurno*, which led to the formation of a new [strong] pres. *spern*), OHG *spornōn* and *spurnan*, OS and OE *spurnan*. Later, by reanalysis of the morpheme boundary, **-nō/na-* or, more accurately, its predecessors **-nah₂/nə₂-* and **-noh₃/nə₃-* became an independent suffix (the vocalization of the zero-grade suffix [i.e. **-nə_{2/3}-* > **-na-* instead of **-ŋh_{2/3}-* > **-un-*] is analogical after the full-grade alternant). A comparable development can be observed in other languages.

The ingressive function of *nō/na-*verbs originates in primary verbs derived from roots with ingressive semantics, cf. Go. *ga-þaursnan* ‘dry up, wither’, OIce. *þorna* ‘become dry’ < PGmc **þurz-nō/na-* ‘become dry, dry up’ from PIE **ters-/*t₂s-* ‘id.’ (also in the transferred meaning ‘become thirsty’), ON *sofna* ‘fall asleep’ < PGmc **suḅ-nō/na-* ‘id.’ from PIE **suep-/sup-* ‘id.’, OIce., MIce., Far. *fúna* ‘rot, decay’ < PGmc **fū-nō/na-* ‘rot, become stinking’ from PIE **peuH-/puH-* ‘id.’. Also in Balto-Slavic, the ingressive function of nasal verbs is best explained in this way, cf. Lith. *bundù* ‘I awake’ : *budėti* ‘wake, be awake’, OCS *vbz-b^hn^hoti* ‘awake’ : *b^hděti* ‘wake’, from the PIE root **b^heud^h-* ‘become awake’ (for a similar explanation, see Meiser 1993: 293 w. lit.). Gorbachov’s claim (2007) that the verbs in question continue a PIE “*h₂e*-conjugation paradigm” is highly improbable.

Some verbs with a geminated voiceless stop such as, e.g., OHG *leckōn*, OE *liccian* ‘lick’ and OHG *zockōn* ‘drag, tug, jerk’, were originally class IV weak verbs. Their geminates arose by the operation of Kluge’s Law, according to which the PIE sequences *VTn^hŷ*, *VDn^hŷ*, and *VD^hn^hŷ* (*T* = tenuis, *D* = media, *ŷ* = stressed vowel) all yielded *VTTV* in PGmc (After a long syllable the geminate *TT* was shortened. Early-PGmc *VRDn^hŷ* [with *VR* < PIE *e/oR* or *R*, and *D* < PIE *T*, *D*, or *D^h*] changed regularly into *VRTV*): PGmc **likkō-*, **tukkō-* < pre-PGmc **liġ^h-náh₂-*, **duk-náh₂-* (cf. Hirt 1931: 91–94, 1932: 164 w. lit.). PGmc **tukkō-* can be compared with the *nā*-present Toch. A *tsāknā-* ‘pull

out' (2sg. *tsäknāt*) < PT **t^sək-na-* < pre-PT **duk-nH-* (on the form of the suffix of present class VI in Tocharian, see Ringe 1996: 95 f. and Pinault 2008: 586 f.).

The reason why verbs like **sub-nō/na-* 'fall asleep' and **wak-nō/na-* 'awake' do not show the effect of Kluge's Law is the generalization of the sequence *-bn-* or *-kn-*, which was regular in the dual and plural (1st and 2nd persons), cf. early PGmc **sub-na-máz*, **wag-na-máz* (the shift of the mediae had still not occurred; note in this connection that in PGmc the shift of the PIE mediae occurred later than that of the tenues and the mediae aspiratae, as the outcome of Kluge's Law shows) > **súb-na-maz*, **wák-na-maz*. Here, the sound law in question did not operate, because the suffix was unstressed (since, as noted above, the operation of Kluge's Law was conditioned by a stressed vowel following the PIE sequences *Tn*, *Dn* and *D^hn*, the PIE sequence *Dn* regularly yielded PGmc *Tn* if no stressed vowel followed). The selection of **sub-n-* and **wak-n-* (and not **supp-* and **wakk-* from **suppō-* und **wakkō-* < **sub-nō-*, **wag-nō-*) was surely influenced by the coexisting verbs **sub-e/a-* 'sleep' and **wak-ai/ja-* 'wake'. (Other, less convincing explanations have been proposed by Kroonen [2011: 95–97, 2012] and Scheungraber [2011: 104 f., 2012].)

6.3.3. Preterite-presents

As already mentioned (in 6.3), the verbs of this class have a mixed inflection. In the present tense, they are inflected as strong preterites (hence their designation), but in the preterite tense as weak verbs. The present is based on the PIE perfect, whereas the preterite is a Gmc innovation. The preterite-presents can be divided into six subclasses according to their root structure (For a detailed discussion of the Germanic preterite-presents, see Tanaka 2011: 105–242. I must admit that I do not agree with some of his suppositions and theories.):

Class I: **wait-/wit-* 'know' (Go. *witan*, etc.), **lais-/lis-* 'know' (Go. *lais*), **aih-/aig-* 'possess, own' (Go. *aih*, *aigun*, etc.).

Class II: **đaug-/đug-* 'suffice, be good for, avail' (Go. *daug*, OE *dēaz/duzon*, etc.).

Class III: **kann-/kunn-* 'know' (Go. *kann*, *kunnun*, etc.), **ann-/unn-* 'love, grant' (ON *ann*, *unno*, etc.), **þarf-/þurb-* 'need' (Go. *þarf*, *þaurbun*, etc.), **dars-/durz-* 'dare, be bold' (Go. *ga-dars*, *ga-daursun*, OE *dearr*, *durron*, etc.).

Class IV: **skal-/skul-* 'owe, shall' (Go. *skal*, *skulun*, etc.), **man-/mun-* 'remember' (Go. *man*, *munun*, etc.).

Class V: **nah-/nug-* 'be sufficient' (Go. *ga-*, *bi-nah*, OE *ze-neah*, *ze-nuzon*, etc.), **mag-/mag-* (later also **mag-/mug-* by analogy) 'have power, be able' (Go. *mag*, *magun*, OSw. *mā*, *magho/mughu*, OE *mæz*, *mazon*, OHG *mag*, *magun/mugun*, etc.).

Class VI: **ōg-/ōg-* 'fear' (Go. *og*, Olce. *óask* ← **ōh* < **ōg*), **mōt-/mōt-* 'have room' (Go. *ga-mot* 'I find room', OE *mōt* 'I may, am allowed', *mōton*, OHG *muoz* 'I may, can', *muozun*).

6.4. The formation of the optative

6.4.1. The present optative of thematic verbs

This optative is formed from the present stem by substituting the suffix *-ai-* for the thematic vowel:

- a) Strong verbs: e.g. **geþ-e/a-* → opt. **geþ-ai-*.
- b) Weak verbs: e.g. **war-eje/a-* → opt. **war-ejai-*.

The suffix *-ai-* derives from pre-PGmc **-oi-* < PIE **-oih₁-*, a combination of the thematic vowel **-o-* + the zero-grade **-ih₁-* of the optative suffix (cf. below).

6.4.2. The present optative of athematic verbs

Due to secondary transformations, this optative is scantily attested in the Gmc languages. Apart from the verb for ‘will’, whose present indicative continues an old optative (cf. Go. *wiljau*, *wileis*, *wili*, *wileima*, *wileiþ*, *wileina*, with the synchronic stem /wil-ī-/), it is best preserved in the verb for ‘be’. In Old English, Old Saxon, and Old High German, the stem form is *sī-* throughout the paradigm (1/3sg. OE, OS, OHG *sī*, 2sg. OE *sī*, OS *sīs*, OHG *sīs*[t], 3pl. OE, OS, OHG *sīn*), which suggests that it had the generalized zero-grade alternant of the PIE optative suffix **-ieh₁/ih₁-* already in PGmc, as is the case in the preterite optative (cf. below). In Gothic and Nordic, on the other hand, the PGmc stem **sī-* (< PIE **h₁s-ih₁-*) has been remodeled under the influence of the thematic optative: 1sg. Go. *sijau*, OIce. *siá* (< **sija* < **sijō* < **sijau*), 2sg. Go. *sijais*, OIce. *ser* (< **sjér* < **sijer* < PN **sijēr* < **sijaiz* < **sij-ai-* ← **sī-*; on the disyllabic optative forms in Old Norse, cf. Jónsson 1921: 260 f., Noreen 1923: 364, and Heusler 1932: 99; for their derivation from the stem **sijai-*, cf. Jóhannesson 1923–24: 113, 321 f. Old English has the by-form *sīe* (pl. *sīen*) (< **sij-ai-*) that, together with the corresponding form *sie* (sg. and pl.) in Old West Frisian, can be assigned to the influence of North Gmc.

6.4.3. The preterite optative

This optative is formed with the suffix **-ī-* (< PIE **-ih₁-*) from the plural stem of the preterite indicative:

- a) Strong verbs: pret. ind. **gab-e* / *gæþ-un(t)* → opt. **gæþ-ī-*.
- b) Weak verbs: pret. ind. **wareðæ(t)*, *-ðō(t)* / *wareðun(t)*, *-ðæðun(t)* → opt. **wareð-ī-* (cf. 3sg. OIce. *verþe*, OE *werede*, OS *weridi*, OHG *weriti*) and **wareðæð-ī-* (cf. 3sg. Go. *waridedi*). PGmc **wareð°* > **warid°* (cf. 6.3.2).
- c) Preterite-presents: pret. ind. **kunþæ(t)*, *-ō(t)* / *kunþun(t)*, *-æðun(t)* → opt. **kunþ-ī-* (cf. 3sg. OIce. *kynne*, OE *cūðe*, OHG *kondi*, MHG *kunde*, *künde*) and **kunþæð-ī-* (cf. 3sg. Go. *kunþedi*).

6.5. The formation of the imperative

The imperative is only formed from the present stem. It existed in the 2nd person singular, dual, and plural and in the 3rd singular and plural. The forms of the 2nd dual and plural are identical with those of the present indicative. The 2nd singular shows the bare stem (the thematic verbs have the stem alternant ending in *-e*), cf. **ḍō* ‘do’, **ḥere* ‘carry’, **ḥafje* ‘lift’, **wurkije*, ‘work’ and **wareje* ‘protect’. The forms of the 3rd person singular and plural go back to the “future imperative” of PIE, cf. Lat. 3sg. *emitō* ‘(s)he shall take’, 3pl. *emuntō* ‘they shall take’ < PIt. **em-e-tōd* and **em-o-ntōd*, respectively (on the future imperative in PIE, see Harðarson 1993: 47–49 w. lit.). After the PGmc endings **-ḍōt* and **-ndōt* had lost their final dental, they were extended by the particle *u*, which is identical with the *u* seen in the active imperative endings of the 3rd person singular and plural in Indo-Iranian and Anatolian, cf. Skt 3sg. *as-tu*, 3pl. *s-antu* (from *as-* ‘be’), Hitt. 3sg. *ēš-tu*, 3pl. *aš-antu* (from *ēš-* ‘id.’) (cf., e.g., Bethge 1900: 375, Prokosch 1939: 215 f., and Krause 1968: 227). The endings of the 3rd singular and plural in Gothic are *-dau* and *-ndau*, respectively. (Brugmann [1921: 39 ff.] preferred to interpret the actually attested Gothic imperatives in *-dau* and *-ndau* [*lausjadau*, *atsteigadau*, *liugandau*] as passive forms. Lühr [1978: 109, 114 f.] assents to that interpretation. But an unbiased examination of the occurrences reveals that it is untenable. The forms are clearly active.)

The derivation of the 3rd person active imperative desinences of Gothic and Old Irish (cf. *bered* ‘let him/her carry’, pl. *berat*) from PIE **-t-ou* and **-nt-ou*, respectively, with an ablaut variant **ou* of the particle **u* (cf. Cowgill 1975: 65, 1983: 81 f. and McCone 1986: 241 f.) is very questionable. For another (and more plausible) account of the Old Irish endings, see Eska (1992).

6.6. The passive voice

PGmc inherited the middle voice from PIE. Gothic is the only Gmc language that preserves it, albeit solely in the present indicative and optative. The Gothic continuants of the PIE middle have exclusively passive meaning. Thus, it is more adequate to assign them to a *passive* category rather than to a middle or mediopassive one. Go. *gamarzjada* (from *gamarzjan* ‘offend’) and *gaurjada* (from *gaurjan* ‘make sorry’) cannot be considered true middles (contra Lühr, this handbook, 4). The synthetic presents (*saei ni*) *gamarzjada* (*in mis*) (Mt 11:6, Lk 7:23) and (*iþ jabai in matis broþar þeins*) *gaurjada* (Rm 14:15) correlate with the analytic passive preterites (*jah*) *gamarzidai waurþun* (*in þamma*) (Mk 6:3) and (*ni unte*) *gauridai wesuþ* (2K 7:9), respectively. This shows that the presents are passive, too. Cf. the opposition of the transitive active *gaurja* and the passive participle *gaurida* in 2K 2:2 (*unte jabai ik gaurja izwis, jah huas ist saei gailjai mik, niba sa gaurida us mis?* ‘For if I make you sorry, who then is there that makes me glad but the one made sorry by me?’; the phrase *sa gaurida us mis* is, indeed, a slavish rendering of Gk ὁ λυπούμενος ἐξ ἐμοῦ, where ἐξ ἐμοῦ denotes the agent of the action).

Apart from Gothic, Nordic and Old English each preserve an isolated relic of the PIE middle, also with passive meaning, i.e. PN **haite**, ON *heite* ‘I am called’ and OE *hätte*

‘is called’ (also used as 1sg. ‘I am called’). The Nordic forms are important insofar as they reflect an old formation of the 1sg. pres. ind. In Gothic and Old English, on the other hand, the form of the 3sg. has been substituted for that of the 1sg.

The development of the verb for ‘be called’ was as follows (on the development of the Gothic paradigm, cf. Lühr 1978: 109–114):

	Go. (unspec.), Nc, OE		PGmc		pre-PGmc	
sg.	1	PN haite , ON <i>heite</i>	<	* <i>haitōi</i>	<	*-o- <i>h₂aī</i>
	2	<i>haitaza</i>	<	* <i>haitazai</i>	<	*-o- <i>soī</i>
	3	<i>haitada</i> = OE <i>hätte</i>	<	* <i>haitadai</i>	<	*-o- <i>toī</i>
pl.	1	<i>haitanda</i>	<	* <i>haitamedai</i>	<	*-o- <i>med^hh₂oī</i>
	2	<i>haitanda</i>	←	* <i>haitadwai</i>	<	*-o- <i>d^h(h₂)uoī</i>
	3	<i>haitanda</i>	<	* <i>haitandai</i>	<	*-o- <i>ntoī</i>

If the Gothic passive is compared with the mediopassive of Sanskrit and Greek, then its restriction to the present indicative and optative is striking. In PIE, the middle voice belonged to both the present system and the aorist system. The perfect system, on the other hand, had no middle voice. The reason why Gmc does not have any passive preterite is apparent: the categories imperfect and aorist were lost in PGmc. The loss of the passive preterite is, thus, linked with the general development of the imperfect and the aorist in PGmc.

In Gothic, the endings of the passive optative have been extended by a particle *u*, cf. 1/3sg. *nimaidau*, 2sg. *nimaizau*, 1–3pl. *nimaindau* (from *niman* ‘take’; it goes without saying that the Gothic forms cited here, as well as elsewhere in this overview, are not all actually attested, but can be securely reconstructed on the basis of other attested forms). The same particle occurs in the imperative endings of the 3rd person singular and plural (see above), cf. 3sg. *nimadau* ‘let him/her take’, 3pl. *nimandau* ‘let them take’ (cf. Osthoff 1881: 256 f. w. lit.).

6.7. The weak preterite

The weak preterite is a special Germanic formation characterized by a dental suffix. Synchronically, it is formed from the corresponding past participle, cf. Go. pret. ind. (1sg.) *nasida*, *salboda*, *habaida*, *waurhta*, *mahta*, pp. (masc. sg. nom.) *nasip̃s*, *salbõps*, *habaĩps* (all with *p* < *đ* before *s*), *waurhts*, *mahts* (from *nasjan* ‘save’, *salbon* ‘anoint’, *haban* ‘have’, *waurkjan* ‘work, make’, *magan* ‘be able’). The indicative paradigm of the weak preterite formed from class I verbs is shown in Table 55.7 (as examples we take the sg. and pl. preterite forms of Go. *warjan*, Olce. *veria*, OE *werian*, OS *werian*, OHG *werien*, *werren* ‘hinder, defend, protect’). Note that in Old High German, though the regular 2sg. ind. ending of the weak preterite is *-tōs* (cf. *weritōs*), the endings *-tēs* and *-dēs* are also attested (cf. *altinotēs* ‘distulisti’, *chiminnerodēs* ‘minuisti’). Also, in Old Saxon, the ending *-dos* (corresponding to OHG *-tōs*) is attested beside *-des*. The OHG dialects South Rhine Franconian and Alemannic have the plural endings *-tōm*, *-tōt*, *-tōn* in place of *-tum*, etc.

Tab. 55.7: The preterite indicative of class I weak verbs in Gothic, Old Icelandic, Old English, Old Saxon, and Old High German

	Go.	OIce.	OE	OS	OHG
sg. 1	<i>warida</i>	<i>varþa</i>	<i>werede</i>	<i>werida</i>	<i>werita</i>
2	<i>warides</i>	<i>varþer</i>	<i>weredes</i>	<i>werides</i>	<i>weritōs</i>
3	<i>warida</i>	<i>varþe</i>	<i>werede</i>	<i>werida</i>	<i>werita</i>
pl. 1	<i>waridedum</i>	<i>varþom</i>	<i>weredon</i>	<i>weridun</i>	<i>weritum</i>
2	<i>warideduþ</i>	<i>varþoð</i>	<i>weredon</i>	<i>weridun</i>	<i>weritut</i>
3	<i>waridedun</i>	<i>varþo</i>	<i>weredon</i>	<i>weridun</i>	<i>weritun</i>

Tab. 55.8: The “endings” of the weak preterite in Gothic compared with the preterite indicative of the verb for ‘do’ in Old English, Old Saxon, and Old High German

	Go.	OE	OS	OHG
sg. 1	<i>-da</i>	<i>dyde</i>	<i>deda</i>	<i>teta</i>
2	<i>-des</i>	<i>dydes(t)</i>	<i>dādi, dedos</i>	<i>tāti</i>
3	<i>-da</i>	<i>dyde</i>	<i>deda</i>	<i>teta</i>
pl. 1	<i>-dedum</i>			<i>tātum</i>
2	<i>-deduþ</i>			<i>tātut</i>
3	<i>-dedun</i>	<i>dydon, dādon</i>	<i>dādun, dedun</i>	<i>tātun</i>

Since the dawn of comparative Indo-European linguistics most scholars have been inclined to connect the dental suffix of the weak preterite with the Gmc verb **dō-* ‘do’ (cf. Lühr 1984: 41 w. lit.). In fact, this connection is very probable, although the details are still controversial. In Table 55.8 the “endings” of the weak preterite in Gothic are compared with the preterite indicative of the verb for ‘do’ in Old English, Old Saxon, and Old High German. (As is well known, the Germanic verb **dō-* is missing in Gothic and North Germanic, both of which instead use the verb *taujan*; in Nordic this verb was later replaced by *gør(v)a*, *ger(v)a*.) The Gothic plural endings, *-dedum*, *-deduþ*, *-dedun*, match completely the corresponding forms of the OHG paradigm, *tātum*, *tātut*, *tātun*. (Note that Go. <e> stands for [e:], which in this case represents the so-called \bar{e}_1 [æ:] in PGmc; this vowel became \bar{a} in both North and West Germanic.) The 3pl. *-dedun* has also its counterparts in OE *dādon* and OS *dādun*, which have been generalized throughout the plural.

Following Rasmussen (1999b: 598 f.), I assume that the weak preterite reflects univerbation of the past participle with the verb for ‘do, make’ (so also Ringe 2006: 167 f.), cf. the Latin expression *missum (aliquem) facio* ‘I make (somebody) go, I send away’ (lit. ‘I make sent’) (see Ernout and Thomas 1972: 278). Originally, the participle was inflected and agreed with the direct object of the verb. Consequently, in pre-PGmc the participle differed in sentences like **im solpātom d^hēd* ‘made him anointed’ (> ‘anointed him’) and **eiām solpātām d^hēd* ‘made her anointed’ (> ‘anointed her’). Later, the acc. sg. neut. of the participle was generalized. Latin shows the same tendency to generalize the neuter singular of the participle, cf. Cic. *Att. VIII, 12b, 2: ut cohortes ... ad me missum facias* ‘that you send me the cohorts’. Such an incongruous use of the neuter

singular is characteristic of Vulgar and Late Latin (cf. Svennung 1935: 263 f.). For the generalization of the acc. sg. neut. of the participle in similar periphrastic constructions, cf., e.g., the *hark*-construction in Hittite (cf. Boley 1984) and the *have*-perfect in the Romance and Germanic languages. A parallel development can be posited for the Oscan *tt*-perfect used by secondary verbs, cf. 1sg. **teremnatom* ‘I limited’ (3pl. *teremnatens*, with *-att-* < **-āt-*) < **(agrom) termenātom fefakom/fēkom* ‘I made (the field) limited’ (cf. Rix 1992: 238 f., 2003: 19 f.; it is possible, although less probable, that the auxiliary in Oscan was the verb for ‘have’, cf. Meiser 2003: 164, who quotes Osc. *prūfatted* ‘probavit’ as an equivalent of Lat. *“*probatum habuit/fecit”*). But the periphrastic construction developed differently in Gmc and Oscan. In PGmc, the participle was univerted with the auxiliary and the sequence **-đaⁿđ-* was reduced to **-đ-* by haplology, resulting in the loss of the participial ending. In Oscan, on the other hand, the auxiliary was lost and the participle in *-om* was reanalyzed as a finite verb form, i.e. as 1sg. This led to the development of a thematic active perfect based on the past participle.

The motive for the use of periphrasis with the verb for ‘make, do’ and the past participle can be seen in the lack of possibility for the secondary (“weak”) verbs to build the categories aorist and perfect. According to this view, the auxiliary verb was primarily used in its aorist and perfect forms, but after the periphrasis had developed the imperfect forms of the auxiliary could also be used. The periphrastic imperfect had the advantage over the simple (synthetic) imperfect in that it was morphologically more marked. This was a major factor contributing to the loss of the latter category. Under these circumstances, we can expect the Gmc weak preterite to reflect forms not only of the aorist and perfect but also of the imperfect of the auxiliary. In fact, this seems to be the case. According to the evidence of the Gmc languages (cf. Table 55.7), the indicative paradigm had the following “endings” in PGmc (on 3sg. PN **talgidai** [Nøvling] [for **talgiđē*], see Harðarson 2005: 226 fn. 64.):

	sg.	pl.
1	<i>*-đōⁿ</i>	<i>*-đume, -đēđume</i>
2	<i>*-đēz, -đōz</i>	<i>*-đude, -đēđude</i>
3	<i>*-đē(t), -đō(t)</i>	<i>*-đun(t), -đēđun(t)</i>

The endings **-đēz, *-đē(t), *-đume, *-đude, *-đun(t)* are most plausibly derived from the root aorist, and **-đēđume, *-đēđude, *-đēđun(t)* from the imperfect and perfect of the root **đē-/đō-* (PIE **d^heh₁-/*d^hoh₁-*). The endings **-đōⁿ, *-đōz, *-đō(t)*, on the other hand, seem to have arisen by analogy with the verb **đō-* ‘do’, esp. its singular preterite forms, cf. the South Rhine Franconian and Alemannic weak plural endings *-tōm, -tōt, -tōn* beside the regular OHG endings *-tum, -tut, -tun* (see above); the former developed under the influence of the singular endings **-tō, -tōs, *-tō (> -ta, -tōs, -ta)* as well as of the whole present paradigm of the verb *tōn (tuōn)*. The distribution of the singular forms **-đōⁿ, *-đēz, *-đē(t)* was surely influenced by the formerly coexistent simple imperfect of secondary *je/o*-verbs, whose singular had the endings **-jom, *-jes, *-jed* in pre-PGmc (cf. Lühr 1984: 48 f.) Hill (2004: 286–292) traces the endings of the weak preterite back to the unaugmented imperfect of the pre-PGmc verb for ‘do’, which he in turn derives from the subjunctive of the PIE root aorist **d^heh₁-*, i.e. **d^heh₁-e/o-*. This approach faces serious problems such as, e.g., the difference in the root structure between the indicative

singular of the present (PGmc **đōmi*, **đōsi/zi*, **đōpi/đi*) and the preterite (PGmc **-đōⁿ*, **-đēz*, **-đē(t)*); furthermore, manifold transformations of the alleged plural forms (**đōme*, **đēpe*, **đōnt*) have to be assumed.

After the periphrasis with the verb for ‘do, make’ and the past participle had been grammaticalized as a preterite of the verb from which the participle was formed it could be adopted by all verbs that lacked an inherited preterite other than the imperfect. Later, when the imperfect was lost, this was the only way for secondary verbs to express the past. Also the preterite-presents, which did not even have an imperfect tense, took advantage of this new verbal category (in forms like **aihtaⁿ đēt* ‘possessed’ and **mahtaⁿ đēt* ‘was able’, which were formed on the model of **waređaⁿ đēt* ‘protected’, the sequence **-taⁿđ-* was later reduced to **-t-* by haplology).

6.8. Non-finite forms of the verb

6.8.1. The infinitive

The Gmc infinitive continues a neuter verbal noun characterized by the PIE suffix **-ono-*. Thus, Go. *bairan* ‘bear, carry’ (<ai> = [ε]) comes from PGmc **beranaⁿ*, which in turn reflects PIE **b^hér-ono-m* ‘bearing, carrying’, cf. Skt *bharaṇam* ‘id.’ < **b^hér-eno-m*. Originally, nouns of this kind were not formed from the respective present stem. In fact, there were no infinitives of the present or the perfect as in Latin and Greek, neither did any middle infinitives exist. In PGmc, the formation of the infinitive was later linked with that of the present stem. Thus, the Go. infinitives *niman*, *hafjan* and *salbon* correspond to the presents *nima* ‘I take’, *hafja* ‘I raise, lift’, and *salbo* ‘I anoint’, respectively.

6.8.2. The present participle

In PIE, the present participle was formed with the suffix **-ent/nt-*. In thematic verbs, the zero-grade **-nt-* was combined with the *o*-grade of the thematic vowel, cf. Gk φέρων, gen. sg. φέροντος, from φέρω ‘carry’. PIE **-nt-* yielded PGmc **-nđ-* according to Verner’s Law. Synchronically, the Gmc present participle is derived from the respective infinitive by substituting the suffix **-nđ-* for **-na-*, cf. inf. **berana-* ‘carry’, **setjana-* ‘sit’, **salbōjana-* ‘anoint’, **đōna-* ‘give’, **wesana-* ‘be’ → ptc. **berand-*, **setjand-*, **salbōjand-*, **đōnđ-* (with analogical *ō* before resonant + consonant, where Osthoff’s Law [cf. 6.3.1, paragraph 4] should operate), **wesand-*.

6.8.3. The past participle

In PGmc, the strong verbs form their past participle with the suffix **-ena-* (> **-ina-*) or **-ana-*. The distribution of these variants cannot be determined with certainty. In Gothic and Old High German, the continuant of **-ana-* is regular, cf. Go. *numans*, OHG *ginoman* from *niman* and *neman*, respectively. But Gothic also shows fossilized relics of

Tab. 55.9: The paradigm of the strong verb **beranaⁿ* ‘carry’ in early PGmc

Active <i>Present</i>			
	Indicative	Optative	Imperative
Sg. 1	<i>*berō</i>	<i>*berajuⁿ</i>	
2	<i>*berēzil/-si</i>	<i>*beraiz</i>	<i>*bere</i>
3	<i>*berēdi/-þi</i>	<i>*berai(t)</i>	<i>*beradō̃(t)</i>
Du. 1	<i>*berōz</i>	<i>*beraiwā/-we</i>	
2	<i>*berataz</i>	<i>*beraitaz</i>	<i>*berataz</i>
Pl. 1	<i>*beramaz</i>	<i>*beraimā/-me</i>	
2	<i>*berēde/-þe</i>	<i>*beraiðe/-þe</i>	<i>*berēde/-þe</i>
3	<i>*berandī/-nþi</i>	<i>*berain(t)</i>	<i>*berandō̃(t)</i>
<i>Preterite</i>			
	Indicative	Optative	
Sg. 1	<i>*bara</i>	<i>*bārīⁿ</i>	
2	<i>*barþa/-ta</i>	<i>*bārīz</i>	
3	<i>*bare</i>	<i>*bārī(t)</i>	
Du. 1	<i>*bāruwe</i>	<i>*bārīwā/-we</i>	
2	<i>*bārutaz</i>	<i>*bārītaz</i>	
Pl. 1	<i>*bārume</i>	<i>*bārīmā/-me</i>	
2	<i>*bāruðe/-þe</i>	<i>*bārīðe/-þe</i>	
3	<i>*bārun(t)</i>	<i>*bārīn(t)</i>	
<i>Passive Present</i>			
	Indicative	Optative	
Sg. 1	<i>*berōi</i>	<i>*beraja</i>	
2	<i>*berazai</i>	<i>*beraiza</i>	
3	<i>*beradai</i>	<i>*beraiða</i>	
Pl. 1	<i>*beramedai</i>	<i>*beraimēda</i>	
2	<i>*beradwai</i>	<i>*beraiðwa</i>	
3	<i>*berandai</i>	<i>*berainða</i>	
<i>Infinitive *beranaⁿ</i>			
<i>Present participle *berand-</i>			
<i>Past participle *burana-</i>			

**-ena-*, cf. *fulgins* ‘hidden’ (beside the synchronic past participle *fulhans* of the verb *filhan* ‘hide, conceal, bury’) and the substantivized neuter *aigin* ‘property’ from **aigins* ‘own’ (cf. ON *eigenn*, OE *āzen*), which belongs to *aigan* ‘have’. Old Saxon, Old English and Old Frisian continue both **-ena-* and **-ana-*. In North Gmc, on the other hand, only

-in(a)- (< **-ena-*) is directly attested, cf. PN **slaginar** ‘killed’, **haitinar** ‘called’, OIce. *bitenn* ‘bitten’, *gefenn* ‘given’, etc. But from participles such as OIce. *brotenn* ‘broken’, *orþenn* ‘become’, and *borenn* ‘borne, carried’, which have undergone *a*-umlaut of the root, it can be concluded that NGmc once possessed the suffix variant **-ana-* beside *-ina-*, and that the latter was later generalized (cf. Harðarson 2001a: 69–73).

The Gmc suffix **-ena-/ana-* comes from PIE **-eno-/ono-*, which bears a similar relation to the formative **-no-* as **-eto-* to **-to-*, **-ero-* to **-ro-*, etc.

The past participle of weak verbs and preterite-presents is formed with the PIE suffix **-to-*, which, depending on phonetic context, gave Gmc **-ða-*, **-þa-* or **-ta-*, cf. Go. *nasips* (*-d-*) ‘saved’ (from *nasjan*), *salboþs* (*-d-*) ‘anointed’ (from *salbon*), *kunþs* (*-þ-*) ‘known’ (from *kunnan*) and *mahts* ‘possible’ (*ni maht ist* ‘is not possible’, from *magan* ‘have power, be able’).

Since the past participles in **-to-* had an oxytone accent in PIE, the PGmc suffix variant **-þa-* is irregular. It indicates that the root (and not the suffix) carried the accent. This root accent, the consequence of which can be observed in a few words such as Gmc **hlupa-* ‘famous’ (in OE *Hloþhere*, etc. vs. **hluða-* in Gmc-Lat. *Hlodericus*, *Chlodomeris*, etc.) and **kunþa-* ‘known’ (in Go. *kunþs*, OIce. *kunnr*, OE *cūð*, OHG *kund*, etc.), has been explained as adjectival oppositive accent. Accordingly, the accent of the participles in question was shifted due to their adjectival meaning (cf. Schaffner 2001: 298–301, where more examples are quoted). In the case of **kunþa-* ‘known’ the generalization of this variant (at the expense of **kundā-*) may have taken place in order to avoid the homonymy with **kundā-* ‘born’ (cf. Lühr 1984: 68 n. 107).

Acknowledgment

I would like to thank Sergio Neri for valuable comments and suggestions on a previous draft of this paper. I am also indebted to one of the editors, Jared S. Klein, for his critical remarks, which led to improvements in the text. Finally, I am grateful to Jón Símon Markússon for having corrected my English.

7. References

- Bammesberger, Alfred
 1990 *Die Morphologie des urgermanischen Nomens*. Heidelberg: Winter.
- Bammesberger, Alfred
 2001 *-um* (> *-on*) as marker of the instrumental singular in Old English and Old Frisian. *Neophilologus* 85: 287–290.
- Barrack, Charles Michael
 1998 *Sievers' Law in Germanic*. (Berkeley Insights in Linguistics and Semiotics 22). New York: Lang.
- Benediktsson, Hreinn
 1968 On the Inflection of the *n*-Stems in Indo-European. *Norsk Tidsskrift for Sprogvidenskap* 22: 7–31.
- Bethge, Richard
 1900 Konjugation des Urgermanischen. In: Ferdinand Dieter (ed.), *Laut- und Formenlehre der altgermanischen Dialekte*. Zum Gebrauch für Studierende dargestellt von Richard

- Bethge, Otto Bremer, Ferdinand Dieter, Friedrich Hartmann, und Wolfgang Schlüter.
Leipzig: Reisland, 345–391.
- Boley, Jacqueline
1984 *The Hittite hark-construction*. (Innsbrucker Beiträge zur Sprachwissenschaft 44). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Brugmann, Karl
1904 *Die Demonstrativpronomina der indogermanischen Sprachen. Eine bedeutungsgeschichtliche Untersuchung*. Leipzig: Teubner.
- Brugmann, Karl
1921 Das gotische *-ada*-Passivum. *Indogermanische Forschungen* 39: 26–62.
- Brunner, Karl
1965 *Altenglische Grammatik*. Nach der Angelsächsischen Grammatik von Eduard Sievers. Dritte, neubearbeitete Auflage. Tübingen: Niemeyer.
- Cowgill, Warren
1959 The Inflection of the Germanic *ō*-Presents. *Language* 35: 1–15.
- Cowgill, Warren
1975 The Insular Celtic Conjunct and Absolute Verbal Endings. In: Helmut Rix (ed.), *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft Regensburg, 9.–14. September 1973*. Wiesbaden: Reichert, 40–70.
- Cowgill, Warren
1983 On the Prehistory of Celtic Passive and Deponent Inflection. *Ériu* 34: 73–111.
- Cowgill, Warren
2006 The *z*-cases of Germanic Pronouns and Strong Adjectives. In: Jared S. Klein (ed.), *The Collected Writings of Warren Cowgill*. Edited with an Introduction by Jared S. Klein. With Contributions by Other Former Colleagues and Students. Ann Arbor: Beech Stave 519–534.
- Darms, Georges
1978 *Schwäher und Schwager, Hahn und Huhn. Die Vřddhi-Ableitung im Germanischen*. Munich: Kitzinger.
- Delamarre, Xavier
2003 *Dictionnaire de la langue gauloise. Une approche linguistique du vieux-celtique continental*. 2^e édition revue et augmentée. Paris: Errance.
- Delbrück, Berthold
1909 Das schwache Adjektivum und der Artikel im Germanischen. *Indogermanische Forschungen* 26: 187–199.
- Di Giovine, Paolo
1996 *Studio sul perfetto indoeuropeo*. Parte II: *La posizione del perfetto all'interno del sistema verbale indoeuropeo*. Rome: "Il Calamo".
- Eichner, Heiner
1992 Anatolian. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: Mouton de Gruyter, 29–96.
- Ernout, Alfred and François Thomas
1972 *Syntaxe latine*. 2^e édition, 5^e tirage revu et corrigé. Paris: Klincksieck.
- Eska, Joseph F.
1992 The Third Person Imperative Desinences in Old Irish. *Historische Sprachforschung* 105: 265–270.
- Falk, Hjalmar and Alf Torp
1909 *Wortschatz der germanischen Spracheinheit*. Unter Mitwirkung von Hjalmar Falk gänzlich umgearbeitet von Alf Torp. (*Vergleichendes Wörterbuch der indogermanischen Sprachen* von August Fick. 4. Auflage. 3. Teil). Göttingen: Vandenhoeck & Ruprecht.

Feist, Sigmund

- 1939 *Vergleichendes Wörterbuch der gotischen Sprache: mit Einschluss des Krimgotischen und sonstiger zerstreuter Überreste des Gotischen*. 3. neubearbeitete und vermehrte Auflage. Leiden: Brill.

Gorbachov, Yaroslav Vladimirovich

- 2007 *Indo-European Origins of the Nasal Inchoative Class in Germanic, Baltic and Slavic*. Unpublished Dissertation. Harvard University.

Griepentrog, Wolfgang

- 1995 *Die Wurzelnomina des Germanischen und ihre Vorgeschichte*. (Innsbrucker Beiträge zur Sprachwissenschaft 82). Innsbruck: Institut für Sprachwissenschaft der Universität.

Harðarson, Jón Axel

- 1989 Die *ōn*-Feminina des Germanischen und der Gen. Plur. anord. *kvinna* / *kvenna*. *Acta Linguistica Hafniensia* 21: 79–93.

Harðarson, Jón Axel

- 1993 *Studien zum urindogermanischen Wurzelaorist und dessen Vertretung im Indoiranischen und Griechischen*. (Innsbrucker Beiträge zur Sprachwissenschaft 74). Innsbruck: Institut für Sprachwissenschaft der Universität.

Harðarson, Jón Axel

- 1998 Mit dem Suffix **-eh₁-* bzw. **(e)h₁-je/o-* gebildete Verbalstämme im Indogermanischen. In: Wolfgang Meid (ed.), *Sprache und Kultur der Indogermanen. Akten der 10. Fachtagung der Indogermanischen Gesellschaft, Innsbruck, 22.–28. September 1996*. (Innsbrucker Beiträge zur Sprachwissenschaft 93). Innsbruck: Institut für Sprachwissenschaft der Universität, 323–339.

Harðarson, Jón Axel

- 2001a *Das Präteritum der schwachen Verba auf -ýia im Altisländischen und verwandte Probleme der altnordischen und germanischen Sprachwissenschaft*. (Innsbrucker Beiträge zur Sprachwissenschaft 101). Innsbruck: Institut für Sprachwissenschaft der Universität.

Harðarson, Jón Axel

- 2001b Review of Paolo di Giovine. *Studio sul perfetto indoeuropeo*. Parts I–III. *Kratylos* 46: 36–44.

Harðarson, Jón Axel

- 2005 Der geschlechtige Nom. Sg. und der neutrale Nom.-Akk. Pl. der *n*-Stämme im Urindogermanischen und Germanischen. In: Meiser and Hackstein (eds.), 215–236.

Harðarson, Jón Axel

- 2009 Review of Antje Casaretto. *Nominale Wortbildung der gotischen Sprache. Die Derivation der Substantive*. *Kratylos* 54: 13–24.

Harðarson, Jón Axel

- 2011 The 2nd Line of the Duenos Inscription. In: Giovanna Rocca (ed.), *Atti del Convegno Internazionale Le lingue dell'Italia antica: Iscrizioni, testi, grammatica / Die Sprachen Altitaliens: Inschriften, Texte, Grammatik. In memoriam Helmut Rix (1926–2004) 7–8 marzo 2011, Libera Università di Lingue e Comunicazione, IULM, Milano. Ἀλεξάνδρεια / Alessandria* 5. 153–163.

Harðarson, Jón Axel

- 2016 Bemerkungen zur Singularflexion der indogermanischen *o*-Stämme. In: Sergio Neri, Roland Schuhmann, and Susanne Zeilfelder (eds.), *dat ih dir it nu bi huldi gibu. Linguistische, germanistische und indogermanistische Studien Rosemarie Lühr gewidmet*. Wiesbaden: Reichert, 151–160.

Heidermanns, Frank

- 1993 *Etymologisches Wörterbuch der germanischen Primäradjektive*. Berlin: Mouton de Gruyter.

- Hellquist, Elof
1957 *Svensk etymologisk ordbok* [Swedish etymological dictionary]. Tredje upplagan. Lund: Gleerup.
- Heusler, Andreas
1932 *Altisländisches Elementarbuch*. 3. Auflage. Heidelberg: Winter.
- Hill, Eugen
2004 Das germanische Verb für 'tun' und die Ausgänge des germanischen schwachen Präteritums. *Sprachwissenschaft* 29: 257–303.
- Hill, Eugen
2009 Das starke Präteritum der Klasse VII in den nord- und westgermanischen Sprachen. Parts I and II. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 6: 49–123, 173–213.
- Hill, Eugen
2012 Die Entwicklung von *u vor unsilbischem *i in den indogermanischen Sprachen Nord- und Mitteleuropas: Die Stammsuppletion bei u-Adjektiven und das Präsens von 'sein'. *North-Western European Language Evolution* 64/65: 5–36.
- Hilmarsson, Jörundur
1989 *The Dual Forms of Nouns and Pronouns in Tocharian*. (Tocharian and Indo-European Studies Supplementary Series 1). Reykjavík: Málvísindastofnun Háskóla Íslands.
- Hilmarsson, Jörundur
1991 On \bar{e}^2 in Germanic. *Acta Linguistica Hafniensia* 23: 33–47.
- Hirt, Hermann
1931 *Handbuch des Urgermanischen*. Teil I: Laut- und Akzentlehre. Heidelberg: Winter.
- Hirt, Hermann
1932 *Handbuch des Urgermanischen*. Teil II: Stammbildungs- und Flexionslehre. Heidelberg: Winter.
- Hoffmann, Karl
1955 Ein grundsprachliches Possessivsuffix. *Münchener Studien zur Sprachwissenschaft* 6: 35–40. [= Karl Hoffmann. 1976. *Aufsätze zur Indoiranistik*. Band 2. Wiesbaden: Reichert, 378–383.]
- Hogg, Richard M. and R. D. Fulk
2011 *A Grammar of Old English*. Vol. 2: *Morphology*. Wiley-Blackwell.
- Hollifield, Patrick Henry
1980 The Phonological Development of Final Syllables in Germanic. *Die Sprache* 26: 19–53 (Part 1), 145–178 (Part 2).
- Jasanoff, Jay
2007 From Reduplication to Ablaut: The Class VII Strong Verbs of Northwest Germanic. *Historische Sprachforschung* 120: 241–284.
- Jasanoff, Jay
2012 Long-vowel preterites in Indo-European. In: Melchert (ed.), 127–135.
- Jóhannesson, Alexander
1923–1924 *Íslenzk tunga í fornöld* [The Icelandic language in ancient times]. Reykjavík: Bókaverzlun Ársæls Árnasonar.
- Jónsson, Finnur
1921 *Norsk-islandske kultur- og sprogforhold i 9. og 10. århundrede* [Norwegian-Icelandic cultural and linguistic relations in the 9th and 10th centuries]. Copenhagen: Høst.
- Klingenschmitt, Gert
1982 *Das altarmenische Verbum*. Wiesbaden: Reichert.
- Klingenschmitt, Gert
1987 Erbe und Neuerung beim germanischen Demonstrativpronomen. In: Rolf Bergmann, Heinrich Tiefenbach, and Lothar Voetz (eds.), *Althochdeutsch*. Heidelberg: Winter, 169–189.

- Klingenschmitt, Gert
1992 Die lateinische Nominalflexion. In: Panagl and Krisch (eds.), 89–135.
- Klingenschmitt, Gert
1994 Die Verwandtschaftsverhältnisse der indogermanischen Sprachen. In: Jens E. Rasmussen (ed.), *In honorem Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Wiesbaden: Reichert, 235–251.
- Krause, Wolfgang
1968 *Handbuch des Gotischen*. Dritte, neubearbeitete Auflage. Munich: Beck.
- Krogh, Steffen
1996 *Die Stellung des Altsächsischen im Rahmen der germanischen Sprachen*. (Studien zum Althochdeutschen 29). Göttingen: Vandenhoeck & Ruprecht.
- Kroonen, Guus
2011 *The Proto-Germanic n-Stems. A Study in Diachronic Morphophonology*. (Leiden Studies in Indo-European 18). Amsterdam: Rodopi.
- Kroonen, Guus
2012 Reflections on the *o/zero*-Ablaut in the Germanic Iterative Verbs. In: Melchert (ed.), 191–200.
- Kümmel, Martin
2000 *Das Perfekt im Indoiranischen. Eine Untersuchung der Form und Funktion einer ererbten Kategorie des Verbums und ihrer Weiterentwicklung in den altindoiranischen Sprachen*. Wiesbaden: Reichert.
- Kuryłowicz, Jerzy
1964 *The Inflectional Categories of Indo-European*. Heidelberg: Winter.
- Lipp, Reiner
2009 *Die indogermanischen und einzelsprachlichen Palatale im Indoiranischen*. Band I: *Neurekonstruktion, Nuristan-Sprachen, Genese der indoarischen Retroflexe, Indoarisch von Mitanni*. Band II: *Thorn-Problem, indoiranische Laryngalvokalisation*. Heidelberg: Winter.
- LIV² see Rix 2001.
- Lühr, Rosemarie
1978 Die Kontinuante der urindogermanischen Medialflexion im Germanischen. *Münchener Studien zur Sprachwissenschaft* 37: 109–120.
- Lühr, Rosemarie
1984 Reste der athematischen Konjugation in den germanischen Sprachen: zu ‘sein’ und ‘tun’. In: Jürgen Untermann and Bela Brogyanyi (eds.), *Das Germanische und die Rekonstruktion der indogermanischen Grundsprache. Akten des Freiburger Kolloquiums der Indogermanischen Gesellschaft, Freiburg, 26.–27. Februar 1981*. Amsterdam: Benjamins, 25–90.
- Magnússon, Ásgeir Blöndal
1989 *Íslensk orðsifjabók* [Icelandic etymological dictionary]. Reykjavík: Orðabók Háskólans.
- Mayrhofer, Manfred
1986 *Lautlehre (Segmentale Phonologie des Indogermanischen)*. *Indogermanische Grammatik*. Ed. by Manfred Mayrhofer. Band I. 2. Halbband. Heidelberg: Winter.
- McCone, Kim R.
1986 From Indo-European to Old Irish: Conservation and Innovation in the Verbal System. In: D. Ellis Evans, John G. Griffith, and Edward M. Jope (eds.), *Proceedings of the Seventh International Congress of Celtic Studies, held at Oxford, from 10th to 15th July, 1983*. Oxford: Oxbow, 222–266.
- Meiser, Gerhard
1993 Zur Funktion des Nasalpräsens im Urindogermanischen. In: Gerhard Meiser (ed.), *Indogermanica et Italica. Festschrift für Helmut Rix zum 65. Geburtstag* (Innsbrucker Beiträge zur Sprachwissenschaft 72). Innsbruck: Institut für Sprachwissenschaft der Universität, 280–313.

- Meiser, Gerhard
2003 *Veni, vidi, vici. Die Vorgeschichte des lateinischen Perfektsystems*. Munich: Beck.
- Meiser, Gerhard and Olav Hackstein (eds.)
2005 *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle an der Saale*. Wiesbaden: Reichert.
- Melchert, H. Craig (ed.)
2012 *The Indo-European Verb. Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles 13–15 September 2010*. Wiesbaden: Reichert.
- Mottausch, Karl-Heinz
2001 Gotisch *-(u)h* und ein vergessenes Lautgesetz. *Northwestern European Language Evolution* 38: 37–47.
- Nedoma, Robert
2005 Urnordisch *-a* im Nominativ Singularis der maskulinen *n*-Stämme. *Northwestern European Language Evolution* 46/47: 155–191. (= *Papers on Scandinavian and Germanic Language and Culture. Published in Honour of Michael Barnes on his Sixty-Fifth Birthday 28 June 2005*.)
- Neri, Sergio
2003 *I sostantivi in -u del gotico. Morfologia e Preistoria*. (Innsbrucker Beiträge zur Sprachwissenschaft 108). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Neri, Sergio
2009 Review of Ringe 2006. *Kratylos* 54: 1–13.
- Neri, Sergio
2011 Wetter. Etymologie und Lautgesetz. Ph.D. Dissertation. Friedrich-Schiller-Universität Jena.
- Noreen, Adolf
1923 *Altnordische Grammatik I. Altisländische und altnorwegische Grammatik (Laut- und Flexionslehre) unter Berücksichtigung des Urnordischen*. Vierte vollständig umgearbeitete Auflage. Halle: Niemeyer.
- Osthoff, Herrmann
1876 *Forschungen im Gebiete der indogermanischen nominalen Stammbildung*. Zweiter Teil: *Zur Geschichte des deutschen schwachen Adjectivum*. Jena: Hermann Costenoble.
- Osthoff, Herrmann
1881 Pronomina, Adverbia, Praepositionen, partikelartige Wörter, Indeclinabilia überhaupt. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen* 4. Leipzig: Hirzel, 222–276.
- Panagl, Oswald and Thomas Krisch (eds.)
1992 *Latein und Indogermanisch. Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg, 23.–26. September 1986*. (Innsbrucker Beiträge zur Sprachwissenschaft 64). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Pinault, Georges-Jean
2008 *Chrestomathie tokharienne. Textes et grammaire*. (Collection linguistique publiée par la Société de linguistique de Paris 95). Leuven: Peeters.
- Prokosch, Eduard
1939 *A Comparative Germanic Grammar*. Baltimore: Linguistic Society of America.
- Rasmussen, Jens Elmegård
1999a Two Phonological Issues in Germanic. In: *Selected Papers on Indo-European Linguistics. With a Section on Comparative Eskimo Linguistics*. Part 1. Copenhagen: Museum of Tusulanum Press, 82–99. [First published in *Acta Linguistica Hafniensia* 18, 1983: 201–219.]
- Rasmussen, Jens Elmegård
1999b On the Origin of the Germanic Weak Preterite. In: *Selected Papers on Indo-European Linguistics*. Part 2. Copenhagen: Museum of Tusulanum Press, 597–603. [First published in *Copenhagen Working Papers in Linguistics* 4, 1996: 161–168.]

- Rau, Jeremy
 2009 *Indo-European Nominal Morphology: The Decads and the Caland System*. (Innsbrucker Beiträge zur Sprachwissenschaft 132). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Rau, Jeremy
 2010 The Derivational History of the Greek Stems in -*ád-*. *Münchener Studien zur Sprachwissenschaft* 64[2004]: 137–173.
- Ringe, Don
 1996 *On the chronology of sound changes in Tocharian*. Vol. 1. *From PIE to Proto-Tocharian* (American Oriental Series 80). New Haven: American Oriental Society.
- Ringe, Don
 2006 *From Proto-Indo-European to Proto-Germanic* (A Linguistic History of English 1). Oxford: Oxford University Press.
- Rix, Helmut
 1986 *Zur Entstehung des urindogermanischen Modusystems*. (Innsbrucker Beiträge zur Sprachwissenschaft. Vorträge und Kleinere Schriften 36). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Rix, Helmut
 1992 Zur Entstehung des lateinischen Perfektparadigmas. In: Panagl and Krisch (eds.), 221–240.
- Rix, Helmut (ed.)
 2001 *Lexikon der indogermanischen Verben. Die Wurzeln und Ihre Primärstambildungen*. Unter der Leitung von Helmut Rix und der Mitarbeit vieler anderer bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp und Brigitte Schirmer. Zweite erweiterte und verbesserte Auflage bearbeitet von Martin Kümmel und Helmut Rix. Wiesbaden: Reichert.
- Rix, Helmut
 2003 Towards a Reconstruction of Proto-Italic: the Verbal System. In: Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe, and Miriam Robbins Dexter (eds.), *Proceedings of the Fourteenth Annual UCLA Indo-European Conference, Los Angeles, Nov. 8–9, 2002*. (Journal of Indo-European Studies Monograph No. 47). Washington, DC: Institute for the Study of Man, 1–24.
- Schaffner, Stefan
 2001 *Das Vernersche Gesetz und der innerparadigmatische grammatische Wechsel des Urgermanischen im Nominalbereich*. (Innsbrucker Beiträge zur Sprachwissenschaft 103). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Schaffner, Stefan
 2003 Der Beitrag des Germanischen zur Rekonstruktion der urindogermanischen Akzent- und Ablautklassen. In: Eva Tichy, Dagmar S. Wodtke, and Britta Irlinger (eds.), *Indogermanisches Nomen. Derivation, Flexion und Ablaut. Akten der Arbeitstagung der Indogermanischen Gesellschaft, Freiburg, 19. bis 22. September 2001*. Bremen: Hempen, 203–218.
- Schaffner, Stefan
 2010 Lateinisch *radius* und Verwandtes. *Glotta* 86: 109–144.
- Scheungraber, Corinna
 2011 Review of Kroonen 2011. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 8: 89–115.
- Scheungraber, Corinna
 2012 Nasal Suffix Verbs in Germanic and Kluge's Law. In: Melchert (ed.), 295–304.
- Schindler, Jochem
 1972 L'apophonie des noms-racines indo-européens. *Bulletin de la Société de linguistique de Paris* 67: 31–38.

- Schindler, Jochem
1973 Bemerkungen zur Herkunft der indogermanischen Diphthongstämme und zu den Eigentümlichkeiten ihrer Kasusformen. *Die Sprache* 19: 148–157.
- Schmidt, Johannes
1889 *Die Pluralbildungen der indogermanischen Neutra*. Weimar: Hermann Böhlau.
- Schulze, Wilhelm
1904 *Zur Geschichte lateinischer Eigennamen*. Berlin: Weidmann.
- Schumacher, Stefan
2005 ‘Langvokalisches Perfekta’ in indogermanischen Einzelsprachen und ihr grundsprachlicher Hintergrund. In: Meiser and Hackstein (eds.), 591–626.
- Streitberg, Wilhelm
1896 *Urgermanische Grammatik*. Heidelberg: Winter.
- Svennung, Josef
1935 *Untersuchungen zu Palladius und zur lateinischen Fach- und Volkssprache*. Uppsala: Almqvist & Wiksell.
- Szemerényi, Oswald
1960 *Studies in the Indo-European system of numerals*. Heidelberg: Winter.
- Tanaka, Toshiya
2011 *A Morphological Conflation Approach to the Historical Development of Preterite-Present Verbs: Old English, Proto-Germanic, and PIE*. (Kyushu University. The Faculty of Languages and Cultures Library 2). Fukuoka: Hana-Shoin.
- Vine, Brent
1985 On the Germanic Type *lūkan*. *Zeitschrift für Vergleichende Sprachforschung* 98: 60–81.
- Weiss, Michael
2009 *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave.

Jón Axel Harðarson, Reykjavík (Iceland)

56. The syntax of Germanic

- | | |
|--|--------------------|
| 1. Overview | 5. Adverbials |
| 2. Word classes | 6. Word order |
| 3. Nominal morphosyntax and adpositional phrases | 7. Sentence syntax |
| 4. Verbal morphosyntax and periphrastic formations | 8. Abbreviations |
| | 9. References |

1. Overview

The syntax of Germanic is mainly reconstructed on the basis of Gothic (= GOTH), Old Norse (ON), and the West Germanic languages Old High German (= OHG), Old Saxon (= OS), Old English (= OE) and Old Frisian (= OFr). But the well known problem of Germanic tradition is that the texts often are translations of foreign originals. In GOTH, for example, bishop Wulfila’s 4th century translation of the Bible has a Greek basis, and

- Schindler, Jochem
1973 Bemerkungen zur Herkunft der indogermanischen Diphthongstämme und zu den Eigentümlichkeiten ihrer Kasusformen. *Die Sprache* 19: 148–157.
- Schmidt, Johannes
1889 *Die Pluralbildungen der indogermanischen Neutra*. Weimar: Hermann Böhlau.
- Schulze, Wilhelm
1904 *Zur Geschichte lateinischer Eigennamen*. Berlin: Weidmann.
- Schumacher, Stefan
2005 ‘Langvokalisches Perfekta’ in indogermanischen Einzelsprachen und ihr grundsprachlicher Hintergrund. In: Meiser and Hackstein (eds.), 591–626.
- Streitberg, Wilhelm
1896 *Urgermanische Grammatik*. Heidelberg: Winter.
- Svennung, Josef
1935 *Untersuchungen zu Palladius und zur lateinischen Fach- und Volkssprache*. Uppsala: Almqvist & Wiksell.
- Szemerényi, Oswald
1960 *Studies in the Indo-European system of numerals*. Heidelberg: Winter.
- Tanaka, Toshiya
2011 *A Morphological Conflation Approach to the Historical Development of Preterite-Present Verbs: Old English, Proto-Germanic, and PIE*. (Kyushu University. The Faculty of Languages and Cultures Library 2). Fukuoka: Hana-Shoin.
- Vine, Brent
1985 On the Germanic Type *lūkan*. *Zeitschrift für Vergleichende Sprachforschung* 98: 60–81.
- Weiss, Michael
2009 *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave.

Jón Axel Harðarson, Reykjavík (Iceland)

56. The syntax of Germanic

- | | |
|--|--------------------|
| 1. Overview | 5. Adverbials |
| 2. Word classes | 6. Word order |
| 3. Nominal morphosyntax and adpositional phrases | 7. Sentence syntax |
| 4. Verbal morphosyntax and periphrastic formations | 8. Abbreviations |
| | 9. References |

1. Overview

The syntax of Germanic is mainly reconstructed on the basis of Gothic (= GOTH), Old Norse (ON), and the West Germanic languages Old High German (= OHG), Old Saxon (= OS), Old English (= OE) and Old Frisian (= OFr). But the well known problem of Germanic tradition is that the texts often are translations of foreign originals. In GOTH, for example, bishop Wulfila’s 4th century translation of the Bible has a Greek basis, and

in OHG the vast majority of texts are translations from a Latin original. Only a few autonomous texts are available. Therefore, to discover genuine Germanic syntactic patterns one has to scrutinize the autonomous versions. E.g., a comparison of the OHG Tatian translation with the Latin version shows that OHG has subject drop and not object drop as in Latin; likewise, absolute participle constructions are not genuine, as the syntactic structure of Germanic suggests that those constructions are sometimes provided with a prepositional head (Lühr 2005a). Word order in particular is a domain where much research has been done during the past decade (Axel 2005; Frascarelli and Hinterhölzl 2007; Petrova and Solf 2008; Petrova 2009), including word order studies of the spoken language (Lühr 2005b). But smaller units have aroused interest, too. Thus, recent studies have examined the nominal phrase, the position of modifier phrases, and the development of the article (Demske 2001; Lühr 1991, 2000b, 2002a, 2002b, 2007c). But other phenomena are nearly unexplored, so that one still has to rely on Behaghel's German syntax (1923–1928).

2. Word classes

The Germanic word classes are inherited from Indo-European: inflectable word classes are nouns, pronouns, adjectives, and verbs; and non-inflectable ones are adverbs, prepositions, conjunctions, particles, and interjections. But the non-inflectable word classes show some peculiarities; e.g. inflected interjections in Goth: *hiri*, dual *hirjats*, plural *hirjiþ* 'here!' or words with the meaning 'or' as sentence adverbials: OHG *ôdo*, *odowân*, *odowar*, *odowila* (Lühr 1996). Furthermore, the Germanic languages have special interrogative particles. There are *nōnne*- and *num*-expressions as in Latin: GOTH *ni*, OHG *ne* ..., *na* or GOTH *þau*, *nibai*, OHG *nū*, respectively; and paralleling Latin *quam* or *an* one finds corresponding words in alternative questions, such as GOTH *þe*, *þau* and particles which are used in indirect assertions in the function of rhetorical questions such as the OHG affirmative particles *thoh*, *thanne*, *ja*, *eno*, *inno*, *nu* (Lühr 1997d). If these particles appear in the leftmost position in a sentence like GOTH *ibai*, they are important for sentence structure, for clitic left-dislocation involves a sentence-external position (Ferraresi 2005; Axel 2005). Sometimes the development of an element into a focus particle can be observed: OHG O 3,23,31 f. *sie fārent thines fērehes / mit **selb** stéinonne* 'they are lying in wait for your life (i.e. to kill you) **even** with stoning'; O 4,13,23 f. *Mit thír bin garo ... / in kárkari zi fāranne joh tóthes **ouh** zi kóronne* 'I am prepared to go with you ... to prison and **even** to choose death'; O 2,4,44 f. *thoh bát er nan **zi nóte** thie stéina duan zi bróte* 'yet he ordered him **urgently** to make the stones into bread' (Lühr 2010a). The old connection between prepositions and adverbs is still clearly seen in GOTH: Mc 8,6 *jah nimands þans sibun hlaibans ... atgaf siponjam seinaim, ei atlagide-deina **faur*** (adv.) 'and taking the seven loaves ... he gave (them) to his disciples to set **forth**'; *jah atlagidedun **faur*** (prep.) *þo managein* 'and they did set (them) **before** the people'. Another archaic feature of GOTH prepositions is a stative use instead of a directional one; cf. with dative continuing an old locative: L 2, 19 *Ip Maria alla gafastaida þo waurda, þagkjandei in hairtin seinamma* (dat.) 'But Mary held fast to all those words, pondering (them) within her heart' (cf. Lühr 2011b). Conjunctions often arose from adverbial expressions as well. Sometimes adverb and conjunction cannot be distin-

guished; cf. GOTH *unte*, OHG *wanta* ‘for, because’ (Krause 1968) and correlatives like OHG *thoh...thoh, so...so* (Lühr 1998b) (cf. 7). An inheritance from Indo-European is also sentence negation. The negation particle *ni* cliticizes to the verb. In GOTH the interrogative particle *-u* is located between *ni* and the verb: Mc 11,17 *ni-u gameliþ ist?* ‘Is it not written?’. A special problem is the interplay of negation and indefinites (Jäger 2005).

3. Nominal morphosyntax and adpositional phrases

Morphosyntax constitutes the set of rules that govern linguistic units whose properties are definable by both morphological and syntactic criteria. In the nominal system of Germanic, the grammatical categories number, gender, and case are of Indo-European origin, as are the three genders masculine, feminine, and neuter. But in comparison with the Indo-European nominal system the category number is not fully preserved. While singular and plural survive in their old functions, traces of the dual are found only in the pronominal system: GOTH *wit, ugkis*, ON *vit, okkar, ok(k)r*, OE *wit, uncer, unc(it)* ‘we both’, GOTH *igqis, igqara*, ON *it, ykkar, yk(k)r*, OE *git, incer, inc* ‘you both’, GOTH *igqar* ‘your’. Remnants of dual pronouns still exist in Bavarian and in Frisian dialects. Also the Indo-European case system was simplified. The nominative, accusative, and genitive cases continued their old forms and functions (for the genitive cf. the *genitivus partitivus*: GOTH J 12,9 *manageins filu Iudaie* ‘large crowds of Jews’, ON *þrír tiger manna* ‘three decades (i.e. thirty) of men’, the *genitivus qualitatis*: ON *mikels háttar maþr* ‘a man of great importance’, in *kenningar* such as *ógnar girðibúð* ‘protection tent of battle’ for ‘shield’ cf. Lühr 2006b). The partitive genitive is documented also with negation: L 1,7 *ni was im barne* ‘οὐκ ἦν αὐτοῖς τέκνον’ ‘they had no child’. The vocative only appears in GOTH (*skalk, þiudan, láisari* ‘teacher’, *þu leiki* ‘you physician’, *juggalauð* ‘young man’, *sunau, magau, sunu, dauþu, talzjand* ‘teacher’, *frijond, frauþinond* ‘ruler’ and in foreign names); in the other Germanic languages the vocative coincided with the nominative. Moreover, in Germanic generally the dative, instrumental, locative, and partially the ablative, were combined into one case, which is called dative from the Germanic point of view. The genitive took over some functions of the ablative insofar as this case did not merge with the dative. An independent instrumental form is documented in older OHG nouns and pronouns (HI *suertu hauwan* ‘to strike with a sword’, *diu, desiu, disiu, hwiu*), sporadically also locative forms occur (e.g. OHG *dorf, holz, hūs*). Inherited grammatical relations indicating the syntactic functions of phrasal categories in a sentence are: the nominative for the subject; it usually functions as the actor or agent and indicates the topic of the sentence, which, in an unmarked sentence, is placed in sentence-initial position. As in Indo-European, the accusative has two different functions: on the one hand, it indicates the direct object in the function of a patient or undergoer in the case of transitive verbs (i.e. accusative object), also with impersonal verbs, as in GOTH J 6,35 *þana gaggandan du mis ni huggreiþ, jah þana galaubjandan du mis ni þaurseiþ hvanhun* ‘the one coming (acc.) to me shall not hunger; and the one believing (acc.) in me shall never thirst’. This type of subjectless construction appears very often in ON (e.g. *skorter þik eige metnoþ* ‘you do not lack reputation’, ‘dir fehlt es nicht an Ansehen’) (Barðal 2004); and with a variety of double accusatives

(OHG O 4,2,27 *odo inan thie ármuati wiht irbármēti* ‘or [that] he took any pity on the poor’, Kelle: ‘Weil ihm vielleicht der Armen Los so überaus zu Herzen gieng’; O 1,4,45 *thaz er gigárawe thie liuti wírdige* ‘that he should make the people worthy’, ON *ek settak þik mikenn mann* ‘I made you into a great man’), the accusative-cum-infinitive construction (OHG O 1,25,15 *then fáter hōrt er spréchan* ‘he heard the Father speak’, ON *hugþo þeir þat vera konungs men* ‘they thought that they were men of the king’), and the *figura etymologica* (ON *sofa suefn* ‘to sleep a sleep’). On the other hand, the accusative expresses that the verbal action possesses an orientation in space and time (cf. 5). Like the accusative, the genitive is used as an object case (as with GOTH *gamunan* ‘to remember’, *gamaudjan* ‘to remember’, ON *minnask* ‘to recall something’, ON *leíta* ‘to seek’). The mixed case dative combines the old personal dative as recipient or benefactor (dative of interest) with instrumental (OHG O 5,20,63 *Hánton joh ouh óugon bigínnent sie nan scóuwon* ‘with their hands and also their eyes they began to look upon him’), ablative, and locative functions (OHG O H 30 *sinen werkon er io kléib* ‘he persisted continuously in his works’). The most often documented relations between verb and dative are motion (OHG O 1,8,19 *ther éngil imo náhta* ‘the angel approached him’), speech (OHG O L 12 *so ih thir zéllu* ‘as I tell you’, showing (OHG O 5,20,3 f. *selbo ... júngoron sinen zéinta, / ... wio ...* ‘he himself ... showed his disciples ... how ...’), help (OHG O L 24 *druhtin hálf imo sár* ‘the Lord helped him immediately’), receiving and taking away (OHG *bringan, beran, geban, geltan*), psychological disposition (OHG *irbelgan, frawōn*), and connections with copular verbs (OHG O 5,7,37 *wárd mir wé mit minnu* ‘pain has come to me from love’ (Schrodt 2004). Above all, the frequency of datives with verbs which in GOTH govern the accusative is a striking attribute of ON (cf. *róa bate* ‘to row a boat’). For the use of the anaphoric personal pronoun instead of the reflexive dative in Ingeveonic cf. Lühr (2006a).

Adpositional phrases with an adposition as head occur as prepositional objects, mostly replacing a dative or genitive object (GOTH K 11,23 *unte ik andnam at frauin þatei jah anafalh izwis* ‘... παρέλαβον ἀπὸ τοῦ κυρίου ...’ ‘and I have received from the Lord that which also I handed over to you’ with denotation of location instead of direction as in Greek [cf. 2]; ON *verða at ógæfu* ‘to have misfortune’, OHG O 4,21,9 *Pilátus wolta sliumo sár fon imo néman tho then wán* ‘Pilate wanted as quickly as possible then to remove the suspicion’, O 4,21,26 *frageta ávur noti bi sinaz héroti* ‘but he asked urgently about his dominion’) and as adverbials. Two subtypes of adverbials are distinguished: optional adverbials (i.e. adjuncts, which indicate time, place, cause, manner, etc. cf. 5), and obligatory adverbials (i.e. complements) (OHG O 1,19,5 *in Aegýpto wis thu sár* ‘Do thou remain above all in Egypt’, ON *búa at Hofi* ‘to live at Hof’, GOTH M 5,25 *in karkara galagjaza* ‘and you will be cast into prison’).

The development of the article in the nominal phrase is characteristic of the Germanic languages. In GOTH the article is missing in expressions with identifiable denotation, i.e. proper names, semantic definites like *gub* vs. Gk. ὁ θεός, phrases with relational nouns like L 4,41 *sunus gudis* ‘ὁ υἱὸς τοῦ θεοῦ’, or with denominations of parts of the body, such as Mc 6,25 *haubiþ Iohannis þis daupjandins* ‘τὴν κεφαλὴν Ἰωάννου τοῦ βαπτιστοῦ’, ‘the head of John the Baptist’, K 12,21 *niþþan mag augo qiþan du handau* ‘οὐ δύναται δὲ ὁ ὀφθαλμὸς εἰπεῖν τῇ χειρὶ’ ‘And the eye cannot say unto the hand’, denominations with local names such as Mc 5,1 *in landa Gaddarene* ‘εἰς τὴν χώραν τῶν Γαδαρηνῶν’, ‘into the country of the Gadarenes’, or phrases with nouns having generic reference, such as Mc 4,32 *fuglos himinis* ‘τὰ πετεινὰ τοῦ οὐρανοῦ’ ‘the birds

of the air'. On the contrary, the article occurs, just as in the Greek version of the Bible, if a generic noun has been previously mentioned: J 11,38 f. *wasuh þan hulundi jah staina ufarlagida was ufaro. ... afnimiþ þana stain* '... λίθος ... τὸν λίθον' 'There was a cave, and a stone had been laid upon it ... Take away the stone'. Here, pragmatic definites are to be assumed (for OS cf. Lühr 2000a). Since GOTH used the definite article only in the case of attributeless generic pre-mentioned concepts, this language is at the beginning of the article cycle. For the anaphoric article is the first step in the development from a deictic element to a definite determiner (Lühr 2005a). As Demske (2001) has shown, OHG behaves in the same way concerning the use of the article quite apart from the fact that OHG possesses the indefinite article. Therefore, only the anaphoric use of the definite article is common to Germanic. Another genuine Germanic feature of the nominal phrase concerns the possessive pronoun. In the oldest Germanic languages this pronoun was an adjective, which could be combined with the definite article, cf. GOTH *so meina laiseins* 'my doctrine', OHG *thia mina fréwida all* 'all my joy'. Later the possessive adjective developed into a possessive article as in New High German or New English. The shift of the preposed attribute to postnominal position and the rise of the new word formation pattern of genitive compounds are related to this development. While preposed inanimate genitives were frequent in Germanic (ON *afreks verk* 'great heroic deeds', OHG I *dhes chrismen salbe* 'anointing oil', OS *uuapnes eggium* 'weapon's edge' and later followed the noun, animate genitives as in OHG *gotes barn* 'God's son' were integrated into the article system, as is documented in German from the time of Notker (Demske 2001). As a consequence of this restructuring of the genitive phrase, prepositional phrases become increasingly frequent; cf. MHG *daz pluot von Abele* 'the blood of Abel'.

The oldest prepositional phrases are attributive local determinations of personal nouns (GOTH *Iosef af Areimathaias* 'Joseph of Arimathaea', ON *Hrútr af Hrútsstoðum* 'Hrútr of Hrútsplace', OS *weroldkesures fan Rumuburg* 'Caesar of Rome', OHG *der chuninc ze berno* 'the king of Bern', of local nouns (OE *ðære cyrican æt Lindesfarena* 'of the church of Lindisfarne', OS *thea burg an Bethleem* 'the city in Bethlehem'), later of other concepts, too (OHG *thaz brot in theme disge* 'the bread in the dish'; cf. such further temporal determinations as OHG *alter dirro werelte fone erist unz in ende* 'the age of this world from the beginning until the end'). Prepositional collocations as in underlying verbal lexemes can be found with *nomina actionis* (ON *ráð á* 'assault on'; cf. *ráða á*, OHG *sagun fone* 'legend about'; cf. *sagēn fon*), sometimes with *nomina agentis*, too (OHG *losare von ubile* 'savior from evil'), and with other prepositions (OHG *minna zi* 'love for', OE *geleafan up* 'faith in' (Behaghel 1923; Schrodt 2004).

4. Verbal morphosyntax and periphrastic formations

The verbal morphosyntax of Germanic involves such form-function correlations as number, tense, mood, aspect, and diathesis. Of the Indo-European number categories, in GOTH the dual is retained, but only in the first and second person active. Therefore the GOTH dual is excluded from narration; cf. M 9,28 *duatiddjedun imma ... ga-u-laubjats þatei magjau þata taujan? qebun du imma ...* 'προσήλθον ... πιστεύετε ...; λέγουσιν' 'they came to him ... Do ye (two) believe that I am able to do this? They said unto

him ...'. The category of tense is reduced as well relative to the Old Indic and Greek system with their present, imperfect, aorist, perfect, pluperfect, and future. Rather, in Germanic the verbal system is characterized by a fundamental opposition [\pm past]. This opposition exists with all finite verbs, but with non-finite verbs only in the present participle and the past participle. Only the present tense and the perfect tense are inherited in Germanic, leading to the Germanic present and the preterite of the strong verb, whereas the weak preterite is an innovation based on an old periphrasis with the verb 'to do' (Lühr 1984). Verbs with a weak preterite generally result from secondary verbal formations such as denominatives and causatives. Because in Indo-European no perfect was built to such verbs, a new formation for these had to be found in Germanic. Another reduction of the Indo-European verbal system is found in the modal system of Germanic. In contrast to the five Old Indic mood categories of indicative, imperative, subjunctive, optative, and injunctive, Germanic languages possess only indicative, imperative, and subjunctive. The last of these continues the Indo-European optative, which expressed possibility (potential function) or a wish on the part of the speaker that either he or another subject carry out some action (cupitive function). In the latter the speaker indicates that he is not directly able to bring about the verbal action. In Germanic this expression of wish was enlarged to a voluntative function.

The question of whether Germanic possesses the grammatical category of aspect, the distinction between perfective and imperfective, has been a matter of debate for over a century. While Indo-European was characterized by aspect, some scholars assume that in Germanic aspect is not grammatically marked, i.e. not a formal property of the language, and that only lexical aspect or *Aktionsart* is expressed as an inherent semantic category. As Streitberg showed (1891; 1920), GOTH indeed has a means of indicating perfectivity or aspectuality in the form of verbal prefixes, especially the prefix *ga-*. Thus, telicity is opposed to atelicity or duration. The same holds true for OHG: *gi-*prefixation conveys aspectuality, for the corresponding *verbum simplex* designates an action without a starting point or an end point (GOTH Mc 4,9 *saei habai ausona hausjandona, gahaus-jai* 'he that has hearing ears, let him hear', J 11,11 f. ... *Lazarus ... gasaizlep ... jabai slepib, hails wairpib* '... Lazarus ... has fallen asleep ... if he is sleeping, he will become whole', OHG T 64,5 *zi hiu giengut üz sehan?* 'why do ye go out to see?' vs. T 57,1 *uuir uuollen fon thir zeichan gisehan* 'we want to see a sign from you', B 194,15 *der farfluahhanan diubil ... keleitta ze neouuehti* 'the accursed devil ... has wrought destruction' vs. B 207,35 *enger uuec ist, der leitit ze libe* 'narrow is the way that leads to life'). Consequently, overall one can say that in Germanic a number of *Aktionsarten* (durative, terminative, momentative, punctual) predominate rather than aspect. But the more infrequently action-type forms are found, the more tense forms expand. Thus, OHG and MHG texts contain several analytic verb forms which were used to express action-type differences, as well. These tense forms developed slowly. Thus, in OHG *habēn/eigan* + past participle constructions arose, occupying an intermediate position between passive, resultative, and perfect. The starting point is the perfective stative (= resultative) function of the past participle as documented in GOTH. This construction belongs to the periphrastic constructions of older Germanic. It shows a variety of stages: In older OHG only terminative transitive verbs are documented (Exh X, 4–5 *ir den christanun namun intfangan eigut* 'qui christianum nomen accepistis' 'you who have taken on the Christian name'), later non-terminative verbs occur, then also transitive verbs without an object are used (Musp 99 *denne der paldet der gipuazzit hapet* 'then can that one be consoled

who has [already] done penance’). The next step is periphrasis with intransitive verbs and finally with intransitive non-terminative verbs (N *so habet er gelogen* ‘thus has he lied’). While the *gi*-prefixation signals accomplishment, the *habēn* + past participle construction represents a state with a completed process. But as the participles are often inflected, these periphrastic constructions are predicative constructions rather than verbal categories (O 3,24,78 *then blinton deta séhentan* ‘the blind did he make see’, cf. 7; Schrod 2004).

Besides these active periphrastic constructions the rise of the periphrastic passive construction can be observed in Germanic. In Indo-European only a synthetic mediopassive occurs, which is well developed in Hittite, Old Indic, and Greek, occurring in different tenses and moods. In contrast, in GOTH, ON, and the West Germanic languages there are only more or less widespread traces of this formation: GOTH possesses a whole paradigm in the present: indicative *-ada*, *-anda*, optative *-aidau*, *-aindau*; but remnants of this paradigm are otherwise found only in a single verb in Old West Norse *heite*, Runic *haf[i]te-ka*, *haite*, *haitika*, Old Swedish *hæti*, OE 1/3sg.ind.pres.pret. *hätte*, 3.pl.ind.pres.pret. *hätton*, Middle Dutch, Middle Low German *hette* ‘is/are/was/were called’. The special semantics of the GOTH mediopassive is that of a typical unaccusative, showing both middle and passive values (verbs of emotion: *gaurjada* ‘grieves’, *gamarzjada* ‘takes offence’, facilitatives: *ingamjada* ‘is easily provoked’ vs. *distahjada* ‘is scattered’). The passive value predominates by a wide margin, however. The first step in the evolution of analytic passive constructions is the elimination of the ambiguity between the functions passive and middle of the synthetic mediopassive. Originally there was a semantic difference between the two constructions. In the passive the original subject is eliminated or put into the periphery and the theta-role theme becomes the subject, while in the middle the subject is preserved, more or less absorbing the theta-role theme. But the inherited synthetic mediopassive had the disadvantage that no preterite corresponding to the active preterite could be formed to it. Therefore, the competitor of the middle became the reflexive construction with reflexive pronoun (GOTH *sik*, ON *-sk* etc.) and that of the passive the *werþan*-, *wesan*-passive as well as the class IV weak verbs in *-na*-. These competitors emerged in Proto-Germanic and acquired a firm position during the development of GOTH. The starting point for the expansion of the reflexive construction is the direct reflexive reading. At first the reflexive pronoun can be found in connection with transitive verbs. The verb assigns a theta-role to its internal argument. It is still transitive but the object is reflexive; cf. with fronted and reinforced *sik*: GOTH E 5,25 *sik silban atgaf faur þo* ‘he gave himself for it’. In the next stage the reflexive pronoun must be adjacent to the verb in postverbal position. The verb has only one theta-role, and the reflexive must be simplex: Mc 4,1 *galesun sik* ‘they were gathered’. The next competitors to the inherited mediopassive, the verbs in *-na*- (*auknan* ‘become/get augmented’, *gabatan* ‘become improved’, *gabblindnan* ‘become blind’ etc.), behave like inchoative ergatives; they have a present participle, but lack a past participle. As one and the same verb can also appear as passive or as a *na*-verb (L 1,20 *usfulljanda* ‘they shall be fulfilled’ vs. M 8,17 *usfullnodedi* ‘it might be fulfilled’), Wulfila apparently made no distinction between the meaning of the passive and that of the *na*-verb. That is because both the passive and the ergative verb assign only an internal theta-role. As for the external theta-role, in a passive verb this role is absorbed, while it is not assigned at all with an ergative verb. Therefore, *na*-verbs never take a *fram*-phrase. But *na*-verbs were not an alternative corresponding exactly to the passive. Moreover, this verbal class

got lexicalized. No productive word formations could be formed from it any longer. However, both formations, the passive variant of the mediopassive and the *na*-verbs, are similar in denoting a change of state, a process. Hence, the passive meaning of the synthetic medio-passive is a processual (or agentive) passive (*Vorgangs-Passiv*; *werden-passive*) and not a stative passive (*Zustands-Passiv*; *sein-passive*); cf. k 3,15 *mipþanei siggwada Moses ...* ‘when Moses is read ...’. As in these instances the stative meaning is lacking entirely, and as GOTH is capable of expressing aspectuality with the dichotomy of *ga*-prefixation and *verbum simplex*, the analytic verbal construction *wisan* + past participle was chosen to denote the stative passive in the present. The model for this construction is provided by pairings of *wairþan* and *wisan* + adjectives, where telicity and atelicity are denoted: Mc 7,27 *let faurþis sada wairþan barna* ‘let the children first be filled’ vs. G 3,3 *swa unfroþans sijub?* ‘Are you so foolish?’. Now, analogous to such patterns sentences with decidedly present character were formed, cf. *wisan* + past participle in a stative meaning: k 1,4 *þizaiei gabrafstidai sijum silbans fram guda* ‘(by the comfort) wherewith we ourselves are comforted by God’. After the creation of a periphrastic passive in the present tense, this formation was extended to the preterite by incorporating the preterite *was* of *wisan*. And a formation for the processual passive was available, too: the preterite *warþ* of the verb *wairþan* ‘to become’ + past participle; cf. with telic and atelic construction L 15,24 *jah fralusans was jah bigitans warþ* ‘and he was lost, and has been found.’ But the circle in the evolution of the passive is not yet completely closed. One periphrastic present passive is still missing, the periphrastic processual passive with *wairþan*. In analogy to the coexistence of the processual passive and the stative passive in the preterite, in the present a processual passive with *wairþan* was created as a counterpart to the stative passive with *wisan*. That this was really the final step in the emergence of a passive category in the GOTH passivization cycle is shown by its limited number of attestations; cf. Mc 9,12 *ei manag winnai jah frakunþs wairþai* ‘that he must suffer much and be despised’. In relation to the great number of synthetic present passives, the few examples with *wairþan*-passives suggest that the synthetic forms were still a living grammatical category for the present passive in GOTH. The difference between the situation observed in GOTH and that obtaining in OHG reflects the well-known tendency to replace synthetic forms with analytic ones. Tense, aspect, and the connection of these grammatical categories are the driving force. A preterite both for the middle and the passive voice must be created in analogy to the active. And aspectuality gave rise to the differentiation between processual and stative passives in present and preterite by means of the *werþan*- and *wesan*-passives (OHG O H 104 *so war sunna lioht leitet so wurtun sie zispreitit* ‘as far as the sunlight extends, so [far] were they scattered’ vs. I 632 *dhazs izs in salomone uuari al arfullit* ‘that it should be all fulfilled in Solomon’, OE Alfred, Or 150,22 *on þæm gefeohhte wæs Antigones ofslagen* ‘in the battle Antigones was slain’). However, the reason why a formal differentiation between the middle and passive became necessary lies in information structure. The passive function of the old mediopassive was separated in order to yield appropriate definite topics. On the contrary, the middle function begot a reflexive construction.

Other periphrastic constructions in Germanic are *wesan* + present participle denoting duration (OHG I 80 *manage sint sohenti* ‘many are seeking’) and *werþan* + present or past participle to express the future; for Germanic had no morphological future tense: OHG T 2,9 *inti nu uuirdist thû suigenti inti ni maht sprehan* ‘and now thou art silent

and cannot speak', Musp 6 f. *sorgen mac diu sela ... / za uuederemo herie si gihalot uuerde* 'the soul must worry ... about which army it will be taken by [as war-booty]'). Modal verbs such as *sculan*, *mugan*, and *wellen* are also employed in future value (OHG I 630 *endi siin hohsetli scal uuesan festista untazs in euun* 'et thronus erit firmissimus in perpetuum' 'and his throne will be most firm unto eternity'), as are other periphrases, such as GOTH *duginnan*, *haban* + infinitive (Behaghel 1924; Lühr 1987; Schrodtt 2004). Modal verbs often provide an alternative to the subjunctive as well (Lühr 1994, 1997abc).

5. Adverbials

Adverbials are adjuncts which normally provide information about time, space, case, manner, etc. But if one differentiates between obligatory, facultative, and free adverbials, they are subject to various scope relationships. Valency-bound adverbials immediately belong to the verb and sentence adverbials refer to the sentence as a whole (cf. 1). Adverbials may take the form of simple adverbs, adverbial cases, prepositional adverbials, and adverbial clauses; cf. local and temporal genitives (GOTH L 15,15 *insandida ina haiþjos seinazos haldan sweina* 'and he sent him into his fields to tend to swine', M 6,30 *gistradagis* 'tomorrow', OHG WS 3 *se uuara se geloufan uualdes ode ueeges ode heido* 'wherever they run, in the forest, or on the road, or in the field', T 218,4 *inti thritten tages arstantan* 'and arise on the third day'), datives (GOTH Mc 4,27 *urreisip naht jah daga* 'he rises night and day'), accusatives (GOTH *rasta aina* 'a mile', *alla naht* 'all the night', ON *einn dag* 'a day', OHG O 1,5,5 f. *Floug er sunnun pad, sterrono stráza, / wega wólkono zi theru itis frono* 'He flew along the path of the sun, the road of the stars, the ways of the clouds to the holy maiden', T 16,2 *uonetun mit imo then tag* 'they remained with him for the day'); instrumental datives (GOTH M 27,60 *faurwalwjands staina mikilamma daurons þis hlaiwis* 'and rolling a great stone to the door of the sepulchre', O 3,18,11 *Bigondun sie ántwurten wórtón filu herten* 'they began to answer with very harsh words'); instrumentals (OHG Hl 40 *nu scal mih suasat chind suertu hauwan, breton mit sinu billiu* 'now will my dear son strike me with the sword, kill [me] with his sword' [cf. 3]); prepositional phrases (GOTH M 6,2 *swaswe þai liutans taujand in gaqumbim jah in garunsim* 'just as the hypocrites do in the synagogues and in the streets', ON *at jólum* 'at the Yule festival' (for adverbial clauses cf. 7).

6. Word order

Some standard generative accounts of the syntax of early Germanic word order assume a uniform SOV basic order which is maintained in subordinate clauses introduced by an overt complementizer in the head of a functional projection CP. In root clauses, the empty position of the complementizer is filled by the finite verb (= V_{fin}), while an optional movement of another constituent to SpecCP yields V2 in the surface. Under this analysis, the early Germanic languages share the syntactic structure of asymmetric SOV languages like modern German or Dutch, with some additional properties which explain apparent violations to this scheme. But new research on word-order variation in

the early Germanic languages has proceeded from a dynamic model of discourse relations as outlined in the Segmented Discourse Relation Theory SDRT of Asher and Lascarides (2003). Thus, Hinterhölzl and Petrova (2011), and Petrova and Solf (2008) studied the interrelation between information structure and discourse organization in the text of the OHG T. Their analysis of genuine OHG structures shows that verb-initial vs. verb-second placement regularly correlates with types of discourse relations attributed to the distinction between coordinating vs. subordinating linkage, whereby the position of the finite verb in the sentence functions as a device for marking the type of discourse relations in this text of the OHG corpus. On the one hand, there is a regular preference for verb-initial structures in sentences establishing new discourse referents, namely in text-initial sentences or presentational contexts: T 2,8 *Uuarun thô hirta in thero lantskeffi* ‘Et pastores erant in regione eadem’ ‘And shepherds were in that region’. On the other hand, sentences continuing an already introduced discourse referent or involving a referent considered accessible via a bridging relation to an already established entity show a regular tendency for verb-second placement against the underlying word order of the Latin original. Here, verb-second seems to be bound to referents that are salient in discourse: T 133,11 *Ih bin guot hirti. Guot hirti tuot sina sela furi siniu scaph* ‘Ego sum pastor bonus. Bonus pastor animam suam dat pro ovibus suis’ ‘I am the good shepherd. A good shepherd gives his life for his sheep’; T 2,1 *inti ira namo uuas Elisabeth* ‘et nomen eius Elisabeth’ ‘And her name was Elizabeth’.

These observations may be summarized as follows for both verb-initial and verb-second sentences:

a) verb-initial

[Vfin....DR_{new}...]FOCUS

b) verb-second

[DR_{giv/acc}]TOP [Vfin.....]FOCUS

Thus, the position of the finite verb serves to distinguish the information-structural domains of topic and focus in sentences of the Old High German period. The sentences with verb-second instances provide more information about a discourse referent already established in the previous discourse. Here, the text function of verb-second instances fits into the rhetorical relation of *elaboration* viewed as the prototype of subordinating linkage of discourse segments. The finite verb in such sentences separates a referential topic constituent from the rest of the utterance as a consequence of functional differentiations on the level of discourse relations. On the contrary, verb-initial sequences appear in contexts establishing a new situation as a basis for further elaboration in subsequent utterances. Such structures are related to the discourse function of *narration* and the coordinating type of linkage. Another new insight into Germanic word order concerns the position of the verb in embedded sentences. Many subordinate clauses involving an overt complementizer display lexical material to the right of Vfin. Such surface orders are analyzed as the result of extraposition to the right of Vfin, in analogy to the same kind of operation found in modern SOV languages. Hence, extraposition in early Germanic seem to affect the same groups of constituents as in modern SOV languages. These are:

(i) PPs and other phonologically “heavy” constituents, e.g. CP-complements or modified NPs; and

- (ii) VPs in clauses with complex predicates (verb clusters) where either the untensed main verb alone or in combination with an argument of the VP occurs to the right of V_{fin}.

In all cases V_{fin} is said to remain in its basic position at the end of the clause. However, this account has been challenged on the basis of previously unnoticed evidence from OE. Pintzuk (1999) provided examples of embedded clauses in which V_{fin} is followed by types of phrases which do not undergo movement to the right in modern SOV languages, so e.g. pronominal objects, “light” adverbs, or verbal particles. Additionally, she found evidence for particles and light elements after V in verb clusters. Such orders undoubtedly point to basic VO structure. This means that, against the standard account, OE cannot be viewed as uniformly head-final in the base. Yet another type of difference between the grammars that yield structural variation in OE is proposed by Fuss and Trips (2002). They observe a mismatch between the distributional properties of pronominal subjects and adverbs in main vs. embedded clauses in OE. While in main declarative clauses, adverbs cannot intervene between pronominal subjects and V_{fin}, they are allowed to do so in embedded clauses. Indeed, it is well known that over a long period of time German also displays a great deal of structural variation in subordinate clauses. This phenomenon has already been accounted for in traditional descriptions, e.g. Behaghel (1932) or Ebert (1978). They show that up to the 17th century, V_{fin} appears to be less rigidly fixed to the final position in subordinate clauses and is followed by types of constituents that never appear postverbally in present-day German. Recent work by Schlachter (2004), Weiß (in press), and Schallert (2006) addresses the situation in earlier records of German, i.e. those of the OHG period, from three rather different perspectives. Weiß discusses word order variation in clauses introduced by *dass* ‘that’ in the so called “minor” texts from the OHG period, Schallert proposes a model according to which OHG is a language with a mixed VO/OV order (cf. Axel 2005). An alternative approach is pursued by Schlachter (2004), who claims that variation in the right periphery of OHG is a correlate of information structure. According to her analysis of word order in *thacz*-clauses in the OHG Isidor, variation on the surface is part of a strategy to arrange the constituents of a message according to their informational value in the context. So, e.g., placing V_{fin} earlier in the clause allows the conveyance of relevant, or focused material in a communicatively favorable position, namely at the end of the utterance. In this respect, the variation in the placement of V_{fin} is used to separate new and relevant information from the domain of given or presupposed material in the clause. The investigation of the OHG T also provides support for the assumption that syntactic variation in the surface order of subordinate clauses is driven by information structure. In many cases, subordinate clausal structures appear as equivalents of different types of subordinate clauses in the Latin original containing conjunctions like *cum*, *ut*, *dum*, *si*, etc. Here, a plethora of different types of phrases occurs in the postverbal domain in OHG non-V-final clauses: PPs, NPs, light NPs, postverbal NPs as complements of copular verbs like *heizzan* ‘be called’, or predicate adjectives: T 44,1 *thaz sie üzvurphin sie* ‘ut eicerent eos’ ‘that they cast them out’; T 19,1 *thie giheizan ist Petrus* ‘qui uocatur Petrus’ ‘who is called Peter’; T 133,3 *daz sie sin blinte* ‘ut ... caeci fiant’ ‘that they be blind’; T 22,6 *thie thar ist giheizan Zelotes* ‘qui uocatur Zelotes’ ‘who is called there Zelotes’. But apart from these data, there are examples which contain diagnostics for a leftward movement of V_{fin} to a position preceding all other constituents of the clause. Thus, V_{fin}

appears to the left of a reflexive pronoun, which is normally situated in the so-called Wackernagel position at the left edge of the VP, immediately after C°: T 13,2 *uuanta nahit sih himilo richi* ‘*adpropinquavit enim regnum caelorum*’ ‘for the kingdom of heaven is drawing nigh’. To explain such word order variation Petrova (2008) drew on information structure again, namely in regard to (i) the distribution of background, or presupposed information vs. novel, or asserted information in the clause, and (ii) the positional distribution of narrow contrastive focus with respect to V_{fin} (cf. Lühr 2010a, 2010c). In OHG subordinate clauses, expressions which refer to given, but also to accessible information in the background domain of the discourse show a regular tendency to appear in the so-called Wackernagel position, i.e. in the position immediately following the subordinating conjunction or the relative pronoun: T 8,4 *thane ir iz findet* ‘*cum inuenertis*’ ‘when you find it’. Furthermore, full lexical phrases are also regularly shifted from postverbal position in the Latin original to Wackernagel position in OHG, when they are discourse-anaphoric or accessible: T 119,10 *Ni santa got sínan sun in uuerolt thaz her uuerolt tuome, uzouh thaz uuerolt si giheilil thuruh inan* ‘*Non enim misit deus filium suum in mundum ut iudicet mundum, sed ut salvetur mundus per ipsum*’ ‘God did not send his son into the world that he should judge the world, but that the world should be saved by him’. A parallel situation is described by Kemenade and Los (2006) and Kemenade (2008) for OE. They observe that discourse-linked material in OE regularly appears in a special syntactic domain situated between the subordinating conjunction and an adverbial *þa* which functions as a discourse partitioner in the clause; cf. a similar function of the OHG adverbial *tho*: T 19,8 *Mit thi u thaz thó gisah Simon Petrus* ‘*Quod cum videret Simon Petrus*’ ‘When Simon Peter then saw that’. But in OHG the role of a discourse partitioner described for *þa* in OE is taken by V_{fin} itself: V_{fin} targets exactly that position in the clause which separates the given, or presupposed, information from the rest of the utterance, which corresponds to the focus domain of the clause. In these cases two basic groups of constituents appear to the right of V_{fin} in subordinate clauses in OHG: (i) arguments of verbs and (ii) non-finite parts of complex predicates: T 22,2 *inti thie thár habetun diuual* ‘*et qui demonia habebant*’ ‘and who there had a devil’; T 94,2 *nibu ir uuerdet giuuentita inti gifremite soso theser luzilo* ‘*nisi conversi fueritis et efficiamini sicut parvuli*’ ‘unless you are changed back and made [to be] like these little ones’.

As opposed to the above, constituents which convey contrastive information in the clause, forming a complementary pair of alternatives with another entity in the discourse, or expressing selection, correction, or emphasis in the clause are associated with the position which immediately precedes V_{fin}: T 33,3 *ni uizze iz thin uuinstra uuaz thin zesuuu tuo* ‘*nesciat sinistra tua quid faciat dextera tua*’ ‘let thy left hand not know what thy right hand does’; T 35,2 *thaz thu mannun ni sís gisehán fastenti, úzouh thinemo fater* ‘*ne videaris hominibus ieiunans, sed patri tuo*’ ‘in order that you should not appear as fasting to men but to your father’. The analysis of Petrova (2008) concludes that OHG is discourse-configurational, i.e., that there is tight correlation between the information-structural value of sentence constituents and their positional realization in the clause. The clause structure derived from OHG is parallel to the situation described by Diesing (1997) for Yiddish:

[C Backgr [_{FocP}ContrF V [_{ArgP} PresF [_{VP} tV]]]].

7. Sentence syntax

As is true of Indo-European languages generally, Germanic has subject-verb agreement. However, a collective in the singular can appear with a verb in the plural: GOTH Mc 3,32 *setun bi ina managei; qeþun þan du imma* ‘And the multitude sat around him, and they said unto him’; vice versa, a plural subject may be combined with a singular verb, if the subject represents a collective concept: OHG O 3,6,55 f. *ward thero áleibo ... / sibun kórbi* ‘of the morsels remaining there were ... seven baskets’. The second agreement attribute, gender, is best seen with the predicate. In Germanic a predicate nominative occurs with special verbs, cf. GOTH *sitan, standan, þugkjan, wairþan, wisan* and in some passive constructions, especially with *haitan*, OHG *sîn, werdan, heizan*, and with verbs like *stân, sizzan, bilîban, haltan, giberan, gangan, faran, queman*, and a predicate accusative appears with verbs like OHG *duan, findan, firlāzan, wizzan*. Sometimes the predicate is inflected, sometimes uninflected: cf. for inflection O 1,4,52 *ált-duam suáráz duit uns iz úrwanaz* ‘harsh old age makes it impossible for us’, O 2,4,9 *Er thar niheina stígilla ni firliáz ouh únfirslagana* ‘he left not even one path there unlocked’ (Schrodt 2004; Fleischer 2007). In older texts natural gender sporadically prevails over grammatical gender: GOTH Tim 3,16 *unsahataba mikils ist gagudeins runa* (= Christ) ‘and incontrovertibly great is the mystery of godliness’, and in the case of pronouns neuter nouns for females are often referred to their natural gender: OHG T 103,1 *uuib, thiu habeta geist unmahti* ‘a woman who had a spirit of sickness’. The third agreement feature, case, can be found with appositions; cf. OHG T 5,8 *Ioseph Dauides sun* ‘Joseph, David’s Son’; OE Beo 426 *ic þe nu ða, brego beorht Dena, biddan wille* ‘I now want to ask you then, lord of the Bright-Danes’.

Other features conforming to Indo-European are the different kinds of sentences: declarative, interrogative, imperative, and exclamatory. Because intonation is generally not denoted in Germanic – an exception is Otfrid (Fleischer 2009; Lühr 2008b) – word order (V1 in imperative sentences), mood (imperative in imperative sentences), or special words (interrogative particles, pronouns in interrogative sentences, interjections in exclamatory sentences) serve as differentiating devices (cf. 1). Also the coexistence of simple, compound, and complex sentences, i.e. parataxis and hypotaxis, is old (for parataxis cf. Lühr 2007b). Compound sentences can be connected asyndetically or syndetically; cf. OS Hel 197 f. *Skred the uuintar forð, / geng thes gêres gital* ‘The winter progressed, the sequence of the year passed’ vs. with copulative relation: GOTH L 17,27 *etun jah drugkun, liugaidedun jah liugaidos wesun* ‘they ate, they drank, they married, and they were given in marriage’; OHG O 1,21,1 *Tho ęrstarp ther kúning Heród, joh hina fúarta inán tod* ‘Then King Herod died, and death led him away’; with disjunctive relation: GOTH L 16,13 *andizuh ainana fijaþ jah anþarana frijoþ aiþþau ainamma andtiloþ, iþ anþamma frakann* ‘for either he will hate the one, and love the other; or else he will hold to the one, and despise the other’; OHG O 1,11,38 *ni méid sih, suntar sje óugti, then gótes sun sougti* ‘she was not ashamed, but she showed that she was nursing God’s son’; with consecutive relation: GOTH M 6,24 f. *ni maguþ guþa skalkinon jah mammonin: Duþþe qiþa izwis* ‘Ye cannot serve God and mammon. Therefore I say unto you’; with causal relation: GOTH R 14,4 *apþan standiþ; mahteigs auk ist frauja gastop[an]an ina* ‘but he shall be held up: for the Lord is able to make him stand’; OHG T 22,13 *Salige sint thie thar sint subere in herzon, uuanta thie gisehent got* ‘Blessed are those who are pure in heart; for they shall see God’; and with adversative relation: ON

Sigurds 32 *Sigmundur ok allir synir hans váru langt um fram alla menn aðra; Sigurður var þó allra framasta* ‘Sigmund and all his sons were longed for by all other men. But Sigurd was the best of all’; late OHG N Bo 141,6 *uuánda ménniskôn mûot ist natúrlichô des uuâren gûotes kër. ... Áber díu míssenómeni. des uuéges. ferléitet sie ze demo lúkken* ‘for the soul of men is eager by nature for the true good. But the mistaking of the way leads them to deception’. An archaic device of connection is GOTH *-uh*: J 6,17 *jah usstigun in skip, iddjedunuh ufar marein in Kafarnaum* ‘And they entered into a ship, and went over the sea into Capernaum’ (cf. Latin *-que*).

Hypotaxis is fully developed in Germanic (Lühr 2011a; Lühr and Zeilfelder 2011). There are subject and object clauses, adverbial clauses and attributive clauses. Until the end of MHG times the law of sequence of tenses applied: With present tense in the matrix clause the subordinate clause appears in the present subjunctive, and with preterite tense in the superordinate clause the subordinate clause has the preterite subjunctive. The most frequent subordinate sentences are relative sentences and clauses corresponding to NE *that*-sentences. But GOTH and ON partially show other means of expressions: So in the case of ‘that’ (Greek ὅτι) GOTH uses a particle *ei*, which can be combined with **þat* to *þatei*, while ON *at* and West Germanic **þat* arise from Proto-Germanic **þat*. Another conjunction for ‘that’ in GOTH is *þei*. The particular conjunction that is used depends on the meaning of the verb. Thus, *ei* occurs with verbs of wishing, commanding, forbidding, requesting, aspiring, requiring, sufficing, or allowing. The mood is the optative expressing volition; cf. L 5,14 *is faurbaud imma ei mann ni qeþi* ‘And he charged him to tell no man’. If the verb has the meaning ‘to hope’ or ‘to mean’ the conjunction *ei* or *þatei* generally appears with the potential optative; cf. Phil 22 *wenja auk ei þairh bidos izwaros fragibaidau izwis* ‘for I trust that through your prayers I shall be given unto you’. With verbs of believing and trusting one finds mostly *þatei* (rarely, *ei*) with the indicative; cf. k 2,3 *gatrauands in allaim izwis þatei meina faheþs allaize izwara ist* ‘having confidence in you all that my joy is the joy of all of you’. *þatei* with the indicative is the rule with verbs of wondering, delighting, knowing, remembering, hearing, seeing, learning, and communicating; cf. Mk 2,8 *jah suns ufkunnands Iesus ahmin seinamma þatei swa þai mitodedun sis ...* ‘And Jesus, immediately perceiving in his spirit that they so reasoned within themselves ...’. Another category of *that*-sentences comprises explicative sentences depending on a noun (Lühr 1992, 1993).

It is generally assumed that the conjunctive value of Proto-Germanic **þat* originates in the employment of a demonstrative pronoun in cataphoric value at the end of a main clause with subsequent shifting of the sentence boundary: OHG O 2,2,8 *joh gizálta in thar tház thiu sálida untar in was* ‘and he told them there that: the blissfulness was among them’ → ‘and he told them there that the blissfulness was among them’. But in Germanic *that*-clauses often have a correlative; cf. OS Hel 855 ff. *Uuissun that thoh managa / liudi aftar them landa, that he uuas an thit lioht cumin* ‘And yet many people in the land knew that, that he had come unto the light’. An investigation of the correlative in OS resulted in three groups of verbs, one with an obligatory correlative, one with a facultative correlative (as *witan* above), and one without a correlative. The verbs without correlative also appear with accusative-cum-infinitive constructions: OS Hel 115 *hiet that fruod gumo foroht ni uuâri* ‘He ordered that the pious man not be afraid’ vs. 317 *hêt sie ina haldan uuel* ‘he ordered that he should care for her well’. These are verbs with S-deletion, meaning that there is no sentence boundary between the matrix verb and the accusative-cum-infinitive construction. With verbs of doing, however, the correlative

licenses the verb to govern a *that*-clause: OS Hel 3320 f. *sô hue sô that ... giduot, / that he ... mâgo gisidli / liof farlêtid ...* ‘whoever does that, that he abandons the dear home of his relatives ...’. Thus the existence or non-existence of a correlative complies with fixed rules; and the widely accepted view that *that* is shifted from the main clause to the subordinate clause is clearly wrong (Lühr 2004a). Another type of subject and object clause comprises indirect questions. Here, the interrogative pronouns, adverbs, and particles are the same as in direct questions, as is mood: cf. with indicative GOTH J 18,38 *hva ist so sunja* ‘what is the truth’ vs. E 6,21 *ei <jah> jus witeiþ hva bi mik ist, hva ik tauja* ‘that ye also may know my affairs (literally ‘what is by me’), how I do’; and with potential optative J 12,27 *hva qiþau* ‘what should I say’ vs. M 6,25 *ni maurneiþ saiwalai izwarai hva matjaiþ jah hva drigkaþ, nih leika izwamma he wasjaiþ* ‘be not anxious for your life, what ye shall eat, and what ye shall drink; nor for your body, what ye shall wear’.

Adverbial clauses are normally introduced by conjunctions in GOTH, ON, and West Germanic. In GOTH the temporal conjunctions mostly show enclitic *ei*: *miþþanei* ‘while’, *sunsei* ‘as soon as’, *faurþizei* ‘before’, *hwanei* ‘when’. But *þan* ‘when’ makes it obvious that in Germanic demonstrative pronouns are originally used as conjunctions as well (cf. also OHG *thar* in *ther thar*). *-ei* is secondary. Other types of conjunctions are GOTH *þande* ‘as long (as), if, because’, *bipe* ‘while, as long (as), after’, and *unte* ‘as long (as), until’. Comparative clauses are introduced with words with the meaning ‘as’: GOTH *swe, swaswe* (versus consecutive *swaei* ‘so that’ besides *swe, swaswe*). Conditional clauses occur both with and without conjunctions in all Germanic languages. For the first type cf. GOTH *jabai, -ba* with indicative for realis conditions (1x), present optative for potential conditions, and preterite optative for irrealis conditions (negated as *nibai, niba*). In OFr and OHG conjunction-less conditionals are used to denote a close connection with the preceding text, expressing a singular instantiation of a general circumstance: OHG O 1,5,53 ff. *Nist in êrdriche thâr er imo io instriche, / noh winkil untar himile thâr er sih ginêrie. / Flîuhit er in then sé, thar giduât er imo wé* ‘There is no place on earth where he (viz. Satan) can ever escape him (viz. Jesus), nor is there a cranny under heaven where he can save himself. (If) he flees into the sea, there he will bring (punishing) trouble upon him’; OFr Rüstr Kûren 7 *Sa hwersa twa sinhigen se and kinda tein hebbath and hiara other forifalle ... sterue thet leste, thet ma thet got inna twa dele* ‘Wherever there are two married people and they have produced children and one of them dies ... (if) the last (child) dies, one should divide the property into two portions’ (Lühr 2007a, 2008c, 2010b). This text-structural function is Proto-Germanic; cf. with the connecting particle *iþ* GOTH J 9,41 *iþ blindai weseiþ, ni þau habaidedeiþ frawaurhtais* ‘if ye were blind, ye should have no sin’. Just as the conditional relation, so expressions for the concessive relation are well developed in Germanic; cf. GOTH *þauhjabai*, OHG *thoh* etc. (Lühr 2003; 2004b).

As for relative sentences, generalizing relative sentences introduced by double ‘so’ are typically Germanic: OHG T 183,2 *so uenan so ih cusse ther ist iz* ‘whoever I kiss, he is it’; cf. GOTH *swa filu swe* ‘as much as’ etc. (Lühr 1998a; for the GOTH type *sahwazuh saei* ‘whoever’ etc. cf. Lühr 2000c).

However, in simple relative clauses West Germanic possesses a different relative pronoun from GOTH and ON. While GOTH uses the relative pronoun *saei* built to the demonstrative *sa* (cf. OHG *ther thâr*) and the enclitic particles *ei, þei* (mostly following

neuter indefinite pronouns), *izei* (*ize*) and *sei*, and ON employs only the uninflected particle *es*, *er* (later *sem*), in West Germanic the demonstrative pronoun appears also as a relative pronoun: OS Hel 4055 f. *nio the sterben ni scal / ... the hêr gilôbid te mi* ‘never shall that one die ... who believes in me’. Because relative sentences and not main clauses are obligatory if they refer to NPs with the definite article, to inherently indefinite determiners, and to predicates (OS Hel 3713 f. *endi frâgodan sân, / hue that uuâri, that thar mid thiû uuerodu quam* ‘and they asked immediately who that might be who came there with the people’), if there is no referent word in the matrix clause (OS Hel 1352 *Than uuôpian thar uuanscefti, thie hêr êr an uunion sîn* ‘then they bemoan there the misery, who previously here are in joy’), and if the subordinate clause appears in the middle of the matrix clause, the demonstrative pronoun cannot be taken over from the superordinate clause. However, the use of the same word in a matrix clause and as a dependent marker can be observed not only with *that*-clauses and relative clauses, but also with adverbial sentences (e.g. OHG *thoh, wanta, sô, nu* [cf. 2]). Probably word order was sufficient to denote hypotaxis in early West Germanic. Generally the finite verb appears in subordinate sentences in a later position than in superordinate sentences (for details cf. 6).

8. Abbreviations

B	Benediktinerregel	Musp	Muspilli
Beo	Beowulf	N	Notker
E	Epheser	N Bo	Notker Boethius
Exh	Exhortatio ad plebem	O	Otfrid
	christianam	O H	Otfrid, Hartumuate
Hel	Heliand	O L	Otfrid, Ad Ludovicum
HI	Hildebrandslied	Or	Orosius
I	Isidor	Phil	Philemon
J	Johannes	R	Römer
K	Korinther I	Rüstr Küren	Rüstringer Küren
k	Korinther II	Sigurds	Sigurdsaga
L	Lukas	T	Tatian
M	Matthäus	Tim	Timotheus
Mc	Markus	WS	Wiener Hundesege

9. References

- Asher, Nicholas and Alex Lascarides
 2003 *Logics of Conversation*. Cambridge: Cambridge University Press.
- Axel, Katrin
 2005 *Studies on OHG syntax: left sentence periphery, verb placement and verb-second*. Amsterdam: Benjamins.
- Barðdal, Jóhanna
 2001 *Case in Icelandic: A Synchronic, Diachronic and Comparative Approach*. Lund: Lund University Press.

- Behaghel, Otto
1923–1932 *Deutsche Syntax. Eine geschichtliche Darstellung*. 4 vols. Heidelberg: Winter.
- Demske, Ulrike
2001 *Merkmale und Relationen. Diachrone Studien zur Nominalphrase des Deutschen*. Berlin: De Gruyter.
- Diesing, Molly
1997 Yiddish VP order and the typology of object movement in Germanic. *Natural Language and Linguistic Theory* 15: 369–427.
- Ebert, Robert P.
1978 *Historische Syntax des Deutschen*. Stuttgart: Metzler.
- Ferraresi, Gisella
2005 *Word Order and Phrase Structure in Gothic. A Comparative Study*. Leuven: Peeters.
- Fleischer, Jürg
2007 Das prädikative Adjektiv und Partizip im Althochdeutschen und Altniederdeutschen. *Sprachwissenschaft* 32: 279–348.
- Fleischer, Jürg
2009 Paleographic clues to prosody? Accents, word separation, and other phenomena in Old High German manuscripts. In: Hinterhölzl and Petrova (eds.), 161–189
- Frascarelli, Mara and Roland Hinterhölzl
2007 Types of Topics in German and Italian. In: Susanne Winkler and Kerstin Schwabe (eds.), *On Information structure, Meaning and Form*. Amsterdam: Benjamins, 87–116.
- Fritz, Gerd and Thomas Gloning (eds.)
1997 *Untersuchungen zur semantischen Entwicklungsgeschichte der Modalverben im Deutschen*. Tübingen: Niemeyer.
- Fuß, Eric and Carola Trips
2002 Variation and change in Old and Middle English: on the validity of the Double Base Hypothesis. *Journal of Comparative Germanic Linguistics* 4: 171–224.
- Hinterhölzl, Roland and Svetlana Petrova (eds.)
2009 *New approaches to word order variation and change in the Germanic language*. Berlin: De Gruyter.
- Hinterhölzl, Roland and Svetlana Petrova
2011 Rhetorical Relations and Verb Placement in Old High German Tatian. In: Christian Chiarcos, Berry Claus, and Michael Grabski (eds.), *Saliency. Multidisciplinary perspectives on its function in discourse*. Berlin: De Gruyter, 173–201.
- Jäger, Agnes
2005 Negation in Old High German. *Zeitschrift für Sprachwissenschaft* 24: 227–262.
- Kemenade, Ans van
2009 Discourse relations and word order change. In: Hinterhölzl and Petrova 2009: 91–120.
- Kemenade, Ans van and Bettelou Los
2006 Discourse Adverbs and Clausal Syntax in Old and Middle English. In: Ans van Kemenade and Bettelou Los (eds.), *The Handbook of the History of English*. Oxford: Blackwell, 224–248.
- Krause, Wolfgang
1968 *Handbuch des Gotischen*. München: Beck.
- Lühr, Rosemarie
1984 Reste der athematischen Konjugation in den germanischen Sprachen: zu ‘sein’ und ‘tun’. In: Jürgen Untermann and Bela Brogyanyi (eds.), *Das Germanische und die Rekonstruktion der indogermanischen Grundsprache. Akten des Freiburger Kolloquiums der Indogermanischen Gesellschaft. Freiburg 26.–27. Februar 1981*. Amsterdam: Benjamins, 25–90.

Lühr, Rosemarie

- 1987 Zu Veränderungen im System der Modalverben. In: Rolf Bergmann, Heinrich Tiefenbach, and Lothar Voetz (eds.), *Althochdeutsch. I. Grammatik, Glossen, Texte*. Heidelberg: Winter, 262–289.

Lühr, Rosemarie

- 1991 Die Determinansphrase aus historischer Sicht. Die Flexion von *der, die, das* in der Funktion als Demonstrativpronomen, Relativpronomen und als bestimmter Artikel. *Beiträge zur deutschen Sprache und Literatur* 113: 195–211.

Lühr, Rosemarie

- 1992 Typen von Explikativsätzen im Althochdeutschen. In: Yvon Desportes (ed.), *Althochdeutsch: Syntax und Semantik. Akten des Lyoner Kolloquiums zur Syntax und Semantik des Althochdeutschen (1–3. März 1990)*. Lyon: Centre d'Études Linguistiques Jacques Goudet, 259–291.

Lühr, Rosemarie

- 1993 Syntaktische Restriktionen bei Abstrakta gestern und heute. *Zeitschrift für deutsche Philologie* 112: 83–104.

Lühr, Rosemarie

- 1994 Zur Konkurrenz von Konjunktiv und Modalverbgefügen im älteren Deutsch. *Nordlyd: Tromsø University working papers on language and linguistics* 22: 116–141.

Lühr, Rosemarie

- 1996 Zum "Etymologischen Wörterbuch des Althochdeutschen": *oder wie, oder wo, oder wann; vielleicht*: ahd. *ôdo, odowân, odowar, odowîla* – die Interdependenz von althochdeutscher Etymologie und neuhochdeutscher Syntax. *Historische Sprachforschung* 109: 144–160.

Lühr, Rosemarie

- 1997a Zur Semantik der althochdeutschen Modalverben. In: Fritz and Gloning (eds.), 159–175.

Lühr, Rosemarie

- 1997b Modalverben als Substitutionsformen des Konjunktivs in früheren Sprachstufen des Deutschen? Die Verhältnisse in der Hypotaxe. In: Fritz and Gloning (eds.), 177–208.

Lühr, Rosemarie

- 1997c Althochdeutsche Modalverben in ihrer semantischen Leistung. In: Yvon Desportes (ed.), *Semantik der syntaktischen Beziehungen. Akten des Pariser Kolloquiums zur Erforschung des Althochdeutschen 1994*. Heidelberg: Winter, 200–222.

Lühr, Rosemarie

- 1997d Altgermanische Fragesätze: Der Ausdruck der Antwort Erwartung. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft, Madrid, 21–24 de septiembre de 1994*. Madrid: Ediciones de la UAM, 327–362.

Lühr, Rosemarie

- 1998a Verallgemeinernde Relativsätze im Althochdeutschen. In: Karin Donhauser and Ludwig M. Eichinger (eds.), *Deutsche Grammatik: Thema in Variationen. Festschrift für H.-W. Eroms zum 60. Geburtstag*. Heidelberg: Winter, 263–281.

Lühr, Rosemarie

- 1998b Konzessive Relationen. In: John Ole Askedal and August Wilhelm Zickfeld (eds.), *Historische germanische und deutsche Syntax. Akten des internationalen Symposiums anlässlich des 100. Geburtstages von Ingerid Dal, Oslo, 27.9.–1. 10. 1995*. Frankfurt: Lang, 165–192.

Lühr, Rosemarie

- 2000a Absolute und relative Begriffe in der Dichtersprache des Heliand. Syntax und Semantik von Abstrakta und Stabreimtechnik. In: Jens Haustein, Eckhard Meineke, and Norbert

- Richard Wolf (eds.), *Septuaginta quinque. Festschrift für H. Mettke*. Heidelberg: Winter, 219–237.
- Lühr, Rosemarie
2000b Zur Wortstellung in der althochdeutschen komplexen Nominalgruppe: Die Stellung von Quantoren. In: Yvon Desportes (ed.), *Zur Geschichte der Nominalgruppe im älteren Deutsch. Festschrift für P.Valentin. Akten des Pariser Kolloquiums, März 1999*. Heidelberg: Winter, 53–69.
- Lühr, Rosemarie
2000c Morphosemantische Remotivierungen bei Quantoren des Gotischen. In: Almut Hintze and Eva Tichy (eds.), *Anusantatyai. Festschrift für Johanna Narten zum 70. Geburtstag*. Dettelbach: Röhl, 161–179.
- Lühr, Rosemarie
2002a “Allgemeine Anaphora”. Zum Artikelgebrauch bei der Fügung “Substantiv und adnominaler Genitiv” im Heliand. In: Matthias Fritz and Suzanne Zeilfelder (eds.), *Novalis Indogermanica. Festschrift für Günter Neumann zum 80. Geburtstag*. Graz: Leykam, 251–281.
- Lühr, Rosemarie
2002b Zur Periodisierung des älteren Friesischen anhand der Syntax: Der adnominale Genitiv. In: Alfred R. Wedel and Hans-Jörg Busch (eds.), *Verba et Litteræ: Explorations in Germanic Languages and German Literatures. Essays in Honor of Albert L. Lloyd*. Newark, DE: Linguatext, 85–100.
- Lühr, Rosemarie
2003 Genitivische Konnektoren im Althochdeutschen. In: Yvon Desportes (ed.), *Konnektoren im älteren Deutsch. Akten des Pariser Kolloquiums, März 2002*. Heidelberg: Winter, 193–211.
- Lühr, Rosemarie
2004a Der Nebensatz in der Westgermania. In: Thorwald Poschenrieder (ed.), *Die Indogermanistik und ihrer Anrainer. Dritte Tagung der Vergleichenden Sprachwissenschaftler der Neuen Länder, Universität Greifswald*. Innsbruck: Institut für Sprachwissenschaft der Universität, 161–179.
- Lühr, Rosemarie
2004b Universale Konzessivsätze. In: Maria Kozińska, Rosemarie Lühr, and Susanne Zeilfelder (eds.), *Indogermanistik – Germanistik – Linguistik. Akten der Arbeitstagung der Indogermanischen Gesellschaft, Jena 18.–20. 09. 2002*. Hamburg: Kovač, 345–359.
- Lühr, Rosemarie
2005a Der Einfluß der klassischen Sprachen auf die germanische Grammatik. In: Gerhard Meiser and Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft. September 2000. Halle*. Wiesbaden: Reichert, 341–362.
- Lühr, Rosemarie
2005b Topikalisierung in Metrik und Prosa im älteren Deutsch. In: Franz Simmler, Claudia Wich-Reif, and Yvon Desportes (eds.), *Syntax. Althochdeutsch–Mittelhochdeutsch. Eine Gegenüberstellung von Metrik und Prosa. Akten zum Internationalen Kongreß an der Freien Universität Berlin, 26. bis 29. Mai 2004*. Berlin: Weidler, 177–191.
- Lühr, Rosemarie
2006a Historische Fallstudie: Altsächsisch. In: Vilmos Ágel, Ludwig M. Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer, and Henning Lobin (eds.), *Dependenz und Valenz. Ein internationales Handbuch der zeitgenössischen Forschung. 2. Halbbd*. Berlin: De Gruyter, 1500–1508.
- Lühr, Rosemarie
2006b Attribute bei den altnordischen Kenningar. In: Georges-Jean Pinault and Daniel Petit (eds.), *La Langue Poétique Indo-Européenne. Actes du Colloque de travail de la Société des Études Indo-Européennes, Paris, 22–24 octobre 2003*. Leuven: Peeters, 253–266.

Lühr, Rosemarie

- 2007a Bedingungssätze in altfriesischen Rechtstexten. In: Rolf H. Bremmer Jr., Stephan Laker, and Oebele Vries (eds.), *Advances in Old Frisian Philology*. Amsterdam: Rodopi, 213–238.

Lühr, Rosemarie

- 2007b Konnektoren im älteren Deutsch. In: Hans Ulrich Schmid (ed.), *Beiträge zur synchronen und diachronen Sprachwissenschaft*. Leipzig: Sächsischen Akademie der Wissenschaften zu Leipzig, 39–51.

Lühr, Rosemarie

- 2007c Die Wiederaufnahme durch den Artikel im Althochdeutschen: Zur Akzentuierung von Definita. In: Yvon Desportes, Franz Simmler, and Claudia Wich-Reif (eds.), *Die Formen der Wiederaufnahme im älteren Deutsch. Akten zum Internationalen Kongress an der Universität Paris Sorbonne (Paris IV) vom 8.–10. Juni 2006*. Berlin: Weidler, 101–116.

Lühr, Rosemarie

- 2008 Die Verbstellungen im älteren Deutsch. Ein bewusster Stellungswandel? In: Jean-Marie Valentin (ed.), *Akten des XI. Internationalen Germanistenkongresses, Paris 2005*. 4: *Germanistik im Konflikt der Kulturen: Integrative Zugriffe auf Phänomene des Sprachwandels*. Bern: Lang, 127–132.

Lühr, Rosemarie

- 2010a Fokuspunkteln im Althochdeutschen. In: Yvon Desportes, Franz Simmler, and Claudia Wich-Reif (eds.), *Mikrostrukturen und Makrostrukturen im älteren Deutsch vom 9. bis zum 17. Jahrhundert: Text und Syntax. Akten zum Internationalen Kongress an der Universität Sorbonne 6.–7. Juni 2008*. Berlin: Weidler, 103–120.

Lühr, Rosemarie

- 2010b Bedingungsstrukturen im Älteren Deutsch. In: Arne Ziegler (ed.), *Historische Textgrammatik und Historische Syntax des Deutschen. Traditionen, Innovationen, Perspektiven*. Berlin: De Gruyter, 157–172.

Lühr, Rosemarie

- 2010c The Role of Information Structure in Language Change: Introductory Remarks. In: Rosemarie Lühr and Gisella Ferraresi (eds.), *Diachronic Studies on Information Structure*. Berlin: De Gruyter, 1–14.

Lühr, Rosemarie

- 2011a Adverbialsatz und Koordination im Althochdeutschen. In: Franz Simmler and Claudia Wich-Reif (eds.), *Syntaktische Variabilität in Synchronie und Diachronie vom 9. bis 18. Jahrhundert*. Berlin: Weidler, 11–24.

Lühr, Rosemarie

- 2011b Prägnante Konstruktionen in den klassischen Sprachen. In: Giovanna Rocca (ed.), *Die Sprachen Altitaliens. Inschriften, Texte, Grammatik. In memoriam Helmut Rix (1926–2004)*. *Rivista di glottologia* 5. 185–198.

Lühr, Rosemarie

- 2012 Informationsstrukturelle Einheiten im Alt- und Mitteldeutschen. In: Michel Lefèvre (ed.), *Syntaktischer Wandel in Gegenwart und Geschichte: Akten des Kolloquiums in Montpellier vom 9. bis 11. Juni 2011*. Berlin: Weidler, 415–436.

Lühr, Rosemarie and Susanne Zeilfelder

- 2011 Zur Interdependenz von Diskursrelationen und Konnektoren in indogermanischen Sprachen: Kontrast und Korrektur. In: Eva Breindl, Gisella Ferraresi, and Anna Volodina (eds.), *Satzverknüpfungen. Zur Interaktion von Form, Bedeutung und Diskursfunktion*. Berlin: De Gruyter, 107–148.

Petrova, Svetlana

- 2009 Information structure and word order variation in the Old High German Tatian. In: Hinterhölzl and Petrova (eds.), 251–279.

- Petrova, Svetlana and Michael Solf
 2008 Rhetorical relations and verb placement in the early Germanic Languages. A cross-linguistic study. In: Cathrine Fabricius-Hansen and Wiebke Ramm (eds.), *'Subordination' versus 'Coordination' in Sentence and Text. A cross-linguistic perspective*. Amsterdam: Benjamins, 329–351.
- Pintzuk, Susan
 1999 *Phrase Structures in Competition. Variation and Change in Old English Word Order*. New York: Garland.
- Schallert, Oliver
 2006 *Hybride OV / VO-Systeme und syntaktischer Wandel zu OV und VO in den germanischen Sprachen*. Diplomarbeit, University of Salzburg.
- Schlachter, Eva
 2004 Satzstruktur im Althochdeutschen. Eine Skizze zur Position des Verbs im Isidor-Traktat des 8. Jahrhunderts. In: Karin Pittner, Robert Pittner, and Jan C. Schütte (eds.), *Beiträge zu Sprache & Sprachen 4. Vorträge der Bochumer Linguistik-Tage*. Munich: Lincom, 179–188.
- Schrodt, Richard
 2004 *Althochdeutsche Grammatik. II: Syntax*. Tübingen: Niemeyer.
- Streitberg, Wilhelm
 1891 Perfective und imperfective Actionsart im Germanischen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 15: 70–177.
- Streitberg, Wilhelm
 1920 *Gotisches Elementarbuch*.^{5/6} Heidelberg: Winter.
- Weiß, H.
 To appear Die rechte Peripherie im Althochdeutschen. Zur Verbstellung in *dass*-Sätzen. In: Rosemarie Lühr and Suzanne Zeilfelder (eds.), *Tagungsakten der Arbeitstagung der Indogermanischen Gesellschaft in Jena*. Köln: Kovač.

Rosemarie Lühr, Berlin (Germany)

57. The lexicon of Germanic

- | | |
|--|---|
| 1. Germanic as a branch of the Indo-European languages | 6 Exclusively Germanic vocabulary |
| 2. Comparisons of vocabulary | 7. Results |
| 3. Degrees of kinship | 8. Neighbours of the Germanic languages |
| 4. Statistical material | 9. Word and stem formation |
| 5. Substratum and superstratum | 10. References |

1. Germanic as a branch of the Indo-European languages

The vocabulary of the Germanic languages is a true continuation of the lexicon of Proto-Indo-European. It is organized into a system of parts of speech, particularly the open classes of nouns, verbs, and adjectives, and it has continually adapted to new demands

<https://doi.org/10.1515/9783110523874-012>

- Petrova, Svetlana and Michael Solf
 2008 Rhetorical relations and verb placement in the early Germanic Languages. A cross-linguistic study. In: Cathrine Fabricius-Hansen and Wiebke Ramm (eds.), *'Subordination' versus 'Coordination' in Sentence and Text. A cross-linguistic perspective*. Amsterdam: Benjamins, 329–351.
- Pintzuk, Susan
 1999 *Phrase Structures in Competition. Variation and Change in Old English Word Order*. New York: Garland.
- Schallert, Oliver
 2006 *Hybride OV / VO-Systeme und syntaktischer Wandel zu OV und VO in den germanischen Sprachen*. Diplomarbeit, University of Salzburg.
- Schlachter, Eva
 2004 Satzstruktur im Althochdeutschen. Eine Skizze zur Position des Verbs im Isidor-Traktat des 8. Jahrhunderts. In: Karin Pittner, Robert Pittner, and Jan C. Schütte (eds.), *Beiträge zu Sprache & Sprachen 4. Vorträge der Bochumer Linguistik-Tage*. Munich: Lincom, 179–188.
- Schrodt, Richard
 2004 *Althochdeutsche Grammatik. II: Syntax*. Tübingen: Niemeyer.
- Streitberg, Wilhelm
 1891 Perfective und imperfective Actionsart im Germanischen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 15: 70–177.
- Streitberg, Wilhelm
 1920 *Gotisches Elementarbuch*.^{5/6} Heidelberg: Winter.
- Weiß, H.
 To appear Die rechte Peripherie im Althochdeutschen. Zur Verbstellung in *dass*-Sätzen. In: Rosemarie Lühr and Suzanne Zeilfelder (eds.), *Tagungsakten der Arbeitstagung der Indogermanischen Gesellschaft in Jena*. Köln: Kovač.

Rosemarie Lühr, Berlin (Germany)

57. The lexicon of Germanic

- | | |
|--|---|
| 1. Germanic as a branch of the Indo-European languages | 6 Exclusively Germanic vocabulary |
| 2. Comparisons of vocabulary | 7. Results |
| 3. Degrees of kinship | 8. Neighbours of the Germanic languages |
| 4. Statistical material | 9. Word and stem formation |
| 5. Substratum and superstratum | 10. References |

1. Germanic as a branch of the Indo-European languages

The vocabulary of the Germanic languages is a true continuation of the lexicon of Proto-Indo-European. It is organized into a system of parts of speech, particularly the open classes of nouns, verbs, and adjectives, and it has continually adapted to new demands

<https://doi.org/10.1515/9783110523874-012>

by a system of word formation, using mainly suffixation and compounding, to a lesser degree prefixation. Many of the idiosyncrasies of Proto-Indo-European are retained, including the use of lengthening (*vridhhi*), of accent shift, and of vowel change (*ablaut*), among others. Semantic subsystems such as kinship terminology, names of animals, terms denoting parts of the body, simple numerals, etc. appear in forms that are regularly developed from PIE. From this perspective the central part of the Germanic vocabulary can be directly derived from its Proto-Indo-European ancestor.

It would be desirable to convert these general remarks into measurable data. To this end an already completed statistical study of a distinct subpart of the Germanic lexicon will be presented in this chapter. It is based upon a survey of the 23rd edition of Kluge's *Etymologisches Wörterbuch der deutschen Sprache*, which records the first attestation of words still used in Modern High German and indicates how widespread these words and their families actually are in the Indo-European languages. This survey therefore does not deal with all Germanic languages but only with German, and even there not with the entire vocabulary of the relevant centuries but only with that vocabulary which has been retained in the present-day language. Nevertheless, it is likely to provide a representative overview of the development of the Germanic lexicon.

2. Comparisons of vocabulary

Before proceeding it will be useful to say a few words about the comparison of vocabularies belonging to different time periods: The lexicon is that part of a language that interfaces most obviously with material cultural, which is notoriously subject to change over time. Consequently, one may expect the lexicon on the whole to change more quickly than other areas of a linguistic system. The older the time period of a given lexicon, the less it will have been subject to the encroachments of cultural change. This is why the vocabularies of early attested languages, such as Greek or Sanskrit, are much more likely to be closer to that of the proto-language than the vocabularies of later attested languages such as Germanic and Baltic.

On the other hand, one cannot assume the proto-language had a completely unified vocabulary. Every language can be subdivided into varieties, which are to a large extent differentiated by vocabulary. But in our very abstract study it makes sense to treat the lexicon of the proto-language as a unity, even if some of the terms belonging to it might have been used only regionally. More particularly, it will suffice for our purposes that a word family is attested in at least one non-European Indo-European language (in most cases this language will be a member of the Indic branch).

3. Degrees of kinship

As a next step, we must consider which other Indo-European languages were especially close to Germanic during the time of and subsequent to its separation from the main stock. Here a western group (Celtic, Italic, Germanic) as well as an eastern group (Germanic, Baltic, Slavonic) can be identified, although the relations to the western group played a more prominent role until the dawn of historic times. Lexical isoglosses which

are confined to one or both of these two groups possibly date back to a comparatively late period. If we move beyond these groups, the next most frequently represented isoglosses are those shared by Germanic and Greek, with these tending to be shared by the other two groups as well. This set is likely to be older than the aforementioned constellations, but they do not necessarily date back to proto-Indo-European times. At the outermost level we encounter general Indo-European isoglosses, which also include non-European Indo-European languages, especially those of the Indian subcontinent.

For our assessment of the Germanic lexicon it will be sufficient to draw upon the evidence of German and Gothic, which may be taken to represent the two regional extremes of the Germanic-speaking territory. More closely related (and thus potentially later) connections are the “late common Germanic” languages, comprising all Germanic languages apart from Gothic, (“North-west Germanic”, except that this designation carries with it other implications which are not relevant to our specific concerns here), the “West Germanic” group (which has to be taken together, but which does not require the assumption of originally shared special features), and finally the “*düdisch*” group (a term I have suggested instead of *deutsch* [German] in order to avoid confusion between “continental Germanic” and “German”).

4. Statistical material

The vocabulary of New High German which is attested for the first time in the 8th, 9th, 10th, and 11th centuries, respectively, and its distribution by source is as follows. The numbers in brackets represent inherited words, assimilated loanwords, and non-assimilated loanwords, respectively:

8 th century	1105 (1014, 80, 11)
9 th century	1111 (934, 147, 30)
10 th century	295 (218, 69, 8)
11 th century	468 (383, 71, 14)
Total (Old High German):	2979 (2549, 367, 63)

By word class, the distribution is approximately substantives : adjectives : verbs = 10 : 3 : 2 (at least in the 8th and 9th centuries; later, verbs become more frequent than adjectives).

I will first identify words which occur in all Indo-European languages (given certain formal concessions). In the 8th century, these number 139 words; in the 9th century, 51; in the 10th century, 7; in the 11th century, 9 for a total of 206 words. Subsequent centuries yield only 4 instances in which words attested in all Indo-European languages occur for the first time (and these are themselves problematic). The following are the substantives, adjectives and verbs in this group, given, where possible, in their New High German form and separated according to semantic groups:

Substantives. Kinship names: 8th c. *Bruder, Mutter, Schwester, Sohn, Tochter, Vater, Vetter, Witwe, (Degen)*; 9th c. *Ahn, Neffe, Schwäher, Schwieger*; 10th c. *Ammе*. Body parts: 8th c. *Ader, Arm, Auge, Fuß, Herz, Horn, Knie, Nacken, Nase, Ohr, Zahn, Zunge, Elle; Atem, Stimme*; 9th c. *Bug, Ferse, Galle, Kinn, Schweiß*; 11th c. *Euter, Wanst. Ani-*

mals and animal products: 8th c. *Aar, Hase* (the meaning is not the same as in the proto-language), *Honig, Hund, Kuh, Lachs, Maus, Ochse, Vieh, Wolf*; 9th c. *Gans, Otter, Sau, Wolle*; 10th c. *Biber*. Plants, nature, landscape: 8th c. *Acker, Dorn, Meer, Nacht, Sonne, Stern, Wasser, Wiese, Wind*; 9th c. *Birke, Boden, Feim, Furt, Mal*; 10th c. *Ampfer, Loh*. Tools and the like: 8th c. *Achse, Joch, Salbe, Tür, Werk*; 9th c. *Ahle, Esse, Helm, Nabe*. Others: 8th c. *Geist, Harm, Jahr, Leumund, Miete, Name, Sieg, Sitte*; 9th c. *Met*; 10th c. *Heim*; 11th c. *Alb, Mast, Mitte*.

Adjectives: 8th c. *äbich, dünn, dürr, eng, hoch, jung, laut, neu, süß, viel, voll, wahn, wohl* (adv.); 9th c. *frei, roh, wert*; 11th c. *schütter*.

Verbs: 8th c. *beißen, binden, bitten, brauchen, brechen, brennen, essen, forschen, gebären, gehen, gießen, kauen, kiesen, kommen, lesen, mahnen, nehmen, schweifen, sehen, setzen, sitzen, sparen, stehen, steigen, suchen, trügen, tun, wachsen, wägen, weben, wehen, wehren, wissen*; 9th c. *bieten, genesen, gewinnen, hängen, hauen, hinken, leihen, liegen, mahlen, meiden, nagen, raten, schwitzen, seihen, speien, stoßen, taugen, weichen, wirken, zeihen, zerren*; 10th c. *gedeihen, lallen, renken*; 11th c. *schneien, spähen*.

Although other word classes have not been included here, this list alone shows that the core vocabulary of the Germanic languages undoubtedly goes back to Proto-Indo-European. Of 1014 native words in the 8th century, 139 or 13.7% of the inherited words attested at the beginning of the transmission go back to the common Indo-European lexicon (for Old High German in its entirety the figure is 206 out of 2549 or 8.1%).

Now let us consider words which can be connected to some subset of Indo-European subgroups outside of Germanic:

Word families with a common Indo-European distribution:	1112
Word families with a European distribution:	404
Word families in the western and eastern group:	137
In the western group:	178
In the eastern group:	176

These figures show that of 2549 inherited terms from the 8th through 11th centuries, 2007 or 78.7% (nearly 4 out of every 5) show Indo-European affinities.

Finally we have to consider those instances where the direct connections are confined to the Germanic languages. The figures for this are:

Common Germanic:	137
Late Common Germanic:	150
West Germanic:	97
“Düdisch”:	158

This makes a total 542 terms, which equals 21.3% of the 2549 units we have investigated. We will have to deal further with these 21.3%, as they may prompt the question of whether Germanic is a purely Indo-European language or whether it comprises a second element which goes beyond simple loans.

5. Substratum and superstratum

5.1. Substratum theories

Theories of this sort are related to the (archaeologically founded) hypothesis that Indo-European members of the prehistoric Corded-Ware Culture immigrated into Denmark and imposed themselves on the indigenous Funnel Beaker Culture. Germanic has been interpreted as the linguistic result of the mixing of these two cultures. A recent version of these theories is Theo Vennemann's hypothesis that after the Ice Age (Central and Northern) Europe was settled by a population whose language, which Vennemann terms "Vasconic", survives only in present-day Basque. But he has also added a second hypothesis to this: the peoples living along the coasts (Germanic and mostly Celtic peoples) were superseded by a people whose archaeological remains are mostly megalithic tombs. Their language, according to Vennemann, was "Semitidic", a form of Afroasiatic. Hence, according to this theory, Germanic would have a Vasconic substratum as well as an Atlantic (Semitidic) superstratum (see Noel Aziz Hanna 2003). A more general approach has been suggested in the etymological explanations of a far-reaching project of Leiden University: These verify the possible substratum (superstratum, adstratum) origins of words which until now have not been etymologized, especially if they show certain characteristic features, such as initial *a*, etc. This is part of a rising tendency to attribute words with an unclear etymology to a substratum (cf. for example Boutkan 2005, especially XIII–XVII).

5.2. Words "without etymology"

The main problem concerning words without an Indo-European etymology is the fact that often the attribution "without etymology" is assigned too quickly. In many cases, words designated "without etymology" can actually be etymologized, even if this is not easy. Three examples of forms with complex etymologies will be illustrated here, all of which may be plausibly analyzed, despite having been characterized as "without etymology". I will take the word *Schwert* to illustrate an example of a "hidden etymology" for which either incredible proposals have been presented or which has been classified "without etymology". The proper approaches for a good etymology of this word have been well known, but their significance has not been favorably assessed: a Greek word for the (magnificent) sword (which later was used in a figurative sense) is the word *áor* (*n*), attested since the Iliad. It is clearly a derivative of the Greek *aeírein* 'to hang', and whoever makes the effort to check the attestations of the word and its family will soon recognize the correlation: *áor* is actually not the sword's blade but the sword as a well-ornamented "work of art", which has been fixed to a holster hanging from the shoulder with valuable rings (often made of gold). As the actual meaning for the Greek word "suspended armour" is given (a one-sided but correct interpretation). The word is traced back to Indo-European **(ə)wer-* (**h₂wer-*) (the *o* goes back to an *o*-grade or to a dialectal zero-grade, which morphologically is more likely). Now, related languages show a word form **swer-* (with "s mobile") for this root, which led Pokorny (1959: 1150 f.) to assume a root **wer-*, also *swer-*. Greek would probably also allow a root

**(ə)swer-*, but the comparison includes also examples from languages which clearly have no *s-*. In Baltic, this root has partly been extended with *d* (< PIE **d*), and Germanic sometimes takes part in similar developments. So we have the root form **swerd-* in a language neighbouring Germanic, and the development of the meaning to ‘sword’ in a Greek derivative, which finally makes the etymology of *Schwert* quite clear: it is actually the sword with all its paraphernalia, the ‘suspended sword’, and hence an Indo-European word.

A second case of the faulty assessment “without etymology” is the word *Schuh*. Here, however, one must reckon with “irregular sound change”. To be sure, the straightforward consonantal skeleton of this word is Germanic **sk h-*. However, if one were to assume a dropped *r-* (cf. New High German *sprechen* versus Modern English *speak*, which is not an isolated case), one finds a group of mostly North Germanic words for ‘leather’, ‘shoe’ and other similar items: included in the (Germanic) word family **skraha-* are Old West Norse *skrá* ‘dried fur, parchment, document’ (as a loan in English this becomes *scrae* ‘old shoe’). The clearly older loanword in Finnish is *raha* ‘furs’; without ‘s mobile’ we find such Celtic words as Old Irish *croiccenn* ‘skin, fur’, etc. Hence, the etymological explanation is Indo-European **skrok-* ‘leather’, **skrōk-* ‘made from leather, shoe’. The assumption of a lost *r* is no worse than the alternative assumption that *Schuh* is a substratum word.

Our third example is the word *Hand*. The most frequently cited etymology of this word takes it to be an agent noun ‘grabber’ built to the root seen in Gothic *frahinþan* ‘take captive’, which itself lacks an accepted etymology. However, agent nouns of this sort (*u*-stem derivatives in *o*-grade with suffix accent) are elsewhere unknown in Germanic. Hence, it is more honest to state that both this verb and its presumed nominal derivative are “without etymology”. However, a semantic path toward an etymology for the noun at least may be opened by noting that the meanings ‘hand’ and ‘fist’ often overlap, leading to the second remark that the meaning ‘fist’ in a number of languages neighbouring Germanic is derived from verbs meaning ‘to hit, to stab’ (no doubt based on the notion of the fist as a weapon employed in fighting): Latin *pugnis* ‘fist’ from *pungere* ‘hit, stab’, Latvian *dūre* ‘fist’ from *dūrt* ‘push, stab’, similarly Old Irish *dorn* ‘fist, hand’ etc. If Gothic *handugs* ‘wise’ originally meant ‘sharp-minded’, then it may be compared to Old High German *hantac* ‘sharp, pointed’, suggesting a noun **handu-* ‘pike’ (**handuga-* ‘having a pike’); and this can in turn be related to Greek *kenteîn* ‘to stab, poke’ together with derived agent noun *kontós* ‘poker’ and of course *kéntron* ‘stinging agent, pricker’, whence ‘fixed point around which one describes a circle’. In this way we find attested all intermediate stages for the etymology of ‘hand’: ‘stab, poke’, ‘sting, stinging agent’, ‘fist’, ‘hand’.

6. Exclusively Germanic vocabulary

I turn now to lexical material which is Common Germanic (if need be, excluding Gothic) but lacks plausible extra-Germanic etymologies (however, forms for which a set of competing Indo-European etymologies have been proposed, however unlikely, have been excluded from this category). Owing to the limited space in this article, I confine myself here to substantives, having verified the material using the most recent version of

KLUGE (24th edition). The list of these forms, arranged by the centuries of their first attestation, is the following:

8th c. *Adel* ‘aristocratic lineage’, *Asch* ‘barge’, *Beute* ‘board’, *Blut*, *Brot*, *Glut*, *Krücke*, *Loch* (from *lōkan* ‘lock’, no comparison possible outside Germanic), *Luft*, *Napf*, *Regen*, *Ross*, *Rüge*, *Säule* ‘pillar’, *Schalk*, *Scham*, *Schatz*, *Schiff*, *Schilling*, *Schwegel*, *Seele*, *Spott*, *Taube*, *Waffe*, *Wamme*, *Wampe*, *Wange*, *Winter*, *Zaine*, *Zauber*; 9th c. *Aland* ‘fish from the carp family’, *Bein*, *Blei*, *Esch*, *Fiedel*, *Glas*, *Harfe*, *Held*, *Kar*, *Klaue*, *Lust*, *Mage*, *Maser*, *Maut*, *Melde*, *Riese*, *Rohr*, *Scholle*, *Schrat*, *Schwalbe*, *Schwelle*, *Tugend*, *Wunde*, *Zeche*, *Zinn*, *Zotte*; 10th c. *Aal*, *Docke*, *Leiste*, *Meise*, *Walm* (from *wölben*, no comparison possible outside Germanic), *Zwerg*; 11th c. *Bast*, *Dachs*, *Nachen*, *Schachen* ‘grove’, *Schauer*, *Schmerl* ‘merlin’, *Schote*, *Stär* ‘ram’, *Zohe*.

Numerous elements in this group can easily be considered loanwords (e.g. *Blei*), and several may represent onomatopoeic words or contain sound symbolism. A likely example of this is *Glut*, the initial *gl-* of which perhaps possesses “phonesthetic” value associated with meanings such as ‘glow’. I do not find the number of terms on this list surprising or problematic. Although I do not want to rule out in principle the possibility that there are relics of “substrata” in the Germanic vocabulary (whatever their actual character might have been), I consider attempts to treat all such elements which have no etymology to be misguided.

7. Results

At this point we may summarize our results so far:

- a) A closed corpus containing 2,979 words has been examined, 367 of which were fully assimilated, 63 not fully assimilated loanwords from well-known languages or known “migrating words”, amounting to 430 words or 14.4% of the total. This part of our results cannot be generalized, as the proportion of loanwords in any language depends on the historic situation existing at the time of the borrowings. Furthermore the classification as a “loanword” in contrast to “foreign word in an indigenous text” often can only be decided on a subjective basis. In what follows we will consider only those words, 2,549 in number, which cannot be clearly shown to have been borrowed (“inherited words”).
- b) In this reduced corpus 1,112 = 43.6% can be connected to common Indo-European word families (direct word comparisons can be made with 206 units = 8.1%). 895 words = 35.1% can be connected to individual Indo-European languages; the total of these two categories is 2,007 words = 78.7%.
- c) 542 words cannot be connected at all to non-Germanic languages. This is 21.3% of the non-borrowed vocabulary. One quarter each of these terms are confined to word families which are only found in German and to word families which are represented in all the Germanic languages; the remaining half comprises word families which are more widespread than German but which are not found in all Germanic languages.
- d) If we may treat this data as representative, we may infer that in Indo-European languages contemporaneous with Old High German and which have not been subject to special outside influences approximately 4/5 of the vocabulary which cannot be clearly shown to have been borrowed will have its origin in the Indo-European proto-

language and its successors. We may further infer that the remaining 1/5 will be of unclear etymology. Furthermore, it is very likely that a large number of these will be unrecognised loanwords; but whether they are interpreted as belonging to special substrata and/or superstrata will inevitably be left to the preferences of individual scholars.

8. Neighbours of the Germanic languages

The fact that Germanic words were also borrowed into neighbouring languages is crucial for the assessment of the history of the Germanic vocabulary. This predominantly applies to the non-Indo-European Finnish language, which has retained older forms whose phonological shapes are important for understanding the absolute and relative chronologies of some Germanic sound laws. There are differing views about the extent of Finnish lexical borrowing from Germanic. The second language which has retained older lexical material is Lithuanian, to which we may add the other Baltic languages in general. However, owing to the fact that these languages are Indo-European, the question of parallel retention vs. borrowing is often difficult to answer (the same applies to loans into Slavic, although these belong to a much later period. Generally, these loanwords are expected to come mostly from Gothic). It would also be crucial to know more about the Scythian language, which bordered Germanic on the east and the south, perhaps also Thracian, but only fragments of these languages (or language groups) are known.

In the west, it was surely the Celts and, according to linguistic reconstruction (which has not been confirmed by archaeologists), the later Italic peoples who were neighbours of the Germanic peoples. Generally, connections to Italic are considered to be time-wise much earlier than the connections to Celtic and are usually treated as the result of common Indo-European heritage. Connections to the Celtic languages are very complex: a portion of the Germanic-Celtic etymologies are certainly due to their joint Indo-European heritage, while just as surely another part comprises borrowed words (well attested in the lexical fields of *Amt* and *Reich/reich*); but in many cases it cannot be decided which of the two options is correct. Further complicating matters, an etymology can sometimes be due to both factors, as when an Indo-European root or form inherited by both branches has been subject to a special development in Celtic, and Germanic has applied that development to its native form. The direction of borrowing almost without exception goes from Celtic into Germanic; the few alleged cases where the direction of borrowing is reversed are very problematic.

9. Word and stem formation

From an Indo-European perspective, one may distinguish primary and secondary formations. The former are built directly to roots, whereas the latter can be divided into derivatives built to primary formations and compounds. The central element of primary formations are (verbal or nominal) roots, from which are built systematic grammatical formations which can be formed as needed and which have a completely regular meaning, such as noun of action, comparative and superlative (for adjectives), or verbal adjec-

tive (tending to become a past passive participle when built to a transitive root). Strictly speaking, these are not part of the lexicon but represent regular morphosyntactic formations which, however, may become specialized (coming to signal, for example, a “passive adjective of necessity” [gerundive] as opposed to, say, just a “verbal adjective”); but they can also develop semantic peculiarities based on their use, which makes them real parts of the lexicon. For example, *Wohnung* is neither the ‘act of living’ nor merely a ‘place of living’, but a ‘separate living area consisting of several rooms’, thereby differentiating it from an *Apartment*, a *Hotel suite*, or a permanent domicile beneath a bridge, etc. The process by which this specification comes about is known as lexicalization.

As noted above, primary derivatives have their origins in verbal roots. The simplest formation type of Proto-Indo-European consists of root nouns functioning as abstract nouns or, when used as the second member of a compound, as agent nouns. Substantives can be modified by adjectives. As primary formations these are often simple active or passive verbal adjectives, normally with suffixes **-to-* or **-no-*, which originally maintained their independence from verbal paradigms, as they still do in the earliest Greek. The other primary types of word formation (abstract nouns, agent nouns, etc.) also occur with various suffixes (**-ti-*, **-tor-*, etc.) and may be subject in individual cases to specialization and lexicalization.

At an outer level of morphological structure are secondary derivations, which originate in existing words and stems, be they primary derivatives of verbal roots or derivatives of primary adjectives or substantives. Among the semantic categories typically signalled by these formations are “material adjectives”, “diminutives”, “appurtenative formations”, etc. as well as “secondary verbs” (denominative verbs, deverbative formations such as causatives, intensives, desideratives, etc.). Originally, secondary verb formations had only a present tense; but in the individual languages they start to develop partial paradigms, often via periphrasis.

In Germanic the central position of verbal roots is continued in the system of “strong verbs”, which form a virtually closed system. Unlike in some other Indo-European languages (e.g. Sanskrit), where a new verbal root emerging from any source may be treated like a primary verbal root, in Germanic such a verb is usually treated like a “weak verb”. A consequence of this is that Germanic has extended the secondary verbs (which have been joined by some primary verbs) to at least two classes of regular and productive verbs (*jan*-verbs and *ōn*-verbs; the *ēn*-verbs and the *nan*-verbs are not as productive as those in *-jan* and are not attested in all Germanic languages).

The development of primary formations in Germanic was initially influenced by the fact that the vowels (and partly also the consonants) of the suffixes and endings were subject to weakening, and furthermore by the fact that the function-changing accent mobility was mostly fixed. Hence the system was concentrated on a few inflection paradigms with unified *ablaut* grades. Among the abstract nouns, root nouns have been replaced by certain *a-*, *i-* and (rarely) *u-*stems with distinct *ablaut* grades and a defined gender. Among consonantal suffixes, feminine *ti*-stem derivatives of strong verbs as well as the corresponding *ni*-stem derivatives of weak verbs were retained for a longer period. The system of agent nouns has been significantly altered: root nouns which were second members of a compound (a type rarely represented in Germanic) became simple *a-* and *-an*-stems, and in German *-ila*-formations occurred much more frequently in this value; but otherwise, the suffix **-arja-* (mostly with a long vowel), borrowed from Latin, has

been taken over in all Germanic languages. The reason for this significant change is probably to be sought in the rising demand for lexical formations with this semantic content, particularly as a more complex society required ever more specialized occupational titles. In the processes of filling this need the basis for new formations was increasingly found among the substantives: in the earliest attested Germanic languages, such as in Gothic, agent nouns are principally denominative formations indicating affiliation (e.g. *fiskja* ‘fisherman’).

Other morphological processes as well played important roles in the derivation of both verbs and substantives. In the former case one such process was prefixation. In Proto-Indo-European, a verb could be semantically determined in a sentence by a particle normally possessing spatio-temporal value, which need not have been adjacent to it. Only if a derivative was formed from “verb + particle” did the particle obligatorily have to precede the verb in direct contiguity, and the two were then joined together to create a new form. These joint formations became increasingly important and finally resulted in prefixation becoming a fundamental word-formation process in the Germanic languages. However, this process did not proceed to completion in all instances, with the result that two different sets of cases ensued: a constantly joined type (the so-called “inseparable verbs”) as well as a type which was only joined together if required by the syntax (the “separable verbs”). This is the situation in Modern German; however, the individual Germanic languages go their own way in the subsequent evolution of this pattern. Thus, the “inseparable verbs” regularly lost their “prefix” in the Nordic languages, and the position of the prefixes of separable verbs was treated differently from one language to the next. Outside of the cases just described, with a few exceptions, verbs can no longer be compounded in the present day Germanic languages.

In the case of substantives, two developments that stand out are the close relationship of active adjectives (participles) to agent nouns as well as the special implementation of *n*-stems in the formation of the weak adjective. The first of these is seen in the fact that active adjectives can occupy the position of agent nouns (*Freund* and *Feind*, etc.). In the second instance we are dealing with the “determined” or definite adjectives. Many languages show that adjectives can be treated differently depending on whether they possess definite or indefinite reference. The difference is often marked syntactically or by certain particles; but in Germanic and in its neighbouring languages (Baltic and Slavic) this is done by using a special inflection (different in Germanic and Balto-Slavic). In Germanic, an *n*-suffix is used, the origin of which is unclear. Possible candidates are 1. a stem suffix *-en-* (which would simply lead to an *n*-inflection), or 2. a secondary suffix *-Vn*, indicating definiteness, which is added to the stem suffix (with vowel contraction where relevant), or 3. – a suggestion made by Hermann Hirt, which is not at all out of date – an added pronoun with *-n-*. This extension of adjectival inflection seems to be connected to the development of the substantive inflection of *n*-stems, with which it is identical in all “regular” cases. In any event, the inflection of substantives in the Germanic languages differentiates a “strong” type (*a, ō, i, u*, i.e. vocalic stem endings) and a “weak” type (*n*-stems). A significant number of *n*-stems in Germanic are agent nouns.

Prominent among secondary noun formations is first of all *vridhhi*, an archaic type of word formation indicating affiliation. Its characteristic feature is that it lengthens the vowel of the first syllable of the base word and moves its accent, while generally also adding a simple suffix to the base word. This formation type is most widely represented in the Indo-Iranian languages, very likely also in Hittite. While not pervasively common in Germanic languages, it is nevertheless unmistakable. All Germanic examples show

the lengthening of Germanic *a* to **ā* or *e* to *ē*. Remnants of the type have been sought in non-Germanic languages (apart from Indo-Iranian and Anatolian); but so far, no convincing examples have come to light. Although this type is extremely archaic, it remained productive for a long while in Germanic. Keynote examples are, as indicated in the title of the ground-breaking treatment by Georges Darms (1978), *Huhn* (< **hōn-n-a-*) from *Hahn* (< **hanen-*) and *Schwager* (< **swēgurá-*) from *Schwäher* (< **swéhura*).

Another group of secondary noun formations, which deserve to be mentioned owing to their wide functional implementation, are patterns employed to produce feminization (which denotes the formation of feminine forms from corresponding masculine forms or generic designations for living creatures). Three formal patterns can be observed, all of which have counterparts elsewhere in Indo-European. The first is the feminine used in adjectival inflection, if it has not been formed after one of the two other forms of feminization. This consists of a relationship **-o* → **-ā* (**-eh₂*), Germanic *-a* → *-ō*, whereby a feminine form of an adjective is derived from a masculine form via replacement of an *a*-stem by an *ō*-stem. This pattern, which is widely found in Indo-European languages in both nominal and adjectival inflection, is in Germanic mostly limited to adjectives. In nominal forms the feminine has undergone extension to an *n*-stem. A comparable derivation seen in Sanskrit is *ásva-* ‘horse, stallion’ → *ásvā* ‘mare’. The second pattern involves, in Indo-European terms, a *yā*-formation with a nominative in *-ī* (**-ī/-yā*-stem [PIE **-ih₂ /-yeh₂-*]), the so-called *devī*-inflection (OI. *devī* ‘goddess’ from *devá-* ‘god’). The final type is a less-widespread inflection in which a nominative in a long *ī* alternates within the paradigm with a stem in **-iy-*, the so-called *ṽkīs*-inflection (OI. *ṽkīs* ‘she-wolf’ from *ṽka-* ‘wolf’). The third formation can be found in relics such as Old West Norse *ylgr* ‘she-wolf’ from *ulfr* ‘wolf’; but the basis for the predominant Germanic form of feminization is the second formation, the *devī*-inflection, which has been retained in Germanic as a *jā*-inflection (with an occasional special nominative). But this formation became accepted as an extension to an *n*-suffix, the origin of which can still be seen in old formations, such as *Henne* (< **han-n-jā*) from *Hahn* (< **hanen*); later it is the type *König* / *Königin* with a suffix **-en-jā*.

Another formation worth mentioning, which stands close to inflected forms, is the collective, which suggestively denotes an “unstructured majority” standing in contrast, partly in competition, partly in complementary distribution with the plural, which denotes a “structured majority”. The associated forms (a likely example is the word for ‘name’, Gothic Sg NA *namo*, Pl NA *namna*, not, as would be regular, **namona*) are likely to have been more widely expressed in earlier times than the attested forms would make us believe, but at no time were they a fixed part of the nominal inflectional paradigm (cf. for instance Bjorvand 1994).

Finally we must say a word about compounds. One generally differentiates determinative compounds (AB is a B, more clearly determined by A), exocentric compounds (AB is someone or something whose B is an A: *Rotkäppchen* is a girl whose hood is red), and copulative compounds (AB is A and B). In Proto-Indo-European these types were probably differentiated by accent. In Germanic, copulative compounds did not play a significant role; and, apart from some moribund forms, exocentric compounds were confined to special formations, chief among them personal names. The productive formation type in all Germanic languages is determinative compounds. But it may be concluded from individual examples that Proto-Indo-European and Early Germanic had a much richer formal and semantic range of compound formations (cf. for instance Seebold 1968).

10. References

- Bjorvand, Harald
 1994 *Holt og holtar. Utviklingen av det indoeuropeiske kollektivum in norrønt* [Wood and woods. The development of the Indo-European collective in Old Norse]. Oslo: Solum.
- Boutkan, Dirk and Sjoerd Michiel Siebinga
 2005 *Old Frisian Etymological Dictionary*. Leiden: Brill.
- Darms, Georges
 1978 *Schwäher und Schwager, Hahn und Huhn. Die Vrdahi-Ableitung im Germanischen*. Munich: Kitzinger.
- Heidermanns, Frank
 1993 *Etymologisches Wörterbuch der germanischen Primäradjektive*. Berlin: Mouton de Gruyter.
- Kluge, Friedrich
 1926 *Nominale Stammbildungslehre der altgermanischen Dialekte*. 3. edn. rev. by Ludwig Sütterlin and Ernst Ochs. Halle (Saale): Niemeyer.
- KLUGE: see Seebold 2002.
- Koivulehto, Jorma: see Ruppel (ed.) 1999.
- Krahe, Hans
 1967 *Germanische Sprachwissenschaft*. III. *Wortbildungslehre* by Wolfgang Meid. Berlin: De Gruyter.
- Meid, Wolfgang: see Krahe 1967.
- Noel Aziz Hanna, Patrizia (ed.)
 2003 *Europa Vasconica-Europa Semitica*. (Trends in Linguistics, Studies and Monographs, 138). Berlin: De Gruyter. (a collection of articles by Theo Vennemann gen. Nierfeld.)
- Pokorny, Julius
 1959 *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.
- Ruppel, Klaas Ph. (ed.)
 1999 *Verba mutuata. Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis relinquerint*. (= Kleine Schriften of Jorma Koivulehto). Helsinki: Finnisch-Ugrische Gesellschaft.
- Seebold, Elmar
 1968 Erhaltung und Schwund des Kompositionsvokals im Gotischen. *Zeitschrift für vergleichende Sprachforschung* 82: 69–97.
- Seebold, Elmar
 1970 *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*. The Hague: Mouton (Diss. Tübingen 1964).
- Seebold, Elmar
 2002 KLUGE. *Etymologisches Wörterbuch der deutschen Sprache*. 24. edn. (23. edn. 1995) Berlin: De Gruyter. [Now: 25. edn. 2011]
- Sütterlin, Ludwig: see Kluge 1926.
- Torp, Alf und Hjalmar Falk
 1909 *Wortschatz der germanischen Spracheinheit*. Göttingen: Vandenhoeck & Ruprecht.
- Vennemann gen. Nierfeld, Theo: see Noel Aziz Hanna (ed.), 2003.

Elmar Seebold, Munich (Germany)

58. The dialectology of Germanic

- | | |
|--------------------------------------|------------------|
| 1. Dialects and historiography | 4. Abbreviations |
| 2. Dialects in the earliest evidence | 5. References |
| 3. Dialects and language grouping | |

1. Dialects and historiography

1.1. Tribal history

Several literary sources of Antiquity describe the subdivision of the Germanic tribes, but they do it in different ways: Pliny lists five subgroups and in some cases he also mentions their seats as well as single tribal names belonging to these groups. The groups are called *Vandili*, *Inguaeones*, *Istuaeones* (near the Rhine), *Hermiones* (in the inner lands) and *Peucini Basternae* (near the Dacians, i.e. near the lower Danube/Black Sea). The *Vandili*, who from other sources are known to have lived in the east of the Gmc area, consist of the *Burgodiones*, *Varinnae*, *Charini*, *Gutones*, names which are likewise attested in the east. The *Inguaeones* include the *Cimbri*, *Teutoni*, and *Chauci*, tribes which were supposed to have lived near the North Sea; and the *Suebi*, *Hermunduri*, *Chatti*, and *Cherusci* belong to the *Hermiones*. For the *Istuaeones* and *Peucini Basternae* there are no further names mentioned.

Tacitus, half a century after Pliny, mentions only three subgroups: *Ingvaeones*, *Istvaeones* and *Herminones*. His account puts the names in a mythical context by tracing them back to the names of three eponymous ancestors, who in their turn are sons of *Mannus* ‘man’. Tacitus is aware, however, that this grouping of the Gmc tribes is not fully accepted and quotes another (anonymous) view of Gmc grouping with entirely different names, i.e. *Marsi*, *Gambrivii*, *Suebi* and *Vandilii*. The importance of this alternative grouping is corroborated by the statement that the latter names are the “true and old names” (*vera et antiqua nomina*). The most interesting fact is, however, that the earlier attested story of Pliny appears to be a synthesis of the two competing versions by Tacitus.

The accounts of Pliny and Tacitus have had much influence on how the history and language of Gmc peoples have been looked upon up to modern times (cf. 3.1). The earliest written sources about Germanic peoples rarely deal with language, but usually with tribal history and culture. This has had consequences for the dialectology of the Germanic languages insofar as scholars, too, have confounded historical, archaeological and linguistic data. Therefore, several theories of Germanic dialect grouping are strongly influenced by those accounts, some of them being even modified or interpreted versions of ethnic arrangements in ancient literature. Many linguistic theories, on the other hand, were rejected because of their seeming incompatibility with historical data (cf. Nielsen 1985: 11 ff.). Moreover, dialectal concepts based on the accounts of Pliny and Tacitus had another problem, namely that their groups only seemed to comprise continental peoples. Both Pliny and Tacitus mention some Scandinavian tribes, but they do not integrate them into their schemes of ethnic relationship. This became an issue for all models that relied on ethnographical information (cf. 3.1).

There are also other authors who mention some kind of grouping or relationship of Gmc tribes, but not as explicitly as Pliny or Tacitus. Some of these statements make their fictitiousness very obvious; Ammianus Marcellinus, for instance, notes that the Burgundians themselves believed that they had originated from the Romans. It is not until the beginning of the Middle Ages that some sources write in depth about the origin and filiation of certain tribes, as for instance the Scandinavian origin of the Goths (Jordanes) and the Langobards (*Origo gentis Landobardarum*, Paulus Diaconus). These stories, however, are not very reliable and, sometimes, they may be considered more or less pointless fiction.

1.2. Methodological aspects

The Goths played an important role in ancient historiography from ca. 100 CE onwards. The Gothic bible translation dating from about 300 years later is ascribed to a small part of the Visigoths called *Gothi minores* ‘smaller Goths’ (the Visigoths themselves being a tribal branch of the Goths), but was later on taken over by all Gothic peoples. Therefore, we can conclude that the *Gothi minores* – minorities subtracted – had a common language, and the Gothic peoples of the 5th century spoke dialectal variants of the same language, which is quite legitimately called the Gothic language. Problems arose when earlier generations of philologists applied this idea to tribal entities without extensive attestations of their language, reconstructing a family called East Gmc which included all neighbouring tribes (and thus languages). The East Gmc character of Burgundian and Vandalic, for example, while plausible, is still a subject of controversy (cf. 3.5). It is even more problematic to postulate a specific Marcomannic or Semnonic language for tribes called *Marcomanni* or *Semnones*. The problem persists if we combine diachronic and diatopic evidence. The name *Langobardi*, for example, is attested around the lower Elbe near Hamburg in the 1st century CE; in the following centuries, this tribe migrated through Eastern Europe until they invaded Italy in 568 CE. In many respects the few attested Langobardic words (most of them from about 1–2 centuries after the invasion of Italy) are quite similar to OHG, but they lack some specific OHG features.

On the basis of the early tribal seats, some scholars attributed the Langobardic language to Ingvaenic (North Sea Gmc), others, however, to Erminonic (Elbe Gmc). The obvious linguistic relationship with OHG, on the other hand, was partly explained as common heritage of Elbe Germanic, whereas others considered it as a result of a secondary OHG-Lang. “sprachbund”. The fact that they had for a while been neighbours of East Germanic tribes, such as the Goths and Gepids, left its marks especially on proper names; yet some scholars even considered Langobardic to be a genuine East Gmc dialect. Their opponents emphasized its North Gmc features and referred to medieval records which claimed Scandinavia as the original homeland of the Langobards (for this discussion cf. Bruckner 1895: 24 ff.; Frings 1932: 32; Maurer 1952: 49 ff.; Schwarz 1951: 233 ff.; Kuhn 1955: 1 ff.).

1.3. Gmc language in ancient sources

If language becomes an issue, ancient sources usually talk about “the Germanic language”, which as a uniform entity seemed to be easily recognizable and was distinguished from non-Gmc languages. Internal variation seemed to be of little interest. Furthermore, most authors (and their sources) obviously did not know enough about the Gmc language to be good judges. Therefore, scholars hesitate to believe that the dialect of the *Aestii*, who belonged to the Gmc subgroup *Suebi* and whose name lives on in modern *Estonian*, was similar to the Brythonic language (*lingua Britannicae proprior*), as Tacitus claims. Rather, it is considered to be a Finnish or Baltic dialect. If there are any hints of dialectal differences at all, they are very well hidden: Tacitus’ statement about the *Marsigni* and *Buri*, who according to their language and culture belong to the *Suebi* (and not to the *Germani*), has been interpreted as referring to a subgroup of the Germanic language. Similarly, he groups the *Canninefates* on the Lower Rhine with the neighbouring Batavians (and not the *Germani*) on account of their language (besides their virtue and origin), which might indicate a dialect of Germanic. The evidence, however, is slender.

Moreover, the fact that in some cases very important linguistic differences were not noticed at all highlights the danger of relying on such kind of evidence. The Cimbrians are usually counted among the Germanic tribes – also by Pliny. Yet, he claims that the Cimbrian name for the ‘dead sea’ in the North is *Morimarusa*, which indeed means ‘dead sea’, but in Gaulish, not Germanic. Pliny also states that the word *sapo* ‘soap’ (**saipjō-*) is Gaulish, although it is most probably Gmc (cf. 2.2).

From the migration period onwards, there is more evidence available, but even then it is often unclear what linguistic entities the sources refer to. For example, Agathias enigmatically claims that in former times the Franks were well known under the name of *Germanoi*, but that in his own time they only distinguish themselves from the Romans by their “strange language”. Agathias seems to include different tribes under Franconian rule. According to Procopius, all “Gothic tribes” (mainly, but not only, Goths, Vandals, Visigoths and Gepids) only differ in their names, and they have the same language, too, which is called the “Gothic language”. It would be daring, though, to conclude from this an EGmc entity.

2. Dialects in the earliest evidence

2.1. Proto-Gmc as an IE language

Gmc belongs to the Western group of IE languages and differs clearly from related language groups such as Italic or Celtic. To this day no agreement has been reached about the question as to where Gmc should be located within the IE family tree, since its historical neighbour languages, i.e. Celtic, Italic, Slavic, and Baltic, are not always regarded linguistically as its closest relatives. Furthermore, the question of what should be considered the main distinctive characteristic of Germanic is debated. The First Consonant Shift (Grimm’s Law) has often been regarded as such a defining criterion, but it probably is only the most obvious one, as it affects often-quoted etymologies like Latin

pater vs. English *father*. Other changes have altered the appearance and especially the system of Gmc much more profoundly (Ringe 2006: 67 ff.).

Important phonological changes are:

- Grimm’s Law
- Verner’s Law
- Loss or vocalization of laryngeals
- Vowel merger $\vartheta, a, o > a$ and $\bar{a}, \bar{o} > \bar{o}$
- Svarabhakti: $j, l, \eta, \eta > ur, ul, um, un$

Accent fixation on the first syllable or stem syllable, respectively

Important morphological changes include:

- Weakening and partial loss of non-initial and especially final syllables
- Incipient convergence of nominal classes, some even disappearing before the transmission of Goth
- Heavy reduction of the categories of verbal inflection
- Emergence of a weak preterite with dental suffixes

One aspect in which Gmc has conserved IE structures quite well is the ablaut system of the strong verb: the present tense primarily continues the IE thematic (non-reduplicated) present stems; the past tense or preterite of the Gmc strong verb, on the other hand, preserves basic structures of the IE perfect.

Gmc has sometimes been referred to as a homogeneous proto-language, which only split up into various dialects at a later time. In view of the data, however, this is rather unlikely, since there are traces of old dialectal differences. There are two principal explanatory models for the historical relationship of languages: the family tree model of August Schleicher and the wave model of Johannes Schmidt. The family tree only considers the genetic background of language history; the wave model, on the other hand, focuses on geographical proximity and linguistic interference between the Gmc subgroups (cf. van Pottelberge 2003). The most adequate way of applying these conceptualizations today is to combine the family tree as a model for visualizing relationship and the wave model for explaining secondary convergence. These two can be complemented by a third mechanism, which is called *drift*, the phenomenon whereby parallel linguistic processes often appear in genetically related languages subsequent to their separation, apparently owing to their common past. This aspect was elaborated into the so-called “*Entfaltungstheorie*” (Höfler 1955–1956). Umlaut processes or *Auslautgesetze*, for instance, occur in the individual languages but are often regarded as consequences of the accent shift. It has always been one of the central problems of Germanic dialect grouping that such models were confused with historical reality; however, they do not explain but only visualize.

2.2. Gmc names and words in ancient literature

The earliest attested records of the Gmc language are words and proper names in classical literature, some of which give indications of Gmc phonology or morphology and can often be linked to lexical units of later times. Some of these items may be traced back

to the Greek seafarer Pytheas, who undertook a journey to the North Sea in the 4th century BCE. Definite evidence, however, begins in the 1st century BCE, when the Romans, beginning with Caesar, established a common border with Germanic-speaking peoples. Some examples of early recorded appellative words are (Green 1998: 182 ff.):

alcis ‘kind of a stag’: **alhiz* (?) this word is related but probably not identical to ON *elgr* (< **algiz*, showing Verner’s Law) and OHG *elahho* < **elhōn-*. The singular form of this word is perhaps preserved in the title of an East Gmc priest, *Alci(s)*, which is recorded by Tacitus. As in Latin, the Gmc word might be an *i*-stem which would make it closest to the ON cognate.

urus ‘aurochs’: **ūran* or **ūraz*, OE *ūr*, ON *úrr*, OHG *ūro*; probably related to German *Auer(-ochs)*.

bruta or *brutes* ‘bride’ (in several inscriptions): probably a loanword from Gmc **brūdiz* ‘bride’ as in Goth *bruþs*, OE *brȳd*, OS *brūd*, OHG *brūt*, ON *brúðr*, all meaning ‘bride, young woman’; heavily influenced by Latin morphology.

glesum or *glaesum* ‘amber’: **glēza* as in MLG *glār* ‘resin’ and perhaps OE *glær* ‘amber’; most of the younger cognates continue a variant **glasa-* without Verner’s Law and without long-grade ablaut vowel.

sualiternicum ‘amber’: is probably a mistake for **sualitervium*, the Gmc form of which is **swali-terwja* ‘burn-resin’ (Neumann 1993). The first element **swali-* contains a verbal stem which is continued in OE OHG *swelan* ‘burn, glow’. The second element **terwja-* is derived from the Gmc word for ‘wood’, **teru-* (cf. NE *tar*).

sapo ‘hair dye’: obviously mistaken from Gmc **saip(j)ō* (thus the early Finnish loanword *saipio*) which is continued in OE *sāpe*, OHG *sei(p)fa* ‘soap’.

ganta ‘goose’: **gantān-* as opposed to **gans(iu)-* of most of the Gmc languages (OHG *gans*, OE OFr. *gōs*).

framea ‘spear’: this might be a derivation **fram-ja-* (a pertinentive derivative of the preposition **frama* ‘forward, from’), but more likely a compound **fram-ij-an-* ‘forward going one’ (with IE **h₁ei-* ‘go’ as second element). In either case, a corresponding verb seems to be continued in ON *fremja*, OHG *fremmen*, OS *fremmian* ‘carry out’ < **fram-jana-*. Semantically, the word is fully comparable to forms found in runic inscriptions like those from Øvre Stabu (*raunijaz* ‘tester’), Kowel (*tilarids* ‘goal rider’) or Dahmsdorf (*ranja* ‘runner’), some of which seem even similar from a morphological point of view.

At least as relevant (and more frequently attested) are proper names, such as personal, tribal, and place names, which give us some insight into the structure of early Gmc. Such information, however, can be misleading, as it depends on assumptions (e.g. on etymology) which often are as uncertain as the desired result. Altogether, these words and names indicate that at least the phonology and most probably the (nominal) morphology as well remained on the level of Proto-Gmc. Yet, as most of these words are recorded in Latin texts, Latin interference in phonology and morphology has to be taken into consideration.

Consonants: The early texts already show the Gmc Consonant Shift as, for example, in *Harigasti*, *Chariovalda*, *Harii*, *Charini* (< IE **korjō-*, **korino-*) and Verner’s Law as in *Venadi* (< IE **uenHtō-*?), the latter perhaps with vocalization of the laryngeal as well (cf. however Müller 2007: 147 ff.). Svarabhakti of IE resonants turns up fully developed (*Burgundii* < **b^hrg^hnt-*), but the loss of nasals before *h* (*Vnh* > *V:h*) does not seem to be completed yet, as the tribal name *Tencteri* is usually interpreted as Gmc **Penhterōz*

(concerning *-anh-* in the Burgundian personal name *Hanhavalduis* cf. 3.5). Since the 4th century ethnic name *Salii* < **Saljōz* shows no traces of WGmc. gemination, it seems likely that this sound change was not completed before the end of the 4th century (Wagner 1989; cf. however 2.3. [*kunni*]). Furthermore, there are no traces of rhotacism during the Roman Imperial Age.

Vowels: IE *o* > *a* has been completed everywhere with the exception of the thematic vowel in composition, often in connection with a second element starting with a labial (e.g. *Lango-bardi* < PIE **dlong^ho-*; but cf. Marchand 1959: 172 f.). On the other hand, Gmc *e* is preserved in all positions, even before *i*, *u* and *nC*, and Gmc *i* as well as *u* are not lowered before *a* (*a*-umlaut). Proto-Gmc *ā* might still be preserved in Caesar's *silva Bacenis* (if this relates to OS OE *bōc*, OHG *buohha* 'beech'; cf. Lat. *fāgus*), and the lowering of Proto-Gmc *ē₁* to *ā* cannot be dated before the 3rd century (Lat. *Suebi* ~ Gmc **Swē₁bōz*; Lat. *Inguioimerus* ~ Gmc **Ingwjamē₁ri/jaz*; cf. 3.2).

Morphology: It is difficult to judge morphology from Gmc words in ancient sources, as in most cases the inflection of foreign names follows the rules of Latin (or Greek) grammar. Therefore, it is quite unclear how much one can trust a Latin *n*-stem, for example, to indicate a Gmc *n*-stem (it is certain in *Gutones* = Gmc **Gutanez* 'Goths'; cf. Pietroassa gen. pl. *gutani* ~ Biblical Goth *gutane*). In some cases, however, Gmc tribal names are recorded with alternating inflectional stems in *-n-* and *-o-*, which perhaps reflects an alternation of Gmc *n*- and *a*-stem inflection (*Burgundii* : *Burgundiones*, *Lugii* : *Lugiones*, *Franzi* : *Francones*). The etymology of these names points to adjectives or participles (**burgund(ja)-* 'being high'; **lugja-* 'mendacious' or 'sworn [companion], confederate'; **franka-* 'aggressive, bold'). The alternation very likely indicates that the Gmc double adjectival declension already existed. A specific problem is posed by masculine personal names in *-a* like *Nasua*, *Catvalda*, *Chariovalda*, which in turn have been interpreted as WGmc *a*-stems with loss of *-z* or as *n*-stems; some scholars have even assumed masc. *ā*-stems as in Lat. *poetā* (cf. 3.2; for a discussion see Marchand 1959; Krause 1971: 19; Boutkan 1995: 49 f.; Nielsen 2000: 166 f.; and Reichert 2003).

Inscriptions from the Lower Rhine bear matron names, i.e. names of female goddesses (*matronae*) with partially Gmc etymologies (Mees 2006). The ending *-ims* on some of them, as for instance *Aflims*, *Vatvims*, alternates with the Lat. *-iabus* (*Aflims* vs. *Afliabus*) and has therefore been interpreted as a Gmc dat. pl. ending < PGmc **-imiz*. If this is correct, these forms indicate that the loss of final *-z* in WGmc is later than the 3rd century (cf. however 2.3 and 3.3 f.), yet the evidence of *ka[m]ba* < PGmc. **kambaz* on the newly-found comb of Fienstedt (ca. 300) indicates loss of *-z* before that date (Schmidt et al. 2011: 141 ff.).

2.3. The early runic language

The first texts written in Germanic are runic inscriptions concentrated in modern Denmark dating back to the 2nd century CE. The language of these inscriptions is very close to Proto-Germanic. It is called Primitive Norse, Proto-Scandinavian or "Urnordisch", since it was originally thought to be the direct ancestor of the medieval and modern Scandinavian languages. This view, however, has been challenged in recent decades. The rather uniform character of this language attested from about 200 to 500 has raised

the question as to whether the idiom of the early runic inscriptions might not be best regarded as a kind of a *Koine* (Makaev 1996). In comparison with Proto-Gmc, however, these inscriptions show some linguistic differences, most of which live on both in later WGmc languages and in Old Norse. The differences particularly concern the vowel system. Thus Gmc \bar{e}_1 appears as \bar{a} in *wajemariz* (Thorsberg around 200 CE, as opposed to Goth *wajamerjan* ‘blaspheme’). Unstressed Gmc *ai* partly becomes \bar{e} as in *tawide* (Illerup, ca. 200) or *woduride* (*a*-stem dat. sg.; Tune ca. 400), but not in *talgidai* (Nøvling, 3rd century) or *hahai* (Möjbro, 5th century); but cf. Nedoma (2005). The ending of *woduride* matches OHG and ON *-e*, but not OE (and partly OS) *-a*. Short *u* is lowered to *o* before mid and low vowels as in *horna* < PGmc **hurnaⁿ* (Gallehus ca. 400); and *holtijaz*, where *o* is due to analogy with **holta*, shows that *o* must already have had phonemic status. Gmc *e*, on the other hand, is partly preserved before *u* and *i*, cf. *leugaz* (Skåang, ca. 500), *erilaz* (several instances). Loss of nasals before *h* might have taken place during this period, judging from records such as *hahai* (Möjbro). If the evidence is trustworthy, the specific Norse innovation of sharpening had not yet taken place (Krause 1971: 32 ff.). The first NGmc sound change to show up in runic inscriptions is the monophthongization of *ai* to \bar{a} before *h* in *fahido* ‘I painted’ (Rö, ca. 400) vs. older *faihido* (Einang, end of 4th century).

Most of the attested forms could at the same time be predecessors of ON and of their WGmc equivalents. That is why the language of the runic inscriptions has been labelled North-West Gmc by some scholars since Kuhn (1955: 24 f.). There are even forms with counterparts in WGmc alone; for example, *asugisalas* (Kragehul, ca. 500) and *godagas* (Valsfjord, ca. 400), which form their gen. sg. in *-as* like the OE *a*-stems. Final *-z* in unaccented position, on the other hand, very likely disappeared quite early in WGmc. or in some of its varieties, as shown above (2.2). In many runic word forms, however, it is well preserved, e.g. *holtijaz*, *leugaz*, *erilaz*. Others look as if the loss of final *-z* had already taken place, for instance in *alugod* (Værløse, ca. 200), *swarta* (Illerup), *lagupewa* (Illerup), and perhaps even in *harja* (Vimose, ca. 160). Admittedly, all these cases are quite doubtful since they can be interpreted as neuters, as weak stems, or as non-nom. sg. forms (cf. Nielsen 2000: 149 f.; Looijenga 2003: 94 ff.; Kortlandt 2006).

In the weak declension, the ending of the masc. nom. sg. has often been regarded as an indicator of dialectal affiliation. Consequently, names with masc. nom. sg. forms in *-a* were regarded as EGmc (Goth *hana*), *-o* as WGmc (OHG OS *hano*) and *-e* or *-i* as ON (Oic. *hani*); for a more recent attempt to explain this alternation within Proto Norse cf. Nedoma (2005). However, it is in many cases difficult to decide whether a name is masculine at all.

Clear WGmc evidence in the early runic corpus is rare: distinct features of WGmc such as the ending of the 2sg. pret. (type OHG OS *nāmi* vs. Goth ON *nami*) cannot be found in the early runic inscriptions; *j*-gemination, however, shows up in *kunni* (Weser bones, probably around 400; for a possible terminus post quem cf. 2.2[*Saliū*]). Obviously WGmc dialect variants can be found from the 6th to the 8th century in the Frisian, Anglo-Saxon and South Gmc corpus, the latter lasting for only a few generations during the 7th century.

There are, however, obvious EGmc inscriptions with word forms that sometimes recall Biblical Goth structures like *tilarids* (Kowel, early 3rd century) < **tilarīdaz* with syncope and “*Auslautverhärtung*”, *gutani* Pietroassa (around 400; cf. 2.4). The EGmc character of other inscriptions is less certain, e.g. *ranja* (Dahmsdorf, 3rd century).

2.4. Biblical Gothic

Apart from the first few runic inscriptions and some proper names, Gothic is the earliest attested Gmc language, but its corpus is significantly larger than that of the runic inscriptions. It consists, above all, of the bible translation by the Visigothic bishop Wulfila (4th century CE). Wulfila created a special script for that purpose, consisting of Greek and runic characters. Besides the bible, there are several smaller texts, fragments and even some runic inscriptions.

While Gothic is, in most respects, less conservative than the earliest runic language, it shows some more archaic features, especially with respect to the vowel system. Gothic preserves \bar{e}_1 , for example, in the plural preterite of strong verb classes 4 and 5 as in *nēmum, sētum* ‘they took, sat’ (OS *nāmun, sātun*), and there is no *a*-mutation (*gulþ* ‘gold’ < **gulþa-* vs. NGmc WGmc **golþa-*). Like runic, Gothic still keeps Gmc *-z*, whereas the medieval languages and probably the runic language after around 700 have *r* (rhotacism). On the other hand, Goth underwent several innovations that characterize its appearance quite noticeably. Every PGmc *e* became *i* except before *r, h, hv*, where it resulted in *ai* [ɛ] (similarly *u* > *au* [ɔ] / *_r, h, hv*). Even more conspicuous is the so-called sharpening (German *Verschärfung*), in which the semivowels *j* and *w* turned into *ddj* (*twaddje* gen. pl. ‘of two’ < **twajje*) and *ggw* (*triggwus* ‘faithful’ < **triwvus*). Word-final syllables are weakened by several processes, including syncope and devoicing of final fricatives (*dags, gasts* < **dagaz, *gastiz*); devoicing occurs also in syllable-final position; therefore, the results of Verner’s Law are levelled in many cases (*hausjan* ‘hear’ < **hauzjana-*). Similarly, Thurneysen’s Law causes voiced spirants in syllable-final position to become voiceless if the initial sound of a syllable is voiced and vice versa. Concerning morphology, a reasonable comparison between Gothic and early runic is not expedient since especially relevant forms are only poorly attested in the runic inscriptions.

By contrast, Gothic is the only Germanic language to have kept the dual inflection of the verb (Fritz 2011: 154 ff.), a synthetic passive (usually called medio-passive), a fully functioning 4th class of weak verbs in *-nan*, as well as reduplication in the 7th class of strong verbs. Furthermore, Gothic still has a vocative case in most noun stems. Syncope, on the other hand, makes the inflectional forms look quite “modern” in comparison with the older runic inscriptions for instance (Goth *gasts* vs. runic *gastiz*; cf. Van Bree 1998). There is one feature of the Gothic verbal system about which there is a strong controversy as to whether it is an archaism or innovation: the long pret. pl. endings *-dedum, -dedup, -dedun* of the weak preterite (Ringe 2006: 167 f.; Hill 2010).

3. Dialects and language grouping

3.1. Tripartite subgrouping

Secondary literature today usually describes an early tripartite grouping into North, West, and East Gmc. Tripartite genealogies, moreover, have had a long tradition since Antiquity and the Middle Ages and can also be found in biblical explanations.

The concrete tripartite model of the Gmc language family, however, was first introduced by Schleicher (1869: 91), the “creator” of the family tree model itself (cf. 2.1). It is

interesting though that in the case of other IE subgroups Schleicher had a strong “binary” component in his family tree (most nodes in his model break up into two branches). The tripartite grouping of Gmc is of course not only due to linguistic features, but also to the mythological genealogy reported by Tacitus (cf. 1.1). Linguistic data, however, were quoted to partly corroborate the assumption of a tripartite family tree. The largest part of the data is taken from the primary corpus languages, some however from less reliable material such as onomastic data.

The East Gmc language branch is assigned to the Gmc tribes that settled in the Eastern part of the Germania during the Roman period. All languages of the East Gmc branch are extinct, their attestations rather sparse. As there is only the Gothic Bible as evidence of any considerable length, EGmc has to be defined on the basis of Gothic, which, in comparison to other Gmc corpus languages, shows some archaic features (cf. 2.4 and 3.5). The main innovations of EGmc are $e > i$ (Goth *giba* ‘gift’ vs. OHG *geba*) and similar heightening tendencies for $\bar{e}_1 > \bar{i}$ and $\bar{o} > \bar{u}$ (the former more pronounced than the latter), sharpening (Goth *twaddje* ‘of two’ vs. OS *tueio*), *Auslautverhärtung*, and extensive weakening of final syllables, resulting in syncopated inflectional forms such as Goth *dags* ‘day’ < **dagaz*.

The North Gmc branch is located in Scandinavia and has a long transmission history, beginning from the earliest runic inscriptions (cf. 2.3). The main language, however, is Old Icelandic, one of the latest attested medieval Gmc dialects. North Gmc is considered to be a rather archaic branch, and its main features developed quite late, i.e. after the extinction of the East Gmc languages (on early $ai > \bar{a} / _ h$ cf. 2.3). Its main characteristics are: loss of initial *w* before *o* and *u* (OHG *wurtum* vs. ON *urðu*), loss of initial *j* (OHG *jār* vs. ON *ár*), preterite participle with *-ina* instead of *-ana* (ON *gefinn* vs. Goth *gibans* ‘given’; cf. however Nielsen 1989: 8 f.) and sharpening (ON *tveggja* ‘of two’ vs. OS *tueio*). Medieval NGmc was strongly exposed to syncope, which, on the one hand, caused several umlaut and breaking rules in stressed syllables and, on the other hand, made the nominative sg. of *a*- and *i*-stems appear quite similar to those of Goth (ON *dagr* ‘day’ vs. Goth *dags*).

The West Gmc branch, finally, is designated for those areas that were closest to the Roman Empire during the Imperial Age, i.e. mainly the Netherlands, Germany, and Southern Jutland. The situation is more complicated than in East or North Gmc, insofar as West Gmc does not have one “reference language”, but at least four, namely OE, OS, OFr, and OHG. These medieval languages are already separated from each other by many significant differences and only kept together by a few common features, the most important of which are: the West Gmc gemination of consonants before *j* (OE *cynn*, OHG *kunni* ‘kin’ vs. Goth *kuni*), replacement of the original 2sg. pret. ending (OHG OS *nāmi* vs. Goth ON *namt*), and the loss of final *-z* in unstressed syllables (OS *dag* vs. Goth *dags*) cf. 4.2 [runic]).

3.2. Bipartite subgrouping

One of the earliest attempts at Gmc dialect grouping can be found in Adelung (1809: 175). Adelung, much earlier than Schleicher, sketched a Proto-Gmc family splitting up into two unequal parts, of which one branch consisted of High German only, whereas

all other languages together belonged to the second branch. The first scholar to express this approach in systematic terms was Rask (1818: 63 ff.), who modified Adelung's view by combining High German with Gothic and opposing them both to the Scandinavian languages. It must be added, though, that these early scholars were more interested in tribes and tribal history than in languages, or at least mixed up both aspects (cf. 1.1).

Theories after Schleicher partly combine the binary aspect of a two-branch tree with the idea of a final tripartite grouping. Schwarz (1951: 47 ff.) grouped EGmc and NGmc into a Gotho-Nordic (*gotonordisch*) branch as opposed to WGmc. According to his theory, the Gotho-Nordic branch later on split up into EGmc and NGmc, while WGmc continued to exist. Traditions about the Scandinavian origin and prehistoric migration from Scandinavia to the Continent – especially of Goths and Langobards – seemed to confirm this theory. Schwarz considered common innovations (e.g. sharpening) as well as common archaisms (e.g. 4th class of weak verbs, nom. sg. marker PGmc *-z) to draw a picture in which Gothic played a quite prominent role compared to Rask's model.

Problems in Schwarz' theory apply to both linguistic and extralinguistic arguments. Apart from the fact that common archaisms do not generally carry conviction, the concrete data are doubtful, too. During the transmission time of Gothic, when *a-* and *i-* stems had already lost their stem vowel before the ending *-s* of the nom. sg., the Scandinavian inscriptions still preserve the full *-az/-iz* ending. Goth and ON sharpening, which Schwarz used as another argument for a Gotho-Nordic branch of Germanic, did in fact not take place at the same time in both languages. Moreover, even though sharpening describes a similar process in both languages, its results are in part different (Goth *twaddje* vs. ON *tveggja*; Petersen 2002). Furthermore, the early medieval literary sources that claim a Scandinavian origin for both Goths and Langobards have been strongly disputed.

Most scholars have therefore rejected the Gotho-Nordic model, the first being Kuhn (1955: 8 ff.), who was of the opinion that the parallels between WGmc and NGmc exceeded those between EGmc and NGmc (or rather Goth and ON) and posited a North-West-Gmc unity attested as such in the early runic inscriptions (cf. 2.3). This model of “*Ausgliederung*” looks like a mirror image of the Gotho-Nordic one: Gothic separated from the PGmc continuum first, leaving some kind of remnant Gmc or North-West Gmc behind. Several scholars have followed this idea, some of them with smaller modifications (Antonsen 1975; Makaev 1996; Penzl 1996; Markey 1976: 24). NWGmc is characterized mainly by the following features: $\bar{e}_1 > \bar{a}$ (OHG *jār*, ON *ár* vs. Goth *jer* ‘year’), rhotacism $z > r$ (OHG OS *mēro*, ON *meiri* vs. Goth *maiza* ‘more’), and phonologically conditioned lowering of *i* and *u* (OE OFr OS OHG *hold*, ON *holtr* vs. Goth *hulþs* ‘gracious’).

There are many arguments in favour of the NWGmc theory, but there are also some objections (e.g. Grønvik 1998: 70 ff.), the most important one being the WGmc loss of final *-z*. This argument, however, is somewhat circular, and the assumption of an early date for this loss remains a topic of debate (cf. 2.2 and 2.3). Other arguments such as, for example, the opposition of the personal pronoun NGmc *ek* vs. WGmc *ik* can be put into perspective by discoveries of NGmc *ik* and WGmc *ek* (Nielsen 2000: 158). Antonsen (1975: 26 f.), as a solution, suggested that NWGmc covered only the WGmc area to the exclusion of Pre-OHG. This was a step in the right direction, as it comes closer to the reality of a dialectal continuum and dismisses the idea of a homogeneous language area extending over large parts of Europe.

Vennemann (1984) presented a completely different view of the language history and dialectology of early Gmc which was, however, similar to the early attempts insofar as it resulted in a division of northern vs. southern Gmc. Vennemann's idea was based on the "glottalic theory", which reconstructed a system of PIE glottalized stops $*p'$ $*t'$ $*k'$ $*k''$ instead of the traditional $*b$ $*d$ $*g$ $*g''$ (Gamkrelidze-Ivanov 1995: 5 ff.). According to Vennemann, there was only one consonant shift instead of two. This single consonant shift "bifurcated", i.e. had different results in pre-OHG and the rest of the Gmc speaking area and therefore did not separate High and Low German, but High and Low Germanic. This theory has met many objections, not least by its creator (Vennemann 2006).

3.3. Multipartite subgrouping

To bypass the problems that resulted from bipartite subgrouping models of Gmc, some scholars tried to find a solution in a multipartite family tree. Grimm was one of the first; he proposed several different views during his scholarly life, like his contemporaries mostly speaking about tribes instead of languages. Most of the newer theories deal with the notorious problem of West Gmc, either denying its existence or modifying the view of its historical reality (cf. 3.4). The main exponent here was Maurer (1952), and his main subject was the relationship between southern Alamanns and northern Scandinavians, which he considered to be closer in an early time than that between the Alamanns and the Frisians, Angles, or Saxons, for example. Maurer explicitly and as part of his methodological concept consulted non-linguistic evidence, i.e. mainly archaeological and historiographical data. Interestingly enough, the family tree depicted by Maurer, is reminiscent of Pliny's description of Germanic tribal grouping (cf. 1.1). This critical view of WGmc is still quite frequent in today's scholarship.

3.4. West Germanic issues

Schleicher's WGmc branch definitely does not represent a uniform subgroup, as there are some old differences between the languages. There seems to be a principal distinction between the northern and the southern part of this group; the demarcation between both parts, however, is a matter of controversy. The northern part, North Sea Gmc or Ingvaenonic, is the larger one, but it is a moot point whether Old Saxon and Old Low Franconian really belong to it, and if yes, to what extent they participate in all its characteristic developments.

North Sea Gmc characteristics are most clearly apparent in OE and OFr (Anglo-Frisian), but it is important to remember that the principal OE dialects – Anglian and West Saxon – sometimes differ considerably. The most important characteristics are (for a comprehensive overview cf. Markey 1976; Nielsen 1985; Krogh 1996: 141 ff.; Nielsen 2001):

1. Fronting of WGmc \bar{a} from Gmc \bar{e}_1 (OFr $j\bar{e}r < *j\bar{a}ra$), which took place only partially in OLF and even less regularly in OS.

2. Fronting of WGmc *a* (OE *dæg* OFr *dei* < **dag*^a), which took place only partially in OLF and OS.
3. Palatalization of WGmc *g* and *k* (OE *circe* OFr *tzierka* < **kirika*), which occurs a few times in OS, but is less certainly attested in OLF.
4. Loss of nasal before *f*, *s*, *þ* with compensatory vowel lengthening (OE OFr *tōþ* < **tanþu*); this development is also quite consistent in OS and somewhat less frequent in OLF.
5. Uniform inflectional forms occur in the acc. and dat. of the 1st sg. pronoun; this also holds true for both OLF and OS.
6. Uniform inflectional forms for the pret. pl., which also occur in OS, but not in OLF.

As a whole, there are arguments for a close relationship between Anglo-Frisian on the one hand and OS and OLF on the other; there are, however, counter-arguments as well. The question as to whether the common features are old and inherited or have emerged by connections over the North Sea is still controversial.

With respect to 1., Anglo-Frisian *ē* looks like a direct continuation of PGmc **ē*, which would be an argument against grouping OE and OFr with WGmc. It has been shown, however, that the Anglo-Frisian *ē* is a secondary product and that all languages had genuine WGmc *ā*.

As to 4., a similar process occurs in the Alemannic dialects in the South-West of the German-speaking area, ("Staub's Law"). Wrede (1924) thought that these dialects originally were closely linked to "Ingvaemonic" and that Bavarian (where loss of nasals does not occur) was secondarily Gothicized. Staub's Law in Alemannic, however, has to be seen as an independent process since it happened about 1,000 years later than in North Sea Gmc. Another link looks more promising, as NGmc also has early loss of nasals, even though it occurs only before *s* and, less regularly, before *f* (ON *gás* OE *gós* vs. OHG *gans* < PGmc **gans*- 'goose'). Moreover, this process seems to be quite early. It is, however, unclear, whether the lack of *n* in early runic *asu*- < **ansu*- (cf. 2.3) is due to phonological change or only to a runic convention that did not allow *n* to be written before homorganic sounds (Nielsen 2000: 247 ff.).

OS shares features of Anglo-Frisian and OHG and almost completely lacks individual characteristics. The linguistic position of OS (and OLF) between Anglo-Frisian and OHG has led to the opinion that both were mixed languages due to Franconian influence that affected them in different ways (Kuhn 1973). Early scholars even reconstructed a German proto-language ("neu-urdeutsch") with OS and OHG as its main branches (Förstemann 1896). In recent years, however, scholars have emphasized the North Sea Gmc character of OS again (Krogh 1996: 398 ff.).

3.5. East Gmc issues

East Gmc tribes such as the Goths or Vandals were the main players in the Germanic migration period. As Jordanes describes the Goths as originating from *Scandza* (which is believed to be a name for Sweden), this tribe was linked with Gotland or with Götland on the Swedish mainland (cf. 3.2). Subsequently, other EGmc tribes were linked to Scandinavia in a similar way, as e.g., the Burgundians from Borgund in Norway or the Danish island Bornholm (ON *Borgundarholmr* OE *Burgendaland*), or the Vandals from

Vendel in Sweden or Vendsyssel in Denmark. Indeed, Gothic in some respects matches East Norse features, e.g. the lack of combinatory lowering (*a*-mutation). It has to be added, however, that these East Norse developments occur at a much later time than the attestation of Gothic (Haugen 1982: 34 ff.; Scardigli 2005).

Gothic is the only corpus language of East Gmc and, therefore, also its referential language. Other supposedly EGmc languages such as Burgundian and Vandalic (not to mention Rugian or the like) are only sparsely attested. “EGmc characteristics” are therefore essentially identical with Gothic characteristics (cf. 2.4). Another language, Crimean Gothic, is attested about 1,000 years later than Gothic, Burgundian, and Vandalic, but, because of forms like *geen* ‘go’ and *goltz* ‘gold’, its EGmc character is not uncontested, despite its form *ada* ‘egg’, which shows a clear EGmc phonology (Stearns 1978; Grønvik 1983).

The early attestations of Burgundian and Vandalic are to a large extent made up of proper names, which as linguistic evidence are quite difficult to handle. Besides the onomastic evidence, there are some legal terms in Latin texts and some loan words. Earlier attempts to prove their EGmc character (Gamillscheg 1936), were rejected (cf. Beck 1978). However, today most scholars again agree with the EGmc character of Burgundian and Vandalic (Francovich-Onesti 2002: 133 ff.; Haubrichs/Pfister 2008).

Two main features are used as arguments in favour of the EGmc character of the proper names: 1. typical lexical elements, such as name constituents; 2. the endings fem. *-o* and masc. *-a* that correlate with the *n*-stem nom. sg. forms of Goth *guma* ‘man’ (masc.) and *qino* (fem.) ‘woman’. Furthermore, one Burgundian personal name from the end of the 4th century, *Hanhavaldu*, has attracted attention: *n* before *h* should have disappeared long before and was suspected to be an archaism and, at the same time, a dialectal feature of Burgundian. More probably, however, it is only a Latin writing for nasalized *q*.

Besides attestations in single words, there are some (very) short texts, three of which are 1. the ‘Domine miserere’, 2. the ‘Gothic epigram’, and 3. the runic inscription on the fibula from Charnay. 1. and 2. are believed to be Vandalic, 3. Burgundian. 3. contains a verb *uþfnþai*, which is interpreted as an equivalent of the Goth subjunctive form *finþai* ‘he/she may find’. 1. *froia arme* ‘Lord, have mercy’ corresponds to Goth **frouja armai*. This expression is spelled with Latin letters and the graphs <o> and <e> seem to indicate a monophthongal pronunciation of what in Goth is <au> and <ai>. However, the evidence of 2. and 3. points the other way, as both the word *eils* (~ Goth *hails*) in 2., written in the Latin alphabet, and *uþfnþai* in 3., written in runes, contain a diphthongal graphic sequence. These records, however, are quite uncertain: especially 1. and 2. might show Vulgar Latin interference.

4. Abbreviations

Gmc	Germanic	NGmc,	
PGmc,		NWGmc	North, Northwest
EGmc,			Germanic
WGmc	Proto, East, West	NE, NHG	New English, New High
	Germanic		German

Goth	Gothic	OS	Old Saxon
OE	Old English	OFr	Old Frisian
OHG	Old High German	OLF	Old Low Franconian

5. References

- Adelung, Johann C.
1809 *Mithridates oder allgemeine Sprachkunde*. Vol II. Berlin: Vosische Buchhandlung.
- Bandle, Oskar
1973 *Gliederung des Nordgermanischen*. Basel: Helbing and Lichtenhahn.
- Bandle, Oskar, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman (eds.)
2002–2005 *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*. Berlin: De Gruyter.
- Beck, Heinrich
1978 Burgunden. In: Beck et al. (eds.), vol. 4., 224–271.
- Beck, Heinrich, Dieter Geuenich, Hans Kuhn, Kurt Ranke, Heiko Steuer, and Reinhard Wenskus (eds.)
1968–2007 *Reallexikon der germanischen Altertumskunde, begründet von Johannes Hoops*. 2nd edition. 35 vols. Berlin: De Gruyter.
- Boutkan, Dirk
1995 *The Germanic 'Auslautgesetze'*. (Leiden Studies in Indo-European 4). Amsterdam: Rodopi.
- Van Bree, Cor
1998 The Strange Simplicity of Gothic. *Northwestern European Language Evolution* 34: 67–75.
- Bruckner, Wilhelm
1895 *Die Sprache der Langobarden*. Strassburg: Trübner.
- Förstemann, Ernst
1896 Alt-, mittel-, neuurdeutsch. *Zeitschrift für vergleichende Sprachforschung* 18: 161–186.
- Francovich Onesti, Nicoletta
1999 *Vestigia Langobarde in Italia (568–774). Lessico e antroponomia*. Rome: Artemide.
- Francovich Onesti, Nicoletta
2002 *I Vandali. Lingua e storia*. Rome: Carocci.
- Frings, Theodor
1932 *Germania Romana*. Halle: Niemeyer.
- Fritz, Matthias
2011 *Der Dual im Indogermanischen. Genealogischer und typologischer Vergleich einer grammatischen Kategorie im Wandel*. Heidelberg: Winter.
- Gamillscheg, Ernst
1936 *Romania Germanica*. Vol. 3. *Die Burgunder*. Berlin: De Gruyter.
- Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov
1995 *Indo-European and the Indo-Europeans: A reconstruction and historical analysis of a proto-language and a proto-culture*, 2 volumes. [Translation of *Indoevropskij jazyk i Indoevropejcy. Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokul'tury*. 1984. by Johanna Nichols]. Berlin: Mouton de Gruyter.
- Green, Dennis. H.
1998 *Language and History in the Early Germanic World*. Cambridge: Cambridge University Press.

- Grønvik, Ottar
1983 *Die dialektgeographische Stellung des Krimgotischen und die krimgotische cantilena*. Oslo: Universitetsforlaget.
- Grønvik, Ottar
1998 *Untersuchungen zur älteren nordischen Sprachgeschichte*. Frankfurt: Lang.
- Harbert, Wayne
2007 *The Germanic Languages*. Cambridge: Cambridge University Press.
- Haubrichs, Wolfgang and Max Pfister
2008 Burgundisch (Burgundian). In: Ulrich Ammon and Harald Haarmann (eds.), *Wieser Enzyklopädie der Sprachen Westeuropas*. Klagenfurt: Wieser, 73–80.
- Haugen, Einar
1982 *Scandinavian Language Structures. A Comparative Historical Survey*. Tübingen: Niemeyer.
- Hill, Eugen
2010 A case study in grammaticalized inflectional morphology: Origin and development of the Germanic weak preterite. *Diachronica* 27: 411–458.
- Höfler, Otto
1955 Stammbaumtheorie, Wellentheorie, Entfaltungstheorie. *Beiträge zur Geschichte der deutschen Sprache und Literatur (Tübingen)* 77: 30–66, 424–476, 78: 1–44.
- Howell, Robert B.
1991 *Old English breaking and its Germanic analogues*. Tübingen: Niemeyer.
- König, Ekkehard and Johan van der Auwera (eds.)
1994 *The Germanic Languages*. London: Routledge.
- Kortlandt, Frederik
2006 The Inflection of the Germanic *n*-Stems. In: *Northwestern European Language Evolution* 48: 3–7.
- Krause, Wolfgang
1971 *Die Sprache der urnordischen Runeninschriften*. Heidelberg: Winter.
- Krogh, Steffen
1996 *Die Stellung des Altsächsischen im Rahmen der germanischen Sprachen*. (Studien zum Althochdeutschen 29). Göttingen: Vandenhoeck & Ruprecht.
- Kufner, Herbert L.
1972 The grouping and separation of the Germanic languages. In: Frans van Coetsem and Herbert L. Kufner (eds.), *Toward a Grammar of Proto-Germanic*. Tübingen: Niemeyer, 71–98.
- Kuhn, Hans
1955 Zur Gliederung der germanischen Sprachen. *Zeitschrift für Deutsches Altertum und deutsche Literatur* 86: 1–47.
- Kuhn, Hans
1973 Altsächsische Sprache. In: Beck et. al. (eds.), vol. 1, 239–241.
- Looijenga, Tineke
2003 *Texts and Contexts of the Oldest Runic Inscriptions*. Leiden: Brill.
- Makaev, Enver A.
1996 *The Language of the Oldest Runic Inscriptions*. Translation of *Jazyk drevnejšix runičeskix nadpisej* (1965) by J. Meredig in consultation with E. H. Antonsen. Stockholm: Kunglinga Vitterhets Historie och Antikvitets Akademien.
- Marchand, James W.
1959 Names of Germanic Origin in Latin and Romance Sources in the Study of Germanic Philology. *Names* 7: 167–181.
- Markey, Thomas L.
1976 *Germanic Dialect Grouping and the Position of Ingvaëonic*. Innsbruck: Institut für Sprachwissenschaft der Universität.

- Maurer, Friedrich
1952 *Nordgermanen und Alamannen*. 3rd edn. Bern: Francke.
- Mees, Bernard
2006 Early Rhineland Germanic. *Northwestern European Language Evolution* 49: 13–49.
- Müller, Stefan
2007 *Zum Germanischen aus laryngalthoeretischer Sicht. Mit einer Einführung in die Grundlagen*. Berlin: De Gruyter.
- Munske, Horst Haider (ed.)
2001 *Handbuch des Friesischen / Handbook of Frisian Studies*. Produced in collaboration with Nils Århammar, Volkert F. Falktings, Jarich Hoekstra, Oebele Vries, Alastair G. H. Walker, and Ommo Wilts. Tübingen: Niemeyer.
- Naumann, Hans-Peter (ed.)
2004 *Alamannen und der Norden. Internationales Symposium vom 18.–20. Oktober 2001 in Zürich*. Berlin: De Gruyter
- Neumann, Günter
1993 Sualiternicum ‘Bernstein’. In: Klaus R. Grunda and Claus-Dieter Wetzel (eds), *Anglo-Saxonica. Festschrift für Hans Schabram zum 65. Geburtstag*. Munich: Fink, 431–439.
- Nedoma, Robert
2005 Urnordisch -a im Nominativ Singularis der maskulinen n-Stämme. *Northwestern European Language Evolution* 46/47 (= *Papers on Scandinavian and Germanic Language and Culture. Published in Honour of Michael Barnes on his Sixty-Fifth Birthday 28 June 2005*), 155–191.
- Nielsen, Hans Frede
1985 *Old English and the Continental Germanic Languages*. 2nd edn. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Nielsen, Hans Frede
1989 *The Germanic Languages: origins and early dialectal interrelations*. Tuscaloosa, AL: University of Alabama Press. (Translation and update of his 1979 *De germanske sprog.*)
- Nielsen, Hans Frede
2000 *The Early Runic Language of Scandinavia. Studies in Germanic Dialect Geography*. Heidelberg: Winter.
- Nielsen, Hans Frede
2001 Frisian and the Grouping of the Older Germanic Languages. In: Munske et al. (eds.), 512–523.
- Nielsen, Hans Frede
2004 Friedrich Maurer and the Dialectal Links of Upper German to Nordic. In: Naumann (ed.), 12–28.
- Penzl, Herbert
1996 Ist das nordisch-westgermanische Runengermanisch das ‘Ur-Altenglische’? *Northwestern European Language Evolution* 27: 137–145.
- Petersen, Hjalmar P.
2002 Verschärfung in Old Norse and Gothic. *Arkiv för Nordisk Filologi* 117: 5–27.
- Van Pottelberge, Jeroen
2003 Die ursprünglichen Fragestellungen hinter August Schleichers Stammbaum-Theorie und Johannes Schmidts Wellenmetapher. *Historiographia Linguistica* 30: 301–364.
- Rask, Rasmus K.
1818 *Undersögelse om det gamle Nordiske eller Islandske Sprogs Oprindelse* [An investigation of the origin of the ancient Scandinavian or Icelandic language]. Copenhagen: Gyldendal.
- Reichert, Helmut
2003 Personennamen bei antiken Autoren als Zeugnisse für älteste westgermanische Endungen. *Zeitschrift für deutsches Altertum* 132: 85–100.

- Ringe, Don
2006 *From Proto-Indo-European to Proto-Germanic*. (A Linguistic History of English 1). Oxford: Oxford University Press.
- Robinson, Orrin W.
1992 *Old English and its Closest relatives. A Survey of the Earliest Germanic Languages*. Stanford: Stanford University Press.
- Scardigli, Piergiuseppe
2005 Nordic-Gothic Linguistic Relations. In: Bandle et al. (eds.), 553–558.
- Schleicher, August
1869 *Die deutsche Sprache*. 2nd edn. Stuttgart: Cotta.
- Schmidt, Christoph G., Robert Nedoma, and Klaus Düwel
2011 Die Runeninschrift auf dem Kamm von Frieñstedt, Stadt Erfurt. *Die Sprache* 49: 123–186.
- Schwarz, Ernst
1951 *Goten, Nordgermanen, Angelsachsen: Studien zur Ausgliederung der germanischen Sprachen*. Bern: Francke.
- Seebold, Elmar
2004 Alemannisch und Nordgermanisch. In: Naumann (ed.), 1–11.
- Stearns, MacDonald Jr.
1978 *Crimean Gothic: Analysis and Etymology of the Corpus*. Saratoga, CA: Anma Libri.
- Vennemann, Theo
1984 Hochgermanisch und Niedergermanisch: Die Verzweigungstheorie der germanisch-deutschen Lautverschiebungen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 106: 1–45.
- Vennemann, Theo
2006 Grimm's Law and Loanwords. *Transactions of the Philological Society* 104: 129–166.
- Vennemann, Theo
1981 *Gothic, Germanic, and Northwest Germanic*. Wiesbaden: Steiner.
- Vennemann, Theo
1992 *Early Germanic Grammar: Pre-, Proto- and Post-Germanic languages*. San Diego: Academic Press.
- Wagner, Norbert
1989 Der Stammesname der Salier und die ‚westgermanische‘ Konsonantengemination. *Zeitschrift für deutsches Altertum* 118: 34–42.
- Wrede, Ferdinand
1924 Ingwäonisch und Westgermanisch. *Zeitschrift für deutsche Mundarten* 19: 270–283.

Ludwig Rübekel, Zürich (Switzerland)

59. The evolution of Germanic

- | | |
|---|---|
| 1. Introduction | 5. Syntactic patterns |
| 2. The range and internal structure of Germanic | 6. Conclusion: Conservatism and paths of innovation |
| 3. Phonological patterns | 7. References |
| 4. Morphological patterns | |

<https://doi.org/10.1515/9783110523874-014>

- Ringe, Don
2006 *From Proto-Indo-European to Proto-Germanic*. (A Linguistic History of English 1). Oxford: Oxford University Press.
- Robinson, Orrin W.
1992 *Old English and its Closest relatives. A Survey of the Earliest Germanic Languages*. Stanford: Stanford University Press.
- Scardigli, Piergiuseppe
2005 Nordic-Gothic Linguistic Relations. In: Bandle et al. (eds.), 553–558.
- Schleicher, August
1869 *Die deutsche Sprache*. 2nd edn. Stuttgart: Cotta.
- Schmidt, Christoph G., Robert Nedoma, and Klaus Düwel
2011 Die Runeninschrift auf dem Kamm von Frieñstedt, Stadt Erfurt. *Die Sprache* 49: 123–186.
- Schwarz, Ernst
1951 *Goten, Nordgermanen, Angelsachsen: Studien zur Ausgliederung der germanischen Sprachen*. Bern: Francke.
- Seebold, Elmar
2004 Alemannisch und Nordgermanisch. In: Naumann (ed.), 1–11.
- Stearns, MacDonald Jr.
1978 *Crimean Gothic: Analysis and Etymology of the Corpus*. Saratoga, CA: Anma Libri.
- Vennemann, Theo
1984 Hochgermanisch und Niedergermanisch: Die Verzweigungstheorie der germanisch-deutschen Lautverschiebungen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 106: 1–45.
- Vennemann, Theo
2006 Grimm's Law and Loanwords. *Transactions of the Philological Society* 104: 129–166.
- Vennemann, Theo
1981 *Gothic, Germanic, and Northwest Germanic*. Wiesbaden: Steiner.
- Vennemann, Theo
1992 *Early Germanic Grammar: Pre-, Proto- and Post-Germanic languages*. San Diego: Academic Press.
- Wagner, Norbert
1989 Der Stammesname der Salier und die ‚westgermanische‘ Konsonantengemination. *Zeitschrift für deutsches Altertum* 118: 34–42.
- Wrede, Ferdinand
1924 Ingwäonisch und Westgermanisch. *Zeitschrift für deutsche Mundarten* 19: 270–283.

Ludwig Rübekel, Zürich (Switzerland)

59. The evolution of Germanic

- | | |
|---|---|
| 1. Introduction | 5. Syntactic patterns |
| 2. The range and internal structure of Germanic | 6. Conclusion: Conservatism and paths of innovation |
| 3. Phonological patterns | 7. References |
| 4. Morphological patterns | |

<https://doi.org/10.1515/9783110523874-014>

1. Introduction

In this article, I focus tightly on the question implicit in the title: How has Germanic as a family developed over time? In particular, this provides an opportunity to concentrate on the shared paths of development within the family. Any linguist who reads these lines presumably already knows the broad sweep of Germanic linguistic history. Compared to Proto-Germanic and more so to Proto-Indo-European, modern Germanic languages and dialects share, for instance, characteristics like these:

Sound system

- All Germanic languages and dialects show conditioned changes in vowel height, and all save Gothic show some forms of the regressive vocalic assimilation process (specifically, fronting) known as umlaut.
- Unstressed vowels are inevitably reduced in some way, even at the earliest stage of attestation, compared to reconstructed ancestors.

Word forms

- Every modern variety of Germanic has reduced the number of overtly marked morphological categories over the attested period.
- The Germanic languages and dialects have expanded the set of “weak” or dental preterit verbs at the expense of “strong” or ablauting verbs.

Sentence structure

- All have increased the use of periphrastic constructions to convey meanings and/or functions which were once marked inflectionally.
- All have moved toward relatively more fixed word order, and verb second (V2) phenomena are found in all of the languages.

After a note on the internal organization of the family (2), this article surveys a handful of selected patterns and tendencies of change from the sound system (3, umlaut and vocalic chain shifts, consonantal chain shifts, and prosody), from morphology (4, case loss and related changes in the nominal system, increase in the number of “weak” verbs at the expense of “strong” verbs), and from syntax (5, consequences of loss of inflection, word order, the rise of determiners, increased periphrasis in the verbal system). I conclude with a note on the broader context of such shared particular evolutionary tendencies (6).

Before moving on, let us note some of the literature on this topic. The immediately preceding articles (of this section) have covered Proto-Germanic, like Prokosch (1938), Hirt (1931–1934) or Ringe (2006). Numerous recent works treat specific branches or languages, like the encyclopedic treatments of North Germanic (Bandle et al. 2002), German (Besch et al. 2002) and English (*Cambridge History of the English Language* 1992–). Some titles suggest coverage of our topic, though the texts pursue other goals: Nielsen’s *The Germanic Languages* provides an accessible treatment of the “breakup” of Germanic, while Meillet’s *Caractères généraux des langues germaniques* offers a survey, but with an eye toward arguing that substrates had a particular role in shaping Germanic. The multi-volume “comparative grammar” project led by Guxman (1962–1966) provides an important overview, though these works are also now dated. Three more recent works provide valuable perspectives for particular audiences: König and van der Auwera (1994) offers historical and synchronic surveys of the standard lan-

guages and some other varieties, including Pennsylvania German and Germanic-lexified creoles. Excellent histories are available for languages less often cited in the Indo-Europeanist and older historical literature, like Roberge (1994) on Afrikaans and Munske (2001) on the Frisian languages. Robinson (1992) may long remain unsurpassed as an introduction to the history and structure of the older Germanic languages.

Many of the above works focus on standard languages, although what happens to have found its way into modern codified languages is only part of the historical trajectories in the family. Books organized around standard languages naturally tend to give less play to the rich array of patterns found in historical and contemporary dialects. From a mostly synchronic perspective, Harbert (2007) offers up-to-date coverage on this count, and provides the most theoretically-informed treatment of syntax. Howell, Roberge, and Salmons (in preparation) will present a directly diachronic handbook. Hutterer (1990) in some sense provides the most direct predecessor of this paper, as “ein Versuch zur historischen Typologie der germanischen Sprachen” (1990: v), especially the last chapter, “Die Strukturmerkmale der germanischen Sprachen” (420–467).

2. The range and internal structure of Germanic

The fundamental organization of Germanic starts with its historical unfolding or *Ausgliederung*. The dialectology of Germanic is treated in detail in Rùbekeil, this handbook, but our interest in similarities across the later Germanic dialects still warrants considering the nature and origin of similarities in light of this question. In terms of tree models, one might simply assume the tripartite structure given below.

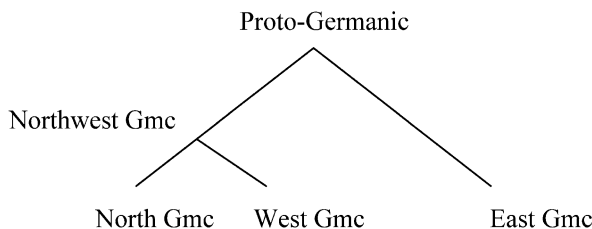


Fig. 59.1: One Germanic family tree

East Germanic, so valuable for understanding early Germanic in general, is ill-attested save for the snapshot provided by Gothic. As a result, it is of limited consequence for the purpose of tracing long-term developments. More generally, *Stammbäume* are notoriously problematic for Germanic and other languages where enduring contact across dialects/languages has been widespread. The illustration below (from Salmons 2012: 85, building on Kufner 1972: 74) offers a better way of portraying contacts among groups, while dividing West Germanic into three groups.

Within these subgroupings, Gothic and Runic evidence inevitably warrant attention as the earliest attested forms of Germanic, and they provide a kind of baseline for our survey (I will not discuss chronology, cf. Polomé 1992 for discussion). Particular attention is due to dialects and languages that appear to lie at ends of the evolutionary contin-



Fig. 59.2: Germanic in a wave model, map by Mark Livengood

ua, i.e. in terms of appearing to be particularly conservative or innovative: Icelandic is rightly regarded as conservative, but another set of West Germanic dialects has been neglected in comparative Germanic: Highest Alemannic and southern Bavarian speech islands show strikingly conservative traits and I will draw on data from them. At the other end of this spectrum, English and some of its dialects show patterns of particularly dramatic innovation (McWhorter 2002), along with mainland Scandinavian and North Sea coast varieties.

Shared ancestry suggests, in some sense, that all the daughters will start along similar trajectories. Contact may promote innovation and diffusion of similarities, not just by borrowing, but also by acting as a catalyst for loss of inflectional categories. This is notable in the core Germanic territory, while varieties spoken on the periphery may have developed more independently, retaining archaisms lost elsewhere or innovating in distinctive ways. Germanic thus supports the common claim in dialectology and areal linguistics that archaisms tend to be found in peripheral areas, innovations in central ones, cf. Bàrtoli (1925), Bonfante (1947), and Chambers and Trudgill ([1981] 1998: 182–183). Let us turn now to our core topic, the paths of development we see across the family.

3. Phonological patterns

We begin with examples of change in the sound systems of Germanic languages (see Stiles, this handbook). A discussion of umlaut will set up a discussion of the propensity to “chain shifting”, followed by notes on the apparently chronic tendency toward “consonant shifts”, and on prosodic developments in the family.

3.1. Umlaut and related changes

The view of umlaut currently presented in handbooks is the position telegraphically given in Twaddell (1938), but it is increasingly rejected by specialists, for reasons that bear directly on the evolution of Germanic. Twaddell’s view implicitly holds that all phonetic umlaut happened in Proto-Germanic. Antonsen makes this explicit, writing (2002: 23): “suffice it to say that by *late* Proto-Germanic, all of the vowel phonemes in root syllables had developed positional variants that were dependent on the following vowel or semivowel” (see also Penzl 1995: 116). Thus, all forms which originally contained a back vowel followed by an *ij* in the next syllable but which do not show umlaut in modern Germanic can only be explained as analogical (also Antonsen 1969). Most unfolding of umlaut across the family comes to be, then, not about sound change, but analogy. While this might be plausible for German, we will see below that it founders on Netherlandic data. The alternative analysis shows that sound changes can account for a far broader set of data as regular. More importantly, this view also places umlaut in the broader context of the unfolding of Germanic vocalism over time and space.

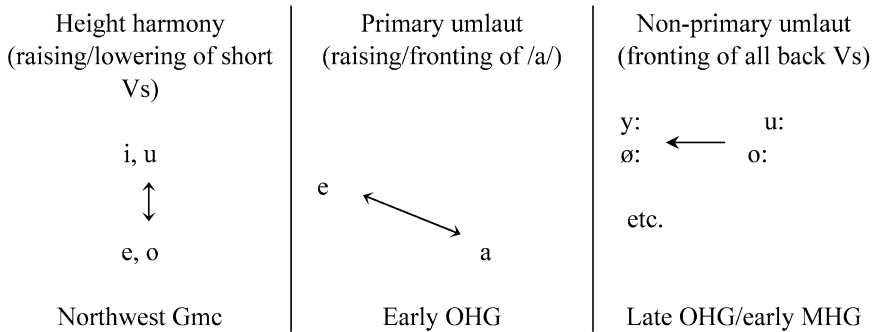
The oldest umlaut-like process is found across all early Germanic, systematic realignments of vowel height. In Gothic, these are consonantly conditioned (cf. Howell 1991), with short high vowels lowering before /h, h^w, r/, while /e/ raises to /i/ elsewhere. Across the family, short /e/ raises to /i/ before a nasal coda. In addition to similar consonantal patterns, in North and West Germanic, the height of (short) stem vowels assimilated partially to the height of following vowels (many posit height harmony for Northwest Germanic, like Nielsen 2000, but others place the process in Proto-Germanic, assuming that later developments obscure its effects in Gothic, like Harbert 2007). This “height harmony” was active in the morphology in some daughters, like OHG, where *geban* ‘to give’ showed raising with high vowels in a personal suffix, e.g. 2/3SG.PRES.IND. *gibis* and *gibit*. Van Coetsem (1994: 88 and elsewhere) treats these as directly related patterns, illustrated here with a pair of back vowels and a pair of front vowels:

- (1) Height harmony, from PGmc
- | | |
|---------------|--|
| Northwest Gmc | u → o / __Ca
*gulpa- > gold, *hurna- > horn |
| Gothic | u → o / __ h, h ^w , r
*gulpa- > gulp, *hurna- > haurn (<au> = [ɔ]) |
| Northwest Gmc | e → i / height harmony and before coda nasal
*benda- > bind, *geba- > OHG geban |
| Gothic | e → i / __ consonants other than h, h ^w , r
*benda- > bindan, *geba- > giban |

For this cluster of changes, Van Coetsem and Buccini (1990) argue that the same rule – raising or lowering of some short vowels – came to be “reconditioned”: the output remains constant while the contexts shaping that output have broadened.

Germanic is better known for another set of regressive V-to-V assimilations, especially *i*-umlaut. Broadly speaking, an *i* or *j* triggers fronting of a preceding back stem vowel. In OHG, it first appears in 8th century texts as “primary umlaut”, the fronting and raising of short /a/ to [ɛ] before *i/j* except when certain consonant clusters intervened (*r*, *l*, *h* + obstruent, variably by dialect). Only later do we find these blocking environments overcome (at least orthographically), and the extension of umlaut from disyllabic to polysyllabic words (that is, not simply *gasti* > *gesti* ‘guest’ but also *magadi* > *mege* ‘maiden’ or *zahari/zahiri* > *zehere* ‘tear’) and the general fronting of all back vowels before *i/j*. This led Buccini (1992) to see primary umlaut as the historical link between height harmony (since *a* is raised) and fronting (since *a* was also fronted), as illustrated below.

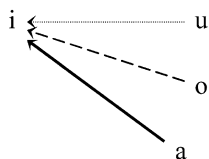
(2) A progression from height harmony to *i*-umlaut in West Germanic



On this view, then, we can follow a chain of changes from consonantal conditioning of vowel height to vocalic conditioning of height harmony, which in turn is related to umlaut by the pivot of primary umlaut, at least in West Germanic.

Howell and Salmons (1997) outline how umlaut patterns across Germanic correlate with how close or distant the trigger – *i/j* for *i*-umlaut – is to its potential “target” in the vowel space: [a] is farthest away from [i, j], so most susceptible to assimilatory change, while [u] is closest and most resistant to assimilation. This “Principle of Maximal Differentiation” is illustrated below with greater susceptibility to assimilation represented by heavier arrows (from Howell and Salmons 1997: 93 ff.):

(3) Maximal differentiation



Most members of the family eventually see umlaut through to completion as a phonological generalization and then a morphological one, but wrinkles have persisted at the

geographical peripheries, the western and southern boundaries of West Germanic and the northwestern-most outpost of North Germanic. The striking examples – not acknowledged in the structuralist tradition – are the coastal dialects of Dutch, where only primary umlaut is found. This includes failure of primary umlaut when *-hC-* clusters intervene, so that blocking environments remain, cf. *machtig* ‘powerful’ vs. German *mächtigt*. From there, the generality of umlaut increases steadily moving eastward, with umlaut of *â* and then of *ô*, for instance. Modern Dutch dialect geography largely recapitulates the unfolding of umlaut in the High German territory for short vowels (but see below on “spontaneous palatalization”). Beyond short vowels, “Standard Dutch knows no umlaut of long vowels” (Schönfeld/Van Loey 1970: 44, my translation). The Netherlandic umlaut facts seem impossible to reconcile with the Twaddellian view.

Even in the German-speaking territory, we find “umlautless residues”, often in the southernmost dialects. This manifests itself typically in the failure of umlaut of *u*, as expected from the above, and especially where geminate and/or velar consonants intervene (data ultimately from Schirmunski 1962: 201–203):

(4) Umlautless residues in Upper German

Southern	Standard	Earlier form	
muck	Mücke	OHG mucka, OS muggia	‘midge’
fduk	Stück	OHG stucki	‘piece’
khuxə	Küche	OHG kuchina	‘kitchen’
lu:gə	Lüge	OHG lugin	‘lie’ (noun)

On this still-emerging view, these areas that did not carry *i*-umlaut through to completion are the ones that best illuminate its history in West Germanic.

While umlaut in North Germanic takes a different path (cf. Iverson and Salmons 2004, 2012 for literature), it too reveals much about the evolution of umlaut. Icelandic *u*-umlaut is still “alive” in some sense, as argued at length by Anderson (1974: 191–195, 200–202). This rule assimilates /a/ to [ö] when /u/ follows underlyingly as in (a), while we see in (b) that it spreads over two syllables (assuming, with Anderson that the medial vowel reduces phonologically from [ö] to [u]). In (c), the rule applies to recent loanwords, but does not spread where the medial vowel has not reduced.

(5) Modern Icelandic *u*-umlaut

- a. jeg kalla ~ við köllum ‘I call ~ we call’
jaki ~ jökull ‘piece of ice ~ glacier’
- b. fatnað ~ fötnuðum ‘suit of clothes, NOM.SG. ~ DAT.PL.’
bakari ~ bökurum ‘baker, NOM.SG. ~ DAT.PL.’
- c. japani ~ japönum ‘Japanese, NOM.SG. ~ DAT.PL.’
almanak ~ almanökum ‘calendar, NOM.SG. ~ DAT.PL.’

This may be the last example of umlaut as what Anderson calls a “persistent rule”, where umlaut remains phonologically active, partially restricted to disyllabics, depending on reduction of medial syllables in longer words.

Umlaut, under the view sketched here, unfolded in an interconnected set of changes from consonantal conditioning of vowel height to general vowel fronting before original

i/j and other assimilations. Parallel developments often go farther, as in the unrounding of front rounded vowels, which happened early in English, but has taken place over most German and Yiddish dialects. While Modern Icelandic still spells *yrki* ‘work’, there too, it has unrounded to [i] (see Bandle 2005: 1120).

3.2. Vocalic chain shifts

Umlaut was a complex process, but in some sense a one-time occurrence – the change began phonetically, and ran its course phonologically, leaving myriad particular effects varying by language or dialect in the sound systems and word forms. In contrast, many independent chain shifts characterize the family from prehistory to the present day. The Great Vowel Shift in Early Modern English is best known, but consequential changes appear to be underway today in American English (Labov 1994 and other works). Such changes are so pervasive and characteristic of Germanic that Stockwell (1978, emphasis in original) wrote:

The vowel shift occurred no more at the usually cited dates than at any other date in the documented history of English. That is, it *did* occur then, and also (equally, I believe) over the past 200 years, or over the 200 years between the birth of Alfred and the death of Aelfric, or any other period of that length. This kind of vowel shifting is a pervasive and persevering characteristic of vowel systems of a certain type.

Sievers (1876) long ago recognized that vowels tend strongly to follow certain paths in chain shifts. Long or tense vowels tend to raise, short or lax to lower and back vowels to move to the front of the vowel space. A variety of explanations have been proposed for these patterns, starting with Sievers’ own notion that longer vowels have more time to reach a fuller articulatory realization, while some recent work like Jacewicz et al. (2006) is exploring the relationship between prosodic prominence and chain shifting.

Labov (1994) gives examples from Yiddish, North Frisian, and Swedish, in addition to a few non-Germanic languages, but such shifts are even more widespread. Take his “Principle III”, according to which tense back vowels move to the front in chain shifts. As with changes underway in the southern and eastern United States today, /u:/ fronting is often the leading edge of such chains. Küspert (1988) presents several examples from western Scandinavian of /u:/ fronting, which is then connected to other changes in back vowels and Wiesinger (1970, elsewhere) does the same across German dialects. In American dialects, mid vowels tend to follow a parallel track forward, but in some German and Scandinavian patterns, we find an apparent drag chain, where /u:/ fronts and /o:/ moves up toward or to the back top corner of the vowel space (cf. Küspert 1988: 285 ff.). In both Dutch and in the southernmost Bavarian speech islands, the very northwest and southeast corners of continental West Germanic, unconditioned /u:/ fronting has taken place, known as “spontaneous palatalization”. In Visperterminen, it triggers a merger of historical /u:, ou, uo/ into /y:/, while /o:/ has not changed (Wipf 1910: 35). According to Wipf (1910: 32), short /o/ continues to show height harmony effects in paradigms. These are frozen alternations, not active phonological processes, since recent loans tend not to show the process.

Even more than umlaut, chain shifts characterize the vowel systems of Germanic standard languages and dialects. The myriad different instantiations of shift in different periods and places support Stockwell's notion of the chronic character of chain shifting in Germanic, while tending to follow a few well-defined paths.

3.3. Consonant shifts

Possibly the best known sound change in linguistics is the First Sound Shift or Grimm's Law. In the barest of traditional terms, it consists of these correspondences (see Stiles, this handbook and Collinge 1985):

(6) Simplified view of Grimm's Law (leaving aside labiovelars, clusters, etc.)

	PIE		Germanic
Voiceless	*p, *t, *k	→	*f, *θ, *x
Voiced	*b, *d, *g	→	*p, *t, *k
Voiced aspirated	*b ^h , *d ^h , *g ^h	→	*b/β, *d/ð, *g/ɣ

On many views (e.g. Ringe 2006: 100), the old breathy stops yielded stops in some positions and spirants in others.

Complex consonantal chain shifts appear to be far less common than vocalic ones, perhaps without known parallel in the languages of the world, although a variety of partial parallels have been proposed, including in Armenian (cf. Macak, this handbook). Still, the Second or High German Consonant Shift partially recapitulates the same changes on the remaining two stop series (as argued again recently, e.g., Iverson and Salmons 2003a). Here, again in the simplest possible form, are patterns of the Second Sound Shift in the strict sense (a) and the related *Medienverschiebung* in (b).

(7) Simplified view of the Second or High German Consonant Shift

a.	p, t, k	→	pf, ts, kx	→	f, s, x
b.	b, d, g	→	p, t, k		

These are phonologically and geographically restricted, but the most advanced forms of shift are found in the south, lessening as one moves north (Lerchner 1971).

Less well known are some modern partial parallels, most importantly one underway in contemporary Liverpool English (Honeybone 2001, 2005, elsewhere). There, /t/ and /k/ affricate or spirantize in most environments (Honeybone uses [θ] to represent a slit, non-sibilant coronal fricative, following earlier scholars):

(8) Obstruent Shift in contemporary Liverpool dialect

t	→	tθ	→	θ	<i>taken</i> = [θ]aken, <i>city</i> = ci[θ]y, <i>alright</i> = alrigh[θ]
k	→	kx	→	χ, x, ç	<i>come</i> = [kx]ome, <i>crackers</i> = cra[x]ers, <i>book</i> = boo[x]

Most scholars have seen these changes as systematic fortition of obstruents, while the recent stream of work by Honeybone puts them into the broader context of lenition and "lenition inhibition". Systematic lenitions are as common crosslinguistically and within

Germanic as fortitions are thought to be rare. Holsinger (2001, 2008) provides detailed and theoretically situated discussions of such patterns, particularly with regard to prosody, a topic to which we now turn.

3.4. Prosodic structure

Particular progress in recent decades has been made in prosody, in the narrower sense of “accent” and the broad sense of strings of sounds “above the segment” – how speech is organized into syllables, feet, and phonological words. The most important research on Germanic historical prosody is synthesized in detail by Lahiri, Riad, and Jacobs (1999: 335–378). As they argue (1999: 340), a simple focus on position of stress might create an illusion that little has changed since Proto-Germanic, but all daughters have seen considerable evolution. While the older languages were trochaic like the modern ones (leaving aside contact-driven deviations from this core stress pattern, such as the Romance component of English, or the Hebrew-Aramaic component of Yiddish [Jacobs 2005: 135 ff.]), in earlier times, the head of a foot required two moras (“Prokosch’s Law”) and the grammar allowed metrical resolution, both features lost in the modern languages. Proto-Germanic constructed feet from left to right, and the rightmost foot within a longer string is normally most prominent in most of the modern languages. That foot is, again, often trochaic, so that the penultimate syllable is often especially prominent. The exceptions to this trend, Icelandic and Faroese, have changed in other ways, for instance now having quantity-insensitive syllabic trochees. Only a few dialects of Swedish, Norwegian, and Swiss German retain the old system of quantity distinctions in consonants. Other notable recent work includes Page (1999) on prosodic change, Murray (2000) and Page (2006) on how quantity changes might be understood in terms of “syllable cut” prosody, and Smith (2007) on the ongoing importance of “prosodic templates” through Germanic phonological and morphological history.

The Germanic accent shift, from the pitch-oriented lexical accent of Proto-Indo-European to what is traditionally described as an intensity-based stress in Germanic (cf. Stiles, this handbook), is a traditional defining feature of Germanic. Controversy long surrounded the question of whether Germanic accent should be seen as initial or as root-oriented. This parallels more recent discussions about whether stress assignment is purely phonological or must make reference to morphological structure (Lahiri, Riad, and Jacobs 1999). Central to this are the complex metrical patterns found with prefixes, especially verbal, which often lack stress in the ancient and modern languages.

Since the earliest work in phonetics (Sievers 1876: 204–206) and IE handbooks (Brugmann 1897: 59–61), the realization of prominence as “expiratory stress” or “intensity” has often been thought responsible for the reduction of unstressed syllables. This terminology in fact does not fit comfortably with some of the most promising interpretations of how stress might lead to reduction. “Intensity” (and perhaps “expiratory” as well) refers to loudness, while concrete appeals are typically made to prosodic prominence realized as duration, where stressed syllables are notably longer. That is, isochrony at the foot level means that longer duration of stressed syllables “robs” unstressed syllables of time for full realization, over time eroding them entirely. Marking prominence by pitch correlates, on such views, with relatively even timing across syllables, stressed

or unstressed, thus not leading to reduction. Various efforts have been made to underpin this view, notably van Coetsem, Hendricks, and McCormick (1981) and van Coetsem (1996), appealing to an ultimately circular notion of “dominating accent”, in contrast to a “non-dominating” accent in languages like Finnish, where fixed initial stress does not correlate with reduction. Despite the vast literature on these questions, only now is this notion being systematically tested: Menz (2010) investigated the acoustics of accented and unaccented syllables in languages which have undergone relatively little historical reduction versus closely related varieties which have undergone much more, such as Highest Alemannic versus Low Alemannic and North Bavarian versus Cimbrian, finding surprisingly little correlation.

4. Morphological patterns

In Proto-Indo-European, most morphological distinctions were historically marked suffixally, so that the foregoing discussion implies that inflectional morphology has generally been reduced over the history of Germanic (cf. Harðarson, this handbook).

4.1. The retreat/loss of whole categories

One inflectional category inherited from PIE was a set of distinct dual forms in the nominal and verbal systems, along with singular and plural. Like in many IE languages, these have faded over time: Howe writes (1996: 118) specifically about the pronouns, “a loss of the dual number ... has occurred now in all the extant Germanic languages”. Gothic possessed paradigms of distinct pronouns and verbal inflection for dual in the first and second persons, though the vagaries of transmission leave us with numerous gaps in attestation (see Sihler 1995: 370–371 and Eichner, this handbook, on problems of reconstructing pronouns in IE; and Ringe 2006: 290–291 on Proto-Germanic pronouns). While they have lost distinct verb forms, many attested West Germanic languages still show some first and second person dual pronouns, for instance, continuing into early Middle English on occasion. Old Saxon has *uuit* ‘we two’ and *git* ‘you two’, but as Holthausen (1921: 113–114) notes, even in the *Hêliand* plural pronouns sometimes refer to pairs, indicating incipient loss of the dual. Some North Frisian dialects retain morphologically distinct dual pronouns today, in the Sylt dialect even for third person subjects, but they are reported to be rare (Howe 1996: 70–73, 194–195). In mainland Scandinavian, dual pronouns persist into the late Middle Ages and in Icelandic until the 15th century (Haugen 1976: 303). Old dual pronouns remain in plural functions in Icelandic (cf. Bavarian *enk* ‘you [pl]’, earlier ‘you [dl]’).

4.2. Nominal inflection

Four cases are robustly attested in Germanic from the earliest texts to some contemporary languages – nominative, accusative, dative, genitive (see Hewson 2006 for detailed treatment of case syntax in the history of Germanic, particularly with regard to prepositional

syntax). Other cases are marginally attested in our earliest documents: Early West Germanic nouns and pronouns show instrumental inflection in some major classes, marked with *-u* or *-o*; Prokosch (1939: 236) interprets some \emptyset -marked datives in Old Norse as traces of instrumental, but see Boutkan (1995: 186–187). In contrast, Gothic shows only the barest remnants of instrumental in two neuter pronouns, demonstrative *þê* and interrogative *hwê* (Braune/Ebbinghaus 1981: 67, elsewhere). Gothic attests a few vocative forms. Many show the same marking as accusative rather than the usual nominative forms, but some *u*-stem nouns like *sunus* ‘son’ show voc. *sunau* ≠ acc. *sunu* (as Jared Klein points out, the *sunau* forms continue the IE vocative). Finally, Braune/Reiffenstein (2004: 185–186) note some early OHG placenames in *-i* that look like locatives, and in the Hêliand’s *Giuuitun im tho te hus thanan* ‘And then they went home’, ‘house’ could reflect an old locative. Still, locative is generally not reconstructed even for Proto-Germanic, cf. Bammesberger (1990), Ringe (2006).

Among the modern standard(ized) languages, a number can be claimed to preserve four cases in some real sense: Icelandic, Faroese, and German. In West Germanic, most case marking is carried on determiners or adjectives rather than the noun itself, while the North Germanic languages noted use more distinct marking on the noun. Compare two words for ‘man’:

(9) Nominal inflection in German and Icelandic

	German		Icelandic	
	Sing.	Plural	Sing.	Plural
nom.	<i>der Mann</i>	<i>die Männer</i>	<i>maðr</i>	<i>menn</i>
acc.	<i>den Mann</i>	<i>die Männer</i>	<i>mann</i>	<i>menn</i>
dat.	<i>dem Mann(e)</i>	<i>den Männern</i>	<i>manni</i>	<i>mönnum</i>
gen.	<i>des Mannes</i>	<i>der Männer</i>	<i>manns</i>	<i>manna</i>

In fact, this picture systematically overstates the distinctions in German on the noun, where the dative singular *-e* is largely gone from spoken and even much written usage, the genitive is increasingly restricted to formal use, and even the dative plural *-n* shows signs of receding. That is, while the paradigm for this word in Icelandic has only two identical surface forms (nom.-acc. plural), in practice, colloquial German is approaching a distinction only based on number on the noun.

While case on nouns continues to retreat slowly, consistent marking of the singular/plural distinction has been steadily expanded (Wegener 2007).

Genitive inflection, widely used as recently as Early New High German, is now lacking in most modern German dialects (Schirmunski 1962: 432–437), including some otherwise conservative varieties. Wipf (1910: 119 ff.) reports that all four cases were in use in Visperterminen (Highest Alemannic), Tyroller (2003: 124–15) notes only “remnants” of genitive today in Lusern (Cimbrian, northern Italy). Similarly, Neil Jacobs (p.c.) describes the Yiddish genitive as “significantly reduced/limited”. Further case syncretism is widespread, typically collapse of dative and accusative or of nominative and accusative. In short, contemporary vernacular German, like some other closely related

West Germanic languages, has less case marking than we find in the standard, and the movement of marking away from the noun to other elements in the phrase, especially determiners, is approaching completion.

It is often observed that compared to nouns, “pronouns tend to be conservative in their inflection” (Harbert 2006: 177). This is illustrated above by the survival of dual pronouns but not of unique dual noun forms, and the survival of traces of instrumental on Gothic pronouns but not on nouns. It is further seen in the retention of the nominative/oblique distinction in most varieties of English pronouns (*I/me, she/her, they/them*, etc., and likewise Afrikaans) but loss of all case marking on nouns and determiners, save for genitive *-s*.

Germanic continues its inherited organization of nouns into stem classes, but very variably. English shows meager traces, in “irregular” plurals like *mice, oxen, deer*. Modern Icelandic can still be reasonably described in terms of those structures, although many particular nouns have changed class (Kress 1982: 55–82). The patterns associated with large masculine/neuter *-a, -ja* and *-wa* classes and small classes such as *-u* and consonant stems remain clear (see Harðarson, this handbook). OHG is described along such lines in handbooks (Braune/Reiffenstein 2004), but a database of nominal forms built at the University of Wisconsin, Project Graff (Luiten et al. 2013), found that handbook paradigms for OHG are often highly archaizing: attested forms show surprisingly advanced reduction, even from early texts.

4.3. Verbal inflection

Weak or dental preterit verbs are a famous Germanic innovation, an inflectional option added to a presumed system of strong or ablauting verbs. Charting the migration of strong verbs to weak is a game played often, by many, and for different purposes, as illustrated by Hare and Elman (1995) and Lieberman et al. (2007). For German, a crude but widely-cited figure (Augst 1975) is that OHG had 349 strong verbs, leaving aside 20 more “mixed” verbs, while MHG had 339 and Modern German has only 169. More recent counts push the number of strong verbs in OHG considerably higher, often to over 450. In OHG, by contrast, there were about 2440 weak verbs (Raven 1963/1967, Bittner 1996: 135, and now Carroll et al. 2012). For more detailed recent coverage of ablaut in Germanic strong verbs and better measures of reductions in complexity within the system, see Bittner (1996), Fertig (2000), Mailhammer (2007a, 2007b).

Icelandic, morphologically conservative in so many ways, has expanded its set of dental preterit verbs at the expense of strong. Modern Icelandic is reported to have about 150 strong verbs in use, according to Ragnarsdóttir, Simonsen, and Plunkett (1999: 592). They estimate that Icelandic, Norwegian, and English all have the same percentage of strong versus weak verbs – 4% strong ~ 96% weak. This is down considerably from ca. 12.5% compared to at least one of the “older dialects”, OHG, based on the above numbers.

Of course, verbs can eliminate many kinds of alternations in addition to or aside from losing ablaut. In Wallis (Wipf 1910: 146), spoken past participles “in merkwürdig weitgehendem Maße [to a remarkably great extent]” are formed with a weak *-t* rather than the typical strong *-en*. Like in most southern German varieties, the simple past is

not used, and some verbs have lost ablaut and become entirely weak (1910: 156), so that the overall result is tremendous simplification of the system's surface complexity. Tyroller (2003: 112–119) reports similar patterns for Lusern (Cimbrian), where *khemmen* 'to come' has the participle *khennt* (with nasal assimilation to the place of the final stop), likewise *sraim* ~ *gesribet* 'to write', *bervan* ~ *geborft* 'to throw' (this development is sometimes attributed to Romance influence in both dialects). Likewise, personal endings have leveled out in various ways, from the uniform present *-er* of Danish to the "*Einheitsplural*" in North Sea Germanic.

At the same time, a trickle of once weak verbs continues to develop ablaut variants, like American English dialectal *dive* ~ *dove* ~ *dove* or *sneak* ~ *snuck* ~ *snuck* (Durrell 2003), or Swedish forms like these (Haugen 1976: 376), from historical weak classes I–III:

(10) Innovative ablaut in Modern Swedish

<i>pipadhe</i>	>	<i>pep</i>	'piped'
<i>knytte</i>	>	<i>knöt</i>	'tied'
<i>hintel/hinde</i>	>	<i>hann</i>	'succeeded'

Ablaut was an important aspect of nominal (derivational) morphology in IE but is almost gone in the nominal system in Germanic. In the verbal system, it has held on far longer even in inflection, but recedes slowly and apparently inexorably.

5. Syntactic patterns

Historical linguistics as a field has seen a great revival in recent years, and the study of syntactic change has been especially vibrant, with progress including a range of theoretical frameworks (see Luhr, Hale, and Keydana, this handbook).

5.1. Syntactic consequences of morphological change?

In comparing English and German, Hawkins (1985, building on Sapir 1921) sees the historical loss of inflectional morphology as having led to less strict mappings between meanings and surface forms. He argues that this phenomenon underlies the "realignment in the mapping between surface form and meaning" (1985: 215):

The morphological and syntactic structures of German are regularly in closer correspondence with their associated semantic representations than those of English. English tolerates greater collapsing of distinct meanings onto common surface forms (whence greater ambiguity and vagueness)

Hawkins traces this through patterns from lexical semantics and word order freedom to raising, extraction, and deletion. Consider the example of differences in subjecthood between the two languages (1985: 53–61), where German continues older patterns:

- (11) Non-agent subjects in German versus English
- a. A dollar doesn't buy much.
*Eine Mark kauft nicht viel. Mit einer Mark kann man nicht viel kaufen.
 - b. This hotel forbids dogs.
*Dieses Hotel verbietet Hunde. In diesem Hotel sind Hunde verboten.
 - c. Tomorrow promises to be cold.
*Morgen verspricht kalt zu sein. Morgen verspricht es kalt zu sein.

While Standard German typically resists allowing non-agentive subjects, English allows a range of locatives, instruments, and so on to serve as grammatical subjects. In maintaining tighter mapping between semantic roles and surface forms, German sometimes lacks overt subjects and it often uses case-inflected forms where the English equivalents would be subjects (1985: 56):

- (12) Case contrasts in English and German
- | | |
|------------------|---|
| I am freezing. | Mich friert.
me.ACC freezes |
| I am warm. | Mir ist warm.
me.DAT is warm |
| I like the book. | Mir gefällt das Buch
me.DAT pleases the book |

While German and other languages sometimes lack overt subjects, Harbert (2007: 221–223) maintains that only Gothic shows pro-drop patterns. Of course Old English, Yiddish, and other Germanic languages past and present show clear elements of pro-drop.

Follow-up work by Loudon (1992) on Pennsylvania German and Shannon (1990) on Dutch has shown that these two West Germanic languages, both of which fall between German and English in terms of retention or loss of inflectional morphology, likewise take an intermediate position with regard to the form/meaning mapping criteria Hawkins develops. Hawkins' core generalization appears robust across Germanic.

5.2. Word order

It is well known that syntactic reconstruction, including word order, confronts the researcher with difficulties not seen in the reconstruction of sounds or word forms. Indeed, it is a matter of longstanding debate whether syntax can be reconstructed at all (cf. Lightfoot 2002 and Campbell and Harris 2002 for an exchange on this issue, and now Walkden 2014). These difficulties beset the study of early Germanic syntax as well, where the reconstruction of word order in particular is rendered problematic by the nature of the textual material that has come down to us. Many works are poetic and reflect highly stylized syntax, like the Old Saxon *Heliand*. Many of our earliest texts are close translations from Latin or Greek, following the word order of their sources. This is true for many early works in OHG, for instance, and perhaps too famously so for Gothic, where some beginning textbooks provide lists of Grecisms, not just in word order

but in basic syntactic patterns, like the use of participial constructions in the function of subordinate clauses (Bennett 1999: 127). Even in Gothic, though, where little can be concluded safely about word order, careful sifting of the texts, especially deviations from the Greek *Vorlage*, has yielded much about Gothic syntax. This is especially true with regard to particles, cf. Harbert (1978), Klein (1992, 1994), Klein and Condon (1993), and Ferraresi (2005), but see also the cautionary arguments of Goetting (2007).

While Runic evidence is valuable for understanding morphological and phonological developments, it provides little evidence on basic word order; Antonsen's catalogue (1975: 24) includes only 34 clauses where the position of a verb can be compared to other elements. While Nielsen finds that the early bracteates show only VO order, he finds only about four instances of VO in other texts in the period of roughly 400–500 CE. Moreover, we have no more than a half dozen OV constructions before 500 CE that include a syntactic object, such as the Einang inscription:

- (13) Einang inscription
 ... dagastiz runo faihido
 '... [I] Dagastiz rune(s) painted'

He concludes, safely it seems, that "OV and VO constructions competed during the fifth century" (2000: 171). This may represent a transition from OV to VO word order. Contemporary Germanic languages show varying degrees of VO or OV patterns, but all show some verb-second or "V2" phenomena, whereby the finite verb must come no later than second in its clause. V2 has been the subject of myriad publications (Platzack 1985, Weerman 1986), and it has been called "the most celebrated feature of Germanic syntax" (Kiparsky 1995: 161).

5.3. Definiteness in nominal syntax

Proto-Germanic lacked definite or indefinite articles, but developed demonstratives from "pronominal" inflections (I use scare quotes following Ringe 2006: 288–289, who distinguishes these from "pronouns proper"). These in turn often evolve into definite articles, though not always – cf. English definite *the* and demonstrative *that*, one from a masculine and the other from a neuter form within the same historical paradigm. As argued by van Gelderen (2007, also 2011) with a wide range of North and West Germanic evidence (including Afrikaans), these constitute a "definiteness cycle". In this cycle, demonstratives, which express definiteness and "location relative to the speech event" lose that latter characteristic in becoming articles. That is, the deictic meaning of the demonstrative is lost and semantic features become formal. In her syntactic terms, this is a reanalysis of elements from a specifier position to head of a phrase.

This example illustrates one of the most popular trends in diachronic syntax, the rise of grammaticalization research, the study of how lexical items and syntactic constructions become functional elements, such as free words becoming inflectional or derivational morphemes. This work came first within "cognitive" or "functionalist" approaches and then has come more recently from "formal" theoretical perspectives. Many recent works pursue such analyses drawing on Germanic data, like Brinton and Traugott (2004) and van Gelderen (2004).

5.4. Periphrasis in verbal syntax

Turning finally to the verbal system, Gothic operated overwhelmingly with the two synthetic tenses reconstructed for Proto-Germanic. Present tense forms covered present and often future meanings, and the preterit covered all “past” tense meanings in some sense (e.g. Wright 1954: 190–192). Even there, however, we see participial/adjectival forms with *wisan* ‘to be’, which at least foreshadow periphrastic structures to come (see Harbert 2007: 293 ff. for accessible discussion of this point in modern terms or Lockwood 1968: 114 ff. for a more traditional treatment). Numerous modern varieties, like Yiddish and southern German dialects, have taken this path to its conclusion, having lost the preterit entirely or almost entirely. Similarly, Standard German and English show continuing retreat of inflected subjunctive forms.

Gothic preserves an inflectional passive (see Faarlund 2004: 126–127 on the rise of a new medio-passive from reflexives in Old Norse), otherwise lost outside of some traces. Even in Gothic, it had a periphrastic variant:

(14) Passives in Gothic

daupjada

baptize-1/3SG.PASS.

‘I am baptized’

(Mark 10:38)

gamēlid

ist

write-PAST.PART. is

‘it is written’

(Luke 2:23, elsewhere)

gaáiwiskōþs

wairþa

make ashamed-PAST.PART. become.1SG.

‘I will be ashamed’

(Philippians 1:20)

Over time, individual languages and dialects have developed new auxiliaries for the passive and future, as laid out in great detail for Swedish by Markey (1969) and much work since. An array of auxiliaries have evolved, often via modals, for marking the future – English *will* and *gonna*, German *werden*, Danish *skal* and *vil*, Icelandic *munu*. At the same time, “present” tense verb forms remain “non-past” in meaning in some languages and dialects, and are readily used for future meanings, e.g. German and Danish.

In syntax too, Germanic varieties show clear lines of parallel development, in the rise of V2 order or of new functional categories and increasing periphrasis in all daughters compared to the proto-language.

6. Conclusion: Conservatism and paths of innovation

No rigid or uniform path of development exists across Germanic, nor are there simplistic connections across all of the trends and features noted above. Nonetheless, this brief

survey suggests that it makes good sense to talk about the “evolution of Germanic”, in the sense of “a process of change in a certain direction” (*Merriam-Webster’s 11th*), despite the inherent difficulties of the “evolutionary metaphor” (cf. Blevins 2004: 17–19).

A fundamentally unresolved question in the history of Germanic is whether these developments are interconnected and, if so, how. To what extent are we dealing with direct genetic inheritance versus instances of parallel (independent) evolution, or convergent evolution (along different pathways but responding to similar pressures)? A full airing of this question will obviously await another forum, but I close by noting some proposals to account for historical similarities in the trajectories of Germanic languages, beginning with drift, a set of changes that are cumulative in some particular direction.

Sapir defines “drift” (1921: ch. 6) as the apparently inexorable movement from more “synthetic” to “analytic” structures, and to increasingly fixed word order, drawing on Germanic examples. He continues (1921: 174):

The mere fact ... that there is a growing tendency to throw the stress automatically on the first syllable of a word may eventually change the fundamental type of the language, reducing its final syllables to zero and driving it to the use of more and more analytical or symbolic methods.

While counterexamples to the role of initial stress are easy to find, similar views abound in the literature, some explicitly connected to “drift” and many others not. While Sapir asserts that “the evidence is overwhelming that this drift has a certain consistent direction” (1921: 171), Hawkins and others have explored how some particular traits do indeed embody these trends and directions, the extent to which they co-occur, precisely how they are interconnected, and so on (see also Keiser 2009).

At the same time, given shared inheritance from Proto-Germanic and constraints on how language changes, strongly parallel outcomes seem highly likely, whether our view of change is anchored in Blevins’ Evolutionary Phonology or the notion of a “life-cycle” of language change from phonetically-driven sound change to phonological generalization, to morphological pattern.

Many others have sought answers to the question in different patterns of change under language contact. O’Neil (1978) lays out a direct correlation between the degree to which a particular language has lost inflectional morphology and the amount of language contact it has undergone. Along similar lines, we have already noted above that many varieties regarded as most conservative are at the geographic periphery of the family and often were long relatively isolated from especially intra-Germanic language contact: Icelandic has been relatively isolated from contact overall, while most varieties of Yiddish have been in heavy contact with non-Germanic languages (especially Slavic) and speech islands in the high Alps, like Highest Alemannic and Cimbrian, are in contact with Romance.

At the other end of the scale, the most familiar answers to the question “what happened to English?” (McWhorter 2002) typically rest on invoking contact in some way, whether in the original mix of speakers who migrated to the British Isles, in contact during the Danelaw or, most often, as a result of the Norman Conquest. Only recently have these proposals come to focus concretely on questions of bilingualism, language acquisition, and transmission. Here too, though, some members of the family swim against the tide. Yiddish is a quintessential case of a language forged in language contact

and shift, and it has been and continues to be largely spoken in diaspora conditions often with widespread bilingualism. Yet it long resisted some of the most characteristic patterns of development associated with contact, such as case loss (see Jacobs 2005: 221–222 for detailed discussion of how contact has and has not shaped Yiddish).

Ultimately, whatever the causes, while we see rich patterns of change and certainly unique patterns of innovation and conservatism in each variety, the picture emerging from this survey is far from anything-goes chaos, as even a few examples remind us. All varieties have undergone vocalic chain shifts along well-defined paths: Time and again, /u:/ has fronted, whether conditioned or not, but front /i:/, y:/ have seldom systematically moved back. All varieties have lost the dual, lost case distinctions, and so on, but none has evolved new case markings or numbers in the nominal system. New ways of marking definiteness in nouns have arisen, and parallel patterns of verbal periphrasis have evolved. While such broad outlines are clear, the particulars of how each community has negotiated change varies, structurally and socially. Despite the vast diachronic literature on Germanic, understanding these interrelations more precisely provides one more engaging challenge for coming generations.

Acknowledgment

I thank the editors for the invitation to write this article, and the following for discussions on this topic and input of various sorts on this project: Elly van Gelderen, Rob Howell, Greg Iverson, Neil Jacobs, Mark Loudon, Monica Macaulay, Tom Purnell, Paul Roberge, and especially Andrea Menz, whose dissertation topic motivated me to explore the questions treated here in the way that they are. The usual disclaimers apply. Some examples and a few short passages closely parallel some of my contributions to Howell, Roberge, and Salmons (in preparation) or Salmons (2012).

7. References

- Anderson, Stephen R.
1974 *The Organization of Phonology*. San Diego: Academic Press.
- Antonsen, Elmer H.
1969 Zur Umlautfeindlichkeit des Oberdeutschen. *Zeitschrift für Dialektologie und Linguistik* 36: 201–207.
- Antonsen, Elmer H.
1975 *A Concise Grammar of the Older Runic Inscriptions*. Tübingen: Niemeyer.
- Antonsen, Elmer H.
2002 *Runes and Germanic Linguistics*. Berlin: Mouton de Gruyter.
- Antonsen, Elmer H.
2003 ‘Weil die Schrift immer strebt ...’: On phonological reconstruction. *North-Western European Language Evolution* 43: 2–20.
- Augst, Gerhard
1975 Wie stark sind die starken Verben? In: Gerhard Augst (ed.), *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr, 231–281.
- Bammesberger, Alfred
1986 *Der Aufbau des germanischen Verbalsystems*. Heidelberg: Winter.

- Bammesberger, Alfred
1990 *Die Morphologie des urgermanischen Nomens*. Heidelberg: Winter.
- Bandle, Oskar, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman (eds.)
2002–2005 *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*. Berlin: De Gruyter.
- Bàrtoli, Matteo
1925 *Introduzione alla Neolinguistica*. Geneva: Olschki.
- Besch, Werner, Anne Betten, Oskar Reichmann, and Stefan Sonderegger (eds.)
2002 *Sprachgeschichte: Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Berlin: De Gruyter.
- Bittner, Andreas
1996 *Starke 'schwache' Verben / schwache 'starke' Verben: Deutsche Verbflexion und Natürlichkeit*. (= Studien zur deutschen Grammatik 51). Tübingen: Stauffenburg.
- Blevins, Juliette
2004 *Evolutionary Phonology*. Cambridge: Cambridge University Press.
- Bloomfield, Leonard
1933 *Language*. New York: Holt, Rinehart, and Winston.
- Bonfante, Giuliano
1947 The Neolinguistic position. *Language* 23: 344–375.
- Boutkan, Dirk
1995 *The Germanic 'Auslautgesetze'*. (Leiden Studies in Indo-European 4). Amsterdam: Rodopi.
- Braune/Ebbinghaus = Braune, Wilhelm
1981 *Gotische Grammatik*. 19th edition by Ernst Ebbinghaus. Tübingen: Max Niemeyer.
- Braune/Reiffenstein = Braune, Wilhelm
2004 *Althochdeutsche Grammatik, I*. 15th edition by Ingo Reiffenstein. Tübingen: Niemeyer.
- Brinton, Laurel and Elizabeth Closs Traugott
2005 *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Brugmann, Karl
1897 *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. 1: *Einleitung und Lautlehre*. Zweite Bearbeitung. Strassburg: Trübner.
- Buccini, Anthony F.
1992 *The development of umlaut and the dialectal position of Dutch in Germanic*. PhD dissertation, Cornell University.
- Buccini, Anthony F.
1995 Onstaan en vroegste ontwikkeling van het Nederlandse taallandschap [The origin and earliest development of the Dutch linguistic landscape]. *Taal en Tongval* 48: 8–66.
- Cambridge History of the English Language*
1992 ff. Cambridge: Cambridge University Press. (Various editors.)
- Campbell, Lyle and Alice C. Harris
2002 Syntactic reconstruction and demythologizing 'Myths and the prehistory of grammars'. *Journal of Linguistics* 38: 599–618.
- Carroll, Ryan, Ragnar Svare, and Joseph Salmons
2012 Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2: 153–172.
- Chambers, Jack K., and Peter Trudgill
1998 [1981] *Dialectology*. Second edition. Cambridge: Cambridge University Press.
- Coetsem, Frans van
1994 *Vocalism of the Germanic Parent Language: Systemic evolution and sociohistorical context*. Heidelberg: Winter.

- Coetsem, Frans van
1996 *Towards a Typology of Lexical Accent: 'Stress Accent' and 'Pitch Accent' in a Renewed Perspective.* (= Monographien zur Sprachwissenschaft 18). Heidelberg: Winter.
- Coetsem, Frans van, Ronald Hendricks, and Susan McCormick
1981 Accent Typology and Sound Change. *Lingua* 53: 295–315.
- Coetsem, Frans van and Anthony F. Buccini
1990 Variation and the reconditioning of phonological rules. *Lingua* 81: 169–220.
- Collinge, N. E.
1985 *The Laws of Indo-European*. Amsterdam: Benjamins.
- Durrell, Martin
2003 From regularity to irregularity in morphology. Ablaut in the West Germanic languages. In: Cornelia Tschichold (ed.), *English Core Linguistics*. Bern: Lang, 125–145.
- Eythórsson, Thórhallur
1994 *Verbal syntax in the early Germanic languages*. Ph.D. dissertation, Cornell University.
- Faarlund, Jan Terje
2004 *The Syntax of Old Norse*. Oxford: Oxford University Press.
- Ferraresi, Gisella
2005 *Word Order and Phrase Structure in Gothic. A Comparative Study*. Leuven: Peeters.
- Fertig, David L.
2000 *Morphological Change Up Close: Two and a Half Centuries of Verbal Inflection in Nuremberg*. Tübingen: Niemeyer.
- Gelderen, Elly van
2004 *Grammaticalization as Economy*. Amsterdam: Benjamins.
- Gelderen, Elly van
2007 The definiteness cycle in Germanic. *Journal of Germanic Linguistics* 19: 275–308.
- Gelderen, Elly van
2011 *The Linguistic Cycle: Language change and the language faculty*. Oxford: Oxford University Press.
- Goetting, Lauren
2007 Greek textual influence on Gothic complex verbs with pleonastic prepositions. *Journal of Germanic Linguistics* 19: 309–347.
- Guxman, M. M., Viktor M. Žirmunskij, E. A. Makaev, and V. N. Jarceva (eds.)
1962–1966 *Sravnitel'naja grammatika germanskix jazykov* [Comparative grammar of the Germanic languages]. 4 vols. Moscow: Nauka.
- Harbert, Wayne
2007 *The Germanic Languages*. Cambridge: Cambridge University Press.
- Hare, Mary and Jeff Elman
1995 Learning and morphological change. *Cognition* 56: 61–98.
- Haugen, Einar
1976 *The Scandinavian Languages*. London: Faber and Faber.
- Hawkins, John A.
1985 *A comparative typology of English and German: Unifying the contrasts*. Austin: University of Texas Press.
- Hewson, John
2006 From Ancient to Modern Germanic. In: John Hewson and Vit Bubenik (eds.), *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam: Benjamins, 274–303.
- Hirt, Hermann
1931, 1932, 1934 *Handbuch des Urgermanischen. Teil 1: Laut- und Akzentlehre. Teil 2: Stammbildungs- und Flexionslehre. Teil 3: Abriss der Syntax*. Heidelberg: Winter.

- Holsinger, David J.
2001 Weak Position Constraints: The Role of Prosodic Templates in Contrast Distribution. In: Tracy A. Hall and Marzena Rochon (eds.), *Zentrum für Allgemeine Sprachwissenschaft Papers in Linguistics* 19. Berlin: Humboldt Universität, 91–118.
- Holsinger, David J.
2008 Germanic prosody and consonantal strength. In: Joaquim Brandão de Carvalho, Tobias Scheer, and Philippe Ségéral, *Lenition and Fortition*. Berlin: Mouton de Gruyter, 273–300.
- Holthausen, Ferdinand
1921 *Altsächsisches Elementarbuch*. Heidelberg: Winter.
- Honeybone, Patrick
2001 Lenition inhibition in Liverpool English. *English Language and Linguistics* 5: 213–249.
- Honeybone, Patrick
2002 *Germanic obstruent lenition: some mutual implications of theoretical and historical phonology*. PhD dissertation, University of Newcastle upon Tyne.
- Honeybone, Patrick
2005 Sharing makes us stronger: Process inhibition and segmental structure. In: Philip Carr, Jacques Durand, and Colin Ewen (eds.), *Headhood, Elements, Specification and Contrastivity*. Amsterdam: Benjamins, 167–192.
- Honeybone, Patrick
Forthcoming *Theoretical Historical Phonology: Lenition, laryngeal realism and Germanic obstruent shifts*. Oxford: Oxford University Press.
- Howe, Steven
1996 *The Personal Pronouns in the Germanic Languages: A Study of Personal Pronoun Morphology and Change in the Germanic Languages from the First Records to the Present Day*. Berlin: De Gruyter.
- Howell, Robert B.
1991 *Old English breaking and its Germanic analogues*. Tübingen: Niemeyer.
- Howell, Robert B. and Joseph C. Salmons
1997 Umlautless residues in Germanic. *American Journal of Germanic Linguistics & Literatures* 9: 83–111.
- Howell, Robert B., Paul T. Roberge, and Joseph Salmons
In preparation *Cambridge History of the Germanic Languages*. Cambridge: Cambridge University Press.
- Hutterer, Claus Jürgen
1990 *Die germanischen Sprachen: Ihre Geschichte in Grundzügen*. 3rd edition. Wiesbaden: Drei Lilien.
- Iverson, Gregory K. and Joseph C. Salmons
2003a Laryngeal Enhancement in Early Germanic. *Phonology* 20: 43–72.
- Iverson, Gregory K. and Joseph C. Salmons
2003b The ingenerate motivation of sound change. In: Raymond Hickey (ed.), *Motives for Language Change*. Cambridge: Cambridge University Press, 199–212.
- Iverson, Gregory K. and Joseph C. Salmons
2004 The conundrum of Old Norse umlaut. *Journal of Germanic Linguistics* 16: 77–110.
- Iverson, Gregory K. and Joseph C. Salmons
2012 Parasitic rule loss in Norse umlaut. *Journal of Germanic Linguistics* 24: 101–131.
- Jacewicz, Ewa, Joseph Salmons, and Robert Fox
2006 Prosodic prominence effects on vowels in chain shifts. *Language Variation & Change* 18: 285–316.
- Jacobs, Neil G.
2005 *Yiddish: A linguistic introduction*. Cambridge: Cambridge University Press.

- Keiser, Steven Hartmann
 2009 When 'speech islands' aren't islands: parallel independent development, drift, and minimal levels of contact for diffusion. *Diachronica* 26: 1–35.
- Kiparsky, Paul
 1995 Indo-European origins of Germanic syntax. In: Adrian Battye and Ian Roberts (eds.), *Clause Structure and Language Change*. New York: Oxford University Press, 140–169.
- Klein, Jared S.
 1992 On the Idiomatic Nature of the Gothic New Testament: A Comparative Study of Prepositional Usage in Gothic and New Testament Greek. *Transactions of the Philological Society* 90: 1–80.
- Klein, Jared S.
 1994 Gothic *paruh*, *panuh* and *-(u)h þan*. *Indogermanische Forschungen* 99: 253–276.
- Klein, Jared S and Nancy L. Condon
 1993 Gothic *-u(h)*: A Synchronic and Comparative Study. *Transactions of the Philological Society* 91: 1–62.
- König, Ekkehard and Johan van der Auwera (eds.)
 1994 *The Germanic Languages*. London: Routledge.
- Kranzmayer, Eberhard
 1956 *Historische Lautgeographie des gesamt-bairischen Dialektraumes*. Wien: H. Böhlau.
- Kranzmayer, Eberhard
 1960 Die Sprachaltertümer in den Mundarten der Tiroler Hochtäler. *Zeitschrift für Mundartforschung* 27: 160–192.
- Kress, Bruno
 1982 *Isländische Grammatik*. Leipzig: Hueber.
- Kufner, Herbert L.
 1972 The grouping and separation of the Germanic languages. In: Frans van Coetsem and Herbert L. Kufner (eds.), *Toward a Grammar of Proto-Germanic*. Tübingen: Niemeyer, 71–98.
- Küspert, Klaus-Christian
 1988 *Vokalsysteme im Westnordischen: Isländisch, Färöisch, Westnorwegisch: Prinzipien der Differenzierung*. (= Linguistische Arbeiten 198). Tübingen: Niemeyer.
- Labov, William
 1994 *Principles of Linguistic Change*. Volume 1: Internal Factors. Oxford: Blackwell.
- Lahiri, Aditi, Tomas Riad and Haïke Jacobs
 1999 Diachronic prosody. In: Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter, 335–422.
- Lerchner, Gotthard
 1971 *Zur II. Lautverschiebung im Rheinisch-Westmitteldeutschen*. Halle: Niemeyer.
- Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A. Nowak
 2007 Quantifying the evolutionary dynamics of language. *Nature* 449: 713–716.
- Lightfoot, David W.
 2002 Myths and the prehistory of grammars. *Journal of Linguistics* 38: 113–136.
- Lockwood, William B.
 1968 *Historical German Syntax*. London: Oxford University Press.
- Louden, Mark L.
 1992 Language contact and the relationship of form and meaning in English and German. In: Rosina Lippi-Green (ed.), *Recent Developments in Germanic Linguistics*. Amsterdam: Benjamins, 115–125.
- Luiten, Tyler, Andrea Menz, Angela Bagwell, Benjamin Frey, John Lindner, Mike Olson, Kristin Speth, and Joseph Salmons
 2013 Beyond the handbooks: a quantitative approach to analysis of Old High German phonology and morphology. *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)*. 135. 1–18.

- Mailhammer, Robert
2007a Islands of resilience: The history of the verbs from a systemic point of view. *Morphology* 17: 77–108.
- Mailhammer, Robert
2007b *The Germanic strong verbs: Foundations and development of a new system*. Berlin: Mouton de Gruyter.
- Markey, Thomas L.
1969 *The verbs varda and bliva in Scandinavian with Special Emphasis on Swedish*. (= *Studia philologiae Scandinavicae Upsaliensia* 7). Uppsala: Almqvist & Wiksell.
- McWhorter, John H.
2002 What happened to English? *Diachronica* 19: 217–272.
- Meillet, Antoine
1917 *Caractères généraux des langues germaniques*. Paris: Hachette.
- Menz, Andrea
2010 *Accent Type and Language Change in Germanic and Baltic-Finnic*. Ph.D. dissertation, University of Wisconsin – Madison.
- Munske, Horst Haider (ed.)
2001 *Handbuch des Friesischen / Handbook of Frisian Studies*. Produced in collaboration with Nils Århammar, Volkert F. Falktings, Jarich Hoekstra, Oebele Vries, Alastair G. H. Walker, and Ommo Wilts. Tübingen: Niemeyer.
- Murray, Robert W.
1988 *Phonological Strength and Early Germanic Syllable Structure*. Munich: Fink.
- Murray, Robert W.
2000 Syllable Cut Prosody in Early Middle English. *Language* 76: 617–654.
- Murray, Robert W. and Theo Vennemann
1983 Sound change and syllable structure in Germanic phonology. *Language* 59: 514–528.
- Nielsen, Hans Frede
1989 *The Germanic Languages: Origins and early dialectal interactions*. Tuscaloosa, AL: University of Alabama Press. (Translation and update of his 1979 *De germanske sprøg*.)
- Nielsen, Hans Frede
2000 *The Early Runic Language of Scandinavia. Studies in Germanic Dialect Geography*. Heidelberg: Winter.
- O’Neil, Wayne
1978 The evolution of the Germanic inflectional systems: A study in the causes of language change. *Orbis* 27: 248–286.
- Page, B. Richard
1999 The Germanic *Verschärfung* and Prosodic Change. *Diachronica* 16: 297–334.
- Page, B. Richard
2006 The diachrony and synchrony of vowel quantity in English and Dutch. *Diachronica* 23: 63–104.
- Penzl, Herbert
1995 Zum *i*-Umlaut in den althochdeutschen Dialekten. *Folia Linguistica Historica* 16: 115–21.
- Platzack, Christer
1985 A survey of generative analyses of the verb second phenomenon in Germanic. *Nordic Journal of Linguistics* 8: 49–73.
- Polomé, Edgar C.
1992 Zur Chronologie des Germanischen. In: Robert Beekes, Alexander Lubotsky, and Jos Weitenberg (eds.), *Rekonstruktion und Relative Chronologie. Akten der VIII. Fachtagung der Indogermanischen Gesellschaft Leiden, 31. August–4. September 1987*. Innsbruck: Institut für Sprachwissenschaft der Universität, 55–74.

- Prokosch, Eduard
1938 *A Comparative Germanic Grammar*. (= William Dwight Whitney Linguistics Series). Baltimore: Linguistic Society of America.
- Raven, Frithjof A.
1963 *Die schwachen Verben des Althochdeutschen. Band I: Lang-, mehr- und kurzsilbige -jan Verba*. Gießen: Schmitz. (Repr. Univ. of Alabama Press, 1964.)
- Raven, Frithjof A.
1963 *Die schwachen Verben des Althochdeutschen. Band II: Die ôn- und ên-Verba*. Gießen: Schmitz. (Repr. Univ. of Alabama Press, 1967.)
- Riad, Tomas
1992 *Structures in Germanic prosody: A diachronic study with special reference to the Nordic languages*. Stockholm: Department of Scandinavian Languages, Stockholm University. (Ph.D. dissertation.)
- Ringe, Don
2006 *From Proto-Indo-European to Proto-Germanic*. (A Linguistic History of English 1). Oxford: Oxford University Press.
- Roberge, Paul T.
1994 *The formation of Afrikaans*. [= SPIL PLUS 23]. Stellenbosch: Stellenbosch Papers in Linguistics.
- Robinson, Orrin W.
1992 *Old English and its Closest Relatives: A Survey of the Earliest Germanic Languages*. Stanford: Stanford University Press.
- Salmons, Joseph
2012 *A History of German: What the past reveals about today's language*. Oxford: Oxford University Press.
- Salmons, Joseph, Robert Fox, and Ewa Jacewicz
2012 Prosodic skewing of input and the initiation of cross-generational sound change. *The initiation of sound change: Production, perception and social factors*, ed. by Maria-Josep Solé and Daniel Recasens. Amsterdam: Benjamins, 167–184.
- Sapir, Edward
1921 *Language: An introduction to the study of speech*. New York: Harcourt Brace.
- Schirmunski, Viktor M
1962 *Deutsche Mundartforschung: Vergleichende Laut- und Formenlehre der deutschen Mundarten*. Berlin: Akademie Verlag.
- Schmitt, Alfred
1924 *Untersuchungen zur allgemeinen Akzentlehre*. Heidelberg: Winter.
- Schmitt, Ludwig Erich (ed.)
1970–1971 *Kurzer Grundriß der germanischen Philologie bis 1500*. 2 vols. Berlin: De Gruyter.
- Schönfeld/Van Loey = Schönfeld, Moritz
1970 *Historische Grammatica van het Nederlands* [Historical grammar of Dutch]. Edited by Adolphe van Loey. Zutphen: Thieme.
- Shannon, Thomas F.
1990 Nederlands tussen Engels en Duits: A typological comparison [Dutch between English and German: ...]. In: Magriet Lacy (ed.), *The Low Countries: Multidisciplinary studies*. Lanham, MD: University Press of America, 45–60.
- Sievers, Eduard
1876 *Grundzüge der Lautphysiologie: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. (= Bibliothek indogermanischer Grammatiken 1). Leipzig: Breitkopf & Härtel. (Later, beginning with the widely-cited 1881 edition, known as *Grundzüge der Phonetik*.)

- Sihler, Andrew L.
1995 *New Comparative Grammar of Greek and Latin*. Oxford: Oxford University Press.
- Smith, Laura Catharine
2007 The resilience of prosodic templates. In: Joseph Salmons and Shannon Dubenion-Smith (eds.), *Historical Linguistics 2005: Selected papers from the 17th International Conference on Historical Linguistics, Madison, 31 July–5 August 2005*. Amsterdam: Benjamins, 351–365.
- Stockwell, Robert
1978 Perseverance in the English Vowel Shift. In: Jacek Fisiak (ed.), *Recent Developments in Historical Phonology*. The Hague: Mouton, 337–348.
- Twaddell, W. Freeman
1938 A note on OHG umlaut. *Monatshefte* 30: 177–81.
- Tyroller, Hans
2003 *Grammatische Beschreibung des Zimbrischen von Lusern*. Stuttgart: Steiner.
- Voyles, Joseph B.
1992 *Early Germanic Grammar: Pre-, Proto,- and Post-Germanic languages*. San Diego: Academic Press.
- Walkden, George
2014 *Syntactic Reconstruction and Proto-Germanic*. Oxford: Oxford University Press.
- Weerman, Fred
1986 *The V2 conspiracy: A synchronic and diachronic analysis of verbal positions in the Germanic languages*. (= Publications in language sciences 31). Dordrecht: Foris.
- Wegener, Heide
2007 Entwicklungen im heutigen Deutschen – Wird Deutsch einfacher? *Deutsche Sprache* 35: 35–62.
- Wiesinger, Peter
1970 *Phonetisch-phonologische Untersuchungen zur Vokalentwicklung in den deutschen Dialekten*. Bd. 1: *Langvokale im Hochdeutschen*; Bd. 2: *Diphthonge im Hochdeutschen*. Berlin: De Gruyter.
- Wiesinger, Peter
1983 Phonologische Vokalsysteme deutscher Dialekte: Ein synchronischer und diachronischer Überblick. In: Werner Besch (ed.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung* 1. Berlin: De Gruyter. 1042–1076.
- Wipf, Elisa
1910 *Die Mundart von Visperterminen im Wallis*. Frauenfeld: Huber.
- Wright, Joseph
1954 *Grammar of the Gothic Language*. 2nd edition by O. L. Sayce. Oxford: The Clarendon Press.

Joseph Salmons, Madison, WI (USA)

X. Armenian

60. The documentation of Armenian

1. Introduction
2. The early period (5th–11th centuries)
3. The middle period (12th–16th centuries)
4. The modern period (17th–21st centuries)
5. References

1. Introduction

Historical Armenia is a large plateau roughly bounded by the Caucasus Mountains, the Black and the Caspian Seas, Iran and Mesopotamia. The highest mountain is Mount Ararat, and the country has three major lakes: Urmia, Van, and Sevan. The most important river is the Araxes.

There are different theories about the origin of the Armenian people. The scholars who consider the steppes in the Southern part of Russia to be the homeland of the Indo-Europeans suppose that the Armenians arrived in their country from the Balkan Peninsula, probably during the second millennium BCE. However, according to other scholars, e.g. Thomas V. Gamkrelidze and Vjačeslav V. Ivanov, the Indo-European homeland roughly coincided with historical Armenia, so the Armenians were just the descendants of the Indo-European tribes settled in this area. A considerably different opinion is held by Mario Alinei. According to his “Paleolithic Continuity Theory”, no Indo-European invasion took place and the Indo-European languages spoken in Europe including, as we can guess, also Armenian, just continue the languages of the inhabitants of the Paleolithic period.

The history of the Armenian language can be divided into three main periods. For further information on this history see Ačařyan (1951); Nichanian (1989); Łazarean (2007). Discussion of toponomastic and anthroponymic issues is beyond our immediate concern.

2. The early period (5th–11th centuries)

2.1. We know nearly nothing about the Armenian language before the 5th century CE, when the Armenian script was invented by Mařtoc', a clergyman also called Mesrop in supposedly later sources. Around 387 CE, Armenia was divided between the Byzantine and the Sasanian states. The part under Byzantine influence was soon annexed by the empire, while the other maintained a sort of independence until 428. The most enlightened persons of the 5th century, such as the patriarch Sahak, Mařtoc', and King Vřamřapowh, were well aware that, with the loss of political independence, the existence of the Armenian ethnos was also threatened. The country had been officially Christianized in the early 4th century (the traditional date is 301), but the Armenian approach to Christianity was not fully acceptable to the Byzantines, whereas the Persians suspected that the Armenians, as a Christian people, certainly were in tacit agreement with the Western

enemy. Thus, the commitment of the Armenian language to writing was first of all necessary for survival and, furthermore, for religious purposes, in order to facilitate the preaching of the Christian faith among the people. Therefore, at the end of the 4th or the beginning of the 5th century, the élite undertook the task of inventing a script for the Armenian language, and it was brilliantly performed by Maštoc' (traditionally, in 405–406).

2.2. After the invention of the alphabet, Maštoc' and his pupils involved themselves intensively in translation. Naturally, the first translated text was the Bible, a critical edition of which remains to this day a desideratum. The old edition by Zohrab (Yovhannēs Zōhrapean) printed in 1805 in Venice and reprinted as a facsimile in Delmar, New York, in 1984, offers only some variant readings, without any indication concerning the manuscript or manuscripts from which they are taken. Between 1985 and 2002 some Old Testament books (the Pentateuch, the Twelve Prophets, Maccabees) have been published in the series *Hay hnagoyn t'argmanakan yowšarjanner / Hay hnagowyn t'argmanakan howšarjanner* ('Oldest Armenian Translated Literary Monuments'), which was printed in Erevan, Ējmiacin (Armenia) or Antelias (Lebanon), while again Deuteronomy, Job, Daniel, and some Pseudepigrapha are available in other series.

As far as the New Testament is concerned, Künzle (1984) offers the transcribed text of the second and third-oldest manuscripts of the Gospels (respectively, Matenadaran 6200, copied in 887, and Matenadaran 2374, copied in 989), while the critical edition of the Acts of the Apostles has just appeared in the *Corpus Scriptorum Christianorum Orientalium* (C.S.C.O.), edited in Louvain/Leuven.

After the Bible, works of Greek and Syriac authors, such as John Chrysostomus, Basil of Caesarea, Eusebius of Caesarea, Gregory of Nazianzus, Ephrem Syrus, Aphraates, and others were also translated into the classical Armenian language. The critical edition of some of these translations is available in C.S.C.O. Later on, probably starting from the beginning of the 6th century, another generation of translators became active. Their aim was to reproduce the original, usually Greek, text as faithfully as possible. In this way they translated many works pertaining to different fields: grammar (the *Tekhnē Grammatikē* by Dionysius Thrax), rhetoric (the *Progymnasmata* by Aelius Theon), philosophy (some works by Aristotle, Porphyry, David the Invincible, perhaps also Plato), as well as a number of books by Philo, Irenaeus, Timotheus Aelurus and others. Though the language of these translations is very artificial, containing numerous syntactic, semantic, and morphological calques, the translated texts of the so-called *Yownaban dproc'* ('Hellenizing school') remained as the reference books of higher learning for centuries. They provided the modern Armenian language with a rich vocabulary and especially with numerous technical terms used in various scientific fields. On the language of the Hellenizing school see Muradyan (2012).

2.3. Starting from the 5th century, original works were also composed in addition to translations. Among those works are the *Elc Alandoc'* ('Refutation of the Sects') by Eznik Kołbac'i, discussing religious subjects, and the *Vark' Maštoc'i* ('Life of Maštoc'') by Maštoc''s pupil Koriwn. These are the most ancient original writings composed in Classical Armenian (also called *grabar*). In the second half of the 5th century, a series of historiographic works were also written, among which we can mention the

Patmowt'iwn Hayoc' ('History of the Armenians') by Agat'angelos; the anonymous *Bowzandaran Patmowt'iwnk'* ('Epic Histories'), also known as the *Patmowt'iwn Hayoc'* ascribed to a certain P'awstos Bowzand; and the *Patmowt'iwn Hayoc'* by Łazar P'arpec'i. These three books form a sort of continuum on the history of Armenia, starting from the account of the events which led to the conversion of the country to Christianity and ending with the rebellion against the Persians in the second half of the 5th century. That rebellion is also dealt with in *Vasn Vardanay ew Hayoc' Paterazmin* ('On Vardan and the Armenian War') by Elišē, a late-5th or 6th-century text. Another important historiographic work is the *Patmowt'iwn Hayoc'* (beginning with the events of earliest antiquity and reaching the year 439) by Movsēs Xorenac'i, traditionally considered a 5th-century author. In the opinion of some scholars, however, this work was composed later. It is impossible to list the editions of these notable texts in this brief overview; for details see Thomson (1995, 2007). It is worth mentioning, however, that original works from the 5th century onwards, the text of which is often based on earlier editions, are collected in the *Matenagirk' Hayoc'/Armenian Classical Authors* series, a work in progress, sponsored by the Calouste Gulbenkian Foundation, the Armenian Catholicosate of Cilicia and, from the 15th volume onwards, also by the Matenadaran of Erevan. Nineteen volumes, encompassing authors from the 5th to the 12th century, have been published so far, since 2003, first in Antelias and then in Erevan. Some of these volumes are also available at: <http://www.matenadaran.am/?id=83&lng=3> (accessed on 2 February 2017).

2.4. Scholars dealing with Armenian literary sources know well that those texts have come down to us through manuscripts copied in much later periods. For instance, the oldest complete manuscripts of the Gospels are codex 1144/86 of the Venice Mechitarist Library (the so-called Queen Mlk'ē's Gospel), copied in or just before 862 and the above-mentioned codex 6200 of the Matenadaran, copied in 887. Thus, since the Gospels were translated into Armenian in the first half of the 5th century, there is a span of four centuries between the translation and the oldest surviving complete copies. The situation is even worse when it comes to original works: the oldest complete witness to the biography of Maštoc' by Koriwn was probably copied between 1675 and 1703 and added as an insert to Matenadaran 2639 (dated 1672); Eznik's treatise is known to us thanks to one manuscript (Matenadaran 1097) dated 1280; the oldest complete manuscript of the *Bowzandaran* is Jerusalem 341, dated 1599, while that of Agat'angelos was copied in the 12th century (Matenadaran 3782); the oldest copy of Łazar P'arpec'i's *History* is Matenadaran 2639, dated 1672, and that of Movsēs Xorenac'i's work is Matenadaran 2865, partly copied in the 14th century and partly in 1567. As we have already stated, all these works were composed in the 5th century, with the possible exception of Movsēs Xorenac'i's *History*. Thus, one can wonder whether a linguistic feature found in a text is really from the 5th century or pertains to a later period of the Armenian language. The limited number of older surviving fragments containing parts of those texts, dated only on paleographical bases, cannot really change this situation, though at times they permit us to evaluate how much a text has been changed during the manuscript tradition (see for instance, the large fragment of Agat'angelos, preserved in a palimpsest of the Mechitarist Library of Vienna, n. 56 according to Tašean's catalogue, and recently studied by Topchyan 2009).

2.5. In addition to these literary texts, the linguist has also a number of early medieval inscriptions at his disposal, the dates of which are usually known, as well as graffiti,

many of them from Sinai. Some of them go back to the first centuries of the history of Armenian, as e.g. the two graffiti of Nazareth and the inscription of Tekor (all three dating to the 5th century). The information drawn from these sources, albeit often reduced to proper names (especially in many graffiti), reflects the oldest phase of the language. The Armenian inscriptions are collected in the *Divan hay vimagrowt'yan/ Corpus Inscriptionum Armenicarum*, 9 volumes of which have been published so far in Erevan (between 1966 and 2012). In this series, also available at: <http://serials.flib.sci.am/openreader/test/index.html> (accessed on 2 February 2017), the inscriptions are presented according to their location. For a synopsis of the early Armenian inscriptions, see Greenwood (2004).

2.6. Loanwords, both into and from Armenian, represent another source of possible linguistic information. As a result of linguistic contacts before or in the 5th century or later, Armenian borrowed a certain number of Iranian, Greek, and Syriac words and proper names. Such loans bear witness to the phonological status of Armenian or, in some cases, of a variety of it. For example, there are two lateral consonants in Armenian, which are represented by separate letters transliterated as <l> and <ł>. In many loans from Greek, the “lambda” is rendered by <ł> (e.g. *hiwl* ‘matter’, *salmos* ‘psalm’, *mlon* ‘mile’, *taland* ‘talent’ and some names: *Atek'sandros*, *Lazar*, *Agat'angelos* etc.), and this is the case in some loans from Syriac as well. All these words and names are attested in 5th-century works; consequently, the way of rendering the Greek lateral should be evaluated in order to establish the exact phonological value of the two lateral consonants in 5th-century Armenian.

Linguistic information can also be obtained through Armenian words preserved in foreign languages. For example, in the so-called *Narratio de Rebus Armeniae*, a history of the Armenian Church composed in Greek by an Armenian at the beginning of the 8th century (text in C.S.C.O., Louvain, 1952), one can find numerous Armenian toponyms and personal names in the Greek script, which could give information about the phonology of at least a variety of the Armenian language.

2.7. Another source of information is provided by some authors who at times speak about peculiarities of Armenian or refer to current opinions about it. For instance, in the *Elc Alandoc'* by Eznik (I, 23) and in the Armenian version of Dionysius Thrax (Adontz 1970: 14), alleged dialectal words are mentioned. Furthermore, the 8th-century commentator on Dionysius, Step'annos Siwnec'i, offers a list of alleged Armenian dialects (Adontz 1970: 187).

2.8. Finally, there are two unique documents from the first period of the history of Armenian which provide information about the language. The first is a papyrus coming from Egypt. It previously consisted of four fragments, which are now attached to one another as if they represent one unbroken text. The papyrus is housed in the Bibliothèque Nationale of Paris (BnF Arm 332). Both the recto and the verso contain texts, a sort of sketch of the Greek language (short conversational phrases, verbal paradigms, lists of words, some stories and maxims) written in Armenian script. Thanks to this document, the corresponding Armenian and Greek letters can be compared and the phonological values of certain Armenian sounds can be clarified. Based on paleographic features, the

papyrus has been dated to the period between the 5th and 7th centuries, making it one of the oldest witnesses to the Armenian script. See the text in Clackson (2000, 2002).

The other important document is a short Latin-Armenian glossary, copied at the end of a manuscript kept in the seminary of Autun (France). The manuscript was supposedly written at the end of the 9th or the beginning of the 10th century, but it seems that the glossary is a copy of an earlier text reflecting an older status of Armenian. The Autun glossary contains ninety entries, all in Latin script, including days of the week, numerals, nouns referring to food and drinks, the sun and stars, parts of the body, and religious concepts. Rendering Armenian by Latin characters, this document gives a dialectal pronunciation of the corresponding words. See the text in Carrière (1886).

3. The middle period (12th–16th centuries)

3.1. Classical Armenian was obviously a literary language, written and also spoken alongside other varieties of Armenian. In the course of time it became only a learned language, while the common people used different vernaculars (different forms of *ašxarhabar*), one of which gradually became the basis of a new literary language in the period in question. Also, two new letters, <ō> and <f>, were added to the Armenian alphabet during this period.

3.2. From the historical point of view, one of the most noteworthy facts in this period is the foundation of an independent Armenian Kingdom in Cilicia, that is to say, outside historical Armenia. Probably starting from the 9th or 10th century, Armenians migrated to this country. Later on a principality was created and finally, in 1198 or 1199, Prince Lewon II acquired royal status, thereby becoming Lewon I and receiving his crown from a legate of Emperor Henry VI. During the Crusades, the Armenian kingdom was an important ally of the Western powers. Within the kingdom itself, a composite society was created, where both Armenians and Europeans (“Franks” according to the terminology of that time) lived and worked side by side. The kingdom ceased to exist in 1375, but what happened in it during a period of less than two centuries is worth noting, also from a linguistic point of view.

3.3. In this period, the split between the classical and the spoken languages was completed, and the former remained as the language of culture, especially religion. On the other hand, the administrative cadres of the new kingdom were people who only partially mastered the classical language. Therefore, a specific literature was created for them, dealing with technical subjects and written in a language more or less close to the spoken one. Such works were both original and translated (from Syriac, Arabic, and French), pertaining to medicine, veterinary science, agronomy, and law. They were written in a language reflecting features of Western Armenian dialects. Among those works are the *Ĵermanc' mxit'arowt'iwn* (‘Consolation of Fevers’), written in 1184 by Mxit'ar Herac'i, the *Datastanagirk'* (‘Law-Code’) by Smbat Sparapet (1208–1276), the *Ansiz Antiok'ay* (‘Assizes of Antioch’), the *Girk' vastakoc'* (‘Book of Farm Labors’), the *Bžškarán jioy*

ew aṙhasarak grastnoy ('Medical Book for Horses and for Beasts of Burden in General').

As we have already stated, Cilician Armenians were in touch with Europeans, especially nobles and merchants, and at times Cilician Armenian kings granted or renewed trade privileges to some of them (Genoese, Venetians, merchants from Catalonia, Provence, Montpellier, and others). Those documents, too, were often written in a language close to spoken Armenian: they are collected in Langlois (1863).

The influence of the "Franks" on Cilician Armenia was evident in many aspects of daily life. It is also documented by many loanwords from European languages, especially French, which penetrated into Armenian, and which provide information on the phonology of both the source and the recipient languages.

3.4. We should also note the Latin cultural and linguistic influence, especially on the church. In this respect, the so-called *Fratres Unitores* (*Elbark' Miabanolk'*) played a significant role. They were a group of friars, both of European and Armenian origin, who helped Bartolomeo di Bologna (or de Podio, †1333) in his missionary activity in Armenia and continued his mission after Bartolomeo's death. These *Fratres*, who became active in the first half of the 14th century, wrote in Classical Armenian (or rather, what was supposed to be Classical Armenian in that period), because *grabar* remained the language of educated people. The *Unitores* also dealt with grammar, and one of them, Yovhannēs K'ṙnec'i (†1347), composed a grammatical work. Unlike previous grammars, it is not a commentary on the Armenian version of Dionysius Thrax's *Tekhnē Grammatikē*. The author shows knowledge of the works of Latin grammarians and devotes a part of his work to syntactic problems. Yovhannēs was also the first Armenian grammarian who gave examples from Middle Armenian, the current language spoken and written in his time.

4. The modern period (17th–21st centuries)

4.1. After the fall of the Cilician Kingdom, no independent Armenian state existed until the 20th century. In this time-span the Armenians were Ottoman or Persian, later on also Russian subjects, often involved in world trade. At the beginning of the 17th century, New Julfa was founded on the outskirts of Isfahan, becoming an important trade and cultural center. The Armenians also had trade colonies in Amsterdam, Venice, Leghorn (Livorno), Marseille, and elsewhere.

4.2. The first Armenian books were printed in Venice, circa 1511, by Yakob Melapart ('The Sinner') who published at least five books. Only one of them, a mass-book, is dated (1513). The others, printed before or after this missal, had more popular contents (prayers and spells, horoscopes, a calendar, and poetry) and were addressed especially to merchants. After Yakob, the Armenians founded other printing houses, again in Venice, but also in Leghorn (Livorno), Amsterdam, Marseille, and other Western and Eastern

cities. Worth mentioning is the Bible printed in 1666–1668 by Oskan vardapet Erewanc'i (1614–1674) in Amsterdam.

4.3. Simultaneously with the more or less ephemeral activity of Armenian publishers, two Western printing houses which also printed books in Armenian, were founded: one in Milan, at the Biblioteca Ambrosiana (a library founded by Cardinal Federico Borromeo and inaugurated in 1609), and the other in Rome, by the Sacra Congregatio de Propaganda Fide, a congregation founded in 1622 by Pope Gregory XV.

These printing houses also published dictionaries and grammars of Classical Armenian, written by both Armenian and European scholars. The novelty of these grammars was that they were not traditional commentaries on the Armenian version of Dionysius Thrax, which until the 17th century was conceived by the Armenians as the only grammar (the above-mentioned work by Yovhannēs K'fneč'i was an exception). They applied the current Western framework (represented by the so-called “extended Latin grammar”, as it was defined by Sylvain Auroux) to their works. We cannot discuss all those works now; let us just mention Francesco Rivola's *Dictionarium Armeno-Latinum*, printed in Milan in 1621 and reprinted in Paris in 1633. The author, putting together Armenian words and loans from other languages, offers an overview of what was regarded as Classical Armenian in the 17th century. Rivola also wrote a *Grammatica Armena* (Milan 1624, reprinted in 1634 in Paris). Among the Propaganda's productions are the grammars written by Yovhannēs Holov (or Ioannes Agop, 1635–1691), printed in 1674 and 1675. All the works just mentioned, particularly those published by the Propaganda, are also notable for another reason: they not only imitate the Latin model in their grammatical description, but also try to mold “Classical Armenian” on Latin, thus creating an artificial language, the so-called *latinatip* or *latinaban hayerēn* (‘Latinized Armenian’) used in both original and translated works.

4.4. In addition to the “classical” language, a new variety of written Armenian was in use in this period: it was the language of the merchants coming from different regions, known as *vačařakanakan hayerēn* (‘merchant Armenian’). This form of Armenian was very close to the so-called *lingua civilis*, about which both Yovhannēs Holov (e.g. in his *Puritas Haygica*, Rome, 1675: 1) and Johannes Joachim Schröder (1680–1756) in his *Thesaurus Linguae Armenicae* (Amsterdam, 1711: 301–302) speak, describing it as a mixed language, halfway between the language of the learned and that of uneducated people, used both by preachers and merchants.

While in the 17th century *vačařakanakan hayerēn* was a kind of unitary language, containing features of Western or Eastern dialects (according to the origin of the writer), in the following century, on the contrary, it progressively split into two varieties: Eastern and Western Armenian. *Vačařakanakan hayerēn* was used to write letters, ledgers or travel journals like the *Ōragrowth'iwn* (‘Diary’) by the merchant Zak'aria Agowlec'i (1630–1691). Also, books for merchants were printed in that language: for example, the *Arhest hamarolowt'ean ambolj ew katareal* (‘The Complete and Perfect Art of Calculation’), a handbook of elementary arithmetic printed in Marseille in 1675, or the *Ganj čap'oy, kšřoy, t'woy ew dramic' bolor ašxarhi* (‘Treasury of Measures, Weights, Numbers and Coins of the Whole World’), printed in Amsterdam in 1699. The dialogues contained in the *Skzbownk' italakani lezowi* (‘A Primer of the Italian Language’), itself a part of the *Girk' aybowbenic' ew kerp owsaneloy zlezown italakan* (‘Spelling-book and a Way

of Learning the Italian Language', Marseille, 1675) are also written in *vačarakanakan hayerēn*.

4.5. As already stated, in the 18th century we find two varieties of the written language, which progressively became quite different from each other. The classical language, too, was still in use, but it was full of loans from several contemporary languages (Turkish, Persian, and Arabic). Besides, morphological and syntactic calques from Latin were also abundant in it; therefore, one of the main goals of the Mechitarist fathers, a congregation founded by Mxit'ar Sebastac'i (1676–1749) in Constantinople in 1700, was to cleanse the classical language of all these “foreign elements”. However, in doing so they stimulated a similar puristic approach towards Modern Armenian. In 1715 Mxit'ar moved to Venice (from 1717 onward he was in San Lazzaro). After his death the Mechitarists split: a group of friars moved to Trieste (in 1773) and then to Vienna (in 1810). We are indebted to both branches of the congregation for numerous studies pertaining to various fields of learning and for many editions of classical Armenian authors, as well as for the thesaurus of Classical Armenian, the two-volume *Nor Baṙgirk' Haykazean Lezowi* ('New Dictionary of the Armenian Language', Venice 1836–1837; reprinted in Erevan, 1979–1981) by Gabriēl Awetik'ean, Xaç'atowr Siwrmēlean and Mkrtič' Awgerean.

4.6. The 18th and 19th centuries are also notable for the publication of periodicals in Armenian. The first of them, *Azdarar* ('The Monitor'), was published monthly in Madras (India) between 1794 and 1796 (18 issues of it were printed). Furthermore, in the 19th century the first studies dedicated to Armenian dialectology and folklore appeared. In the second half of the century, a discussion (the so-called *grapyk'ar* 'language struggle') started among learned people concerning the form of language (either classical or vernacular, or a mixture of both) which the Armenians were to use for teaching and writing.

In summary, the written varieties of both Western and Eastern Armenian had finally been formed in the late 19th century, more or less influenced by the classical language, which was still in use. As for the spoken language, in addition to the two literary varieties, local dialects were also certainly spoken.

4.7. The main historical events of the 20th century were the genocide of 1915–1917 and the annexation of Armenia to the Soviet Union. In consequence of the former, the Armenian-speaking population of the Ottoman Empire was drastically reduced and the survivors were forced to find refuge abroad, so that today Western Armenian is almost exclusively spoken, as a second language, in the Armenian diaspora. As to the second event, it took place in 1921: Armenia, previously a part of the Russian Empire, was incorporated into the Soviet Union after a short period of independence (1918–1920). In the new Socialist republic, an orthographic reform was carried out in order to make the orthography of Eastern Armenian adhere more closely to its pronunciation. The reform started in 1922 and was completed in 1940. Later on it was also adopted for Eastern Armenian used in other Soviet republics, as well as for Western Armenian used in Romania and Bulgaria, while Western Armenian in the diaspora and Eastern Armenian in Iran and India were still written according to the old orthography. We should also add that Eastern Armenian underwent the influence of Russian, mostly in vocabulary and at times in syntax as well. Today an influence of English can be noticed too.

On the other hand, Western Armenian, being spoken only as a second language, was influenced by the main language of the given country, so that there are differences between the varieties spoken, for instance, in France and in the U.S.A. Furthermore, after 1989, the mass emigration from the (former Soviet) Armenian Republic to the countries already having Armenian communities resulted in a new situation. The newcomers, whose mother tongue is Eastern Armenian, sometimes form a larger part of the Armenian-speaking population than the descendants of the genocide survivors, who speak Western Armenian only as a second language. This may presage a continuous influence of Eastern Armenian upon Western Armenian and, sooner or later, the death of the latter.

5. References

- Ačařyan, Hračya
 1951 *Hayoc' lezvi patmowt'yown. II mas* [History of the Armenian Language. Part II]. Erevan: Haypethrat.
- Adontz, Nicolas
 1970 *Denys de Thrace et les commentateurs arméniens*, transl. from Russian by René Hottebeex. Louvain: Imprimerie Orientaliste.
- Carrière, Auguste
 1886 *Un ancien glossaire latino-arménien*, Paris: Imprimerie Nationale.
- Clackson, James
 2000 A Greek Papyrus in Armenian Script. *Zeitschrift für Papyrologie und Epigraphik* 129: 223–258.
- Clackson, James
 2002 New Readings on the Armeno-Greek Papyrus BnF 332 Arm. *Zeitschrift für Papyrologie und Epigraphik* 141: 116.
- Greenwood, Timothy
 2004 A Corpus of Medieval Armenian Inscriptions. *Dumbarton Oaks Papers* 58: 27–91.
- Künzle, Beda O.
 1984 *Das altarmenische Evangelium/L'Évangile arménien ancien*. Teil I/I^{ère} partie: Edition, Teil II/II^e partie: *Lexikon/lexique*. (Europäische Hochschulschriften, Reihe XXI, Linguistik und Indogermanistik 33). Bern: Lang.
- Langlois, Victor
 1863 *Le trésor des chartes d'Arménie ou Cartulaire de la chancellerie royale des Roupéniens*. Venice: Typographie Arménienne de Saint-Lazare.
- Łazarean, Serob Łazari
 2007 *Hayoc' lezowi patmowt'iwn* [History of the Armenian Language], 3rd edn. Antelias: Meci Tann Kilikiyoy Kat'olikosowt'iwn.
- Muradyan, Gohar
 2012 *Grecisms in Ancient Armenian*. (Hebrew University Armenian Studies 13). Leuven-Paris-Walpole, MA: Peeters.
- Nichanian, Marc
 1989 *Agés et usages de la langue arménienne*. Paris: Entente.
- Thomson, Robert W.
 1995 *A Bibliography of Classical Armenian Literature to 1500*. Turnhout: Brepols.
- Thomson, Robert W.
 2007 Supplement to *A Bibliography of Classical Armenian Literature to 1500*: Publications 1993–2005. *Le Muséon* 120: 163–223.

Topchyan, Aram

- 2009 The Vienna Palimpsest of Agathangelos' *History*. In: V. Somers, (ed.), *Palimpsestes et éditions de textes: les textes littéraires. Actes du colloque tenu a Louvain-la-Neuve (septembre 2003)*. Louvain-la-Neuve: Institut Orientaliste.

Alessandro Orenco, Pisa (Italy)

61. The phonology of Classical Armenian

- | | |
|------------------------------------|--|
| 1. Introduction | 7. Morphologically conditioned dissimilation |
| 2. Alphabet | 8. Historical phonology |
| 3. Orthography and transliteration | 9. Consonantism |
| 4. Traditional pronunciation | 10. Vocalism |
| 5. Phonemic inventory | 11. Abbreviations |
| 6. Prosody | 12. References |

1. Introduction

Armenian is a living branch of Indo-European with fairly rich inscriptional and substantial textual attestation. The earliest inscriptions in *Old Armenian* (OA) date from the period after the creation of the Armenian alphabet by Mesrop Maštoc' in ca. 406 CE. The textual attestation of the so-called *Grabar* (lit. 'literary [language]'), or *Classical Armenian* (CA) in its broad sense, consists of more than 30,000 extant manuscripts dating from 862 (*Gospels of Queen Mlkē*) to ca. 1700 (Stone et al. 2002: 42¹⁵⁹, 118; Stone 2006: 467 f., 487 f.).

The terms OA and CA are often used interchangeably. This is when the term CA is used in its narrow sense to refer to the form of OA codified before ca. 450 CE, the period of the so-called *Golden Age* reflected in the Bible translation and the writings of the earliest Armenian authors such as Eznik and Koriwn (ca. 406–ca. 450). The distinction between this "classical" form of the language and that of the later manuscript tradition was first recognized by the Viennese Mekhitarist grammarians Č'alōxean and Aytənean (1885).

The extant CA manuscripts exhibit traits classified by Ĵahowkyan (1969) as *Post-CA* (ca. 450–ca. 700) and *Pre-Middle Armenian* (ca. 700–ca. 1100); however, Jungmann and Weitenberg (1993: 4) point to the insufficiency of the linguistic criteria used for this conventional periodization. The term *Middle Armenian* is synonymous with *Medieval Cilician* (ca. 1100–ca. 1350) which coexisted with CA as a literary language (cf. Karst 1901). The intermediate stage between OA and modern Armenian vernaculars spoken in Armenia proper up to ca. 1700 is conventionally referred to as *Medieval Armenian* (cf. Weitenberg 1995: 7).

Topchyan, Aram

2009 The Vienna Palimpsest of Agathangelos' *History*. In: V. Somers, (ed.), *Palimpsestes et éditions de textes: les textes littéraires. Actes du colloque tenu a Louvain-la-Neuve (septembre 2003)*. Louvain-la-Neuve: Institut Orientaliste.

Alessandro Orengo, Pisa (Italy)

61. The phonology of Classical Armenian

- | | |
|------------------------------------|--|
| 1. Introduction | 7. Morphologically conditioned dissimilation |
| 2. Alphabet | 8. Historical phonology |
| 3. Orthography and transliteration | 9. Consonantism |
| 4. Traditional pronunciation | 10. Vocalism |
| 5. Phonemic inventory | 11. Abbreviations |
| 6. Prosody | 12. References |

1. Introduction

Armenian is a living branch of Indo-European with fairly rich inscriptional and substantial textual attestation. The earliest inscriptions in *Old Armenian* (OA) date from the period after the creation of the Armenian alphabet by Mesrop Maštoc' in ca. 406 CE. The textual attestation of the so-called *Grabar* (lit. 'literary [language]'), or *Classical Armenian* (CA) in its broad sense, consists of more than 30,000 extant manuscripts dating from 862 (*Gospels of Queen Mlkē*) to ca. 1700 (Stone et al. 2002: 42¹⁵⁹, 118; Stone 2006: 467 f., 487 f.).

The terms OA and CA are often used interchangeably. This is when the term CA is used in its narrow sense to refer to the form of OA codified before ca. 450 CE, the period of the so-called *Golden Age* reflected in the Bible translation and the writings of the earliest Armenian authors such as Eznik and Koriwn (ca. 406–ca. 450). The distinction between this "classical" form of the language and that of the later manuscript tradition was first recognized by the Viennese Mekhitarist grammarians Č'alōxean and Aytənean (1885).

The extant CA manuscripts exhibit traits classified by Ĵahowkyan (1969) as *Post-CA* (ca. 450–ca. 700) and *Pre-Middle Armenian* (ca. 700–ca. 1100); however, Jungmann and Weitenberg (1993: 4) point to the insufficiency of the linguistic criteria used for this conventional periodization. The term *Middle Armenian* is synonymous with *Medieval Cilician* (ca. 1100–ca. 1350) which coexisted with CA as a literary language (cf. Karst 1901). The intermediate stage between OA and modern Armenian vernaculars spoken in Armenia proper up to ca. 1700 is conventionally referred to as *Medieval Armenian* (cf. Weitenberg 1995: 7).

The synchronic details of phonetics and phonology conventionally associated with CA in its broad sense are treated in sections 2 to 7. This is followed by a treatment of the diachronic sources of CA segmental phonology in sections 8 to 10.

2. Alphabet

α β γ δ ε ζ η θ ι κ λ μ ν ξ ο π ρ σ τ υ φ χ
 Ա Բ Գ Դ Ե Զ Է Ը Թ Ժ Ի Լ Խ Ծ Կ Զ Ը Ժ Մ Յ Զ Շ Ո Չ Պ Ջ Ռ Ս Վ Շ Ի Փ Ք Օ Ֆ
 ա բ գ զ ե շ ը թ ժ ի լ խ ծ կ չ ճ զ ճ մ յ ն շ պ ջ ռ ս վ ա ղ ր ճ ց ւ փ ք օ ֆ
 a b g d e z ē ə t' ž i l x c k h j l č m y n š o č' p' j' r' s v t r c' w p' k' o' f

The ordering and shape of the bulk of the characters is apparently inspired by the Greek cursive alphabet (Feydit 1982: 36 ff.). The uncial form of the script, or *Erkat'agir* 'Iron script', is attested in the inscriptions and earliest monuments up to ca. 1200. From early on, the uncials were used mixed with the minuscules. This full set was presumably referred to as *Bolorgir* 'whole script', but this term was later narrowed down to refer to just the minuscules. This type set has been used in printing since ca. 1500 (Stone 2006: 503 ff.).

3. Orthography and transliteration

The transliteration of Armenian orthography in this chapter follows the principle of one-to-one correspondence recommended by Schmitt (1972) with minor modifications of certain characters (ē instead of ê; õ instead of ô). The phonetic transcription follows IPA (2005).

The orthography of CA is often characterized as phonemic (cf. Benveniste 1966: 24); however, this ideal is complicated by the distribution of the representations ու ow, վ v, and ւ w for /u/ (3.1, 4.6); ի i and յ y for /i/; and by the alternation of ու ր for ր /r/ before ի /n/ (5.1).

3.1. The digraph ու ow represents the phoneme /u/ (cf. Gk. ου), which is most readily seen in the transcription of foreign names, e.g. Յիսուս *Yisows* [hi'sus] 'Jesus' (4.7) rendering Gk. Ἰησοῦς. Crucially, ու ow transcribes an entity which behaves phonologically just like ի i with respect to the regular vocalic alternations (due to pretonic reduction, 6.3): եմուս *e-mowt* traditionally pronounced [jɛ.mut] '(s)he entered': մտի *mt-i* [mə.ti] 'I entered', cf. իլիք *e-lik'* [jɛ.li.kʰ] '(s)he left': լքի *lk'-i* [lɔ.kʰi] 'I left'.

3.2. The grapheme է ē renders the phoneme /ɛ/ (4.2), which represents a relatively recent outcome of the monophthongization of the PA diphthong *eĭ (10.2.1).

3.3. A wedge (˘) indicates a palatal articulation of the corresponding non-palatal grapheme, e.g., յ (j) renders /dʒ/, the palatal counterpart of an alveolar affricate ճ (d) /dʒ/, etc.

3.4. A left half-ring indicates aspiration: *t'* (**թ**) /tʰ/, *č'* (**չ**) /tʃʰ/, *c'* (**ց**) /t͡sʰ/, *p'* (**փ**) /pʰ/, and *k'* (**ք**) /kʰ/; the omission of the diacritic indicates unaspirated voiceless consonants (5).

3.5. Both **օ** *ō* and **ֆ** *f* are post-CA additions; **օ** /*ō*/ continues CA **աւ** /aw/ (10.2), while **ֆ** /*f*/ made its way into Armenian via borrowing, e.g. **Քրեղք** *frer-k'* ‘crusaders’ < OFr. *frere* ‘brother(s)’.

3.6. The most frequent ligatures used in Armenian texts are **ւ** for **եւ** *ew* and **մն** for **մն** *mn*.

4. Traditional pronunciation

As CA became fixed as a literary language, its pronunciation gradually drifted away from that of the originally spoken OA and came to reflect that of the evolved spoken vernacular. Today, a scholar of CA is exposed to the traditional pronunciation, which represents phonological accretions essentially spanning a period from ca. 800 up to the immediate precursors of the modern Armenian standards (cf. Minassian 1976: 25; Weitenberg 1995: 4).

4.1. The mid vowels **ե** /*e*/ and **օ** /*o*/ are word-initially pronounced with a homorganic onglide, i.e. [jɛ-] and [vɔ-], respectively, except in enclitics: **ես եմ** *es=em* [ˈjɛsɛm] ‘I am’. The early explicit spellings **Յեսէ** *Yessē*, **Յեկիթայեայ** *Yep'tayeyay*, or **Յեքոնիա** *Yek'onia* rendering Gk. Ἰεσσαί, Ἰεφθάε, and Ἰεχονιάς, respectively (cf. **Ենովք** *Enovk'* rendering Ἐνωχ) suggest that the ‘pre-yodization’ was not a feature of OA. Therefore, spellings such as **Երուսաղէմ** *Erowsalēm* or **Երեմիա** *Eremia* (Gk. Ἰερουσαλήμ, Ἰερεμία, respectively) likely represent high-frequency words orthographically redacted at a later period (Ritter 1996: 19).

4.2. The vowels **ե** /*e*/ and **է** /*ē*/ (< **e*j; 3.2) are both pronounced as [ɛ]; however, word-initially, **է** -*ē*- does not receive an onglide: **ես** /*es*/ [jɛs] ‘I’, but **էշ** /*ēš*/ [ɛʃ] ‘donkey’. Some modern dialects of Armenian also represent the original opposition in non-initial stressed syllables, in which original /*e*/ “breaks” into [jɛ], while original /*ē*/ remains [ɛ] (Adjarian 1909: 2). This feature is already present in the Armenian-Latin glossary of Autun (ca. 1100), our earliest document testifying to significant dialectal divergence: *III' eriec*, cf. CA **երեք** *ere-k'* [jɛ.rɛkʰ] ‘three’, *facies' eriesc*, cf. CA **երեսք** *eres-k'* [jɛ.rɛ.sɔkʰ] ‘face’; but *presbiter' eresce*, cf. CA **երէց** *erēc'* [jɛ.rɛtsʰ] ‘elder’, etc. (cf. Weitenberg 1983).

4.3. The graphic sequence **ոյ** *oy* is ambiguous. As a diphthong, i.e., when it alternates with a pretonic **ու** [u] (6.3), **ոյ** /*oy*/ is pronounced [uɨ], e.g. **լոյս** *loys* [luɨs] ‘light’, oblique **լուսոյ** [lu.sɔ] (on the quiescence of final -յ -y, see 4.7 [end]); otherwise, **ոյ** *oy* is a composit of **ո** *o* and **յ** *j*, e.g. **սոյ** *xoy* [xɔj] ‘ram’, oblique **սոյի** [xɔ.ji].

4.4. Traditionally, η / λ / is the voiced counterpart of β / x /, i.e. uvular fricative (or approximant) [λ] (or [λ]). Speaking for the uvular place of articulation are the later, post-CA, renditions of a foreign uvular plosive [q] by η / λ /: Azer. *Qarabağ* [qarabaγ] ‘Karabakh’ > *Ղարաբաղ* *Ķarabat* [ʁaraˈbax] (Job 1995: 295¹⁵); Arab. [qurʔa:n(u)] ‘Quran’ > *Ղուրան* *Lowran* [ʁuˈran]. The voiceless uvular fricative β / x / [χ] frequently alternates with η / λ / even in the earliest texts: *ածուղ acut* ~ *ածուխ acux* ‘coal’ (cf. Martirosyan 2010: 19–20). However, in the alphabet (2), η / λ / stands in the place of Greek λ , and in the earliest loans, it is used to render a voiced alveolar lateral approximant *[l] (Hübschmann [1897] 1972: 327): *եկեղեցի ekelec’i* ‘church’ < Gk. ἐκκλησία, *եղիշէ Elišē* < Syr. *Eliša*’, *Երուսաղէմ Erowsalēm* ‘Jerusalem’. There are, however, a few exceptions to this generalization: (no doubt very early) Middle Iranian *Պահլավ Pahlav* (also *Պալշաւ Palhav*) ‘Parthian’, *դահլիճ dahlič* ‘hall; closet’, and (the majority of) the biblical names in *-էլ -ēl*, e.g., *Շ(ա)մուէլ Š(a)mowēl* ‘Samuel’ (Sköld 1927: 781, 788). Vacillation between η / λ / ~ ι / l / in numerous loans begins ca. 600 (Sköld 1927: 783): *ողոմպիաղ olompiad* ~ *ոլոմպիաթ olompiat* < Gk. Ὀλυμπιάδ- (cf. Hübschmann 1897: 368).

4.5. The phoneme \bar{r} / $r̄$ / is pronounced as a trilled alveolar [r̄]; the pronunciation of r / r / falls between an alveolar flap [r] and a palato-alveolar approximant [ɹ] (except before \bar{u} / n /, 7). Phonological evidence suggests that \bar{r} / $r̄$ / patterns as a [–cont] (or [–approximant]), while r / r / patterns as a [+cont] (or [+approximant], cf. / t / → [ɹ] in Tehran Armenian [Hacopian 2003: 73–78]): (i) r / r / is used to render a foreign [+cont] alveolar *[ð], e.g. *Հրեայ Hreay* ‘Jew’ < Syriac *ihūdāyā*, *սպարապետ sparapet* ‘general’ < Iran. *spāda-pati- (cf. Pahl. *spāhrpat*), *տիպար tipar* ‘example; type’ < Gk. τυπάδ-; (ii) r / r / fully assimilates to a following [+cont] segment, e.g. *պարհել parhel* ~ *պահել pahel* ‘to observe’ *[pa(ɹ).hɛl]; *արժան aržan* *[a(ɹ).zɛn] ‘worthy’: dial. *աժան ažan* ‘cheap’; (iii) r / r / tends to be (acoustically?) enhanced into a fricative before a [–cont] segment: *երթալ ert’al* ‘go’: dial. *եճթալ ešt’al*; *մարդ mard* ‘man’: dial. *մաջթ mašt*’ (cf. Feydit 1982: 52); similarly, the seemingly aberrant 2sg. reduplicated aor. subj. of *անես- aṛ-n-e-* ‘do’, *արասցես ar-as-c’-e-s* ← **/ar-ar-c’-e-s/* ‘thou shalt do’; (iv) \bar{r} / $r̄$ /, in turn, dissimilates before another [–cont], e.g. *բարբառ bar-baṛ* ‘dialect’ ← **/baṛ-baṛ/*; (v) Loanwords with a trilled geminate [r̄:] are rendered by means of \bar{r} / $r̄$ /, not **րր* / rr /; e.g. *քառք k’aṛ-k’* < Lat. *carrus* [kar:us] ‘wagon’; *փառք p’aṛ-k’* ‘glory’ < MIran. **far:ah-* (MP *farrāh*); the orthographical geminate **րր* / rr / is traditionally pronounced in two distinct articulations, e.g. *տարր tar-r* [ta.r̄r] ‘element’.

4.6. The graphemes ν and w are in complementary distribution: ν occurs word-initially and after n o while w occurs elsewhere: *վաւաշ vawaš* ‘lewd’, *վեւ vev* (the name of the letter) ν ; cf. *բերանով beran-o-ν* ‘mouth’-them.-Isg. but *ձեռնաւ jeṛ-a-w* ‘hand’-them.-Isg. There is also a special digraph *ով ov* for orthographical Gk. ω: *Մովսէս Movsēs* ‘Μωϋϑς’, *Յակովբոս Yakovbos* ‘Ἰάκωβος’. The digraph νw , which represents /u/ [u] (3.1), also renders [v] (< *[w]) when post-consonantal after the reduction of the “weak” vowels /u/ or /i/ (6.3), e.g. *անիւ aniw* [a.ˈniv] ‘wheel’: *անուոյ anow-o-y* [an.ˈvɔ] ‘id. (GDAbsg.)’ < PA **aniβ-ó-ii*o; *թուել tōw-e-l* [tʰɔ.ˈvɛl] ‘to count’, denom. of *թիւ t’iw* [tʰiv] ‘number’. On the other hand, w represents an allomorph of the stem morpheme /-i-/ in the original *-i(i)o-stem declension, e.g. *գինի gin-i* ‘wine’ (NAsg.): *գինուոյ gin-w-oy* [ginˈvɔ] ‘id. (GDAbsg.)’ < PA **γ^hein-i-óii*o. An orthographical νw may also indicate the sequence /-uɥ-/ from PIE *-uɥ- (9.6.12) or PA *-uβ- (< *-ub^hV-, *-ɔb^hV-,

*-upV- or *-ōpV-; 9.6.6, 9.6.7) or from PA *-mu-/*-um- (< *-mō-/*-ōm-; 9.6.8, 9.6.9), e.g. *անուն* *an-own* [a.'nun] 'name' < PA */anuṽ-n/ (< *Hnō-mn), cf. *անուան* *anow-an* [an.'van] 'id. (GDLsg.)' ← */anuṽ-an-; *մեղու* *mel-ow* [mɛ.'ku] 'honey (Isg.)' < PA *mel-ū-βi, (ultimately from *med^hu-b^hi). The traditional pronunciation treats postvocalic word-final -ւ -w as [-v], exactly as is the case with -ով -o-v [-ɔv]. Word-final -իւ -i(-)w is traditionally pronounced [-iv], e.g. *բանիւ* *ban-i-w* [ba.'niv] 'word (Isg.)'. Pre-consonantal ւ *iw* is pronounced as a rising diphthong [ju], e.g. *ձիւն* *jiwn* [dʒiun] 'snow', *իւղոյ* *iwt-oy* [ju.'ɛɔ] 'oil (GDABsg.)', etc. These facts suggest that ւ *w* and վ *v* represented /u/ when syllabified consonantly as a bilabial OA *[w] (> CA [v]), while ու *ow* represented either a vocalic [u] or an underlying sequence /Uu/ → *[(ə)w].

4.7. Word-initial յ- *y-* is pronounced the same way as հ *h*, i.e. as a voiceless glottal fricative [h]. However, the evidence of the dialects that underwent Adjarian's law demonstrates that յ- *y-* was formerly a "breathy" *[ɦ] (< OA *[j]). The traditional formulation of Adjarian's law states that a vowel becomes fronted after a (word-initial) voiced stop (cf. Vaux 1992). The conditioning segments have been more recently redefined as "breathy" (Garrett 1998) or [+spread vocal folds, -stiff vocal folds] which also subsumes *[ɦ]. The fronting had occurred before the dialects merged the pronunciation of */h-/ (from CA հ /h/) and */ɦ/ (from CA յ- /y-/ *[j]) into the voiceless [h]: *Լարաթ հանձի* [hændi] 'in the pasture' < *[ɦændi] < *[fiandi], cf. OA *յանդի* *y-and-i* *[jan.'di] 'in'-field'-Lsg. (Martirosyan 2010: 74). Medially, -յ- *-y-* is realized as a palatal glide [j]: *նայիմ* *nay-i-m* [na.'jim] 'I observe', *այլ* *ajl* [ajl] 'other'. In final position, -յ- *-y* ceased to be pronounced except in monosyllabic nouns and adjectives: *գոյ* *goy* [gɔ] 'exists', *արքայ* *ark'ay* [ar.'kʰa] 'king', but *Հայ* *Hay* [haj] 'Armenian', *խոյ* *xoy* [χɔj] 'ram'. By convention, յ *y* is not written in CA after ու *ow*, ե *e*, and ի *i*, cf. *աղայի* *al-a-j-i* 'grind'-them.-impf.-1sg. but *բերեի* *ber-e-i* 'carry'-them.(-impf.)-1sg.; *աղայ* *al-a-y* 'grind'-them.-3sg. but *հեղու* *hel-ow* 'fill, flow'-them.-3sg.

4.8. The graphic sequence *բա ea* is ambiguous. It may represent an underlying sequence (*/e-i-a/, e.g. *մարգարեանամ* *margare-an-a-m* [mar.ga.rɛ.ja.'nam] 'I prophesy', denominative of *մարգարէ* *margarē* 'prophet' ← */^oeṽ#/ (3.2); *բերեաք* *ber-e-ak'* [bɛ.rɛ.'jak^h] < */ber-e-i-ak'/ 'carry'-them.-impf.-1pl. (recall that -յ- *-y* is not written after ե *e*). In such instances *բա ea* reflects a pretonic (unstressed) *b* /e/, which in the later manuscript tradition (ca. 1100) tended to be redacted to է /ɛ/; thus, *բերեաք* *berēak'* for earlier *բերեաք* *bereak'* and *մարգարեանամ* *margarēanam* for earlier *մարգարեանամ* *margareanam*. In other contexts, *բա ea* represents a composite (coalesced) phoneme (*/i-a/ (> /ea/), which is under stress pronounced as a rising diphthong [ia] and subject to vocalic alternations (6.3), e.g. *քրիստոնեայ* *k'ristoneay* [k^hɛ.ris.tɔ.'nɪa] '(a) Christian' ← */^o-ni-ai/; cf. *քրիստոնեական* *k'ristone-a-kan* 'Christian-', i.e. [k^hɛ.ris.tɔ.nɛ.ja.'kan] ← */^o-ni-ai-a-kan-, in which the underlying /ea/ (from the "leftmost" */-i-a/) is reduced pretonically to [ɛ] (cf. later *քրիստոնէական* *k'ristonēakan* 'id.').

5. Phonemic inventory

The Armenian phonological system is characterized by a three-way VOT (voice-onset time) opposition of stops in all positions including word-finally, where this phenomenon

is typologically rare (Hacopian 2003). This opposition is phonologically best conceptualized as “marked” /D/:/T/ vs. “unmarked” /T^h/ based on phonological control of the laryngeal features (Vaux and Samuels 2005). The /D/:/T/ opposition is characterized by control of the VOT: in /D/ (or [D̥] in some varieties), the VOT is negative; in /T/, voicing is simultaneous with the release. The significant VOT, i.e. aspiration, in /T^h/ is thus a phonetic result of the lack of specification for laryngeal control.

That the so-called “aspirated” series is “unmarked” can be additionally seen from: (i) the patterns of neutralization, including intervocalic and post-/r/ devoicing of /D/ to [T^h]: NEA /grabar/ [gəɾɑp^hɑɾ] ‘the literary language’, /ordi/ [vɔɾt^hi] ‘son’, /ergel/ [jɛɾk^hɛl] ‘to sing’. Conversely, processes in which /T^h/ might be analyzed as “neutralized” into *[D] or *[T] do not occur; (ii) the last point is also evident in the diachronic distribution of stops within Armenian dialects. While the diachronic reanalysis of the control of the laryngeal features in both */T/ and */D/ series results in an array of stop inventories, all dialects uniformly preserve the continuity of the PA */T^h/ series, which is therefore diachronically “stable”; and in no dialect has it merged into the other series. In contrast, the other series not infrequently merged into it (cf. Pisowicz 1976: 73–86); (iii) the degree of aspiration on /T^h/ is subject to phonetic variation in the varieties in which /T/ is realized “with a concomitant tightening of the glottis” (cf. Fortson 2010: 394), i.e., the phonological opposition and its respective realization is /T/ [T²] vs. /D/ [D] vs. /T/ [T^h]~[T].

Because of their realization in some Armenian dialects, an ejective articulation of the voiceless occlusives has been plausibly argued to go as far back as OA (Fleming 2000; Holst 2009: 24 ff.). Whether it was inherited from PIE (cf. especially Kortlandt 2003: 20–25, 126–128; 2010: 57–61) remains controversial. In the NEA system of stops, it is by no means a norm but seems to be a feature of at least the voiceless affricates /c/ [tʰ²] (Ϸ) and /č/ [tʃ²] (ճ) (Khachatrian 1996: 187).

Consonants:	voiceless	voiced	aspirated
Stops	/p/ պ	/b/ բ	/pʰ/ փ [p ^h]
	/t/ տ	/d/ դ	/tʰ/ թ [t ^h]
	/k/ կ	/g/ գ	/kʰ/ ք [k ^h]
Affricates	/c/ ժ [tʰ]	/j/ ճ [dʒ]	/cʰ/ ջ [tʰ ^h]
	/č/ ճ [tʃ]	/j/ ջ [dʒ]	/čʰ/ չ [tʃ ^h]
Fricatives	/s/ ս	/z/ զ	
	/š/ շ [ʃ]	/ž/ ժ [ʒ]	
	/h/ հ		
	/x/ խ [χ]		
Resonants:			
Liquids		/l/ լ [l]	/lʰ/ ղ [ɫ] < *[l]
		/r/ ր [r] < *(*)[ɾ]	/rʰ/ ռ [ɾ]
Nasals		/m/ մ	/mʰ/ մ
Glides		/w/ վ , ւ [v] < *[w]	/y/ յ [h-] < *(*)[h-] < *[j-]
Vowels:			
Monophthongs	/i/ ի	/e/ ե	/u/ ու
		[ə] ը	/o/ ո
		/a/ ա	

Diphthongs /ē/ է [ɛ] < *[eɪ] /oy/ ոյ [uj] /ea/ եա [ia]
 Traditionally also ay այ aw աւ (~ Post-CA /ō/ օ) ew եւ iw իւ [iv] ~ [iu] < *[iɥ]

5.1. \bar{r} -neutralization

The phonemes r /r̄/ and r /r/ (4.5) are phonetically neutralized before \bar{v} /n/: աբարի *ar-ar-i* [a.ra.'ri] 'I did' but առնեմ *ar-ne-m* [ar:nɛm] 'I do' ← /ar-ne-/; դրունք *dr-own-k'* [dɔ:ruŋkʰ] 'door (Npl.)' but դուռն *dowr-n* [du:rɛn] 'id. (Nsg.)' (i.e. /dur-n/). The apparently exceptional cases of surface $-r\bar{v}-r\bar{n}-$ [-rɛ-] represent sequences of /-rUn-/ with an underlying high vowel deleted in pretonic position (6.3): կորնչիմ *kornč'im* 'I perish', traditionally pronounced (4) [kɔ:r.nə.'ʃʰim] ← /kori-nč'-i-/; կորեայ *koreay* [kɔ.'rjɑ] ← /kori-a-i/ 'I perished (Aor.)'; վերնոյ *vernoy* (traditionally) [vɛr.'nɔ] 'upper (GDAbsg)' ← /verin-o-i/, cf. վերին *verin* 'id. (NAsg.)'. It has been suggested that the source of these exceptions should be sought in the relative chronology of the phonological processes involved, i.e. \bar{r} -neutralization preceded vowel reduction (6.3). However, the OA pronunciation may have differed from the traditional one in this respect: [vɛr.'nɔ] < OA *[vɛ.ɹə.'nɔj], [kɔ:r.nə.'ʃʰim] < OA *[kɔ.ɹən.'ʃʰim] (cf. Clackson 1994: 38; Hübschmann 1906: 475). Loans, such as պոռնիկ *pořnik* 'whore' from Gk. πορνικός, նեռն *neřn* [nɛ.rɛn] 'antichrist' from Gk. Νέρων 'Nero (PN)', are most likely gratuitous and not probative, since r /r̄/ renders Greek ρ (2) across the board at this stage.

6. Prosody

The fundamental phonological processes contributing to CA prosody are: (i) *vowel epenthesis*, or the insertion of [ə] which mediates syllabification (6.1); (ii) *oxytonesis*, or the placement of the stress on the final (non-epenthetical) nucleus (6.2); (iii) the reductions of the diphthongs and of the "weak" high vowels in pretonic positions resulting in the morphophonological patterns of *vocalic alternations* (6.3): է /ē/ [ɛ] → ի [i] (< *[eɪ/; 3.2); ոյ /oy/ [uj] → ու [u]; եա /ea/ [ia] → ե [ɛ]; ի /i/ [i] → Ø; ու /u/ [u] → Ø.

6.1. Vowel epenthesis or schwa insertion

Probably the most important feature of the traditional pronunciation is the production of systematic patterns of syllabification by a rule-governed insertion of [ə], which is generally not represented in the orthography. The grapheme \bar{v} ə is rarely written since it is phonologically predictable; it is primarily used to indicate a lexically determined syllabification, e.g. ընկեր *ənker* 'friend', i.e. /Unker/ [əŋ.'kɛr]; սուրնջեան *tow-ənj-ean* 'day (GDLsg.)' /tiu-Unj-ean/ [tə.vən.dʒjan], cf. տիւ *tiw* 'day (NAsg.)'. It is fairly consistently used to indicate a lexical boundary in composition, e.g. հատընտիր *hat-əntir* 'select,

choice, fine' (*Հատ hat* 'a cut[ting]' + *ընտիր ɛntir* 'chosen'), etc., or a [ə] if the word spans a break in the line, e.g. *ծն ընդեան ԵՆ անձեան* [tsə.nən.djan] 'birth (GDLsg.)' (Matenadaran 355), with () indicating a line break, for *ծննդեան ԵՆՆԵԱՆ*. Occasionally, it appears in certain salient contexts, e.g. *ի սկզբանէ i skəzban-ē* [is.kəz.ba.nɛ] 'from (the) beginning' in the historiated initial words of Genesis 1 : 1 (The British Library, ms. Or. 8833, f. 3r) for text-internal *ի սկզբանէ i skzbanē*.

A descriptive illustration of how the rule mediates syllabification in the practical selection of word-initial (orthographical) consonant clusters is the following (cf. Thomson [1975] 1989: 116–121): CCV- → [C^o.CV]: *ծնունդ ɛnownd* [tsə.'nund] 'birth'; CCCC- → [C^o.C^o.CV]: *բժշկելիք շէկել* [bə.ʒɛf.'kɛl] 'to heal'; CRCCV- → [C^oRC.CV]: *քրտնիլ կ'րտնիլ* [k^hɛrt.'nil] 'to sweat'; CFCV- → [C^oF.CV]: *դժուար Ժժուար* [dəʒ.'var] 'difficult'; SPCCV- → [S.P^o.CV]: *սկզբանէ skzbanē* 'beginning (Absg.)' [əs.kəz.ba.'nɛ]; SFV- → [S^o.FV]: *սխալ ԵՄ ԵՄ* [sə.'xal] 'mistake', etc.

The traditional syllabification seems to reflect the following overarching principles: (i) "no complex margins" (i.e. */CC-/ → [CəC]); (ii) "have an onset" (within the domain of a prosodic word), and "a coda may be a resonant or a fricative" (i.e. [+cont]). Deviations are due to later developments, e.g. *կ'րտնիլ* [k^hɛrt.'nil] < OA *[k^hɛrt.tə.'ni.{}], or more dominant prosodic principles, e.g. initial /s-/ is syllabified as a phrasal element, i.e., outside of ω, cf. Vaux and Wolfe (2009). The different treatments of epenthesis word-initially are likely due to sonority-driven syllable contact principles: /sP/ → [s^o.P], but /sF/ → **[^os.F] → [s^o.F], /sR/ → **[^os.R] → [s^o.R].

6.2. Oxytonesis or final stress system

The primary stress in CA is assigned to the final non-epenthetical (6.1) vowel in the word, e.g. *կերակուր ԿԵՐԱ ԿՐ* [kɛ.ra.'kur] 'food', but *մեղր ԵՄ ԵՄ* ['mɛ.ɛər] 'honey'; *թիւն ԵՄ ԵՄ* ['t^hi.vən] 'number'-def.; *զկեանսն ԶԿԵԱՆՍՆ* [əz.'kjan.sən] prep.-'life'-acc.pl.-def., etc. The accented texts point to two exceptions: (i) non-final stress in certain adverbials: *այնպէս ԵՄ ԵՄ* 'so', *որպէս ԵՄ ԵՄ* 'as', *ահա՛ւասիկ ԵՄ ԵՄ* 'behold, lo!' (cf. Meillet [1905] 1977: 328; Jensen 1959: 19); (ii) vocatives are often accented on the first syllable: *Դ՛ազարե Լճառ Լճառ* 'Lazarus!', *ռա՛բբի Բճբբի* 'Rabbi, master!' (Gk. ραββί), *Հա՛յրիկ Կճրիկ* 'father!'; or they forfeit their accent onto a vocative particle: *ո՛վ (ա՛յ) մարդ ԵՄ ԵՄ* 'O man' (cf. Martirosyan 2013: 90; Künzle 1984: 93).

6.3. Vocalic alternations

The shift of stress in inflection and derivation causes reduction of the diphthongs *է /ē/* [ɛ], *ոյ /oy/* [uɪ] and *եա /ea/* [iɑ] in destressed syllables, e.g. *լոյս ԼՅՍ* ['luɪs] 'light': *լուսոյ ԼՅՍՅ* *լոս-ո-յ* [lu.'sɔ] 'id. (GD Absg.)', *էջ ԵՄ* ['ɛʃ] 'donkey': *իջոյ ԵՄ ԵՄ* [i.'ʃɔ], *լեան ԼԵԱՆ* [l'jɑ.rən] 'mountain': *լերին ԼԵՐԻՆ* [lɛ.'rin] 'id. (GDLsg)'. The alternation *եա /iɑ/* ~ *ե /ɛ/* occurs frequently in inflection and lexical composition, since the diphthong *եա /ea/* arises from phonemic composition of any front vowel with *-a-* (10.2.10): /sɛr-e-ac'-Ø/ 'love'-them.-aor.-3sg.act. → *սիրեաց ԵՄ ԵՄ* [si.'rjats^h] '(s)he loved': /sɛr-e-ac'-i/ 'love'-

them.-aor.-1sg.act. → սիրեցի *sirec'i* [si.rɛ.'tʰi] 'I loved'; բարի *bari* 'good' + -ա- *-a-* (linking vowel) + կամ *kam* 'wish' → ^(*)/bari-a-kam/ → բարեկամ *barekam* [bɑ.rɛ.'kam] (lexicalized) 'friend'; երկեամ *erkeam* 'of two years, biennial' ← /erk-i-/ 'two' (< *d̥u-i-) + /am/ 'year' (9.5.2); որդեակ *orde-ak* 'dear child (Vsg.)' ← /ord-i-/ +/ak-a-/ (dimin.) < MIran. *aka- (cf. նավ *nav* 'ship' → նավակ *nav-ak* 'boat'), etc.

The high vowels *h* /i/ and *u* /u/ are also subject to reduction: սիրտ *sirt* [sirt] 'heart' but սրտի *srt-i* [sər.'ti] 'id. (GDLSg.); գլուխ *glowx* [gə.'luχ] 'head' but գլխոյ *glx-o-y* [gəl.'χɔ] 'id. (GDLSg.); հինգ *hing* [hing] 'five' but հնգասան *hnge-tasan* [həŋ.gɛ.ta.'san] 'fifteen'. Diachronically, *h* /i/ and *u* /u/ in pretonic syllables go back to reduced diphthongs (8.1), which may, however, become opaque synchronically. Thus, surface high vowels are not reduced in pretonic positions if they enter the derivation unstressed: ուրախ *owrax* [u'raχ] 'happy' → ուրախութիւն *owrax-owt'iwɛn* [uraχu'tʰjun] 'joy' (cf. քունե- *k'own-e-* 'sleep' < *k^hou₃n-; 10.2.8).

7. Morphologically conditioned dissimilation

The paradigm of the aorist subjunctive is characterized by deaffrication of -*g* /-c'/ (aor. formant), when in contact with the pretonically reduced allomorph of -*h**g* /-ic'/ (subj.): սիրեցես *sir-es-c'-es* [si.rɛs.'tʰɛs] 'thou shalt love' (← /^oeac'-ic'-es/ aor.-subj.-2sg.act., cf. սիրեցից *sir-ec'-ic'* [si.rɛ.'tʰiɛtʰ] 'I will love'). There are, however, systematic exceptions to this morphologically conditioned dissimilation: լացցես *la-c'-c'-es* 'thou shalt weep' (← /l-a-c'-ic'-es/, cf. լացից *la-c'-ic'*); ածցես *ac-c'-es* 'thou shalt lead' (cf. ածից *ac-ic'*); this exceptional (but perhaps only apparent) retention of the affricate cluster is traditionally associated with or even attributed to the "monosyllabicity" of the aor. stem (Meillet 1913: 95); cf. also ստացցիս *sta-s-c'-i-s* [əs.tas.'tʰis] (6.1) 'thou shalt obtain (med.)' (cf. ստացայց *sta-c'-a-yc'* 'I will obtain [med.]').

However, phonologically it seems more plausible that the prosodic shape of the overall morphological structure played a role in the preservation or deletion of the pretonically reduced high vowel in the subjunctive morpheme. Thus, perhaps, a prehistoric parsing into left-headed binary feet, prior to the complete loss of the final syllable, may account for the observed phenomenon: OA [si.rɛs.tʰɛs] < *(si.rɛts^h)(tʰɛ.ɛ.si), not **si(rɛ.tʰɛ)(tʰɛ.ɛ.si), < *[si.rɛ.tʰɛ.tʰɛ.ɛ.si]; OA [əs.tas.'tʰis] < *(əs.tats^h)(tʰi.ɛ.si), not **əs(ta.tʰɛ)(tʰi.ɛ.si), < *[əs.ta.tʰɛ.'tʰi.ɛ.si]; but OA ^(*)[la.tʰɛ.'tʰɛs] < *(la.tʰɛ)(tʰɛ.ɛ.si), not **lats^h(tʰɛ.ɛ.si), etc.; cf. 9.2.5.

8. Historical phonology

In the following discussion and throughout the rest of this chapter forms preceded by an asterisk indicate PIE reconstructions unless indicated otherwise. Forms preceded by a question mark indicate possible but uncertain reconstructions. Derivations which arguably lead to the attested Armenian forms are, regardless of the time depth or reference to other concepts such as "Consonant shift", labeled PA. Mesropian orthography indicates CA forms. A vertical line above a nucleus indicates the reconstructed PA intensity

accent to differentiate it from the PIE musical pitch accent, e.g. *septḗn > PA *h̑eṓtʰan > *բւթն ewt'n* ‘seven’.

8.1. Diachrony of final syllables

The rhymes of PIE final syllables are generally not preserved in OA. This process is referred to as *apocope* and is ascribed to the shift of the original prosodic system with syllables characterized by mobile pitch differences to the system with intensity accent (cf. Meillet 1936: 19; contra Pedersen 1904 [= 1982: 3 ff.]).

As for the original final codas specifically: (i) PIE obstruents are lost presumably quite early; Holst (2009: 80 f.) sees here an isogloss that Armenian shares with the “Balkan” languages: Greek, Albanian, and Phrygian: *ելիք e-lik* ‘(s)he left’ < *é-lik^u-e-t (Ved. *áricat*, Gk. *ἔλιπε* ‘id.’); (ii) Original syllabic resonants are believed to be preserved as such until relatively late (at least word-finally): *տասն tas-n* < *dékṃ (Ved. *dása*); *ոտն ot-n* ‘foot’ < *ród-ṃ (Gk. *πόδα*); *գելուճի gel-owmn* ‘twist, contortion’ < *uelū-mṃ (Lat. *volūmen* ‘scroll’, Gk. *εἰλῶμα* ‘wrapper’); *աղբիւր albiwr* ‘source’ < *b^hrē-ur (Gk. *φρέαρ*); (iii) Original non-syllabic liquids are preserved: *այր ayr* ‘man’ < *h₂nér (Gk. *ἀνήρ*); *աստի ast-t* ‘star’ < *h₂stél (Gk. *ἀστήρ*) (if Olsen [2010] is right, the Armenian lateral may be inherited); (iv) Non-syllabic nasals are traditionally assumed to be preserved in monosyllables: *k^uéh₂-m > *քան k’an* ‘than’ (Lat. *quam*); *dóm > *տուն town* ‘house’ (Hom. *δῶ*, Av. *dqm-i* [Lsg.]; cf. Stempel 1990); also *ձիւն jiwn* ‘snow’ (Gk. *χιών* ‘id.’, Lat. *hiem-s* ‘winter’), which would, however, have had to be understood as a monosyllable at the PA level, i.e. *ǵ^hiǵóm > PA *[dʒiùn] or *[dʒiun] (cf. Ravnæs 1991: 100). It would seem that in words of more than one syllable only *-Vn is preserved, while *-Vm is lost (so Pisani 1951: 47 f.): *urh₁-ēn > *ղառն gař-n* ‘lamb’ (Gk. *ἀρήν*), *հարսն hars-n* ‘bride’ (< ?*p^hk-ōn; based on its inflection; cf. Hamp 1988; Godel 1975: 100 f.), but *h₁ék_u-o-m (Asg.) yields *էջ ēš* ‘donkey’ (Gk. *ἵππον*, Ved. *ásvam*, Lat. *equum*). Kortlandt (1984a: 97 f. = 2003: 45 f.) assumes that all final *-VN# sequences develop into nasalized vowels; he considers *ղառն gař-n* and *հարսն hars-n* to be based on the original Asg.; so also Pedersen (1905: 216 f. [= 1982: 72 f.]), who, however, also derives *տուն town* from *dómṃ (Gk. *δῶμα*) and *հարսն harsn* from *p^hk-nó-m (cf. *բեռն beř-n* ‘load’ < *b^her-néh₂- cognate to Gk. *φερνή* ‘dowry’) and compares *ղառն gařn* to Skt. *uraṇam* ‘wether’ (< *-én-ṃ).

The most economical solution is to assume that nasals are preserved in final long vowel + nasal sequences (*-vN#) and lost in short vowel + nasal sequences (*-v̄N#). In the latter, the vowel was nasalized, subsequently reduced and lost (*-v̄N > PA *-v̄ > *-ə > -Ø, while *-v̄N > PA *-Vn). This assumption accounts not only for the monosyllables cited above, which invariably contain *-v̄N, but also for one monosyllabic form that has not been part of the discussion, the preposition *(h₁)en > PA */in/ *[i] > *ի i* (sandhi-variant *j- y-*) ‘in’, where the nasal must have been lost prior to the stage at which it was reanalyzed as an “inflectional prefix” with two allomorphs, i.e. PA */i/ → {i}~{j-}.

8.2. Origin of the final stress system

The newly stressed penultimate nucleus acquired prosodic salience, which led to the reduction of the final post-tonic syllable, e.g. **élik^uet* (Ved. *áricat*, Gk. *ἔλιπε*) > PA **[ɛ.li.k^hə]* > *Էլիք e-lik'* [*ɛ.li.k^h*] '(s)he left'. The shifts of stress onto the final syllable in inflection and derivation (6.2) triggered pretonic reduction of high vowels and certain diphthongs. Pisani (1950: 168) assumes that final *-e disappears before the period in which the (reflex of the) original **k^u* was palatalized (9.2.18). Ravnæs (1991: 55) bases himself on the proposal of Dressler (1976: 305) that the palatalization before **i* was chronologically earlier than that before **e* and suggests that between these two stages the final, post-tonic **e* was reduced to **[-ə]* (vel sim.), i.e. a vowel which did not cause palatalization. The first part of this scenario finds typological parallels in modern French, where only high front segments condition palatalization of the coronals but not mid front vowels: *dit* [dⁱi] 'said', *du* [dⁱy] 'of the', *tu* [tⁱy] 'thou', *tiens* [tⁱjɛ̃] (~ [tⁱɛ̃ɛ̃]) 'here you go' vs. *de* [də], *des* [de] 'of', *deux* [dø] 'two', *terre* [tɛːʀ] 'earth'; *génitif* [ʒɛni'tif] 'genitive' vs. *né(e)* [ne]. Alternatively, others assume PA **[ɛliŋ^hɛ]* → **[ɛlik^hɛ]* by analogy to the forms of the present (**lik^h-ŋ-* > *Լքանիք lk'-an-e-*) or hypothetical forms with **o* vocalism of the thematic vowel, such as **é-lik^h-o-m* (cf. Beekes 2003: 177 and note the identical phenomenon, with leveling of the labial outcome, in Greek *ἔλιπε*).

The outcome of these developments was the attested final stress system (6.2) often compared to the development of final stress in words with original penultimate stress in the history of French, e.g. Lat. **[sa'lu:tem]* > OFr. **[sa'lyt]* (modern French *salut* [saly]). As a rule of thumb, the nucleus which receives the PA stress is part of the original PIE penult. However, since the fixation of the stress occurred at a period after the lenition and loss of certain consonantal onsets of inherited final syllables, namely **s*, **i*, and possibly **u* (Viredaz 2001–2002a), the position of the stress might on occasion correspond to the original antepenult: **méh₁d-es-ŋ₂* (Gk. *μήδεα* 'plans') > **médēha* > PA **[mi.dɛ.a]* (vel sim.) > **mi.ta-* > → *Միտք mit-k'* 'intention; intellect'; *(s)kor-éje-ti ([iter.] of **[s]ker-*) (Gk. *κείρει*, ON *skerr* 'cuts, shears') > **[k^hɔ:rɛ.θi]* > **[k^hɔ:lɛi]* > **[k^hɔ:lɛ]* > *քորէ k'or-ē* [*k^hɔ:rɛ]* 'scratches'.

The patterns created by the various reductions just noted became phonologized into the synchronically productive process of vocalic alternations (6.3): **pénk^ue* (Gk. *πέντε*) > PA **[pɛŋ.gə]* > *հինգ hing* 'five', but **pénk^ue-dek^umt* (Ved. *pāñcadaśa*) > → PA **[pəŋ.gɛ.ta.sa.ni]* > *հնգասասան hnge-tasan* 'fifteen'; **uójn-o-* (Gk. [F]οἶνος) > PA **γ^uɛjn-o-* > **[γ^ui:nɔ]* → **o-ijo-* > *գինի gin-i* 'wine'. This process seems to have been productive even at the time of the latest loans: *լսիր lt-er*, GDLsg. of *լիտր lit-r* 'pound' (< Gk. *λίτρα* 'weight').

9. Consonantism

The contrast in the original PIE three series of stops is preserved in Armenian. The original palatal series develops into sibilants or affricates: **k̑*, **g̑*, **g^h* > *u* /s/, *ɬ* /c/ [tʰ^(ʔ)], *ɬ* /j/ [tʰ], respectively. The non-palatal series are characterized by a modification of their original laryngeal features, an epiphenomenon conveniently referred to as the "Armenian consonant shift". The "shift" is most cogent in its word-initial outcomes (and especially

in the PIE voiced series); thus, *p, *t, *k^(u) become (?*p^h or *φ >) < /h/ (or Ø), [θ /t/, ɸ /k'/, respectively (i.e. PIE *T- > CA T^h-); *b, *d, *g^(u) > ɸ /p/, ɸ /t/, ɸ /k/, respectively (i.e. PIE *D- > CA T-). The PIE *D^h series is traditionally taken to yield normal voiced ɸ /b/, ɸ /d/, ɸ /g/, respectively; however, voiced aspirated stops are recognized as reflexes of this series in modern dialects by Sievers (1893), Pedersen (1906: 336–342 [= 1982: 112–120]), Adjarian (1909), Allen (1951: 200), Benveniste (1959), and Vogt (1938: 327, 1958). Garrett (1998) adduces phonological arguments for the feature [+slack vocal folds, –stiff vocal folds] for this series (cf. 4.7), and these arguments are experimentally confirmed in Schirru (2012: 435–458). The reflexes of this series in some modern dialects of Armenian and additional phonological considerations (such as the mechanism of Adjarian's law) strongly indicate that at least for PA the traditional “voiced” series should be reconstructed as “murmured” and will be recognized as such in the discussion that follows (i.e. *b^h, *d^h, *g^h, *g^{(u)h} > PA * ɸ̣, *ɸ̣, *ɸ̣, *ɸ̣, respectively).

Medially, PIE voiceless stops undergo a series of profound changes: (i) Lenition to total loss (*-T- > CA (*)-F- > Ø): *b^hh₂-tí-s (Dor. φάτις) > PA *[ɸ̣áθ^hiḥ] > *[ɸ̣á.θ̣i] vel sim. > ɸայ bay [baj] ‘speech’; *pede-tó-s > PA *ɸeteθ^hóh > hetéβo > (այս-ով) հետեւ (ays-ow-)het-ew ‘*from (this) step (on)’ > ‘henceforth’ (Ved. patsu-tás lit. ‘from [a place] at the feet’; Lat. funditus ‘(*from the bottom > completely’). Conditioned reflexes of original intervocalic *t are assumed by Pisani (1951: 68), Klingenschmitt (1982: 98 f.), Job (1988: 28), Olsen (1999: 151), and Matzinger (2005 passim): *-t- > PA *-θ̣i-, a palatalized dental fricative (> *-ð̣i- > -j- -y- [9.6.1] or -Ø-) and *θ̣u-, a labialized dental fricative (> *φ > *β, a bilabial approximant > ɸ ~ ɸ̣ /ɸ̣/ [9.6.5]) conditioned by the [±back] features of the following vowel; also *-t- > PA *-θ̣u-, before consonantal */r/ and probably */l/. In their phonetic exegesis all the scholars just cited go well beyond Meillet (1936: 33) and Schmitt (1981: 59), who simply posit *t > *y regardless of environment; (ii) After liquids and nasals, PIE voiceless stops merge with the inherited voiced aspirates (*-{R/N}T- > *-{R/N}Ḍ-): *H₁-tú- > PA *árđ-u- > արդ ard ‘shape, order’ (Ved. irtú- ‘appointed time, order’), cf. *é-d^heh₁-t > PA *é-đi > Էդ ed ‘(s)he put’ (Ved. ádhāt ‘id.’). In addition, original clusters of a (plain or aspirated) voiced stop plus resonant undergo metathesis (i.e. *[-JD^(h)R]- > *-RD/T-): *b^hrátēr > PA *[ɸ̣rájɾ] > Եղբայր elbayr ‘brother’ (Ved. bhrātā ‘id.’); *még^h(s)ri > PA *médz.ri > մերձ merj ‘near’ (Gk. μέγρι ‘id.’); *suid-ro(s)-(m) > PA *[ḥɸ̣id.ran] > քիրտն k’irt-n ‘sweat’ (Gk. ἰδρῶ ‘id. [Asg.]’ < PGk. *-óh-α).

A characteristic feature of satem languages, the merger of *K and *K^u, seems to have taken place within the history of PA itself. The inherited labiovelar is affected by so-called “u-epenthesis” after a nasal (i.e. *-NK^u- > PA *-aŋK-). After the merger of *K and *K^u, a preceding labial segment “satemizes” the velar so that the latter shows reflexes identical to those of the original PIE palatovelar *Ķ (*-aŋK- > PA *-aŋĶ-; 9.2.2, 9.3.2): աւժ awc ‘(s)he anointed’ < PA *aŋĶ- < *h₂ŋĶ- (Lat. unguen ‘ointment’), cf. ածէ ac-ē ‘drives’ < *h₂éĶ-e-ti (Gk. ἄγει, Lat. agit ‘id.’); (9.2.1).

Another feature often regarded as a peculiarity of Armenian consonantism is the resistance of all the original plain velars and *g^u to palatalization (but see 9.2.7): *g^uéñh₂- becomes կին kin ‘woman’ (Ved. jāni-, MPers. zan, OCS žena, TB šana ‘id.’) as opposed to քիւմ քիւմ քիւմ ‘warm’ from *g^uer-mó- (Gk. θερμός, Ved. gharmá-, OE wearm ‘id.’). Opponents of this view include Pedersen (1906: 392 = 1982: 171 f.) and Kortlandt (1975: 43 f. = 2003: 10 f.), who attribute the absence of palatalized reflexes to the workings of paradigmatic leveling, thus, e.g., *g^uén(e)h₂- was originally palatalized to *ǰina

but subsequently restored to *կին* *kin* based on the plural *կանայ-* *kan-ay-* (< *g^uñh₂-), etc.

In the following sections we shall examine each member of the phonological inventory of CA from the point of view of its sources.

9.1. Stops

/p/ պ	*b _{9.1.1} ??*{s}p- _{9.1.2}	/b/ բ	*b ^h _{9.1.9}	/p/ փ	(*p ^h <) ?*sp- _{9.1.15} ??*-h ₁ p- _{9.1.16}
/t/ տ	*d _{9.1.3} *{s}t _{9.1.4} *{k}t _{9.1.5} PA *{uk}t _{9.1.5; cf. 9.3.2}	/d/ դ	*d ^h _{9.1.10} *{R/N}t _{9.1.11}	/t/ թ	*t- _{9.1.17} ?*pt- _{9.1.18} ??*-h ₁ t- _{9.1.19}
/k/ կ	*g ^(u) _{9.1.6} *{d}u _{9.1.7} *{-st}u- _{9.1.8}	/g/ զ	*g ^{(u)h} _{9.1.12} *u _{9.1.13} *{R/N}k ^(u) _{9.1.14}	/k/ կ	*k ^(u) _{9.1.20} *su _{9.1.21} *tu _{9.1.22} (?)*-s _{9.1.23}

9.1.1. պ *p* < *b: *ստիպե-* *stip-e-* ‘compel’ < *stéjb-e- (Gk. στειβω ‘I trample’); *ըմպէ* *əmp-ē* ‘drinks’ < */imipēj/ < PA *imj-φib-ē-θi < *(h₁en-)pi-b(h₃)é-ti (Ved. *pibati*, Lat. *im-bibit*).

9.1.2. սպ- *sp-* may reflect *sp, parallel to *st > ստ *st* (9.1.4), but only two plausible etyma are available: *սոսսպել* *ař(-)a-spel* ‘fable’ < *př(H) *spel-eh₂ (Goth. *spill*, MHG *bī-spel* ‘parable, example’, Alb. *fjalë* ‘word’); *սպառնա-* *spař-na-* ‘threaten’ < *spř-n-H- (Lat. *spernō* ‘I despise’, ON *sperna* ‘kick away’, Ved. *sphurāti* ‘kicks away’); alternatively, *sp- > փ- *p’-* (9.1.15).

9.1.3. տ *t* < *d: *տասն* *tasn* ‘ten’ < *dékṃ (Ved. *dása*, Gk. δέκα, Goth. *taihun*); *տուն* *town* ‘house’ < *dóm (Gk. δῶ[μα], Lat. *dom-us*, Av. *dəm[i]* ‘at home’); *տու* *e-tu* ‘I gave’ < PA *e-tù-i ← *é-deh₃-m (Ved. *ádām* ‘id.’).

9.1.4. ստ *st* < PA *st < *st, */s/d (*[z]d): *ստերջ* *sterj* ‘barren’ < *ster-ǰ(e)h₂- (Ved. *star-ī-h* ‘barren cow’, Gk. στειρα ‘infertile’, Goth. *stairo* ‘id.’); *աստղ* *astl* ‘star’ < *h₂stél (Gk. ἀστῆρ, Ved. *stř-bhih*, Lat. *stēlla*); *ոստ* *ost* ‘branch’ < *h₂o/s/do- (Gk. ὄχος, Goth. *asts*); *նիստ* *nist* ‘sit down!’ < PA *n(ih)iste < *ni *sí-/s/d-e (Gk. ἵζε, Lat. *sīde*, Ved. *nī-šīda* ‘sit down!’).

9.1.5. ստ *st* < *k̂t, *{u}k̂t (see 9.3.2): *երաստանք* *erastan-k’* ‘buttocks’, generally thought to be related to Gk. πρωκτός ‘anus’ from *prōktó-, although the absence of CA *u* from *ō gives pause (cf. Ved. *přšthá-* ‘mountain ridge’; YAv. *paršta-* ‘spine’, if these are related); *դուստր* *dowstr* ‘daughter’ < PA *dūstir < *d^huktēr (Lith. *duktė*, OCS. *dъstъ* ‘id.’), which seems to result from the *-PH.CC- > *-P.CC- rule driven by sonority principles in PIE medial codas: *[d^hug]_σh₂[tr-]σ > *d^huk.tr- (NIL 127^{1,5}; Byrd 2015: 123).

9.1.6. *q k* < *g^(u): *կալալ* *kal-a-w* ‘seized’ < *g_lH-t-ó (Lith. *galiù* ‘*I have in power > I can’, RCS *golėmь* ‘*powerful > huge’, MW *gallaf* ‘I can’ (cf. LIV² 185 f., Klingenschmitt 1982: 269 f.); alternatively, *կալալ* *kal-a-w* is from *g^uh₁-tó (Hom. βλήτο) and is related to *կեղ* *kel* ‘wound’ < *g^uélh₁-os- (Gk. βέλος ‘missile’; de Lamberterie 2005: 333 f.); *անկիւն* *ankiwn* ‘corner’ < *h₂(e)ng- (Lat. *angulus*, OCS *огуль* ‘id.’, OIc. *ekkja* ‘ankle’); *կռուկ* *kř-ownk* ‘crane’ < PA *kirH-un- < *g^{erh}₂-ōn- (Gk. γέρανος, Lat. *grūs*, Lith. *gėrvė*); *կին* *kin* ‘woman’ < *g^uēnh₂- (Gk. γυνή, OIr. *ben*, OE *cwēn*); *կով* *kov* ‘cow’ < *g^uou- (Ved. *gāv-*, Gk. βό[F]-).

9.1.7. *q k* in PA cluster *{r}k < *(-)dū-: *երկար* *erkar* ‘long-(lasting)’ < *dūéh₂-ro- (Gk. δηρός ‘id.’, Ved. *dū-rá-* ‘distant’, Lat. *dū-dum* ‘formerly’ < *duh₂-, cf. *տեւ* *tew* ‘duration’ < *deuh₂-, RCS *davē* ‘recently; yesterday’, OCS *dav-ьnъ* ‘ancient’ < *dōuh₂-); *երկն* *erk-n* ‘labor pains’ < *h₁éd-uōn (Hitt. *idāl-u-* ‘evil < *that which bites’, TB *yolo* ‘id.’ < *h₁ed-uōl-[ōn-]; cf. Gk. ὀδύνη ‘pain’); *երկու* *erk-ow* ‘two’ < *dū-ó-h₁ (Ved. *dvā*, Go. *twai* ‘id.’); *երկես* *erke-a-w* ‘feared’ < PA *e-erkī-á-βo < *é-dūi- (Hom. ἔ-δει-σ-ε[v] ‘was startled’ [with orthoepical -δδ- for *-δF- in some mss.], Av. *duuaē-θā-* ‘threat’); cf. Ved. *dvēs-ti* ‘is hostile; hates’ < *dūéi-s-. *dū > *րկ* *rk* (“Meillet’s law”, cf. Meillet 1924: 1 f.) is rejected by Pedersen (1906: 398 f. [= 1982: 196 f.]) and Kortland (1980: 99 f.; 1989: 50 [= 2003: 28¹; 94]), who posit *dū > *q k* instead and propose new etymologies or different reconstructions, e.g., *երկ-* *erk-* ‘fear’ < (?*herk- <) *preg^(u)- (Goth. *faurh-t-jan* ‘id.’ < *prg^(u)-t-jo-, TB *parskam* ‘they fear’ < *prK-skō-nti; LIV² 491); Gk. ὀδύνη < *h₃(e)d-un-, i.e. a different root, cf. (ἀ)νώδυνος ‘painless’ < *h₃-h₃d^o (cf. Lith. *úodas* ‘gnat < *stinger’ < *h₃od-o-); *dūō > PA *ku → *երկու* *erkow* ‘two’ with *er- adopted from *երեք* *erek* ‘three’. In addition, the phonetic mechanics of *dū > *րկ* *rk* remains a stumbling block to many; cf. the “feature metathesis” analysis of DeLisi (2013).

9.1.8. *q k* < *{-st}u-; the only citable example is *ոսկր* *osk-r* ‘bone’ < PA *(h)ós(t)-kir ← *Host- + *uēr, cf. Ved. *ásthi-*, Av. *ast-*, *as[-ca]*, Lat. *os(s)*, *ossua*, Gk. ὀστέον ‘id.’ The form *Host-uēr, originally analyzed as *ostu-er, i.e. *Host-u- + *-r/-n- (Meillet 1936: 51; Lat. *ossu-a*, Gk. ὀστέF-ov), is a transposed projection and hence very likely a reflection of the original root noun (?)*H_{2/3}ost- which adopted the (reflex of the original) suffix *-uer/n-; cf. Skt. *ásthi* < either *Host-i/n- or *Host-H₁?, Gk. ὀστέον < *Hosth₁-jo- or *Host-eu-?, Hitt. *haštāi* < *Host(h₁)-ōi, Lat. *oss-u-* < ?*Host-u-, os[s] < ?*Host(H); OCS *kostь* < ?*Host(-i-). In any case, it is not clear at which stage the projected sequence *-stu- reflects historically real linguistic input and in which form; perhaps PA */ost-uir/ * [os]_σ[tuir]_σ > * [os]_σt[*vir*]_σ > * [os]_σ [*vir*]_σ > * [oskir] and shows *s preserved in *st (9.1.2), rather than *{s}tū > *{s}k^h (9.1.22).

9.1.9. *p b* < *b^h; *p b* continues an original *b^h word-initially and after a resonant; (*-Vb^hV- > *l w* or *q v*, 9.6.7, 4.5), e.g. *բամբ* *b-a-m* ‘I say’ < *b^héh₂-mi (Dor. Gk. φάμι, Lat. *fā-tum* ‘prophecy’); *հարբ* *har-b* ‘father (Isg.)’ < PA *hárbi < *φά(θ)arbi < *ph₂tr₂-b^hi; *դամբ* (*արման*) *damb-(ar)an* ‘grave’ < *d^hmb^h- (YAv. *daxma-* ‘id.’ < *daf-ma-, Hom. τάφος ‘funeral rites’, OPr. *dambo* ‘ground’; cf. Clackson 1994: 120 f.; Beekes 2010: 534; LIV² 143 f.); *ամբ* *amb* ‘cloud’ < ?*ṅb^h-és- (Ved. *ámphas-* ‘water’, cf. Lat. *imber* ‘rain shower’ < ?*ṅb^h-es-, Ved. *abhrá-* ‘cloud’, Av. *aβra-* ‘rain’ < *ṅb^h-ró-); cf. Clackson (1994: 133); for the proposed connection of the arguably older variant *ամսյ*

amp ‘id.’ with Skt. *ambu-* ‘water’, *ambara-* ‘sky; garment’, and Gk. ὄμβρος ‘(rain-) shower’, see Martirosyan (2010: 51).

9.1.10. $\eta d < *d^h$; in all positions except perhaps $*-Vd^hV-$ (9.4.8), e.g. $\eta\mu\eta\eta$ *dowr-k* ‘door’ $< *d^h\acute{u}r-(e)h_2-$ (Gk. θύραι, Lat. *forēs*, Goth. *daur*); $\eta\eta\eta\eta$ *gind* ‘earring; necklace’ $< *uend^h-eh_2-$ (cf. Goth. *windan* ‘to wind’, Ved. *vandhúr[a]*- ‘charioteer’s seat’), etc.

9.1.11. $\eta d < *-\{R/N\}t-$; i.e. sonorization after $*R$, e.g. $\omega\eta\eta$ *ard* ‘now’ $< *h_2\acute{e}rti$ (Gk. ἄρτι ‘just now’, Lith. *arti* ‘near’); $\mu\omega\eta\eta$ *mard* ‘human’ $< *m\acute{r}-t\acute{o}-$ (Ved. *mṛtá-* ‘*having died > dead,’ Gk. βροτός ‘mortal’, OCS *mṛtvŭ*); $\rho\eta\eta\eta\eta$ *ander-k* ‘bowels’ $< ?*h_1\acute{e}n-ter-h_2$ (Gk. ἔντερα, Ved. *ántara-* ‘interior’, *āntrá-* ‘intestine’), etc.

9.1.12. $\eta g < *g^{(u)h}$: $\eta\eta\eta\omega\eta$ *gel-awn* ‘song’ $< *g^h\acute{e}l-$ (ON *gala*, OHG *naht-gal-a* ‘*night-singer > nightingale’); $\eta\eta\eta\eta\eta$ *getj-k* ‘glands’ $< *g^h\acute{e}l\acute{g}^h-$ (Lith. *gėležuonys*, OCS *žlězy*); $\mu\eta\eta\eta$ *mēg* ‘fog’ $< *h_3mei\acute{g}^h-o-$ (Ved. *meghá-* ‘cloud’, Gk. ὁμίχλη ‘mist’, Lith. *miglà* ‘id.’); $\eta\omega\eta$ *gan* ‘blow; wound’ $< *g^{uh}\eta-(ti)-$ (cf. Hitt. *kuen-zi*, Ved. *hán-ti*, Gk. θείνει ‘kills’).

9.1.13. $\eta g < *u$: $\eta\eta\eta\delta$ *gorc* ‘work’ $< *u\acute{o}r\acute{g}-o-m$ (Dor. *Fέργον*, OHG *werk*, Av. *varəza-*); $\eta\eta\eta$ *gin* ‘price’ $< *ues-n\acute{o}-$ (Ved. *vasná-*, Lat. *vēnum*); $\omega\omega\eta\eta\eta$ *taygr* ‘brother-in-law’ $< *deh_2iu\acute{e}r$ (Ved. *devár-*, Lith. *dieveris*); $\delta\eta\eta\eta\eta\eta$ *cowng-k* ‘knees’ $< *g\acute{e}on\eta-(h_2/i)h_1-$ (Gk. γούνα, Lat. *genū*, TA *kanwē-ṃ*); $\eta\eta\eta\eta$ *kogi* ‘butter’ $< *g^u\acute{o}\eta-(i)\eta-$ (Ved. *gávya-* ‘cow-’, YAv. *gaoiia-*, Gk. -βουιος).

9.1.14. $\eta g < *-\{R/N\}k^{(u)}$; i.e. sonorization after $*R$: $\omega\eta\eta\eta\eta$ *argel* ‘obstacle’ $< *h_2(e)rk-$ *el-* (Lat. *arcula* ‘casket’, OHG *rigil* ‘bolt’; cf. Gk. ἀρκέω, Lat. *arceō* ‘I ward off’); $\eta\eta\eta\eta$ *erg* ‘song’ $< *h_1erk^u-o-$ (Ved. *arká-*, Hitt. *arku-* ‘chant’, TB *yarke* ‘worship’); $\zeta\eta\eta\eta\eta$ *hing* ‘five’ $< *pénk^ue$ (Ved. *pánca*, Gk. πέντε, Lat. *quīnque*, Lith. *penkì* ‘id.’).

9.1.15. $\eta p < (?*p^h <)*sp-$: $\eta\eta\eta\eta\eta$ *p’oyt* ‘zeal’ $< ?*speud-to-$ (Gk. σπουδή, Lith. *spáusti* ‘squeeze’, *spūdà* ‘urgency’, ?Alb. *punë* ‘business’ $< *spud-nā-$); $\eta\eta\eta\eta\eta$ *p’ownd* ‘container’ $< *p^hon(-)d^h-o-$ (OCS *spodŭ* ‘bushel’, Latv. *spanda* ‘strap’; MLG *span* ‘frame [of a ship]’); $\eta\omega\omega\omega\omega$ *p’ast* ‘evidence’ $< ?*sp\acute{e}k-ti-$ (Ved. *á-spaṣ-ṭa-* ‘invisible’, Lat. *speciō* ‘I see’, Gk. σκέπτομαι ‘I view; consider’ $< *σπεκ-$) (but see 9.1.2).

9.1.16. $\eta p < ??*h_1p-$; voiceless aspirates may continue an inherited PIE sequence $*h_{(1)}+*T$ (cf. 9.1.19): $\eta\eta\eta\eta$ *t’owp* ‘bush’ $< ?*tuHp-$ (Gk. τύφη ‘reed mace’, OE *pūf* ‘thicket’). However, this view (see Olsen 1999: 773–775, 2010: 40³⁴) has not met with general approval (cf. e.g. Martirosyan 2010: 323). The items generally involved in the discussion of voiceless aspirates frequently represent peripheral (substratal?) vocabulary. In addition, Olsen’s proposal implies an $*[h]$ -like realization of $*h_1$, a decidedly minority view.

9.1.17. $\eta t < *t$: $\eta\omega\omega\omega\omega$ *t’afam* ‘withered’ $< *t\acute{r}s-$ (Goth. *ga-paursana* [Asg.], Ved. *tṛṣāná-* ‘thirsty’, Gk. τέρσομαι ‘I wither’); $\eta\omega\omega\omega$ *t’ana-* ‘dampen’ $< *t-\eta-h_2-$ (Gk. τή-κω ‘I melt’, OCS *ta-j-etŭ* ‘melts’, Lat. *tā-b-eō* ‘I waste away’); $\eta\eta\eta\eta\eta$ *-owt’iwn* (abstract noun suff.) $< ?*eu-ti-h_3on-$ (Lat. *-tiō[n-]*, OIr. *-tiu*, Goth. *ra-hjo* ‘account, number’).

9.1.18. $\beta t' < ?*pt-$: $\beta b\eta t' er$ 'leaf; petal; side' ($< *'wing'$) $< *pter-$ and $\beta\eta n-$ $ti\bar{r}-$ 'flight; fly (vb.)' $< *pt\bar{e}r-$ (cf. Gk. $\pi\tau\epsilon\rho\acute{o}\nu$ 'feather, wing'; $\pi\tau\epsilon\rho\nu\acute{\xi}$ 'bird'; $\pi\tau\acute{\epsilon}\sigma\theta\alpha\iota$ 'to fly').

9.1.19. $\beta t' < ??* -h_1t-$ (cf. 9.1.16): $\beta n\eta\eta\beta$ *lowrt* 'shiny, light-blue' $< PA$ $*\hat{k}l\bar{u}t^hro-$ $< *k\bar{l}uh_1-tro-$ (Goth. *hlutrs* 'clear, pure').

9.1.20. $\beta k' < *k^{(u)}$: $\beta\eta\eta\eta k' or-\bar{e}$ 'scratches; itches' $< *(s)kor-\acute{e}je-ti$ (Alb. *harr* 'clips', Gk. $\kappa\acute{\epsilon}\iota\pi\epsilon\iota$ 'shears', Hitt. *iškāri* 'pricks', OHG *sceran* 'to shear'); $\beta\eta\eta\eta\eta^o k' ani$ 'how many?' $< PA$ $*k^h an-tjo-$ $\leftarrow *k^u\acute{e}h_2-(\eta)e nt-$ (Lat. *qua-nt-us*, Ved. *kī-vant-*, Av. *cuuant-* 'id.'): $\eta\beta o-k'$ 'anyone; *whosoever' $< *k^u\acute{o}s-k^ue$ (Ved. [*yá- (...)*] *kás ca*, Lat. *quis-que*, Goth. *huaz-u-h*).

9.1.21. $\beta k' < *s\bar{u}-$: $\beta\eta\eta\eta k' own$ $< *s\bar{u}\acute{o}p-no-$ (Ved. *svápna-*, ON *svefn*, Lat. *somnus*); $\beta\eta\eta\eta k' oyr$ 'sister' $< *s\bar{u}\acute{e}s\acute{o}r$ (Ved. *svásā*, Goth. *swistar*, Lat. *soror* 'id.').

9.1.22. $\beta k' < *t\bar{u}-$: $\beta\eta k' o$ 'your' $< *t\bar{u}o-(s)jo$ (Ved. *tvá-*, Gk. $\sigma\acute{o}\varsigma$, OCS *tvo[jb]*); $\beta\eta\eta n-$ *ka\bar{r}-* 'four' (an allomorph of $\beta\eta\eta\eta -\acute{c}or-$ from $*k^u\acute{e}t[u]or-es$) $< *(k^u)t\bar{u}\bar{r}-$, cf. Gk. $\tau\rho\acute{\alpha}\pi\epsilon\zeta\alpha$ 'table' ($< *'of four-legs'$ $< PGk. *t[u]r-ped-ja$); Av. *ā-xtū'rī-m* 'four times', Ved. *turīya-* 'fourth' ($< *k^u\acute{t}ur-$). For the possibility that $*\{s\}t\bar{u}$ may yield $\{s\}k$, cf. 9.1.8.

9.1.23. $-\beta -k' < *-(e)s$: $\beta\eta\beta\eta\beta e-re-k'$ 'three' $< *tr\acute{e}j-es$ (Ved. *tráyah*, Gk. $\tau\rho\epsilon\acute{\iota}\varsigma$, Lat. *trēs*); $\beta\eta\eta\eta\beta \acute{c}or-k'$ 'four' $< *k^u\acute{e}t\bar{u}or-es$ (Ved. *catvārah*, Dor. $\tau\acute{\epsilon}\tau\omicron\rho\epsilon\varsigma$); $-\beta\eta\beta -m-k'$ (1pl.) $< *me-s$ (Ved. *-mās*, Dor. *-με\acute{s}*, Lat. *-mus*, OCS *-mъ*); $-\beta\eta\beta -\gamma-k'$ (2pl.) $< *te-s$ (2du.?) (cf. Lat. *-tis* 'id.', Ved. *-thah* [2du.], Goth. *-ts* 'id.'): $-\beta\eta -k'$ (Npl./Ipl.) $< *Vs/*(-b^hi-)s$, e.g. $\beta\eta\eta\eta\beta k' or-k'$ 'sisters (Npl.)' $< *s\bar{u}\acute{e}sor-es$ (Ved. *svásār-ah*, Lat. *sorōr-ēs* 'id.', Gk. $\acute{\epsilon}\omicron\rho-\epsilon\varsigma$ 'relatives' [Hes.]); $\beta\eta\eta\eta\beta\beta har-b-k'$ 'fathers (Ipl.)' $< *p\eta_2t\bar{r}-b^his$ (Ved. *pitṛ-bhih*).

9.2. Affricates

$/c/ \delta$	$*\hat{g}$ 9.2.1 $*\{u\}g$ 9.2.2 $*ds$ 9.2.3 $??*d\bar{i}$ 9.2.4	$/j/ \acute{d}$	$*\hat{g}^h$ 9.2.8	$/c'/ g$	$*s\hat{k}, *sk, *k\hat{s}$ 9.2.14 $*t\hat{k}, ?*d^h\hat{g}^h, ?*d^h\hat{g}^{uh}$ 9.2.15 $?*t\bar{i}$ 9.2.16 $?*k\bar{i}$ 9.2.17
$/č/ \acute{d}$	$(?)*\hat{g}\bar{i}-$ 9.2.5 $(?)*\hat{d}\bar{i}$ 9.2.6 $??*g^{(u)}\{e, \bar{i}\}$ 9.2.7	$/j/ \acute{z}$	$*g^{uh}\{e, \bar{i}\}$ 9.2.9 $*-d^h\bar{i}-$ 9.2.10 $*-\{R\}\bar{i}-$ 9.2.11 $(?)*\bar{i}-$ 9.2.12 $??*-\{R\}t\hat{k}$ 9.2.13	$/č'/ \acute{z}$	$*k^u\{e, \bar{i}, \bar{i}\}$ 9.2.18 $?*(-)k\bar{i}-$ 9.2.19 $?*\bar{t}\bar{i}-$ 9.2.20

9.2.1. $\delta c < *g\hat{c}$: $\delta\eta\eta cer$ 'old (man)' $< PA$ $*iser-o(h)-$ $\leftarrow *g\acute{e}r\eta_2-s$ (Gk. $\gamma\acute{\epsilon}\rho\alpha\varsigma$ 'honor $< ?*maturity$ '; cf. Ved. *járant-*, Gk. $\gamma\acute{\epsilon}\rho\omicron\nu\tau-$, Oss. *zæronđ* 'old', OCS [*sŭ*]/*zbrĕti* 'ripen'); $\delta\eta\delta mec$ 'big' $< *m\acute{e}g-\eta_2-$ (Gk. $\mu\acute{\epsilon}\gamma\alpha\varsigma$, Lat. *mag-nus*, YAv. *maz-nt-*, Ved. *māhi*, Alb. *madh*).

9.2.2. $\delta c < *-\{u/\underline{u}\}g-$, i.e. “satemization” of [labio]velars after $*u/\underline{u}$ (cf. 9, 9.3.2): լուծ lowc ‘yoke’ ← PA $*\text{ju}g\text{-o-}$ < $*\text{jug-ó-m}$ (Gk. ζυγόν, Lat. *iugum*, Goth. *juk*, OCS *igo*); presumably contaminated by $\text{լուծանե- lowc-an-e-}$ ‘loosen, dissolve; unyoke’ (cf. Martirosyan 2010: 316); -բոյժ -boyc ‘nourishing’ < $*b^h\text{óug-o-}$ (Ved. *bhóga-* ‘enjoyment’, Lat. *fu-n-g-i-* ‘to occupy o.s. with’); ործա- orc-a- ‘vomit’ < PA $*(\text{φo- +}) *aruc-$ < $(*h_2p\acute{o} +) *h_1reug-e-$ (Gk. ἐρεύγομαι ‘I belch’, Lat. *ē-rūgere* ‘disgorge noisily’, Lith. *r(i)áugėti* ‘belch’; OCS *otb-rigati* ‘*belch out’ > ‘disgorge [words]’), etc. (cf. 9.3.2. for the parallel $u s < *-\{u/\underline{u}\}k$).

9.2.3. $\delta c < *ds$: աղծ- alc- ‘filth; salt’ < $?*s\ddot{a}l-d-s$ (cf. աղա alt ‘dirt; salt’ < $?*s\ddot{a}ld-i-$) (Goth. *salt*, OHG *sulza* ‘silt; sediment’); անիծ anic ‘nit’ < $*[(s)k]onid-s$ (Gk. κνίδ-, Alb. *thëni*, OE *hnitu*).

9.2.4. $\delta c < ?*-d-i-$; this development stands (or falls) on morphological considerations. Some scholars (e.g. Scheffelowitz 1905: 30; Godel 1965: 25 [= 1982: 23], 1975: 82; Polomé 1980: 21; Klingenschmitt 1982: 194–195; Olsen 1999: 88, 92, 811, etc.) consider δc to be the reflex of the original cluster $*-d-i-$ in present tense $*-i\ddot{e}/o-$ formations to roots in final $*d$ ($>$ PA $*t$). These forms generally surface as nasal presents: $\text{մուծանե- mowc-an-e-}$ < $*m\ddot{o}ud-i\ddot{e-}$ ‘introduce’ < ‘make enter’ (causative of the *sōpire* type), cf. *բխալու e-mowt* ‘entered’; $\text{անիծանե- anic-an-e-}$ ‘curse’ < $*h_3n\ddot{o}id-i\ddot{e-}$ (iter.) (cf. Goth. *ga-nait-ja-* ‘revile’). Other scholars maintain that $*-d-i-$ phonetically yields $\delta \check{c}$ (9.2.6) and that the forms with δc are better explained by assuming that the Armenian nasal presents were built on the reflexes of original *s*-aorists (i.e. $\delta c < *-d-s-$; 9.2.3), and therefore surface as their synchronic counterparts: անիծ anēc ‘(s)he cursed’ < $*h_3n\ddot{e}id-s-$ (YAv. *nāist* ‘id.’ < $*n\ddot{a}id-s-t$), հեծալ hec-a-w ‘(s)he mounted (a horse)’ < $*s\ddot{e}d-s-$ (OCS *sědv* ‘I sat [aor.]’, Ved. *sátsi* ‘sit down! [aor.impv.]’), cf. հեծանի- hec-an-i- ‘sit’.

9.2.5. $\delta \check{c} < (?)*-g_i-$; some examples adduced for this development involve root etymologies: աճուկ ačowk ‘groin’ < $?*pag-i\text{-o-}$ (Ved. *pājasyá-* ‘belly’, *?pakśá-* ‘wing; side’, Russian *pax* ‘loins’); աճիւն ač(-)iwn ‘ash’ < $?*azg-i\text{-o-}$ (Gk. ἄζω ‘I dry up’, Goth. *azgo*, OHG *asca* ‘id.’). However, the two most viable possibilities are աճե- ače- ‘grow’ and կոճ koč ‘log’. The former is possibly related to աւճան awč-an ‘help’, and they both may go back to PA $*auj-$ from (virtual) inner-Armenian $*-i\ddot{e}$ formations $*h_2eug-i\ddot{e-}$ (cf. Lat. *augeō*, Goth. *aukan*, Lith. *áugti*) and $*h_2eug-i\text{-o-}$ (cf. Lat. *auxilium* ‘help’ < $*aug-s-$), respectively. The expected offglide in աճե- ače- (< $*auj\ddot{e-}$) is also missing in the pretonic syllable of աճուկ acul ‘coal’ from PA $?*a\ddot{u}ts\ddot{u}t\text{-o-}$, PIE $*h_1\ddot{u}g^h\text{-ó}l-$ (cf. OCS *oglb*, Ved. *ángāra-* ‘id.’, Martirosyan 2010: 43). կոճ koč ‘log; [chopping] block’ (< $?*go(\underline{u})g-i\text{-o-}$; cf. also կոճի կոճ-ի ‘beam , $\text{կոճակ կոճ-ակ ‘button}$; կոճքե- կոճ-կ-ե- , $\text{կոճուքե- կոճ-օք-ե-}$, $\text{կոճուքի- կոճ-օտ-ի-}$, etc., all ‘beat; break’) seems to be related to cognate forms (cf. Lith. *gūzas* ‘crop’, *gūžė* ‘cabbage head’, *gugà* ‘button’, OHG *kuocho* ‘cake’, Old Czech *hýže* ‘thigh’ as well as Gk. γογγύλος ‘round’, ON *kokkr* ‘lump’, etc.), none of which are $*-i\ddot{e}/o-$ formations. Therefore, it is formally more likely that $\delta \check{c}$ reflects a conditioned development of PA $*-ts-$ (or $*-dz-$ before the “sound shift”) from either $*\check{g}$ (9.2.1) or a ‘satemized’ reflex of the inherited $*g$ (9.3.2) which interacted with productive PA verbal or nominal $*i$ -formants (i.e. $*\{u\}g- >$ PA $*\{u\}\acute{g}- > *-ts- \Rightarrow *-ts-i-$ [or perhaps $*\{u\}\acute{g}- > *-dz- \Rightarrow *-dz-i-$ > $*-d\check{z}-$] > $*-j-$ > $-\delta-$ $-\check{c}-$). Cf. կոճ- koc- ‘beat;

lament' from $?*go(u)g-/ *goġ-$ (apparently related to $կոճ-$ *koč-* 'block', cf. Martirosyan 2010: 756), which served as a base for an inner-Armenian reduplicated $*-je-$ present $կոչկոճե-$ *koš-koč-e-* 'beat, break, chop' presumably from PA $*kots-kots-je-$ via $*kots-kofj-è-$ ($*-ts-ĭ-$ > $ճ$ \check{c}) with shibilant assimilation to $*kofj-kofj-è-$ (cf. $ճանաչե-$ *čan-ač-e-* 'know', aor. $ճանա-$ *cane-a-* 'knew' < $*ġnh_3-$) and with dissimilation of $[\pm cont]$ across a syllable boundary to $[kof.ko.ʃjè-]$; cf. $բարբառ$ *bar(-)bař* 'dialect' < $*/bař-bař-/$ (4.5); $վեշտասան$ *veš-tasan* '16' < $*uēʃ^h-t^o$, $վաթսուն$ *vat-sown* '60' < $*uēʃ^h-sun-i-$ (10.1.12), in each instance due to a synchronic rule $*[\alpha cont] [\alpha cont] \rightarrow [\beta cont] [\alpha cont]$ (which points to a shibilant realization of the intermediary reflex of $*ks$ [$> g$ c' , 9.2.14], i.e. $վեց$ *vec'* '6' < $*uēʃ^h[ʃ^h]$ < $*s(u)uēks$). This rule seems to have been part of the Armenian phonological system for some time, cf. aor. subj. $սիրբոցես$ *sir-es-c'e-s* 'wilt love' as opposed to $լացցես$ *lac'-c'-es* 'wilt weep' (7).

9.2.6. $ճ \check{c}$ < $(?) *dj-$; this development is proposed based on phonetic parallelism with the preceding development rather than etymological comparison (see Martirosyan 2010: 718 referring to Pedersen 1906: 396–397 [= 1982: 174–175]): $որոճե-$ *oroč-e-* 'chew' < PA $?*(\varphi) *rōd-je-$ ← $?*red-$ (Lat. *rōdō* 'I gnaw', Ved. *rādati* 'digs'); unfortunately, the cited word equations are difficult to assess.

9.2.7. $ճ \check{c}$ < $? *g^{(u)}\{e, \check{i}\}$; i.e. palatalization of a plain velar or $*g^u$. A few forms in which $ճ \check{c}$ may be derived from an inherited $*g^{(u)}$ undergoing palatalization before a front vowel are encountered in the literature: $ճեղ$ *čel* 'bald' < $? *g^{(u)}el-no-$ (OHG *kalo* 'id.' < $*g^{(u)}ol-uo-$, OCS *golb* 'naked') and the related $ճիղ$ *čilm* 'young, unripe' (< $? *smooth$) < $*g^{(u)}ēl-mo-$; $ճղճիլ$ *člčim* 'meager; stingy'; $ճղճիլ-$ *člčm-i-* 'be stingy; be scarce; be wrinkled' (OE *clingan* 'shrink, wither', MHG *klōk* 'wise <(*)thrifty'). Another item mentioned is $ճլլե-$ *čm-l-e-* 'squeeze' < $? *gem-l-$ (OCS *žeti* 'reap, harvest', Oic. *kumla* 'wound', Mir. *gemel* 'fetter', Gk. Cypr. [Hsch.] $\check{\upsilon}\gamma\text{-}\gamma\epsilon\mu\text{-}\omicron\varsigma$ $\sigma\upsilon\lambda\lambda\alpha\beta\eta$). However, palatalization of the inherited plain velars comes at a price of assuming extensive analogy in the formative period of PA morphology (cf. 9). Thus, $ճեղ$ *čel* 'bald' could also be connected with Lat. *calvus*, Skt. *-kulva-*, YAv. *kauruua-* 'thin haired' < $*k\check{I}H-(e)uo-$ (Derksen 2008: 176). In most cases, $ճ \check{c}$ is of loan origin, e.g. $ճիւ$ *čiw* 'shank; leg' < Iran. $*čīva-$ (Av. *ascūm* 'shin [Asg.]' < $*ast-čīwa-$ [Martirosyan 2005; cf. Vogt 1958: 159]).

9.2.8. $ճ j$ < $*ġ^h$: $ճերն$ *jeř-n* 'hand' < $*ġ^hēs-r-n$ [Asg.] (Gk. $\chi\epsilon\iota\rho\text{-}\alpha$, Ved. *hás-ta-*, Av. *zas-ta-*); $անձուկ$ *anj-ow-k* 'narrow' < $*h_2(e)nġ^h-u-$ (Ved. *amhú-*, OCS *ozb-kb*, Goth. *aggwus*).

9.2.9. $ջ \check{y}$ < $*g^{uh}\{e/i/?\check{i}-\}$, i.e. palatalization of $*g^{uh}$: $ջեր$ *jer* 'warmth' < $*g^{uh}ér-os-$ (Gk. $\theta\epsilon\rho\varsigma$ 'harvest; summer', Ved. *háras-* 'flame, heat'); $ջիլ$ *jit* 'sinew, tendon' < $*g^{uh}iH-(s)leh_2-$ (Lat. *filum* 'thread', cf. Lith. *gýsla* 'vein', OCS *žila* 'vein, sinew', SCr. *žila*, which might, however, formally also go with Gk. $\beta\acute{\iota}\omicron\varsigma$ 'bow-string' [$< *g^{uh}iH-ó-$]); $ջերմ$ *jerm* 'warm' < $*g^{uh}er-mó-$ (Gk. $\theta\epsilon\rho\mu\acute{\omicron}\varsigma$ 'id.', Skt. *gharmá-*, Thrac. $\Gamma\epsilon\rho\mu\alpha\varsigma$ TN '[*]hot springs').

9.2.10. $ջ j$ < $*-d^h\check{i}-$: $մէջ$ *měj* 'middle' < $*méd^h-jo-$ (Ved. *mádhya-*, Gk. $\mu\acute{\epsilon}\sigma\sigma\omicron\varsigma$, Goth. *midjis* 'id. '); $ննջե-$ *nnje-* 'nod off; fall asleep' < PA $*ni-nud\check{z}-è-$ < $*-snud^h-je-$ (Lith.

snáudžiu ‘I doze, slumber away’, cf. also Gk. νυστάζω [**(s)nudh-tázō*], Kölligan 2007); *ղէջ* *gēj* ‘moist’ < **g^heǰ^h-jo-* (cf. Russ. *židkij* ‘watery’, SCr. *židak* ‘id.’).

9.2.11. *ջ* < **-{R}j-*: *վերջ* *verj* ‘edge, end’ < **(s)uper-jo-*; Skt. *upári* ‘over’, Lat. *s-uper* ‘id.’; *ողջ* *olj* ‘whole; safe’ < **h₂ol-jo-* (OIr. *huile*, Lith. *ali-aĩ* ‘completely’). The double treatment of **Rj* (> *Rj/jR*; 10.2.5; cf. Schmitt 1980: 427–430) has been connected with the presence or absence of (*)*a*-vocalism in the preceding syllable, i.e. **aRj* > *ajR* (e.g. Godel 1982a: 60 f.; Olsen 1999: 795 ff.). Viredaz (2001–2002b) suggests that the process of epenthesis operated only after the vocalization of the syllabic resonants (i.e. PA **Rj* > *aRj* before PA **aRj* > **ajR*).

9.2.12. *ջ-յ-* < (?)*(*H*)*j-*: *ջուր* *jowr* ‘water’ < **juH-r-* (Lith. *júrės*, OPruss. *iūrin* [Asg.]); *ջան* *jan* ‘effort’ < **ieh₂-ni-*; cf. Dor. ζᾶλος ‘zeal’, Ved. *yámi* ‘I beseech’; *ջով* *jov* ‘sprout’ < ?**ieuh₁-o-*; cf. Ved. *yáva-* ‘barley’, Hitt. *ewa-* ‘id.?’; Lith. *javaĩ* ‘grain’; cf. Gk. ζεῖα ‘spelt’); *ջորի* *jori* ‘mule’ < ??**ieuo-ro-* (Lith. *jáutis* ‘ox’; cf. Ved. *yuváti* ‘yokes’, *ni-yú-t* ‘team’); cf. Lat. *iūmentum* < **ieug-s-mŋ-to-* ‘pack-horse’ (√**ieu[-]g-* ‘to yoke’). The evidence for this development is not universally accepted (cf. Kölligan 2012: 138); alternatively, **j-* > Ø (cf. Kortlandt 1998 [= 2003: 122–124]; Martirosyan 2010: 706). The most convincing evidence for this would be the relative pronoun *որ* *or* ‘who, which’ under the assumption that this form comes from *(*H*)*jo-*. Cf. *ուր* *owr* ‘where?’, *ով* *ov* ‘who?’, *որ* *or* ‘which’ < *(*H*)*jo-* (Ved. *yá-*, cf. West Slavic **ja-kъ* ‘which[?]’).

9.2.13. *ջ* < ??**{R}tḱ-*: *արջ* *arj* ‘bear’ < **h₂tḱo-* (Hitt. *har-tág-ga-aš*, Ved. *ṛkṣáh*, Gk. ἄρκτος). The sources I consulted are uniformly dismissive of the possibility that *արջ* ‘bear’ could be derived from **h₂tḱo-* by regular sound change (cf. e.g. Martirosyan 2010: 743). This dismissal seems to be primarily based on two assumptions: (i) a comparison with the reflexes of **tḱ* > *g c’* (9.2.15) as they surface in *other* positions (cf. e.g. Olsen 1999: 184), and (ii) the traditional view that the development of the thorn clusters had to involve metathesis and fricativization parallel to Gk. or Skt., i.e. **tḱ* > **kḅ* (vel sim.), which has, however, recently been questioned by Lipp (2009). Based on the development of {**tḱ* =} **sḱ*, i.e. {**θtḱ* =} > **htḱ* > **hʃ* > **ʃ^h* > *g c’* [*t^h*], cf. {**t₂* (9.1.22)=} **s₂* (9.1.21), i.e. {**θ₂* =} > **θ₂* = **θ₂* > **hx* > **x*: > **x* > *g k’* [*k^h*] (fortition; cf. 9.3.15), the development of **-Rtḱ-* may reflect **Rθtḱ* > **Rðḱ* (sonorization) > **Rddḱ* (fortition) > *Rj*. Furthermore, **d^hg^h* seems to reflect **d^hg^h* (dissimilation of aspiration) > **ddḱ* > **tdḱ* (“shift”) = **tts^h* > *g c’* (9.2.15) (cf. **tḱ*, **ts* > *g c’*).

9.2.14. *g c’* < **sḱ*, **ks*, **sk*: *բարց* *e-harc’* ‘asked’ < **é-p₂(k)-skē-t* (Ved. *áp₂rechāt*, Av. *pərəs-*, Lat. *poscere*); *վեց* *vec’* ‘six’ < **s(u)ueks* (Ved. *ṣát*, Gk. ἕξ, Av. *xšuuas*); *-ից* *-i-c’* (Gpl.) < **-isk-o-*, in e.g. OCS (*ljud*)-*ьskъ* ‘human, (typical) of men’, Lith. (*lietu̯v*)-*iška-s* ‘Lithuanian, of the Lithuanians’, Goth. (*þiud*)-*isk-s*, OHG (*diut*)-*isc* ‘ἐθνικός’. Here may perhaps be also included cases of the relatively early voicing assimilation of **sg^(u)* > **sk^(u)*: *երէց* *e-rēc’* ‘elder’ < PA **φρέisk-u-* < ?**pre(i)s-g^u-* (Gk. *πρεσβύς*, Lat. *priscus*).

9.2.15. *g c’* < **tḱ*, ?**d^hg^h*, ?**d^hg^{uh}*; i.e. so-called “thorn” clusters: *ցին* *c’in* ‘vulture’ < **tḱiH-ino-* (Gk. *ἰ-κτινός*, *ἰ-κτιν* ‘kite’); cf. Ved. *śyená-*, Av. *saēna-* (< ?*[t]ḱiēH-ino-); *ցամաք* *c’amak’* ‘(dry) land’ < **d^hg^hḱ-* (Ved. *kṣáh*, Gk. *χθών*, Hitt. *tēkan*, TA *tkam*).

ձուկն *jowk-n* ‘fish’ is in some accounts considered a counter-example to the development of this specific thorn-cluster configuration because it is understood as a reflex of *d^hǵ^huH- (Gk. ἰχθῦς ‘id.’) or *d^(h)ǵ^huH-(m) (Beekes 2003: 196, 2010: 606; cf. Lith. *žuvīs* ‘id.’). It is, however, impossible to determine whether **ձ** *j* continues *d^(h)ǵ^h- or whether in this particular word, the cluster simplified to *ǵ^h-word-initially (9.2.8).

9.2.16. **ց** *c* < ?*tj: **երիցս** *eric’ (-)s* ‘three times’ < *tri-tjo-, cf. Lith. *trėčias* ‘third’, OP *θritiya-*.

9.2.17. **ց** *c* < ?*k̄j: **լուցե-** *lowc’-e-* [aor. (originally impf.)] ‘lit’ < PA *lojts^hé- < *louk̄-je- < ?*lōuk̄-je- (Lith. *láukti* ‘expect < ?*allow to be seen clearly < *illuminate’, presumably with significant change in verbal valence, cf. *louk̄-éje- ‘light up’ > Lat. *lūcēre*, Ved. *rocáyati*; cf. OCS *po-lučiti* ‘reach, get < *intend’, again with complex change in syntactic frame). According to Klingenschmitt (1982: 265–266), the PIE causative/iterative type R(ó)-je- (e.g., *k^uólh₁-jo-h₂ai > Gk. *πωλέο[μ]αι* ‘I wander to and fro’; *súóp-je- > Lat. *sōp-t-re* ‘lull to sleep’ LIV² 23) is at the root of the Armenian causative formation in **-ոյց-** *-oyc’* (**-ոյգ-** *-oyz-*, **-ոյս-** *-oys-*, and marginally lexical **-ոյծ-** *-oyc’*): *h₁óuk̄-je-t ‘caused to learn, taught’ > *ouk̄je(t) > *oyc’ → ows-oyc’ **ուսոյծ** ‘id.’ (i.e. the causative *oyc’ reinforced by the present stem *oys- < *ouk̄- < h₁éuk̄- ‘learn’; cf. OCS *uč-i-ti* ‘teach’ < *h₁ouk̄-éje-; Ved. *ucyási* ‘you get used to’ < *h₁uk̄-je-si).

9.2.18. **չ** *č* < *k^u{e/i/ĩ}-, i.e. palatalization of labiovelars (except *g^u, 9, 9.2.7), e.g. **չորք** *čor-k’* ‘four’ < *k^uet(ü)or-es (Ved. *catvárah*, Dor. *τέτορες*, Lat. *quattuor*); **չոգան** *čog-a-n* ‘they went, moved’ < *k^uiéü-nt-o, cf. Ved. *cyávante* ‘they move’, Gk. (Hes.) *σῦθι· ἔλθε* ‘come!’, etc.

9.2.19. **չ** *č* < ?*(-)k̄j: **-իչ** *-ič’* (noun of agent suffix) < *-k-ĩã- (OCS *-ičb*, *-ačb*, Alb. *-s*); **փչե-** *p’č’e-* ‘blow’ < ?*p^hük̄-je-; cf. **փուք** *p’owk’* ‘breath’ < *p^hü-ko-; (Gk. *φῦσα* ‘bellows’ < *p^hüt-ja, OCz. *puxati* ‘to swell’ < *p^hü-s-); cf. Klingenschmitt (1982: 69¹³). Alternatively, **փչե-** *p’č’e-* < ?*p^hüt-je-; **փուք** *p’owk’* < ?*p^hüt-ko-, cf. Skt. *phūt-karoti* ‘makes the sound *phūt*’.

9.2.20. **չ** *č* < ?*-tj- (cf. **ց** *c* < *tj: 9.2.16): **կոչե-** *koč’e-* ‘call, name’ < ?*g^uot-je- (Goth. *qipan* ‘say, express’ < *g^uét-e/o-); however, the word could have been remade based on **գոչե-** *goč’e-* ‘to shout’ < *uok^u-je- (Ved. *vāc-*, Lat. *vōx*, Gk. *ὄπ-* ‘sound’; Ved. *ávocat*, Gk. *ἔπιε* < *é-ue-uk^u-e-t).

9.3. Fricatives

/s/ ս	*k̄ 9.3.1	/z/ զ	?*-ǵ ^h - 9.3.9	/h/ հ	*p- 9.3.12
	{u/ü}k 9.3.2		??-d ^h - 9.3.10		*s- 9.3.13
	*Ns 9.3.3				?*H{a}- (?<*h ₂ e-) 9.3.14
	*-SS- 9.3.4				?*H{o}- (?<*h ₃ e-) 9.3.14
	*s{t}, *z{d} 9.3.5				?*k ^u - 9.3.15
	*ps-, *s{p}- 9.3.6				

Hitt. *uarša-* ‘mist’, Gk. ἔερωση ‘dew’). Meillet (1898) invoked this rule to account for the shibilant in *վեշտասան* *veštasan* ‘16’ and the affricate in *արջ* *arj* ‘bear’; however, both of these might be explained otherwise (cf. 9.2.5, 9.2.13). Pedersen (1905: 208 [= 1982: 70]) assumes that all but the last two items just cited show underlying sequences *-{r}sj-. With the articulation shifted toward the palate, *[r]/s/[i] could well have resisted the assimilation of *rs to PA *-rr-, whence *n*, the regular development (9.4.9), which must have been motivated by the homorganic (alveolar) articulation of the cluster. The evidence for *{k}s is even more tenuous. Thus, *գիշեր* *gišer* ‘night’ has been compared to Middle Welsh *ucher* and Bulgarian (Vinga dial.) *ušcer*, perhaps from ?**ueksperos*, a contamination of **uesperos* ‘evening’ (Gk. ἔσπερος, Lat. *vesper*) and **ksep* ‘night’ (Ved. *kšáp*, Hitt. *išpant-* [*ksp-ént]). Furthermore, *ուշի* *uši* ‘storax, holm oak’ (Gk. ὄξυα ‘beech, spear’, Lith. *úosis* ‘ash-tree’) can be derived from a virtual **h₃ek-s-ieh₂-*. Finally, the evidence for *{k}s comes only from Armenian dialects, where one finds *հաշնիկ* *hašnik* ‘wedding’ (Nor Naxijewan, Sivrihisar) as opposed to *հարսանիկ* *harsanik* (**prk-s-*). Conceivably, however, this development is the result of the following *n* (cf. YAv. *frašna-* ‘question’ beside OAv. *frasā* ‘id.’). On these forms, as well as on “ruki” in general see Martirosyan (210: 709–710). The putative development of ruki *š to *r* posited by Olsen, this handbook (4.4, 7.3, 7.4) (following Pedersen 1905: 231 [= 1982: 93]) should be treated with caution.

9.3.9. *q z* < *-*ǵ^h*-; a native origin of *q z* has been assumed for several intervocalic developments of inherited aspirated stops. The most credible of these involves the voiced aspirated palatal velar *-*ǵ^h*-, which word-initially and adjacent to a consonant yields PA **ǵ* (> *ǵ j* [dʒ], 9.2.8). If the forms listed below are not of Iranian provenance, they would show intervocalic spirantization of the affricate: *դէզ* *dēz* ‘heap’ < PA **dēiǵo-* < **d^héiǵ^ho-* (Av. *daēza-* ‘wall’, Gk. τεῖχος ‘id.’); *մէզ* *mēz* ‘urine’ < **h₃moiǵ^ho-* (Av. *maēza-*, Skt. *meha-*, Gk. ὀμείχω ‘I pee’); *ոզնի* *ozni* ‘hedgehog’ < PA **oǵini-a-* < **h₁oǵ^hi-Hn-* (i)ieh₂ (Gk. ἐχῖνος, cf. Lith. *ežỹs*, OHG *igil*); *լիզ* *liz-ow-*, *լիզ(ան)ե-* *liz-(an)-e-* ‘lick’ < **léiǵ^h-e-* (Gk. λείχω, OCS *ližō*; cf. YAv. *riz-*, Ved. *réh-*).

9.3.10. *q z* < ?*-*d^h*-; This development is rejected by Godel (1975: 130) and Greppin (1980: 131 f.) but supported by Normier (1980: 19 f.), Viredaz (2005: 85), and Kortlandt (2003: 80): *ելուզանե-* *el-uz-an-e-* (a synchronic causative of *ելանե-* *el-an-e-* ‘go out/up’) < ?**h₁louǵ^h-eǵ-* (Hom. ἦλυθον ‘I came’, Ved. *áruhat* ‘went up’); *սուզ(ան)ե-* *sowz-(an)-e-* ‘dive; hide’ < ?**keu(-)d^h*- (Gk. κεύθω, Goth. *huzd* ‘treasure, hoard’ < **kud^h-tó-* ‘hidden’). If this development is correct, one might compare it to its palatalized analog *ǵ j* from *-*d^hi-* (9.2.10). Otherwise, *q z* is frequent in Iranian loans, e.g. *գաւազան* *gawaz-an* ‘ox-goad’ < Iran. **gaw-āz-* (Av. *gauuāza-*) with the Iranian development of *-*az-* from *-*h₂eǵ-*; cf. *ածե-* *ace-* < **h₂eǵ-e-*.

9.3.11. *š ž* < ?*-*g^{(u)h}*-; There is a possibility that medial *š ž* develops from an inherited intervocalic *-*g^{(u)h}*-: *իշ* *iž* ‘snake’ < **h_{1/3}éǵ^{(u)h}i-* (Gk. ὄφις, ἔχις, Ved. *áhi-*, Av. *aži-*). Otherwise initial *š ž* is invariably of foreign (Middle Iranian) origin, as are numerous instances of this phoneme in other environments: *ժաման* *žaman* ‘prompt’, *ժամ* *žam* ‘hour’ < MParth. *jm’n* [zama:n] ‘time’; *ժիր* *žir* ‘busy; diligent’ < Parth. *jyr* [ži:r] ‘wise’ (Av. *jira-*, Ved. *jirá-* ‘lively’). At least in one case, initial *š-ž-* is the result of a language-

internal distant shibilant assimilation: *ժոյժ չօյժ* ‘endurance’ < *z-oyž; cf. *յոյժ y-oyž* ‘very’, *ոյժ օյժ* ‘power’ < Pahl. *ōž* (Av. *aojah-*, Ved. *ójas-*).

9.3.12. *հ-* < *p-: *Հինք* *hing* ‘five’ < *pénk^he (Ved. *pāñca*, Gk. πέντε); *հուր* *howr* ‘fire’, (*հնոց հո-ոց* ‘furnace’) < *puH-r-ó- (*pū-n-)(cf. Gk. πῦρ, Hitt. *pahhur*, Gsg. *pahhuenas*, Olc. *fýrr*, Goth. *fon* ‘id.’); *հերու* *heru* ‘last year’ < *péruti (Gk. πέρυσσι, Ved. *parút* ‘id.’); *հարսն* *harsn* ‘bride’ < *přk- (cf. Lat. *procus* ‘suitor’, Lith. *piřšti* ‘to propose’, Ved. *praśná-* ‘question’).

9.3.13. *հ-* < *s-: *Հին* *hin* ‘old’ < *sén-o- (Ved. *sána-*, Lith. *sėnas* ‘id.’, Gk. ἔνος ‘of last year’).

9.3.14. *հ-* < ?*H-, i.e. *հո-* < ?*h₃e-, *հա-* < ?*h₂e-; many scholars treat the evidence for *հ-* from an initial laryngeal with reservation: *հոտ* *hot* ‘smell’ < *h₃ed-es- (Lat. *odōs*, Gk. ὄδμη ‘id.’); *հովիւ* *hoviw* ‘shepherd’ < *h₃eui-peh₂- (Ved. *ávi-pā-lá-* ‘id.’, CLuw. *hāui-* ‘sheep’, Gk. ὄϊς ‘id.’); *հաւ* *haw* ‘grandfather’ < *h₂euh₂o- (Hitt. *huhha-*, Lat. *avus* ‘id.’, Lith. *avýnas* ‘maternal uncle’); *հաւ* *haw* ‘bird; rooster’ < *h₂éu-i- (Lat. *avis*, Ved. Gsg. *véh* [< *h₂u-ej-s] ‘id.’), cf. Gk. αἰετός ‘eagle’ (< *aFt-) (but note *Ho- > *n-* *o-*, e.g. *ոռ օր* ‘rump’ [Gk. ὄρρος, OHG *ars*], etc.). This development is supported by Kortlandt (1983: 12 [= 2003: 42], 1984b: 42 f. [= 2003: 55]) and Beekes (2003: 181 f.). It neatly accounts for the positive evidence, most of which comes from the dialects, e.g. *հարաւուկ* *harawown-k’* ‘arable land’ < *h₂erh₃-uon-; dial. **hand* ‘cornfield’ (cf. *անդ* and ‘id.’). In other cases Kortlandt resorts to ad hoc transpositions for the sole purpose of saving the rule (cf. Olsen 1999: 766), thus, *արաւր* *arawr* ‘plow’ < *h₂řh₃trom; *ածե* *ace-* ‘lead’ < *h₂ġ-es- (Lat. *gerō* ‘I carry’).

9.3.15. *հ-* (< PA ?*x-) < *k^h-, cf. *հիմ* *him* ‘why, for what?’ alongside *իմ* *im* ‘id.’ (dial. *xi, xe*) < *uē-sm-ōj (YAv. *cahmāi*, OCS *čemu* ‘id.’); *ինչ* *inč’* ‘something’ (dialectal *հինչ* *hinč’*) (Ved. *kím cit* [< *k^him = k^hid], Lat. *quis-quam* ‘anyone’, etc.). Some scholars find this particular development of initial *k^h problematic (cf. e.g. Martirosyan 2010: 299). There is, however, some possibility that the development of *k^(u) to *ք* *k’* passed through a stage of lenition (i.e. via PA *x, i.e. a velar fricative distinct from a uvular *խ* /x/ [χ], 4.4, 9.3.16) and underwent a subsequent fortition to the attested aspirated velar stop, i.e. *k^(u) > PA *x > *ք* *k’*. This way the reflexes of *s merged word-finally (i.e. *-Vs > PA *-Vh# > *-x, 9.1.23) with the reflexes of *k^(u). It is conceivable that PA *x in *հիմ* *him* and *ինչ* *inč’* does not show the regular reflex but has been lenited further to the attested *հ-* and even further to *Ø-* due to the same specifics of “weak” sentential prosody that may be assumed for *tū > PA *θū, lenited to *ðu and subsequently fortified to *դու* *du* ‘thou’ (cf. OE *bū* > archaic NE *thou* [ðau] ‘id.’), not ***t’u*, 9.1.17; similarly, *[-]to- > *դա* *da* ‘this (by you)’, [*ay-/d* ‘id.’, etc. (cf. NE *the* [ðə] < OE *þ-* [θ-], Goth. *ga-* *[γa-], Dutch *ge-* [χə-] < PGmc. *[χa-] < *kom-).

A few words have forms with initial *հ-* alongside forms without it: *հիմ* ~ *իմ* ‘for what?’, *հաւայ* ~ *սկայ* *hskay* ~ *skay* ‘giant’, etc. In certain instances, the initial *հ-* clearly lacks etymological justification, cf. *հալով* *halowē* < Gk. ἅλων, *հիրիկ* *hirik* < Gk. ἱρίς, *հոկտեմբեր* *hoktember* ← < Lat. *October*, *Հռոմ* *Hřom* < Lat. *Roma*. Jerejian (1953) further argues that the *h-zero* alternation not only reflects the age of the composition, cf. presumably older *լեզուատ* *lezow-at* ‘dumb’ vs. more recent *լեզուահատ* *lezow-a-hat*

‘one whose tongue is cut’, but may also be charged semantically, i.e. the forms with 𐎠-*h-* denote specificity, result, or intensity, cf. 𐎎𐎗𐎒 *ogi* ‘breath’: 𐎎𐎗𐎒𐎗 *hogi* ‘spirit’, 𐎗𐎎𐎗𐎗𐎗 *arac* ‘plague’: 𐎎𐎗𐎎𐎗𐎗𐎗 *harac* ‘scurf’, 𐎎𐎗𐎎𐎗𐎗𐎗 *Astowac* ‘God’: 𐎎𐎗𐎎𐎗𐎗𐎗𐎗 *hastowac* ‘creation’, 𐎎𐎗𐎒𐎗 *erk* ‘work’: 𐎎𐎗𐎒𐎗𐎒 *herk* ‘furrow’, 𐎗𐎎𐎗𐎒𐎗 *arbil* ‘to drink’: 𐎎𐎗𐎎𐎗𐎒𐎗 *harbil* ‘to be intoxicated’, 𐎎𐎗𐎒𐎗 *elk* ‘tepid’: 𐎎𐎗𐎒𐎗𐎒 *helk* ‘lazy’, 𐎎𐎗𐎒𐎗𐎒 *olot* ‘overflow’: 𐎎𐎗𐎒𐎗𐎒 *helet* ‘flood’, etc.

9.3.16. *lʷ x* < *(-)kh₂-: *lʷaxʷlθb-* *xayt²-e-* ‘stab’ < *kh₂ej-t- (Lat. *caedō* ‘I hew’ [*< *kh₂ej-d-*], ?Ved. *khidāti*); *gʷlʷ* *c²ax* ‘branch’ < ?*kōk-h₂- (with dissimilation) (Ved. *śākhā*, Lith. *šakà*, OCS *soxa* ‘pole’, Goth. *hoha* ‘plow’). Phonetically, *lʷ* /x/ is a uvular segment, i.e. [χ] (4.4); this realization may be understood as a result of assimilation to the following *h₂ with its pharyngeal place of articulation, i.e. *[h̠] (vel sim.). Since in PA *-kh₂- represented two segmental articulations, it was a fortis sound (perhaps at some point even a geminate). The place of articulation and the fortis nature of the sound have made it distinct from PA velar (non-geminate) *[x] from PIE *k^(u) (9.3.15).

9.3.17. *uʷlʷ sx* < (??*sk^h-) ??*sG^h-: *uʷlʷlθb-* *sxal-e-* ‘go astray’ < ?*sk^{uh}al- < ?*(s)g^{uh}al- (Ved. *skhālati* ‘stumbles’, Gk. σφάλλω ‘I make stumble’ [*-λ-j-], Lat. *fallō* [*< ?*fal-nō*] ‘I lead astray’).

9.4. Liquids

The original liquids are preserved as such even in PA unaccented final syllables (8.1), e.g. 𐎎𐎗𐎒𐎗 *hayr* ‘father’ < PA *φáthir < *p_h₂tér (Gk. πατήρ); 𐎗𐎎𐎗𐎒𐎗 *albiwr* ‘source’ < PA *(ə)ḫriuar < *b^{hr}éhi-ur̥ (Gk. φρέαρ); 𐎗𐎎𐎗𐎒 *astl* ‘star’ < PA *ástil < *h₂stél (Gk. ἀστήρ), etc.

<i>l</i> / <i>l</i>	*l- 9.4.1 *Ol- 9.4.2	<i>r</i> / <i>r</i>	*-r- 9.4.6 *Tr- 9.4.7 ?*d ^h - 9.4.8 *(-)d{u}- (see 9.1.7)
<i>l</i> / <i>ŋ</i>	*-l{C}- 9.4.3 *-ln- 9.4.4 *r{...r} 9.4.5	<i>r̄</i> / <i>ŋ</i>	*rs, *sr 9.4.9 ?*rH{V} 9.4.10 (*r{n} 9.4.11

9.4.1. *l- l-* < *l-: *l^hqb-* *liz-e-* ‘to lick’ < *léjō^h-e- (Gk. λείχω, cf. OCS *ližō*, OIr. *ligim*, Lat. *lingō* ‘id.’); *lnqaw²u-* *log-an-a-* ‘to bathe’ < *leu_h₃- (Lat. *lavō* ‘id.’, Gk. λε-λου-μένος ‘having washed’); *lnj²u* *loys* ‘light’ < *léuk-os- (Av. *raocah-* ‘id.’); *l²w-* *l(-)a-* ‘to weep’ < *leh₂- (Ved. *rāyati* ‘barks’, OCS *lajō*, Lat. *lātrō* ‘id.’).

9.4.2. *l- l-* < *Ol-: *ln²u* *low* ‘flea’ < *Blus- (Ved. *plūṣi-*, Gk. φύλλα < *psul-ja < *plus-ja, Lith. *blusà*, OCS *bluxa*, Lat. *pūl-ex* < *pusl- < *plus-, OHG *floh*, Alb. *plesht* ‘id.’); *ln²u* *low* ‘hearing; heard, known’ < *klu-tó- (Ved. *śrutá-*, Gk. κλυτός, OIr. *cloth* ‘id.’); *ln²u-* *lnow-* ‘to fill’ < *plē-n- ← *p_{l̥}-né-h₁- (Ved. *pṛṇāti* ‘fills’, Gk. πλήτο ‘got full’, Lat. *-plēvit* ‘filled’).

9.4.3. $\eta l < *{-}IC{-}$: **գեղձք** *gelj-k* ‘glands’ < *g^helǵ^h- (OCS *žlěza*, Lith. *gėležuonys* ‘id.’); **ողբ** *olb* ‘wail’ < *H₍₃₎olb^h- (Gk. ὀλοφύρομαι ‘I wail’ [?← *ὀλοφ-]; Lith. *ulbúoti* ‘call; sing’); **աղա** *alt* ‘filth’ < *(?s)ǵd- (ON *úldna* ‘to rot’ < *Hǵ-d; cf. Dutch *uilig* ‘mouldy’, MLG *ulm* ‘rotteness’; OIr. *sal* ‘dirty’ < *sal-, OHG *salō* ‘id.’ < *sal-uo-).

9.4.4. $\eta l < PA *{-}ln{-}$: **Հեղու-** *hel-u-* ‘pour; fill’ < PA *pel-nu- < *pelh₁-; **սաղ** *tal* ‘fable; poem’ < *dl̥₁-(s)ni-; **սող** *tol* ‘line, rank’ < *dol-(s)ni- (Oİc. *telja* ‘to tell’, *tal* ‘number; narration’); **Աղիւն** *Aliwn* TN < PA (?) *Alniβna, cf. Αβαλιβνα (Ptolemy; ca. 150 CE).

9.4.5. $\eta l < PA *r\{\dots r\}$: **եղբայր** *elbayr* ‘brother’ < *b^hrātēr; **աղբիւր** *albiwr* ‘spring’ < *b^hreh₁-ur.

9.4.6. $-r- -r- < *{-}r{-}$: **արաւր** *arawr* ‘plow’ < *h₂érh₃-tro-m (Gk. ἄροτρον, OIr. *arathar* ‘id.’, etc.); **Ներու** *herow* ‘last year’ < *peruti (Gk. [Dor.] πέρυτι, Ved. *parút* ‘id.’).

9.4.7. $-r- -r- < *Tr{-}$: **երեք** *erek* ‘three’ < *tréjes; **երեւի-** *erew-i-* ‘appear’ < *prep- (Gk. πρέπω ‘am manifest’); **երէց** *erēc* ‘elder’ < *prejsko- (Lat. *priscus*, Gk. πρέσβυς ‘id.’).

9.4.8. $-r- -r- < ?*{-}d^h{-}$: **գերե-** *ger-e-* ‘capture; lead into captivity’ < *ued^h-e- (Lith. *vedù* ‘I lead’, cf. YAv. *vādaia-* ‘lead’) beside **վարե-** *var-e-* ‘lead’ borrowed from Iranian, cf. Praust (2005); **այրե-** *ayr-e-* ‘burn (tr.)’ < *h₂ejd^h-e- (Gk. αἴθειν ‘id.’); $-r- -r$ [2sg. Imper.], e.g. **լուր** *low-r* ‘listen!’ < *klu-d^hi (Ved. *śrudhí*, Gk. κλῶθι ‘id.’); cf. Jasanoff (1979: 145 f.).

9.4.9. $n \bar{r} < *sr$, *rs: **ձեռն** *jeṛn* ‘hand’, *jeṛ-k* ‘id. (Npl.)’ < *ǵ^hés-r-ṛn, *ǵ^hés-r-es (Gk. χεῖρα, χεῖρες, Hitt. *kiššeran* ‘id.’); **առու** *aṛow* ‘stream’ < *sru-tí- (Ved. *srutí-*, Gk. ῥύσις, cf. OIr. *sruaim* ‘id.’); **քեռ** *keṛ* ‘sister Gsg.’ < *suešros (Goth. *swistrs* ‘id.’); **ոռք** *or-k* ‘buttocks’ < *Hors-o- (Gk. Ion. ὄρσος, Hitt. *arra-*, OHG *ars* ‘id.’); **թառամ** *taṛ-am* ‘withered’ < PA *t^hárs- < *tṛs- (Goth. *ga-þaurs-ana* [Asg.] ‘id.’).

9.4.10. $n \bar{r} < *{-}rH{-}$: **առ** *aṛ* ‘at, by, before’ < *pṛH- (e.g. Gk. παρά); **գառն** *gaṛ-n* ‘lamb’, *gaṛ-in-k* ‘id. (Npl.)’ < *uṛh₁-én, *uṛh₁-én-es (Gk. Φαρήν [Gortyn], [πολύ-]ρρηγ-ες ‘having many lambs’ [-ρρηγ- < *{-}Fρρηγ- < *{-}uṛh₁-n-]); but cf. **հարաւ** *haraw* ‘south’, if from *pṛH-uo- (OCS *prǐvŭ* ‘first’).

9.4.11. $n \bar{r} < PA *{-}r\{n\}$: **առն** *aṛ-n* [a.rən] ‘man (Gsg.)’ < PA *ánro- < *h₂ṛ-r-ós (Gk. ἀνδρός, Ved. *naráh* ‘id.’); cf. Nsg. **այր** *ayr* < *h₂nér (Gk. ἀνήρ, Av. *nā* ‘id.’); **ջերմու-** *jeṛmow-* ‘get warm’ < PA *dʒer-nu- ← ?*g^{uh}ṛ-nu- (Skt. *ghṛ-ṇo-* ‘burn’, Goth. *brinnan* ‘id.’), cf. **ջեր** *jeṛ* ‘warmth’ < *g^{uh}er-os- (Gk. θέρος, Ved. *háras-* ‘id.’).

9.5. Nasals

Original syllabic nasals are realized, as are syllabic resonants generally (10.1.11), with a preceding prop vowel, i.e. *N̥ > PA *aN̥, ultimately reflected as **-աւ-** *-am-* (9.5.2) or

-ան- -an- (9.5.5). The original non-syllabic nasals are generally preserved as such (especially medially), except: (i) immediately preceding an original labiovelar, i.e. *NK^u > PA *uK̄ (9.6.4); (ii) before a high vowel (9.6.3); (iii) immediately preceding an *s, i.e. *Ns > (PA ?*ss >) *u s* (9.3.3); (iv) word-finally (8.1).

/m/ <i>մ</i>	*-N-p/b ^h - 9.5.1	/n/ <i>ն</i>	*-N/N _o # 9.5.4
	*մ 9.5.2		*ն 9.5.5
	*(s)m 9.5.3		*(s)n 9.5.6

9.5.1. *մ m* < *-N-p/b^h-: *ամուլ amowl* ‘barren’ < PA *am-φ^ul-o- < *ն̄-pōl-o- (Gk. πῶλος ‘foal’; cf. *ուլ owl* ‘(goat) kid; fawn’ < *pōl-o-) or perhaps from *ն̄-putl-o- (Ved. *a-pútra-* ‘without sons’; Olsen 1999: 35); *ամոք amok* ‘soft; agreeable (of food)’ < *սմ-pok^u-ó- (cf. Skt. *sam-pak-va-* ‘thoroughly cooked; tender’).

9.5.2. *մ m* < PA *(a)m- < *-մ-: *ամ am* ‘year’, *ամառն amar-n* ‘summer’ < PA *սմH-ā- ← < *sem- ‘summer; season’ (Ved. *sāmā-* ‘year; season’ < *sém-eh₂-; OIr. *sam* ‘id. < *սմ-h₂-ó-).

9.5.3. *մ m* < *(s)m: *մի mi* ‘one; a(n)’ < PA *smijó- (Gk. μία ‘one [fem.]’ < *sm-ih₂-; cf. Lat. *semel* ‘once’); *մի mi* (modal negation) < *méh₁ (Ved. *mā́*, Gk. μή); *մեծ mec* ‘great’ < *meġ-h₂(-) (Ved. *máhi*, Gk. μέγα[ς] ‘id.’).

9.5.4. *-ն -n* < *-N/N_o#; original consonantal *-m# and *-n#, together with the result of vocalization of *-մ# and *-ն# merge into PA *-n# (8.1): *գառն gar-n* ‘lamb’ < PA *gārH-in- < *գրհ₁-én (Gk. ἄρην, Ved. *úrā*); *եւթն ewt-n* ‘seven’ < PA *hèp^han < *sept^h (Gk. ἑπτά). Thus, a final *-մ-m* may only reflect an original medial *-m- from a period before final syllable reduction: *եմ e-m* ‘I am’ (← PA *i-mi) < *e(h)mi < *h₁és-mi (Ved. *ásmi*, Gk. εἶμι, OCS *jesmb* ‘id.’); cf. *ġ^hi-óm (Nsg.) > PA *ġzijun > *ճիւն jiw-n* ‘snow’ (Gk. χιών ‘id.’), but *ġ^hi-m-ós (Gsg.) > PA *ġzim- > (dial.) *jm-ayr ‘snow blindness’.

9.5.5. *ն n* < PA *(-){a}n- < *ն̄: *ան- an-* (privative prefix): *անկար an-kar* ‘impossible’; *կանա-ց kan-a-c* ‘woman (GDpl.)’ (an earlier form of *կանանց kan-an-c* ‘id.’) < PA ?*gana-sk- ← *g^unh₂- (Boeot. βανά ‘id.’). The stem /kan-an-/ of the later form is no doubt imported from *կանամբք kan-am-b-k* (Ipl.), which is itself secondary (cf. *կնաւ kn-a-w* [Isg.] < *kiná-ḅi ← < *g^uneh₂-b^hi-; Ved. *gnábhīh*, OIr. *mnáib* ‘id.’). Medial *-ն -n-* may also reflect original *-m/մC-, with assimilation of the nasal to the place feature of the following consonant: *հանգիւմ hang-ċ-i-m* ‘I rest’ < PA *haŋ-gi- < *սմ-k^uih₁- (Lat. *quiescō* ‘I rest’, OCS *po-čijō* ‘id.’); cf. 9.5.1.

9.5.6. *ն n* < *(s)n: *նաւ naw* ‘ship’ < *néh₂u-s (Gk. ναῦς, Ved. *náu-*, Lat. *nāvis*, OIr. *nau*, Olc. *nór* ‘id.’); *հնոց hn-oc* ‘furnace; (lit.) fire-place’ < PA *fun- ← *péh₂-u, *ph₂-uén- (Goth. [fon], funins [Gsg.] < *p(h₂)un- [pace Kroonen 2013: 151]; Gk. πῦρ, πῦρός, Umbr. *pír* < *pūr-); *նու now* ‘daughter-in-law’ < *snus-ó- (Ved. *snusá-*, OCS *snъxa*, Gk. νούς, Lat. *nurus* ‘id.’); *նեարեք near-d-k* < *sneh₂-uḡ-t- ‘sinews’ (Gk. νεῦρον, Lat. *nervus* [< *neur-o-], Skt. *snāvan-*, Hitt. *išhunawar* ‘string; sinew’).

9.6. Glides

/y/ ɟ *-{V}t{ĩ, ē̃}- 9.6.1 *-sĩ- 9.6.2 *-n(T){ĩ, ē̃}- 9.6.3	/w/ ʌ *-n{K ^u } 9.6.4 *-t{o/r}- 9.6.5 *-p- 9.6.6 *-b ^h - 9.6.7 PA *-m{u}- 9.6.8 ?PA *-{u}m{n} 9.6.9 ?*-C{T}- 9.6.10. *-k̂{r}u, ? k̂{l} 9.6.11	[v] ʎ < *[w] *uṽ- 9.6.12 *-{o}b ^h - 9.6.13
---	--	--

9.6.1. ɟ y < *-{V}t{ĩ, ē̃}-: բայ *bay* ‘speech’ < *b^hh₂-ti (Gk. φάσις ‘id.’); -է -ē (< *-ē-ji < PA *-e-θⁱ), -այ -a-y, -ու -u (< *-ù-ji < PA *-u-θⁱ), go-y (3sg. pres. ind.) < PA *-V-θⁱ (Skt. -ti, Gk. ἐσ-τί ‘is’, OLat. *iouesa-t* ‘iūrat’); հայր *hayr* ‘father’, մայր *mayr* ‘mother’, եղբայր *elbayr* ‘brother’ < *-tēr (Lat. *pater, māter, frāter*, Ved. Asg. *pitāram, mātāram, bhrātāram* ‘id.’, Gk. πατήρ, etc.); հերու *heru* ‘last year’ < PA *qerùji < *peruti (Gk. [Dor.] πέρυτι, Ved. *parút* ‘id.’).

9.6.2. ɟ y < *-sĩ-, cf. -ոյ -o-y (GDabs.) < PA (?)*-o-ji < *-o-sjo (Gsg. *o-stem ending) (Hom. -οιο, Ved. *-asya* ‘id.’).

9.6.3. ɟ y < *-N(T){ĩ, ē̃}-, so-called “*i*-epenthesis”: այր *ayr* ‘man’ < *á^hn(i)r or *an(i)r < PA *ánir < *h₂nēr (Gk. ἀνήρ, Av. *nā* ‘id.’, Lat. *Nerō* PN); ար *ayr* ‘cave’ < PA *án(n)ir < *antēr (Gk. ἄντρον ‘id.’ ← *ántrh₂: ἄντρα); cf. Hom. ἄστῆρ (sg.): ἄστρα (coll.) → ἄστρον (sg.); de Lamberterie (1978: 243 ff.).

9.6.4. ʌ w (< PA *m^u) < *-N{K^u}-: աւժ *awc* ‘grease’ < *áũĝ- < PA *amĝ < *h₃ŋĝ^u- (Ved. *āñj-as-*, Lat. *unguen* ‘id.’, OIr. *imb* ‘butter’, OHG *ancho* ‘id.’); աւճ *awj* ‘snake’ < *áũĝ^h- < PA *ámĝ^h- < *h₂ŋĝ^{uh}-i- (Lat. *anguis*, OHG *unc*, Lith. *angis*, Ved. *áhi-*), աւճիք *awji-k* ‘collar’ as if from *h₂ŋĝ^{uh}- (cf. Aeol. ἄμωφην ‘neck’ [αὐχίην ‘id.’ < *áũĝ^hēn] < *h₂mĝ^{uh}-ēn, cf. Pronk 2010).

PA *m^u represents a shorthand for an apparently necessary intermediate stage in which the labial feature of the original labio-velar stop is reflected in the featural composition of the preceding nucleus. The traditional interpretation of this stage invokes a so-called “*u*-epenthesis” followed by the deletion of the nasal, i.e. *(V)nK^u > *(V)n^uK > *V^unK > *V^uK > *VwĶ. While this is of course conceivable, a “simpler” assimilation seems more likely. The (unordered) labial feature of the labio-velar was realized on the preceding nasal in the form of a (bi)labial approximant (or fricative) with nasal resonance phonetically similar to the lenited nasal in the history of Celtic, e.g. *dōm(h₂)-o- ‘belonging to the household’ (vrddhi derivative of *dom[h₂]- ‘house’; cf. Matasović 2009: 88f.) > OIr. *dám* [da:μ] ‘retinue; dining party’ > NIrish *dámh*, Munster [dɔ:f], Ulster [daw]), i.e. PIE *ŋK^u > PA *amK^u (assimilation) > amK (*K and *K^u merge) > aũĶ (phonetically less marked).

A notorious counter-example to this development is *անկանի-* *ank-an-i-* ‘fall’ (Goth. *sigqan* ‘sink [Intr.]’ < *séng^u-e-, LIV² 531; cf. *ընկենու-* *enke-now-* ‘hurl < *cause to fall’ < *song^u-éje-; Goth. *sagq-ja-n* ‘sink [tr.]’). For Winter (1962: 258), *անկանի-* *ank-an-i-* contains a syllabic nasal, i.e. *sŋ^u-n(é)-g^u- (LIV² *ibid.*) which would

have blocked the development. Klingenschmitt (1982: 181 f.) assumes the opposite: the development specifically involves a syllabic nasal, i.e. $*\eta K^u > *wK^u$ while $*VnK^u > *VnK$. Klingenschmitt's view is supported by $*p\acute{e}nk^ue > \text{Հինգ } hing$ 'five', while Winter's finds support in $\text{Հանգիմ } hang-\acute{c}-i-m$ 'I rest' $< *sm-k^u h_1-$. As for $\text{անկանի- } ank-an-i-$, the evidence for a labio-velar in this root is debatable: Hom. $\acute{\epsilon}\alpha\phi\theta\eta$ 'crashed down (of armor)' (Il. 13,543; 14,419), presumably $\leftarrow *ξαπτο < *\acute{\epsilon}-sng^u\text{-to}$ (cf. LIV² *ibid.*²) is formulaic and could equally well go back to $*\acute{\epsilon}-sng^u\text{-to}$ '*resounded; clang' (cf. Goth. *siggan* 'sing, resound'; LIV² 532^{1a}); Goth. *siggan* may be related to PGm. **sakk/gōn* 'to drop; sag' and analyzed as an inner-Germanic thematized nasal present, i.e. $*/se-n-k\text{-nu-}\acute{e}ntil/ \rightarrow */senk\text{-nu-e-}/ > *senkkwe-$ (Kroonen 2013: 423), which alternated with **senkwe-* (the geminate simplified in superheavy syllables) (Kroonen, email Jan. 20. 2014). Thus, Goth. *siggan* 'to sink' in and of itself may not be as strong a piece of evidence for the presence of a labio-velar in the root as previously assumed.

9.6.5. $\text{ } \text{ } w < PA * \theta^u < * -t\{o, r, ?l\}$ -, e.g. $\text{-} \text{ } w$ (aor.med.3sg. ending) $< * -t\text{-o}$, cf. *βηλε ele-w* 'was; became $< *turned$ (out)' $< * \acute{\epsilon}\text{-}k^u\text{-e-to}$ $\leftarrow *k^u[h_1\text{-}t\acute{o}$ (Gk. $\acute{\epsilon}\pi\lambda\epsilon\tau\omicron$, Alb. [OGeg] *cleh* 'id.', Lat. *coluit* 'colonized $< *went$ around'); *Հաւր hawr* 'father (Gsg.)' $< *ph_2\text{tr-}\acute{o}s$ (Gk. $\pi\alpha\tau\rho\acute{\varsigma}$, Lat. *patris* 'id. '); *արաւր arawr* 'plow' $< *h_2\acute{e}rh_3\text{-tro-m}$ (Gk. $\acute{\alpha}\rho\omicron\tau\rho\nu$, Lat. *arātrum*, OIr. *arathar* 'id. '). Original $*t$ may be continued by $\text{ } \text{ } w$ also before a lateral, if *ամուլ amowl* 'barren' reflects $*\acute{h}\text{-putlo-}$ (Ved. *a-pútra-* 'having no sons'; Olsen 1999: 235).

9.6.6. $\text{ } \text{ } w < * -p$:- *եւ ew* 'and' $< *h_1\text{epi}$ (Ved. *ápi* 'also', Gk. $\acute{\epsilon}\pi\acute{\iota}$ 'upon, in addition to', Goth. *if-tumin daga* 'on the following day' [$< *h_1\text{ep-t}\eta\text{-}?\eta$]); *եւթն ewt'n* '7' $< *sept\eta\eta$ (Gk. $\acute{\epsilon}\pi\tau\acute{\alpha}$, Ved. *saptá*, Lat. *septem* 'id. '); *քուն k'own* 'sleep' $< PA *k^h\text{ou-no-}$ $< *s\acute{u}op\text{-no-}$ (Ved. *svápna-*, Lat. *somnus, sopor*, Gk. $\acute{\upsilon}\pi\nu\omicron\varsigma$ 'id. '); *արծուի arcowi* 'eagle' $< PA *arci\acute{u}io- < *h_2\acute{r}\acute{g}ip(i)\acute{i}o-$ (Ved. *ṛjipyá-*, Av. *ərəzifīia-* 'id. ').

9.6.7. $\text{ } \text{ } w < * -b^h$:- $\text{-} \text{ } w$ (Isg. ending) $< * -V\text{-}b^h$, cf. *բայի bay-i-w* 'by a word' $< *b^h_2\text{-ti-b}^h$; *-աւոր [-a-]wor* 'carrier, carrying' $< * -b^h\text{or}\acute{o}$ -, cf. *լուսաւոր lows-a-wor* 'bright $< *bringing$ light' $< *le\acute{u}k\text{-o-b}^h\text{or}\acute{o}\text{-s}$ (cf. Lat. *lūci-fer* 'id.', Ved. *vājam-bhará-* 'bringing booty').

9.6.8. $\text{ } \text{ } w < PA * -m\{u\}$ - $< PIE *m\{\acute{u}, \acute{o}, oN\}$: *աւր awr* 'day' $< PA *ámur < *ámōr$ (Hom. $\acute{\eta}\mu\alpha\rho < *ám\text{-}r$, Myc. *a-mo-ra-ma* /ámōr-āmar/ 'day by day'); *-ուք -ow-k'* (aor. subj. 1 pl.) $< PA *-\acute{o}mu\text{-}k^h \leftarrow < ? * -o\text{-}m\acute{o}$ (Ved. *-mā*, Goth. *-ma* [1pl. Opt.] $< * -m\acute{o}/\acute{e}$, Jasanoff 1979: 136); *-իւն -iwn* (verbal abstract/action noun suffix) $< PA * -i\text{-}\mu\text{un} < ? * -e\text{-}m\acute{o}n$ (Ritter 1985).

9.6.9. $\text{ } \text{ } w < PA * \{u/\acute{o}\}m\{n\}$; i.e. dissimilation to the following $*(-)n$: *-ուն -own* ([fossilized] Med. Ptp.) $< PA *-\acute{o}\text{-}\mu\text{no-} < * -o\text{-}m(h_1)n\acute{o}$ - (Av. *-mna-*, Gk. $\acute{o}\mu\epsilon\nu\omicron\text{-}\zeta$, Ved. *-áná-*, TB *-mane*); *անուն anown* 'name' $< PA *an\acute{u}m\eta < *H_{1/3}n\acute{o}\text{-}m\eta$ (Gk. $\acute{o}\nu\omicron\mu\alpha$, Ved. *nāma-* 'id. ').

9.6.10. $\text{ } \text{ } w < ?? * -C\{T\}$:- *գիտ giwt* 'discovery' $< (?) *uid\text{-}ti$ -, cf. *git* 'id.', *եգիտ e-git* 's/he found' (Ved. *á-vit-ti-* 'poverty, lit. not finding'); *Հաւս hawt* 'flock of sheep' $< ? *p(e)h_2\text{-}d\text{-}ti$ - (MP *pāda-* 'id. ').

9.6.11. $\iota w < *-\tilde{k}\{r\}u-$, perhaps also $*-\tilde{k}\{l\}-$: մաւրուք *mawru-k'* ‘beard’ $< *smokru-$ (Ved. *śmásru*-‘id.’, Lith. *smãkras* ‘chin’); possibly արտաւար *artawsr* ‘tear’ if from PA $*arta\tilde{u}(s)ru$ (?) $< *dra\tilde{k}ru$ (with PA $*-s-$ leveled from plural արտաւարք *artasow-k'* $< PA *artas\tilde{u}a-$ $< *dra\tilde{k}u_2$ [Kortlandt 2003: 60]); possibly գիւղ *giwl'* ‘village’ if from PA $*gi\tilde{u}li-$ $< *ui\tilde{k}l\bar{i}-$ $< *ui\tilde{k}(s)l-i\tilde{h}_2-$ (Ved. *vís-* ‘settlement’, Lat. *vīcus* ‘district’, *vīlla* ‘dwelling’, OCS *vъsb* ‘village’, Goth. *weihs* ‘id.’).

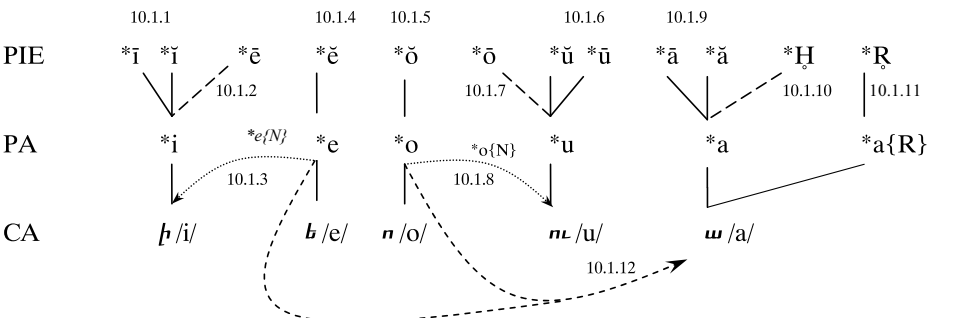
9.6.12. $\psi-$ $v-$ $< *u\bar{u}-$: վեց *vec'* ‘six’ $< PA *(h)u\bar{u}e\bar{s}^hV-$ $< *s^u\bar{u}e\bar{k}s$ (Lindeman variant) (Dor. Ἐξ , YAv. *xšuuas̄*, MW *chwech*); (h) վեր (*i*) *ver* ‘up(wards)’ $< PA *(en) (h)u\bar{u}eri$ $< *up-ér-i$ (Ved. *upár-i*, Gk. *ὑπέρ*, Lat. *[īn-]s-uper*). Otherwise, words with an initial $\psi-$ $v-$ are exclusively of foreign pedigree, except perhaps for վայ *vay* ‘woe!’ (Lat. *vae*, Goth. *vai*, Mlr. *fae*), which, presumably on account of its expressive value, is believed to have exceptionally escaped the change of $*\bar{u}- > \bar{q}-$ $g-$ (9.1.13) (cf. Godel 1982b: 9; Martirosyan 2010: 591). Alternatively, վայ *vay* is a Semitic loanword (cf. Gk. *οὐαί*, Beekes 2010: 1123).

The PA medial sequences $*-u\bar{u}-$ and $*-\bar{u}u-$ are also reflected as OA /u/ $*[u]$ (i.e. not as \bar{q} /g/, 9.1.13; cf. Eichner 1978: 148 f.): լուացա *low-ac-a* ‘washed (oneself)’ $< PA *lu\bar{u}-a-s\bar{k}-$ $\leftarrow *lu\tilde{h}_3-$ (Gk. *λού-ω*, *λού-ο-μαι* ‘id.’; cf. OIr. *lóathar* ‘basin’, ON *lauðr* ‘soap; lather’), cf. լոգա(ցա) *log-a(c-a)* ‘id.’ $< PA *lou\bar{u}-a(s\bar{k})-$ $< *leu\tilde{h}_3-$ (10.1.13).

9.6.13. $-\psi$ $-v < *-\{o\}b^h-$: մարդով *mard-o-v* ‘man (Isg.)’ $< *m\bar{r}\bar{t}\bar{o}-b^hi-$.

10. Vocalism

The quantitative contrast characteristic of the original PIE vocalism is lost. Distinct reflexes of the original mid vowels ($*\bar{e}$: $*\bar{e}$ [10.1.4, 10.1.2] and $*\bar{o}$: $*\bar{o}$; [10.1.5, 10.1.7]) imply that the contrast based on length gave way to one based on quality (or timbre) distinctions.



10.1. Monophthongs

The peripheral vowels, PIE *ī, *ū and *ā are reflected as *h* /i/ (10.1.1), *ow* /u/ (10.1.6; 3.1), and *u* /a/ (10.1.9), respectively. The original long mid vowels merge with the reflexes of the original high vowels: *ē merges with *ī into *h* /i/ (10.1.2) and *ō merges with *ū into *ow* /u/ (10.1.7). The original short mid vowels *ĕ and *ĝ are during the PA stage raised before nasals, where they too merge with *h* /i/ (10.1.3) and *ow* /u/ (10.1.8), respectively. Elsewhere, original short *ĕ and *ĝ are generally preserved as such (but cf. 10.1.12), especially in the final stressed syllables of Armenian citation forms and inflectional stems (cf. 10.1.4, 10.1.5).

10.1.1. *h* *i* < *ī: *h*₁*h*₂ *e-lik* ‘has left’ < *é-lik^u-e-t (Gk. ἔλιπε, Ved. *áricat* ‘id.’); *h*₁*h*₂ *jil* (~*h*₁*h*₂ *jil*) ‘tendon’ < *g^{uh}iH-(s)leh₂ (Lith. *gýsla* ‘vein’, SCr. *žila* ‘id.’, Lat. *fīlum* ‘thread, string’).

10.1.2. *h* *i* < *ē: *h*₁*h*₂ *sirt* ‘heart’ < *kér-d(-) (Gk. κῆρ, Ved. *hárđi*, HLuw. *zārza* ‘id.’); *h*₁*h*₂ *mi* (prohibitive particle) < *méh₁ (Gk. μή, Ved. *mā́*, Av. *mā*, Alb. *mo-* ‘id.’).

10.1.3. *h* *i* < *e{N}: *h*₁*h*₂ *hing* ‘five’ < *pénk^ue (Ved. *pāñca*, Gk. πέντε); (*h*₁*h*₂)*im* ‘to what?’ < *himu < PA *g^hemui₁ < *k^ué-sm-ō₁ (YAv. *cahmāi*, OCS *čemu* ‘id.’; cf. Umbr. *esmei* [dem. pronoun Dsg.], OPruss. *st-esmu* ‘id.’). Contrast *h*₁*h*₂ *ē-r* ‘of what?’ < PA *hēi₁o-r < *k^ué-sjo (OAv. *cahiīā*, OCS *česo* ‘id.’); cf. 10.1.8.

10.1.4. *h* *e* < *e; as a general rule in stressed syllables: *h*₁*h*₂ *het* ‘footprint’ < PA *pét-o- < *ped-ó-(m) (Ved. *padá-*, OIc. *fet* ‘id.’, Gk. πέδον ‘ground’); *h*₁*h*₂ *jerm* ‘warm’ < *g^{uh}er-mó- (Gk. θερμός ‘id.’, Thrac. Γέρμας TN ‘[*]hot springs’). For PA *e > *u* *a* cf. 10.1.12.

10.1.5. *h* *o* < *o; as a general rule in stressed syllables: *h*₁*h*₂ *otn* ‘foot (NAsg.)’ < PA *pót-an < *pód-ŋ (Asg.) (Gr. πόδ-α, Ved. *pád-am*); *h*₁*h*₂ *orb* ‘orphan’ < *Horb^h-o- (Gk. ὀρφ-ανός, Lat. *orbis* ‘childless, bereft of parents’, OCS *rabъ* ‘slave’). For PA *o > *u* *a* cf. 10.1.12.

10.1.6. *ow* *u* < *ū: *ow*₁ *now* ‘daughter-in-law’ < *nū.o < PA *hnúhoh < *snusós (Gk. νύος, Lat. *nurus*, Ved. *snušā́*); *ow*₁*ow*₂ *mow-kn* ‘mouse’ < *mús (Ved. *múh*, Gr. μῦς, OCS *mysъ*).

10.1.7. *ow* *u* < *ō: *ow*₁ *owl* ‘kid’ < PA *pūlo- < *pōlo- (Gk. πῶλος ‘foal’); *ow*₁*ow*₂ *towr* ‘gift’ < *deh₃-ro- (Gk. δῶρον, OCS *darъ*; cf. Ved. *dā-ná-*, Lat. *dō-t-* ‘gift, dowry’).

10.1.8. *ow* *u* < *o{N}: *ow*₁*ow*₂ *cown-r* ‘knee’ < *gón-u- (Gk. γόνυ, Ved. *jānu-*); *ow*₁*ow*₂ *-sown* ‘-ty’ < *-(d)kóm-t- (Gk. -κοντ-α, OIr. [*tri*]-cho ‘[thir]-ty’); the raising of original *o to *ow* *u* before a nasal must be later than the weakening of original *s (> *h > Ø, 9.3.13); *ow*₁*ow*₂ *-owm* DLsg. (pronominal ending) < *-um(m)u < PA *ōhmui₁ < *ō-sm-ō₁ (Ved. *kásmāi*, Av. *kahmāi*, OCS *komu* ‘id.’; cf. also Goth. *þamma* [< *to-sm-]); cf. 10.1.3.

10.1.9. ω a < * \tilde{a} : $\omega\eta$ *at* ‘salt’ < PA * $h\acute{a}l$ < * $s\tilde{a}l$ -s (Gk. ἅλ-ς, Lat. *sāl* ‘id.’; cf. OCS *solb*); $\mathfrak{d}\omega\mathfrak{y}\mathfrak{p}$ *mayr* ‘mother’ < PA * $m\acute{a}\theta\acute{i}r$ < * $m\acute{a}t\acute{e}r$ (Dor. $\mu\acute{\alpha}\tau\eta\rho$, Lat. *māter*, OIr. *máthir* ‘id.’).

10.1.10. ω a < * \mathfrak{H} : $\omega\rho\omega\rho$ *arawr* ‘plow’ < PA * $ar\acute{a}\theta\acute{r}o$ - < * $h_2\acute{e}r\acute{h}_3$ -tro-m (Gk. ἄροτρον, OIr. *arathar* ‘id.’); $\omega\mathfrak{y}\mathfrak{p}$ *ayr* ‘man’ < PA * $\acute{a}n\acute{i}r$ < * $h_2n\acute{e}r$ (Gk. ἀνὴρ, Phryg. *anar*, Ved. *sū-nāra*-).

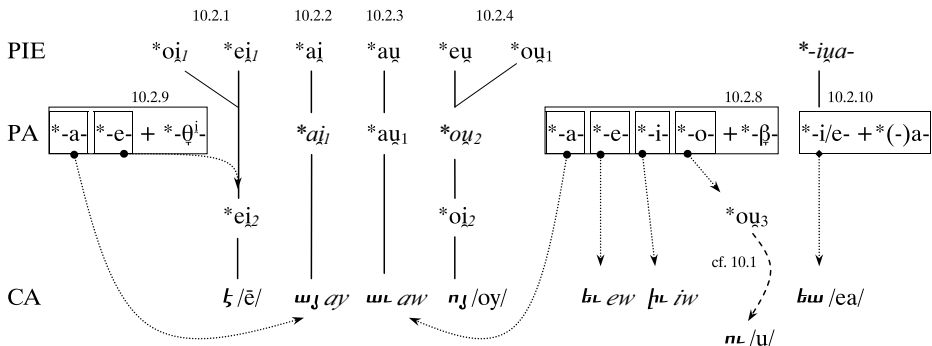
10.1.11. ω $a\{R\}$ < * \mathfrak{R} : $\mathfrak{d}\omega\rho\eta$ *mard* ‘human’ < PA * $m\acute{a}r\acute{d}$ -o- < * $m\acute{r}$ -tó- (Ved. *mṛ-tá*- ‘*having died > dead,’ Av. *mərəta*, Gk. βροτός ‘a mortal’, OCS *mṛtvь*, OIr. *marb* ‘dead’). For cases where * \mathfrak{R} = * \mathfrak{N} , cf. 9.5.2, 9.5.5.

10.1.12. In some cases, Armenian shows ω /a/ for an expected */e/ or */o/. It seems that unstressed PA *o was lowered to ω /a/ in open initial syllables (Grammont 1918: 223 f.; Kortlandt 1983: 10 [= 2003: 40]; Morani 1994); e.g. $\omega\mathfrak{h}\mathfrak{p}$ *alík* ‘grey hair’ < PA * $[\varphi.o.li]$ (Gk. πολίος, Lith. *pálvas* ‘pale’; cf. Ved. *palitá*- ‘id.’ < *pel-). However, in one notable instance this phenomenon seems to have occurred in stressed syllables: $\omega\mathfrak{z}\mathfrak{p}$ *ač-k* ‘eyes [Npl.]’ from PA * $[\acute{o}.ki]$ ‘id. [NAdu.]’ (Gk. ὄσσε, OCS *oči*, Lith. *aki* ‘id.’ < * h_3ok^u -ih₁). Ambiguous is $\omega\mathfrak{u}\mathfrak{p}$ *as-r* ‘fleece’, which could be from PA * $[(\varphi)\acute{o}.k\acute{u}]$ (Gk. πόκος ‘id.’, Myc. *po-ka* ‘shorn wool’); however, Greek also shows an e-grade form (see below).

It has been also proposed that *e is lowered to ω /a/ when the following syllable contains /u/ (Meillet 1936: 55; de Lamberterie 1978: 271; Ravnæs 1991: 13; Clackson 1994: 126). The most notable examples are $\mathfrak{v}\mathfrak{w}\mathfrak{d}\mathfrak{u}\mathfrak{n}\mathfrak{l}\mathfrak{v}$ *vat-sown* ‘60’ < * $v\acute{e}t^h$ -sun-i- (Gk. ἐξήκοντα, Lat. *sexāgintā* ‘id.’), cf. $\mathfrak{v}\mathfrak{h}\mathfrak{p}$ *vec* ‘6’ and $\mathfrak{v}\mathfrak{h}\mathfrak{z}\mathfrak{u}\mathfrak{w}\mathfrak{u}\mathfrak{w}$ *veš-tasan* ‘16’ from a Lindeman variant * $s^u\acute{e}k^s$ -; $\mathfrak{d}\omega\mathfrak{h}\mathfrak{p}$ *cał-r* ‘laughter’ < PA * $t^s\acute{e}lHu$ -r < * $g\acute{e}lh_2$ -ōs- (Gk. γέλως); cf., $\omega\mathfrak{u}\mathfrak{p}$ *as-r* ‘fleece’ if, alternatively, from PA * $[(\varphi)\acute{o}.k\acute{u}]$ (Gk. πέκος ‘id.’ [cf. πόκος above], Lat. *pecus* ‘sheep’ < * $peku$ -). However, there are clear counterexamples to this rule: $\mathfrak{u}\mathfrak{h}\mathfrak{b}\mathfrak{u}\mathfrak{n}\mathfrak{l}\mathfrak{p}$ *skesur* ‘mother-in-law’, *henum* ‘weave’ (beside zero-grade *hanum* ‘id.’, Lith. *pinti* ‘pleat’); but the most notable of these is the frozen adverb $\mathfrak{z}\mathfrak{h}\mathfrak{r}\mathfrak{u}$ *herow* ‘last year’ < * $peruti$ (Dor. $\pi\acute{e}p\upsilon\tau\iota$, Ved. *parút*, OIc. *i fjorð*).

10.1.13. In addition, the change of *e to ω conditioned by a following * \mathfrak{u} and a [+back] vowel (frequently a prop vowel from the following syllabic resonant or a laryngeal) seems to be regular, i.e. * $e\mathfrak{u}$ - $\mathfrak{H}/\mathfrak{R}$ - > PA * $o\mathfrak{u}\mathfrak{a}$ - (or PA * $^?e\mathfrak{y}^u\mathfrak{a}$ - > * $o\mathfrak{y}^u\mathfrak{a}$ -) > $-\mathfrak{n}\mathfrak{q}\mathfrak{w}$ -*oga*-, e.g. $\mathfrak{r}\mathfrak{n}\mathfrak{q}\mathfrak{w}\mathfrak{g}\mathfrak{w}$ - *log-ac-a* ‘washed (oneself)’ < PA * $lo\mathfrak{u}\mathfrak{a}$ -(sk- ←) < * $le\mathfrak{u}\mathfrak{h}_3$ - (Lat. *lav-ō* < * $le\mathfrak{u}\mathfrak{h}_3$ -e/o-), cf. $\mathfrak{r}\mathfrak{n}\mathfrak{w}\mathfrak{g}\mathfrak{w}$ - *low-ac-a* ‘id.’ (9.6.12); $\mathfrak{z}\mathfrak{n}\mathfrak{q}\mathfrak{w}\mathfrak{u}\mathfrak{w}$ *čog-a-n* ‘they went’ < PA * $\mathfrak{y}^h\acute{e}\mathfrak{u}\mathfrak{a}\mathfrak{n}$ < * $k^u\acute{i}\acute{e}\mathfrak{u}$ -nt-o (Ved. *cyávante* ‘id.’); cf. $\mathfrak{z}\mathfrak{n}\mathfrak{l}\mathfrak{b}$ - *čow-e* ‘id.’ < * $k^u\acute{i}\acute{e}\mathfrak{u}$ -/* $k^u\acute{i}\mathfrak{u}$ - (Ved. *cyávam* ‘I shall move’).

10.2. Diphthongs



The primary, i.e. PIE, diphthongs $*ai$ and $*au$ are preserved as $\omega_{\text{J}} ay$ and $\omega_{\text{L}} aw$ into the CA period (10.2.2; 10.2.3). Secondary $\omega_{\text{J}} ay$ and $\omega_{\text{L}} aw$ arise as a result of so-called *i*-epenthesis and *u*-epenthesis, respectively (cf. 10.2.5; 10.2.6). As early as ca. 800, $\omega_{\text{L}} aw$ [au] monophthongizes before a consonant (Weitenberg 1996: 95). Beginning with ca. 1100, the monophthong thus formed is represented by a new character $\circ \bar{o}$ (3.5) and in the traditional pronunciation merges with $n /o/$ everywhere but in word-initial position: $n\text{r}$ [vɔɾ] ‘who’ (relative pronoun) but $o\text{r}$ (= $\omega_{\text{L}}\text{r}$) [ɔɾ] ‘day’. There are apparent exceptions to the monophthongization which are subject to specific explanations, such as $\omega_{\text{r}}\eta_{\text{L}}\omega_{\text{L}}\text{ɬ}$ *alawni* ‘dove’ and $\text{ɬ}\omega_{\text{L}}\text{ɬ}$ *nawt* ‘crude oil’, traditionally pronounced [aβavnɪ] and [naft^h] (not $**[a\beta\alpha\text{v}\text{ɔ}\text{ni}]$ and $**[n\alpha\text{t}^{\text{h}}]$), respectively. The latter is a loanword showing post-CA $\omega_{\text{L}} aw$ (cf. Farsi *naft*, Gk. *váφθα*), the former represents OA $*[a.la.w\bar{o}.ni]$ (10.2.8), reflected in dial. $*a\text{v}\bar{\alpha}\text{v}\bar{\alpha}\text{ni}$ ‘id.’ with syncope of medial unaccented *-a-* which postdated the preconsonantal monophthongization (ca. 1100; cf. Martirosyan 2010: 29).

PIE $*ei$ and $*oi$ merge into the secondary PA $*ei_2$. This diphthong surfaces as a monosegmental $\text{ɬ} / \bar{e} /$, but synchronically still behaves like an underlying diphthong (3.2). PIE $*eu$ and $*ou_1$ merge into the diphthong $\text{ɲ}_{\text{J}} / \text{oy} /$ (presumably via secondary $*ou_2$), traditionally pronounced [u̠] (4.3). New diphthongoidal $b_{\text{L}} ew$, $\text{ɬ}_{\text{L}} iw$, $\omega_{\text{L}} aw$ and $n_{\text{L}} ow$ (= $*ou_3$) arise as a result of the spirantization of $*b^{\text{h}}$ and $*p$ (> PA $*\beta > w > \text{L } w$, traditionally [v]) (10.2.8). The diachronically tertiary diphthong $*ou_3$ subsequently merges with the monophthongal $n_{\text{L}} ow /u/$.

The only “true” diphthongs of CA are $\text{ɬ} / \bar{e} /$, $\text{ɲ}_{\text{J}} / \text{oy} /$, and $b_{\text{L}} / ea /$. The phonological value of the other traditional “diphthongs” is at best ambiguous: prevocally, the off-glide behaves as an onset of the following syllable, e.g. $\omega_{\text{L}}\text{b}_{\text{L}}\text{ɬ}$ *aweli* [a.vɛ.'li] ‘more’, $\omega_{\text{J}}\text{ɲ}$ *ayo* [a.jɔ] ‘yes’, $\text{ɬ}_{\text{r}}\text{ɬ}_{\text{L}}\text{ɬ}$ *erewim* ‘I appear’ [jɛ.rɛ.vim]; $\text{ɬ}_{\text{L}}\text{ɬ}_{\text{L}}\text{ɬ}$ *iwik* [i.vik^h] ‘any (Ipl.)’, $\text{ɬ}_{\text{r}}\text{ɬ}_{\text{L}}\text{ɬ}$ *govem* ‘I praise’ [gɔ.vɛm], etc.; word-finally, the value is ambiguous, but the off-glide could be understood as the coda of the same syllable: $\text{ɬ}\omega_{\text{L}}\text{ɬ}$ *naw* [nav] ‘ship’, especially since final $\omega_{\text{L}} aw$ never monophthongizes into post-CA $\circ \bar{o}$. Taking into consideration the insight of Godel (1975: 9) that VwC and VjC are phonologically indistinct from VrC and VnC , the segmental sequence $\omega_{\text{L}} aw$ must have been phonologized as a diphthong at a period from CA to post-CA prior to its monophthongization. The stipulation “before a consonant” is gratuitous, since early-post-CA $\omega_{\text{L}} aw$ behaves as a diphthong

exclusively in this environment. A parallel fate later befell the dialectal (even secondary) **այ** *ay*, e.g. CA **արժան** *aržan* ‘worthy’ > **ayžan* > dial. *ēžan* ‘cheap’ (Feydit 1982: 52).

Finally, instances in which **-ւ -w** “diphthongs” appear as reflexes of PA obstruents before what seems to be the PA avatar of the PIE abstract noun suffix */-ti-/ (9.6.10) pose great difficulties both phonetically and chronologically (cf. Winter 1962: 261; Clackson 1994: 40, 155; Martirosyan 2010: 732 f.). The glide in this position is the product of the PA syllable coda, which seems to have undergone lenition. Factoring in subsequent devoicing of *D to T (9), the laryngeal features of the following dental suffix correspond to those of the original precursors of **ւ w**: **գիւտ** *giwt* [gʲiut] ‘finding’ < **[γiβ̥.ti]* < PA **γ^hid-di-* (PIE *√uid*; Ved. *á-vit-ti-* ‘not finding’), cf. **եգիտ** *e-git* ‘(s)he found’ < **é-uid-e-t* (Ved. *ávidat* ‘id.’); **արաւտ** *arawt* ‘pasture’ < PA **θradz-di-*, cf. **արածե-արաւ-ե-** ‘graze’ < **tréh₂g̥-e-* (cf. Gk. *τρώγω* ‘I chew’ [*← *troh₂g̥-*], *τράγος* ‘he-goat’); **բուտ** *bowt* [but] ‘fodder’ < **[b̥uβ̥.ti]* (4.6) < PA **ḫudz-di-* (PIE *√b^hug*; cf. Skt. *bhukti-* ‘consumption’), cf. **բուծանե** *bowc-an-e-* ‘nourish; fatten’ < **b^heug-* (9.2.2); **ծանալի** *canawt* ‘known; acquaintance’ < **[tsa(n):naβ̥.t^hi]* < PA **tsanHaβ̥^h-t^hi-*, cf. **ծանալե-ճանաճ-ե-** ‘know’ (< **ts...β̥^h*; cf. aor. **ծանեա-** *cane-a-* ‘knew’) < **ǵ^hh₃-sk-ǵé-*. It should be noted, however, that the developments presupposed by this scenario contradict the general IE tendency (a notable exception is Bartholomae’s Law) for voicing assimilation to be regressive.

10.2.1. **է ē** < (PA **e₁₂* <) **e₁₁*, **o₁₁*: **դէզ** *dēz* ‘heap’ < **d^héiǵ^h-os-* or **d^hóǵ^h-o-* (OAv. [*pairi-*]*daēza-* ‘enclosure’, Gk. *τείχος*, *τοιχος* ‘wall’, Goth. *daigs* ‘dough’); **դէտ** *gēt* ‘skillful, knowing; sorcerer’ < **u₀ǵid-ā-* (cf. Ved. *véda* ‘I know’, Gk. *οἶδα*, Goth. *wait*, OCS *vedě* ‘id.’ < **u₀ǵid-h₂e*).

10.2.2. **այ** *ay* < **ā_i*: **այծ** *ayc* ‘goat’ < **h₂eǵ(-ih₂)-* (Gk. *αἴζ*, Gk. Lacon. *αἴζα*); **սայր** *tayg-r* ‘son-in-law’ < **deh₂i-uēř* (Gr. *δῆῖρ*, Ved. *devár-*, Lith. *dieveris* ‘id.’).

10.2.3. **աւ** *aw* < **ā_u*: **աւիտ** *awt(-oc)* ‘garment’ < PA **a_u-to-* < **h₂e_u(H)-* (Lith. *aūtas* ‘foot-cloth’, Latv. *āuts* ‘cloth’); **աւիտ** *awt* ‘lodging’ < PA **a_u-ti-*, cf. **աղաւ** *ag-a-w* ‘spent the night’ ← **áγ^hi* < PA **a_uist* < **é-h₂uēř(-s)-t* (Ved. *ávātsīt* ‘id.’, Hom. *ἄεσα* ‘I spent the night’).

10.2.4. **ոյ** *oy* < (**o₁* < PA **ou₂* <) **e_u*, **ou₁*: **լոյս** *loys* ‘light’ < PA **lóu₀ko-* < **léuk-os-* (OP *raučah-* ‘day’; cf. Ved. *róka-*, Gk. *λευκός* ‘white, bright’, Lat. *lux*, OCS *luča* ‘ray’); **բոյճ-բոյց** ‘feeding’ < PA **b^hou₀g-o-* < **b^hou₀g-o-* (Ved. *bhóga-* ‘delight’); **մոյլ** *joyl* ‘molten (mass)’ < **ǵ^he_u(-t)lo-* (Ved. *hotrá-*, Av. *zaoθra-* ‘libation’). Apart from the merger of the primary diphthongs **e_u* and **ou₁*, PA **ou₂* may reflect the result of inner-Armenian contractions: **քոյր** *k’oyr* ‘sister’ < **xo₁r* < PA **h^hu^hehur* < **suésōr* (Ved. *svásā*, Lith. *sesuō*, OIr. *siur*, Lat. *soror* ‘id.’); **ունի-** *own-i-* ‘have; capture’ < **o₁n-* < PA **hehun-* < **se-sónh₂-* (Ved. *sasána* ‘has obtained’, OHG *sann* ‘has striven’; cf. **ողջոյն** *oljoyn* ‘greetings!’ ← *olj* ‘safe, hale’ + **oyn*, the imperative of an intransitive ‘have’, de Lamberterie 2005: 338). The assumption of an intermediary **o₁* is based on compounded demonstrative pronouns such as **նոյն** *no-yn* ‘the same’ < PA **no-* + **-in*, cf. **ինքն** ‘he himself’, etc. (*inkⁿ + in* ‘the self-same s/he, it’).

10.2.5. Secondary **այ** *ay*, so-called “*i*-epenthesis”: This phenomenon is perhaps best conceptualized as a transfer of the palatal feature over a sonorant consonant. The trigger seems to be the presence of **ա** in the preceding syllable: **այր** *ayr* ‘man’ < *^hajir < PA *anir < *^h2nēr (Gk. ἀνῆρ ‘id.’); **այլ** *ajl* (~**այլ** *ajl*) ‘other; but’ (< ?*ajlo < ?*álllo) < PA *aljo- < *^h2él-jo- (Gk. ἄλλος, Lat. *alius*, OIr. *aile* ‘other’, cf. Gk. ἀλλά ‘but’, Goth. *alja* ‘id.’); **սայր** *sayr* ‘edge’ < *sári < PA *sari- < *^h3-ri- (Lat. *cō-ti-* ‘whetstone’, *ca-tu-s* ‘clever < *sharp[ened]’, Ved. *śi-tá-* ‘sharpened’, OIr. *cath* ‘wise’, OE *hān* ‘[whet]stone, hone’ [^hke^h3-i-]), cf. also inner-Armenian morphophonemic variant **սուր** *sowr* ‘sharp’ < *^hk{e/o}h3-ro-; **փայլ** *phayl* ‘splendor’ < *^hajl-i- < PA *^hal-i- < *(s)ph^hIH-(i) (?Skt. *sphuliṅga-* ‘spark’, ?(-)*sphu{l/r}i-ta-* ‘flashed; *erschienen*’), cf. in any case inner-Armenian morphophonemic variant **փալիփ(աղ)իմ** *phal-p’(al)-i-m* ‘I glitter’.

Epenthesis has also been posited if the trigger was any low or back vowel. The relevant etyma for *o* are: **թոյլ** *toyl* ‘permission’ < ?*tol^h2-i-; cf. **թողու-** *tōlow-* < ?*tolnu- ← *t^hne-h2- ‘allow’ (Klingenschmitt 1982: 243); **բոյլ** *boyl* ‘assembly’ < *^holi-, cf. **բոլոր** *boror* ‘all’. However, the epenthesis did not materialize in several *i*-stems: **բան** *ban* ‘word’ < *^heh2-ni- (Oic. *bón* ‘request’), **սալ** *sal* ‘anvil’ < *^h3-li- (Skt. *śi-lā* ‘rock’), or in other potentially qualifying forms: **անիւ** *aniw* ‘wheel’ < *^h3nē^h-o/eh2- (?Ved. *nābh-i-* ‘nave’, cf. Gk. ὀμφαλός ‘navel’ < *^h3nb^h-l-).

10.2.6. Secondary **աւ** *aw*, so-called “*u*-epenthesis”: **աւճ** *awj* ‘snake’ < *^hu^gh-i- < PA *a^hg^h-i- < *^h2-^hg^{uh}-i- (OHG *unc*, Lat. *anguis*, Lith. *angis*, cf. Ved. *áhi-*, Gk. ὄφις [^h < *Hóg^{uh}-i-]) (9.6.4); **արտաւար** *artaw-s-r* ‘tear’ < *^hdra^hku-; cf. **արտաւարք** *artasow-k’* ‘id. (Npl.)’ (Gk. δάκρυ, Ved. *ásru-*, TB *akrūna*) (9.6.11).

10.2.7. Secondary **է** *ē* < PA **e* before palatals: **մէջ** *měj* ‘middle’ < PA *[médʒ:ɔ] < *^hméd^h-jo- (Ved. *mádhya-*, Gk. μέσσος, Goth. *midjis* ‘id.’; 9.2.10); **էջ** *ěš* ‘donkey’ < PA *eš(š)o- < *(h)ék^hmos (Ved. *ásva-* ‘horse’, Gk. ἵππος, Lat. *equus*; 9.3.7); **գիշեր** *gišer* ‘night’ < PA *geš(š)éro- < *^hek^hspero- (vel sim.)(Gk. ἕσπερος, Lat. *vesper*, OCS *večern* ‘id.’) (9.3.8); (cf. Pedersen 1905: 205 [= 1982: 67]; Bonfante 1937: 27; de Lamberterie 1978: 264–266).

10.2.8. Tertiary **աւ** *-aw-*, **եւ** *-ew-*, **իւ** *-iw-*, **ու/ով** *-ow/ov-* from *-Vβ- (which apparently results from lenited reflexes of PA *-b- [< *-b^h-, 9.6.7.]), PA *-φ- (< *-p-, 9.6.6), and PA *-θ^u- (< *t{*o*, *r*, ?}l, 9.6.5):

աւ *aw* < *-aβ-: **տաւն** *tawn* ‘feast’ < *táβn-o- < PA *taφ-no- > *^hd^h2p-nó- (Lat. *daps* ‘sacrificial meal’, *damnum* ‘expense’ [^h < *^hd^h2p-nó-], ON *tafn* ‘victim’; cf. Gk. δάπτω ‘I devour’); **աղաւրի** *alawr-i* ‘mill(er)’ < PA *[Hal.Ha.θri.ja] < *^h2lh1-tr-ij^h2 (Gk. ἄλετρις ‘id.’ [f.]); possibly **աղաւնի** *alawn-i* ‘dove’ < ?PA *[qal.Ha.βu:n-ija-] ← < ?*p^hh2-b^h-ōn- (cf. at some remove Lat. *palumbēs* ‘wood-pigeon’).

եւ *ew* < *-eβ-: **եւթն** *ewt’n* ‘seven’ < PA *hèφ^han < *sept^hn (Gk. ἑπτά, Lat. *septem*); **եղեւ** *el-e-w* ‘happened; turned out’ < PA *e-xl-e-θ^uo < *é-k^hl-e-to (Gk. ἐπλετο ‘id.’).

ու /u/ < *ou³ < PA *-oβ-: **քուն** *k’own* ‘sleep’ < *s^huop-no-; **ութ** *owt’* ‘eight’ < PA *optō ← *ok^htō; cf. Elean Gk. ὀπῶ for ὀκτώ). The monophthongization and the subsequent merger of *ou³ with */u/ likely postdates the pretonic reduction of high vowels (6.3). This is seen in relic forms with unreduced, non-alternating pretonic **ու** *ow* which continues *ou³: **քունեմ** *k’own-em* [k^hu:nem] ‘I sleep’, synchronically analyzable as a denom. of **քուն** *k’own* [k^hun] ‘sleep’ < *k^hou³n-o- < *s^huop-no- (Ved. *svápna-* ‘id.’);

contrast գրեմ *grem* [gə:rɛm] ‘I write’, denom. of գիր *gir* ‘letter; writing; manuscript’; դլամ *dlam* [də:lɑm] ‘I cease’, denom. of դուլ *dowl* ‘pause; truce’. The pretonic reduction in the noun քուն *k’own* [kʰun] itself (: oblique քնոյ *k’n-o-y*), which is in this analysis historically irregular, is readily explained by the mechanics of the synchronically highly productive process of vocalic alternations, which involve pretonic reductions of paradigmatically de-stressed vowels (i.e., the alternations are the result of a derived environment effect; cf. Khanjian 2009: ուրախ *urax* [u:rɑχ] ‘happy’ → ուրախանա- *owrax-an-a-* ‘re-joice’, ուրախութիւն *owrax-owt’-iwn* ‘happiness’, not **[ə.rɑχ- ɔ]). Thus, once */ou₃/ (*[ɔβ] ~ *[ou] vel sim.) became phonetically realized as *[u(ɔ)], and its reflex was subject to alternations involving shifts of stress, it was reanalyzed and fully merged with */u/; i.e. */kʰou₃n-o-/ → *[kʰun]: *[kʰu.nɔj] (< PA *[hʷɔφ.nóh.jo] < *suópn-o-sjo; Ved. *svápnasya* ‘id.’) was leveled to */kʰun-o-/ → քուն *k’own*: քնոյ *k’noy* [kʰə.nɔ] ‘sleep (GDAbsg.)’. However, the monophthongized */ou₃/ was never destressed in the verbal paradigm (since no form of the paradigm carries stress on the root), and its reflexes (i.e. *[ou] >*[u]) thus technically merged with (the reduced variant of) /oy/. This divorce was no doubt aided by mismatches in semantics as well, since the noun քուն *k’own* means only ‘sleep’, but the verb քունեմ *k’ownem* at some point also acquired the meaning ‘copulate’ (presumably based on the euphemistic ‘sleep with’). The linguistic result of this is that the verb քունե- *k’own-e-* is analogically reduced on its way to modern Armenian, becoming քնե- *k’n-e-*, precisely in the meaning ‘to sleep’, because of its morpho-semantic connection with the nominal քուն /kʰun-/ ‘sleep’. However, its phonologically regular etymological doublet քունե- *k’un-e-* has only the specialized obscene meaning ‘*futuere*’.

10.2.9. Tertiary -այ- *-ay-*, -է- *-ē-* < PA *-V- +*-θⁱ- (< *t{i, e}, 9.6.1):

այ *ay* < PA *-aθⁱ-: մայր *mayr* ‘mother’ < PA *maθⁱr < *mátēr (Gk. μήτηρ, Lat. *māter*, OIr. *máthir*, TB *mācer* ‘id.’); կայ *kay* ‘station, place’ < *g^uh₂-tí- (Gk. βάσις ‘base’ could in principle represent both *g^uh₂-tí- and *g^uh₃-tí-).

է *ē* < *e_i < PA *-eθⁱ-: բերէ *ber-ē* ‘carries’ < *b^{er}-e_i < PA *b^{er}-eθⁱ < *b^hér-e-ti (Ved. *bhárati*, OCS *beretǔ*); -է *-ē* (Absg. ending) < ?*éti (Gk. ἔτι ‘also; further’, Ved. *áti* ‘over’).

10.2.10. եա *ea* < PA *-i(-)a- and *-e(-)a-: կեա- *kea-* ‘live’ < *g^uéih₃-C- or ?*g^ui_h3-C- (Hom. βέομαι ‘I shall live’ < *g^uéih₃-e- (aor. subj.); cf. Ved. *jivá-*, Gk. ζῶς, OCS *živǔ* ‘alive’ < *g^ui_h3-ǔó-); -եա- *-e(-)ac-* (aor. formant): սիրեաց *sir-e-ac* ‘love’-them.-aor.(-3sg.) ‘(s)he loved’.

11. Abbreviations

Languages:

Alb.	Albanian	Fr.	French
CA	Classical Armenian	Gk.	Greek
CLuw.	Cuneiform Luwian	Goth.	Gothic
Dor.	Doric Greek	Hitt.	Hittite

Hom.	Homeric Greek	OHG	Old High German
Lat.	Latin	OIc.	Old Icelandic
Latv.	Latvian	OIr.	Old Irish
Lith.	Lithuanian	OIran.	Old Iranian
MIran.	Middle Iranian	ON	Old Norse
MLG	Middle Low German	OP	Old Persian
MP	Middle Persian	Parth.	Parthian
MParth.	Middle Parthian	PGm.	Proto-Germanic
Myc.	Mycenaean Greek	PIE	Proto-Indo-European
NEA	Modern Eastern Armenian	RCS	Russian Church Slavic
OA	Old Armenian	Skt.	Sanskrit
OAv.	Old Avestan	TB	Tocharian B
OCS	Old Church Slavonic	Ved.	Vedic
OE	Old English	YAv.	Young Avestan
OFr.	Old French		

Segmental phonology:

C	consonant	S	sibilant
V	vowel	F	fricative
U [+high]	vowel	R	resonant
T	voiceless stop,	L	liquid
T ^h	voiceless aspirated stop	G	glide
D	voiced stop	N	nasal
D ^h	voiced breathy stop	B	labial
Ḍ	murmured stop	K	velar
P	stop	K ^u	labio-velar
O	obstruent	[±cont]	feature [continuant]

Historical phonology:

φ/β	voiceless/voiced bilabial fricative (or approximant)	θ ^u	labialized voiceless alveolar approximant
μ	nasalized bilabial fricative	◌̤	consonantal or semi-vocalic segment
θ/ð	voiceless/voiced dental (or alveolar) fricative (or approximant)	◌̇	(PA) intensity stress
θ ⁱ /ð ⁱ	palatalized voiceless/voiced alveolar approximant	◌̈	(PIE) pitch accent.

Morphology:

A	accusative	def.	definite article
Ab	ablative	denom.	denominative
act.	active	dimin.	diminutive
aor.	aorist	du.	dual
caus.	causative	G	genitive
D	dative	I	instrumental

impf.	imperfect	red.	reduplication
inf.	infinitive	sg.	singular
impv.	imperative	subj.	subjunctive
iter.	iterative	them.	thematic (used both in the traditional Indo-Europeanist sense to designate specifically the vowel *e/o as well as in the more general sense ‘stem vowel’)
L	locative		
med.	medio-passive		
N	nominative		
pl.	plural		
pres.	present		

Other:

→	“serves as a stem of” or “is phonologically realized as”	> →/← <	“becomes/is derived from by a combination of phonological and grammatical change”
←	“is derived from” or “is a phonological realization of”	dial.	dialectal
>/<	“becomes/is derived from by regular phonological change”	mss.	manuscripts
		PN	personal name
		TN	toponym

12. References

- Adjarian, Hrachia
1909 *Classification des Dialectes Arméniens*. Paris: Champion.
- Allen, William Sidney
1951 Notes on the phonetics of an Eastern Armenian speaker. *Transactions of the Philological Society* 1950: 180–206.
- Beekes, Robert S. P.
2003 Historical Phonology of Classical Armenian. In: Frederik Kortlandt. *Armeniaca: Comparative Notes*. With an appendix on the historical phonology of Classical Armenian by Robert S. P. Beekes. Ann Arbor, MI: Caravan, 133–211.
- Beekes, Robert S. P. with the assistance of Lucien van Beek
2010 *Etymological Dictionary of Greek*. 2 vols. Leiden: Brill.
- Benveniste, Émile
1959 Sur la phonétique et la syntaxe de l’arménien classique. *Bulletin de la Société de linguistique de Paris* 44: 46–68.
- Benveniste, Émile
1966 *Problèmes de linguistique générale*. Paris: Gallimard.
- Bonfante, Giuliano
1937 Les isoglosses gréco-arméniennes. I. Faits phonétiques. In: Hjelmslev, Louis (ed.). *Mélanges linguistiques offerts à M. Holger Pedersen, à l’occasion de son soixante-dixième anniversaire, 7 avril 1937*. Copenhagen: Universitets forlaget i Aarhus, Levin & Munksgaard, 15–33.
- Byrd, Andrew Miles
2015 *The Indo-European Syllable*. Leiden: Brill.
- Čaləxean, Vrt’anēs H. and Arsēn Aytənean
1885 *K’erakanut’iwn haykaznean lezui* [Grammar of the Armenian Language]. Vienna: Mk’it’arean Tparan.

- Clackson, James
1994 *The Linguistic Relationship between Armenian and Greek*. (Publications of the Philological Society 30). Oxford: Blackwell.
- DeLisi, Jessica
2013 Feature metathesis and the change of PIE *d̥u to Classical Armenian *-rk-*. *Diachronica* 30: 469–491.
- Derksen, Rick
2008 *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden: Brill.
- Dressler, Wolfgang
1976 Das Altarmenische und die Phonologietheorie. *Handēs Amsōreay* 90: 301–318.
- Eichner, Heiner
1978 Die urindogermanische Wurzel *H₂reu- ‘hell machen’. *Die Sprache* 24: 144–162.
- Feydit, Frédéric
1982 *Considérations sur l’alphabet de Saint Mesrop et recherches sur la phonétique de l’arménien*. Vienna: Mechitharisten-Buchdruckerei.
- Fleming, Harold C.
2000 Glottalization in Eastern Armenian. *Journal of Indo-European Studies* 28: 155–196.
- Fortson, Benjamin W. IV.
2010 *Indo-European language and culture. An introduction*. 2nd edition. Chichester, UK: Wiley-Blackwell.
- Garrett, Andrew
1998 Adjarian’s law, the glottalic theory, and the position of Armenian. In: Benjamin K. Bergen, Madelaine C. Plauché, and Ashlee C. Bailey (eds.), *Proceedings of the 24th annual meeting of the Berkeley Linguistics Society (February 14–16, 1998): Special session on Indo-European subgrouping and internal relations (February 14, 1998)*. Berkeley: Berkeley Linguistics Society, 12–23.
- Godel, Robert
1965 Les origines de la conjugaison arménienne. *Revue des Études Arméniennes* 2: 21–41.
- Godel, Robert
1975 *An introduction to the study of Classical Armenian*. Wiesbaden: Reichert.
- Godel, Robert
1982a *Linguistique arménienne: études diachroniques*. Paris: Fondation des frères Ghoukasiantz.
- Godel, Robert
1982b Une ‘loi phonétique’ bien difficile à énoncer: *w > arm. w (v) / g / zéro. *Revue des études arméniennes* 16: 9–16.
- Grammont, Maurice
1918 Notes de phonétique générale. Arménien classique. *Mémoires de la Société de linguistique de Paris* 20: 213–253.
- Greppin, John A. C.
1973 *Initial vowel and aspiration in Classical Armenian*. Vienna: Mechitharisten-Buchdruckerei.
- Greppin, John A. C.
1980 The origin of Arm. z from IE *gh and *dh. In: Greppin (ed.), 131–137.
- Greppin, John A. C. (ed.)
1980 *Proceedings of the first international conference on Armenian linguistics*. Delmar, NY: Caravan Books.
- Hacopian, Narineh
2003 A three-way VOT contrast in final position: data from Armenian. *Journal of the International Phonetic Association* 33: 51–80.

- Hamp, Eric P.
1988 On the essentiality of the Armenian nasal declensions. *Annual of Armenian Linguistics* 9: 19–20.
- Holst, Jan Henrik
2009 *Armenische Studien*. Wiesbaden: Harrassowitz.
- Hübschmann, Heinrich
1972 [1897] *Armenische Grammatik. I Theil: Armenische Etymologie*. Hildesheim: Olms.
[Leipzig: Breitkopf & Härtel.]
- Hübschmann, Heinrich
1906 *Armeniaca. Indogermanische Forschungen* 19: 457–480.
- IPA (= *International Phonetic Alphabet*, revised to 2005).
2005 Available online <http://www.internationalphoneticassociation.org/sites/default/files/IPA_chart_%28C%292005.pdf> [Last accessed 30 December 2014.]
- Ĵahowkyan, Gevorg B.
1969 *Hayoc' lezvi zargac'owmæ ew kařowc'vac'ə* [The development and structure of the Armenian language]. Erevan: Petakan Hamalsaran.
- Jasanoff, Jay H.
1979 Notes on the Armenian Personal Endings. *Zeitschrift für vergleichende Sprachforschung* 93: 133–149.
- Jensen, Hans
1959 *Altarmenische Grammatik*. Heidelberg: Winter.
- Jerejian, Armen V.
1953 The h-zero alteration in Classical Armenian. *Word* 9: 146–151.
- Job, Michael
1988 Armenian *anjrew* 'rain'. *Annual of Armenian Linguistics* 9: 27–34.
- Job, Michael
1995 Zum Lautwandel im Armenischen: Probleme einer relativen Chronologie. In: Wojciech Smoczyński (ed.), *Kuryłowicz Memorial Volume, Part One*. (Analecta Indoeuropaea Cracoviensia 2). Cracow: University Press, 291–311.
- Jungmann, Paul and J. J. S. Weitenberg
1993 *A reverse analytical dictionary of Classical Armenian*. Berlin: Mouton de Gruyter.
- Karst, Yosef
1901 *Historische grammatik des kilikisch-armenischen*. Strassburg: Trübner.
- Khachatryan, Amalia
1996 The Armenian sound system from typological aspect. In: Sakayan (ed.), 177–189.
- Khanjian, Hrayr
2009 Stress dependent vowel reduction. Unpublished paper presented at the 36th Meeting of The Berkeley Linguistics Society. The University of California, Berkeley, CA.
- Klingenschmitt, Gert
1982 *Das altarmenische Verbum*. Wiesbaden: Reichert.
- Kölligan, Daniel
2007 Arm. *nĵem* 'to sleep'. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 4: 25–35.
- Kölligan, Daniel
2012 Regular sound change and word-initial */j/- in Armenian. In: Philomen Probert and Andreas Willi (eds.), *Laws and Rules in Indo-European*. Oxford: Oxford University Press, 134–146.
- Kortlandt, Frederik
1975 A note on the Armenian palatalization. *Zeitschrift für vergleichende Sprachforschung* 89: 43–45.
- Kortlandt, Frederik
1980 On the relative chronology of Armenian sound changes. In: Greppin (ed.), 97–106.

- Kortlandt, Frederik
1983 Notes on Armenian historical phonology III: h-. *Studia Caucasica* 5: 9–16.
- Kortlandt, Frederik
1984a Proto-Armenian case endings. In: Gevorg B. Jahukyan and Andrzej Pisowicz (eds.), *International Symposium on Armenian Linguistics: Erevan, September 21–25, 1982: Reports*. Erevan: Armenian Academy of Sciences, 97–106.
- Kortlandt, Frederik
1984b PIE. *H- in Armenian. *Annual of Armenian Linguistics* 5: 41–43.
- Kortlandt, Frederik
1989 The making of a puzzle. *Annual of Armenian Linguistics* 10: 43–52.
- Kortlandt, Frederik
1998 The development of *y- in Armenian. *Annual of Armenian Linguistics* 19: 15–18.
- Kortlandt, Frederik
2003 *Armeniaca: Comparative Notes*. With an appendix on the historical phonology of Classical Armenian by Robert S. P. Beekes. Ann Arbor, MI: Caravan.
- Kortlandt, Frederik
2010 *Studies in Germanic, Indo-European and Indo-Uralic*. Amsterdam: Rodopi.
- Kroonen, Guus
2013 *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- Künzle, Beda O.
1984 *Das altarmenische Evangelium/L'Évangile arménien ancien*. Teil I/I^{ière} partie: Edition, Teil II/II^e partie: *Lexikon/lexique*. (Europäische Hochschulschriften, Reihe XXI, Linguistik und Indogermanistik 33). Bern: Peter Lang.
- de Lamberterie, Charles
1978 Armeniaca I–VIII: études lexicales. *Bulletin de la Société de Linguistique de Paris* 73: 243–285.
- de Lamberterie, Charles
2005 Le verbe arménien unim/kalay. In: Günter Schweiger (ed.), *Indogermanica: Festschrift Gert Klingenschmitt; indische, iranische und indogermanische Studien dem verehrten Jubilar dargebracht zu seinem fünfundsechzigsten Geburtstag*. Tübingen: Schweiger VWT, 333–358.
- Lipp, Reiner
2009 *Die indogermanischen und einzelsprachlichen Palatale im Indoiranischen 2*. Heidelberg: Winter.
- LIV² = Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
2001 *Lexikon der indogermanischen Verben: die Wurzeln und ihre Primärstambildungen*. 2nd, corrected edn. Wiesbaden: Reichert.
- Martirosyan, Hrach K.
2005 On Armenian čiw ‘shank’. In: Garnik Asatrian (ed.), *Iran and the Caucasus* 9.1. Leiden: Brill, 81–84.
- Martirosyan, Hrach K.
2010 *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.
- Martirosyan, Hrach K.
2013 The place of Armenian in the Indo-European language family: the relationship with Greek and Indo-Iranian. *Journal of Language Relationship* 10: 85–137.
- Matasović, Ranko
2009 *Etymological dictionary of proto-Celtic*. Leiden: Brill.
- Matzinger, Joachim
2005 *Untersuchungen zum altarmenischen Nomen: die Flexion des Substantivs*. (Münchener Studien zur Sprachwissenschaft, Beiheft 22). Dettelbach: Röhl.

- Meillet, Antoine
1898 Étymologies arméniennes. *Mémoires de la Société de linguistique de Paris* 10: 274–282.
- Meillet, Antoine
1913 *Altarmenisches Elementarbuch*. Heidelberg: Carl Winter.
- Meillet, Antoine
1924 Remarques étymologiques. *Revue des Études Arméniennes* 1/4: 1–6.
- Meillet, Antoine
1936 *Esquisse d'une grammaire comparée de l'arménien classique*. 2nd edn. Vienna: Mekhitaristes.
- Meillet, Antoine
1977 De quelques évangélistes arméniens accentués. In: *Études de linguistique et de philologie arméniennes II*. Louvain: Imprimerie Orientaliste, 322–357. [Originally published in *Mémoires publiés par l'École Nationale des Langues Orientales Vivantes*. 1905. 133–168.]
- Minassian, Martiros
1976 *Manuel Pratique d'arménien ancien*. Paris: Fondation des Frères Ghoukassiantz.
- Morani, Moreno
1994 Sui casi di IE *o > Arm. a. In: Palmira Cipriano, Paolo Di Giovine, and Marco Mancini (eds.), *Miscellanea di studi linguistici in onore di Walter Belardi* 1. Roma: Il Calamo, 385–398.
- NIL = Wodtko, Dagmar S., Britta S. Irslinger, and Carolin Schneider (eds.)
2008 *Nomina im Indogermanischen Lexikon*. Heidelberg: Winter.
- Normier, Rudolf
1980 Beiträge zur armenischen Etymologie I. *Annual of Armenian Linguistics* 1: 19–22.
- Olsen, Birgit Anette
1999 *The Noun in Biblical Armenian: Origin and Word-formation: with Special Emphasis on the Indo-European Heritage*. Berlin: Mouton de Gruyter.
- Olsen, Birgit Anette
2010 *Derivation and composition: two studies in Indo-European word formation*. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Pedersen, Holger
1904 Hin hayeren lezui sesta [The accent in the old Armenian language]. *Handēs Amsoreay* 18: 131–133.
- Pedersen, Holger
1905 Zur armenischen Sprachgeschichte. *Zeitschrift für vergleichende Sprachforschung* 38: 194–240.
- Pedersen, Holger
1906 Armenisch und die Nachbarsprachen. *Zeitschrift für vergleichende Sprachforschung* 39: 334–485.
- Pedersen, Holger
1982 *Kleine Schriften zum Armenischen*. Ed. by Rüdiger Schmitt. Hildesheim: Olms.
- Pisani, Vittore
1950 Studi sulla fonetica dell'armeno. *Ricerche Linguistiche* 1: 165–193.
- Pisani, Vittore
1951 Studi sulla fonetica dell'armeno. *Ricerche Linguistiche* 2: 47–74.
- Pisowicz, Andrzej
1976 *Le développement du consonantisme arménien*. (Polska Akademia Nauk. Oddział w Krakowie. Prace Komisji Językoznawstwa 43). Wrocław: Polish Academy of Sciences.
- Polomé, Edgar C.
1980 Armenian and the Proto-Indo-European laryngeals. In: Greppin (ed.), 17–34.

- Praust, Karl
 2005 Bedeutung und Vorgeschichte von armenisch *gerem* (mit einem Exkurs zu griechisch ἀίχμαλωτεύω, ἀίχμαλωτίζω). *Die Sprache* 45: 134–159.
- Pronk, Tijmen
 2010 On Greek ἀρχμός ‘drought’ and ἀρχήν ‘neck’. *Glotta* 86: 55–62.
- Rasmussen, Jens E. and Thomas Olander (eds.)
 2009 *Internal Reconstruction in Indo-European: Methods, Results, and Problems. Section Papers from the XVIth Conference on Historical Linguistics. Copenhagen, 11th–15th August, 2003*. Copenhagen: Museum Tusulanum.
- Ravnæs, Erling
 1991 *The chronology of the sound changes from Proto-Indo-European to Classical Armenian*. University of Oslo unpublished Doctoral thesis.
- Ritter, Ralf-Peter
 1985 Zum Wandel uridg. *mn > arm. wn. *Münchener Studien zur Sprachwissenschaft* 45: 197–199.
- Ritter, Ralf-Peter and García Martínez Francisco Javier
 1996 *Introducción al armenio antiguo*. Madrid: Clásicas.
- Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
 2001 *Lexikon der indogermanischen Verben: die Wurzeln und ihre Primärstambildungen*. 2nd, corrected edn. Wiesbaden: Reichert.
- Sakayan, Dora (ed.)
 1996 *Proceedings of the Fifth International Conference on Armenian linguistics, McGill University Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan.
- Scheffelowitz, Isidor
 1905 Zur altarmenischen Lautgeschichte. *Beiträge zur Kunde der indogermanischen Sprachen* 28: 282–313.
- Schirru, Giancarlo
 2012 Laryngeal features of Armenian dialects. In: Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen, and Jens Elmegård Rasmussen (eds.), *The sound of Indo-European: phonetics, phonemics and morphophonemics*. (Copenhagen Studies in Indo-European 4). Copenhagen: Museum Tusulanum, 435–457.
- Schmitt, Rüdiger
 1972 Empfehlungen zur Transliteration der armenischen Schrift. *Zeitschrift für vergleichende Sprachforschung* 86: 296–306.
- Schmitt, Rüdiger
 1980 Lautgeschichte in Abhängigkeit von Etymologie: das Armenische. In: Manfred Mayrhofer, Martin Peters, and Oskar E. Pfeiffer (eds.), *Lautgeschichte und Etymologie: Akten der VI. Fachtagung der Indogermanischen Gesellschaft, Wien, 24–29. September 1978*. Wiesbaden: Reichert, 412–430.
- Schmitt, Rüdiger
 1981 *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Sievers, Eduard
 1893 *Grundzüge der Phonetik: zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf & Härtel.
- Sköld, Johannes Evelinus
 1927 Die lateralen Konsonanten des Armenischen. *Handēs Amsoreay* 41: 779–794.
- Stempel, Reinhard
 1990 Die Entwicklung von auslautendem *m und das Problem sekundärer n-Stämme im Armenischen. *Indogermanische Forschungen* 95: 38–62.

- Stone, Michael E.
2006 *Apocrypha, pseudoepigrapha and Armenian studies: collected papers* Vol. 2. Leuven: Peeters.
- Thomson, Robert W.
1989 [1975] *An Introduction to Classical Armenian*. 2nd edn. Delmar, NY: Caravan.
- Vaux, Bert
1992 Adjarian's Law and Consonantal ATR in Armenian. In: John A. C. Greppin (ed.), *Proceedings of the Fourth International Conference of Armenian Linguistics*. Delmar, NY: Caravan, 271–293.
- Vaux, Bert and Bridget Samuels
2005 Laryngeal markedness and aspiration. *Phonology* 22: 395–436.
- Vaux, Bert and Andrew Wolfe
2009 The Appendix. In: Eric Raimy and Charles Cairns (eds.), *Contemporary views on architecture and representations in phonology*. Cambridge, MA: MIT Press, 101–143.
- Viredaz, Rémy
2001–2002a Contractions et place de l'accent en ancien Arménien. *Revue des Études Arméniennes* 28: 1–11.
- Viredaz, Rémy
2001–2002b Sur le traitement arménien des sonantes voyelles. *Slovo* 26–27: 24–36.
- Viredaz, Rémy
2005 Notes on Armenian historical phonology I. *Annual of Armenian Linguistics* 24–25: 85–103.
- Vogt, Hans
1938 Arménien et caucasique du sud. *Norsk Tidsskrift for Sprogvidenskap* 9: 321–338.
- Vogt, Hans
1958 Les occlusives de l'arménien. *Norsk Tidsskrift for Sprogvidenskap* 18: 143–159.
- Weitenberg, Jos J. S.
1983 Armenian dialects and the Latin-Armenian glossary of Autun. In: Thomas J. Samuelian and Michael E. Stone (eds.), *Medieval Armenian Culture: Proceedings of the Third Dr. H. Markarian Conference on Armenian Culture*. (University of Pennsylvania Armenian Texts and Studies 6). Chico, CA: Scholars Press, 13–28.
- Weitenberg, Jos J. S.
1995 *New approaches to medieval Armenian language and literature*. Amsterdam: Rodopi.
- Weitenberg, Jos J. S.
1996 On the early development of the Armenian dialects. In: Sakayan (ed.), 93–118.
- Winter, Werner
1962 Problems of Armenian Phonology III. *Language* 38: 254–262.
- Wodtko, Dagmar S., Britta S. Irslinger, and Carolin Schneider (eds.)
2008 *Nomina im Indogermanischen Lexikon*. Heidelberg: Winter.

Martin Macak, Athens, GA (USA)

62. The morphology of Armenian

- | | |
|---------------|----------------------|
| 1. Summary | 5. Pronouns |
| 2. Nouns | 6. Personal pronouns |
| 3. Adjectives | 7. Verbs |
| 4. Numerals | 8. References |

1. Summary

The development of the Armenian morphological system from its Indo-European roots is fundamentally marked by a series of dramatic phonological changes, above all the process of apocope regularly deleting a great deal of the inherited inflectional endings. The noun is characterized by a complete loss of gender distinction, but a differentiated case system is maintained, partly by means of added suffixes and postpositions. The verb has undergone a considerable restructuring, eliminating such categories as the optative and the perfect, and the set of demonstrative pronouns and adverbs excels in a systematic distinction between first, second, and third person deixis. More detailed historical surveys may be found in Meillet (1936), Solta (1963), Godel (1975), Schmitt (1981), and de Lamberterie (1988/1989), beside the monographs by Pedersen (1905) on the demonstrative pronouns, Klingenschmitt (1992) on the verb, Olsen (1999) on the noun, and Matzinger (1995, 2005) on nominal declension.

2. Nouns

2.1. By the beginning of the Armenian written tradition the declensional system of the protolanguage had already been subject to substantial simplifications. The dual number was extinct and, more importantly, the category of grammatical gender had been eliminated. The case system, on the other hand, was more resistant, and the distinction between seven cases, the nominative, accusative, genitive, dative, locative, ablative, and instrumental, is in principle maintained; only an independent vocative has not survived. However, there are several instances of case syncretism, and a given noun never has more than four separate forms in the singular and four in the plural. The genitive and dative are distinguished only in pronouns, and the nominative and accusative singular only in the personal pronoun.

Nouns are inflected according to various stem classes roughly mirroring the inherited system, e.g. **mṛtos* > *mard* (o-st.) ‘man’, **sm(h)ah₂* > *am* (a-st.) ‘year’, **b^hah₂nis* > *ban* (i-st.) ‘word’, **-uēstus* > *z-gest* (u-st.) ‘dress’, **uṛh₁ēn* > *gaṛn* (n-st.) ‘lamb’, **d^huḡ₂tēr* > *dustr* (r-st.) ‘daughter’, though secondary adjustments have to some extent blurred the picture. Thus, neuter *s*-stems, like **h₁erk^(w)os* > *erg* (o-st.) ‘song’, have fed the *o*-stems, a partial transfer from *i*- to *u*-stems was triggered by a regular sound change **-ui-* > **-uu-* (Schindler apud Matzinger 1992 and 2005: 58), e.g. **h₂reui-* > *arew* (u-st.) ‘sun’, a new category of *l*-stems was created on the basis of the inherited *astl* ‘star’, supported

by stray loanwords like *arkl* ‘chest’ (cf. Lat. *arcula*), and the synchronically inconvenient root nouns have been adapted to the *n-* or *i-*stem paradigm.

Since Meillet (1913) a major distinction in the description of Armenian noun declension has been made between the vocalic stems, where the form common to the nom.acc.sg. serves as an invariable stem throughout the paradigm, and the variable stems in sonorants, which exhibit ablaut alternation.

2.2. The inflection of the four vocalic stem types, *o-*, *a-* (or *a/i-*), *i-*, and *u-*stems, may be exemplified by *mard* ‘man’, *am* ‘year’, *ban* ‘word’ and *zgest* ‘dress’:

sg.	o-st.	a-st.	i-st.	u-st
nom.acc.	mard	am	ban	zgest
gen.dat.	mardoy	ami	bani	zgestu
loc.	mard	ami	bani	zgestu
abl.	mardoy	amē	banē	zgestē
inst.	mardov	amaw	baniw	zgestu
pl. nom.	mardkʻ	amkʻ	bankʻ	zgestkʻ
acc.loc.	mards	ams	bans	zgests
g.d.abl.	mardocʻ	amacʻ	banicʻ	zgestucʻ
inst.	mardovkʻ	amawkʻ	baniwkʻ	zgestukʻ

The historical background of the case endings may be envisaged as follows:

As a consequence of the regular loss of final syllables the nom. and acc.sg. **-Vs*, **-Vm* end up with zero ending, and the same goes for the nom.acc.sg. of original neuter *o-*, *i-* and *s-*stems, feminines in **-ah₂*, **-ah₂m* and the locative of *o-*stems in **-oj*. The fate of final nasals is a matter of some controversy, but it seems likely that they were only regularly lost after the open vowels *-o-* and *-a-* (Olsen 1999: 5 f.). If this is the case, the zero ending in the acc.sg. of the originally animate *i-* and *u-*stems must be analogical.

The *o-*stem gen.sg. *-oy* < **-osjo* (cf. Skt. *-asya*, Hom. *-oio*, OLat. *Valesiosio*) is also used for the dative and ablative where the old endings **-ōj* (< **-o-ej*) and **-o-ad* (Klingenschmitt 1992: 94) would regularly have undergone apocope. The secondary formal identity between genitive and dative singular is shared and no doubt influenced by the other vocalic stems.

In all the vocalic declensions the instrumental in **-b^hi* is regularly continued as *-w/* *-v* (*-v* after *-o*; *u-*stems **-uw* > *-u*) with the corresponding inst.pl. *-wkʻ*, adding the general plural marker *-kʻ*, while *-b* (pl. *-bkʻ*) is found with consonant stems. The *o-*st. plural *-ovkʻ* for **-ōjs* is analogical.

The much-debated plural marker *-kʻ*, sometimes considered a supplementary particle whether of Indo-European or foreign origin, is preferably analyzed as a sandhi variant of **-s*, also yielding zero (Klingenschmitt 1992: 23 f.). The starting point of this semantic specialization, restricting *-kʻ* to the plural and zero to the singular, e.g. **mrtos* > *mard* vs. **mrtōs* > *mardkʻ*, must have been paradigms with an asigmatic nom.sg. (*a-*, *n-*, *r-*

and *l*-stem), possibly furthered by the consistently asigmatic vocative. The assumption of a uniform development of postvocalic **-s* > *-k'* (Kortlandt 1984: 89 f., following Pedersen 1905: 71 ff.), implying an analogical replacement of the nom.sg. by the acc.sg. in **-Vm* is less likely, leaving the general merger of the neuter *s*-stems (nom.acc.sg. **-os*) and the *o*-stems unexplained.

The acc.pl. in *-s*, the regular outcome of **-Vns*, has merged with the loc.pl. with the original ending **-su* in all stem classes, apparently by analogy with the *s*-stems (**-Vs-su* > *-Vs*) and possibly the *n*-stems with secondary restitution of the nasal (e.g. *gařins* 'lambs', *jkuns* 'fishes' for **gařis*, **jkus*); in the vocalic declensions **-s-* would have been lost in intervocalic position, and in the *r*-stems the cluster **-rs-* would have yielded **-rš-* or **-ř-*.

Finally, the original ending of the gen.dat.abl.pl. in *-Vc'* was most likely *-ic'* (Selde-slachts 1991: 261), going back to the adjective of appurtenance in **-isKo-*, whence the analogical *o-*, *a-* and *u-*stem forms *-oc'*, *-ac'*, *-uc'*. Here the primary case function was that of the genitive, e.g. **mardisKo-* → *mardoc'* 'belonging to the group of men', i.e. 'of the men', and the secondary transfer to the dative and ablative is explained by analogy from the singular where these three cases merged in the *o*-stems.

As a consequence of the loss of final syllables the inflection of the inherited *ā*-stem paradigm would have been restricted to two forms in the singular, the instrumental **-āb^hi* > *-aw* as opposed to zero as the regular continuation of nom. **-ā*, acc. **-ām*, gen.abl. **-ās*, and dat.loc. **-āi*. Thus, the maintenance of a differentiated inflectional pattern would necessitate analogical influence from other stem classes and/or the addition of postpositions to clarify the various case functions. As for the genitive and dative, the ending *-i*, common to the *ā*- and the *i*-stems, is regular with the hysterodynamic *i*-stems in **-iḡos*, **-iḡei*, eventually merging with the *vrkīḡ*-type in **-ih₂os*, **-ih₂eḡi*, cf., e.g., the Armenian *i*-stem *hac'* 'bread' < **p_ek^wtih₂s* 'Gebäck' rather than **p_ek^wtis* 'Backen'. Another potential source of a gen.dat.sg. in *-i* and model for extension to the *a*-stems is a Sievers-variant of the *devī*-type **-iḡah₂-s*, **-iḡah₂-ai* or a later contamination of the *vrkīḡ*- and the *devī*-types (Matzinger 2005: 101). The *u*-stem gen.dat.sg. ending *-u* apparently reflects a Sievers-variant of the hysterodynamic **-uḡos*, **-uḡei*.

However, the locative in *-i* of the *a*- and *i*-stems is not so easily explained on the basis of an old case ending. The *a*- and *i*-stem ablative is *amē*, *banē*, not **ami-ē*, **bani-ē* (*u*-stems do show a variant in *-uē*, however: *zgestē* beside e.g. *zarduē* 'ornament'), despite the fact that the ablative ending is otherwise added to the original locative, e.g. *i hawrē* 'from the father' (lit. 'from in the father') vs. nom. *hayr*, and thus it seems more likely that a new locative marker *-i*, derived from the postposition **h₁en* 'in', was added to the *a*- and *i*-stems at a time when apocope had made the old ending opaque. Eventually the formal identity of the *a*- and *i*-stem locative facilitated the extended use of the *i*-ending in the genitive and dative of the *a*-stems as well.

No doubt the facultative locative ending *-i* of the *o*-stems, only used after the preposition *i* where it was essential to distinguish between the meanings 'in' (loc.) and 'into' (acc.), is historically identical with that of the *a*-stems. In the consonant stems, on the other hand, where accusative and locative were differentiated by stem alternation, the clarifying postposition was unnecessary, thus *i tun* (acc.) 'into a house' vs. *i tan* (loc.) 'in a house'.

In contrast to the *o*-stems, the *a*-, *i*-, and *u*-stems have a separate ablative ending *-ē* which is probably best explained as reflecting the postposed adverb **eti* > Skt. *āti* 'over,

beyond', Gk. *ἔτι* 'also, further', added to underline an otherwise unclear ablative function (Godel 1975: 105); the traditional derivation of *-ē* from **-etos* (Meillet 1936: 73), allegedly matching such examples as Skt. *mukhatāh* 'from the mouth, in front', Lat. *funditus* 'from the bottom', runs afoul of the development of **-eto-* > **-ew-* (cf., e.g., *elēw* 'became' < **e-k^wl[h₁]eto*). It is peculiar that the only stem type where the protolanguage did have an abl.sg. distinct from the genitive, viz. the *o*-stems, is the one where an early analogical process, occurring before the restructuring of the *a*- and *i*-stem paradigms, produced a merger between genitive and ablative.

The *wo*-stems (gen.dat.abl. *-woy*, dissimilated from **-iiojō* < **-iiosjō*, inst. *-wov*; gen.dat.abl.pl. *-woc'*, inst. *-wovk'*), e.g. *kogi* 'butter' < **g^woujō-* = Skt. *gāvya*, Av. *gaoiia-*, Gk. *-βοιος*, Toch.B. *kewiye*, are variants of the *o*-stems containing the suffix **-io-*. Similarly, the *ea*-stems are variants of the *a*-stems with considerable analogical influence from the *wo*-stems: gen.dat.sg. *-woy*, loc. *-woj* < **-wo-* + *-d^hi-V-* with the same adverbial element as Gk. *ὀπavόθι* 'in the sky', abl. *-woy* on the model of the *o*-stems or *-wojē* (loc. + ablative postposition *-ē*), inst. *-eaw*, gen.dat.abl.pl. *-eac'*, inst. *-eawk'*.

A special subtype of the *a*-stems with gen.dat.loc.sg. *-ay*, inst. *-aw*, gen.dat.loc.pl. *-ac'*, inst. *-awk'* mostly occurs with proper names, e.g. *Vardan*, gen. *Vardanay*. The inflection is influenced by the *o*-stems: *-ay* beside *-oy* (Weitenberg 1989).

2.3. In the sonorant declensions, the form of the nom.acc.sg. differs from the oblique stem, as exemplified by the paradigms of the *n*-stems *gaṛn* (< **uṛh₁ēn*) 'lamb' and *mukn* (< **muhsnt-?*) 'mouse', the *r*-stem *dustr* (< **d^huḡə₂tēr*) 'daughter' and the *l*-stem *astl* (< **h₂stēl* for **h₂stēr*) 'star':

sg.	n-st	n-st.	r-st.	l-st.
n.a.	gaṛn	mukn	dustr	astl
g.d.l.	gaṛin	mkan	dster	astel
abl.	gaṛnē	mkanē	dsterē	astelē
inst.	gaṛamb	mkamb	dsterb	astelb
pl.n.	gaṛink'	mkunk'	dsterk'	astelk'
acc.loc.	gaṛins	mkuns	dsters	astels
g.d.abl.	gaṛanc'	mkanc'	dsterac'	astelac'
inst.	gaṛambk'	mkambk'	dsterawk'	astelawk'

N-stems exhibit a number of archaic features, preserving traces of the old apophonic variation. The elements of the inherited alternating *n*- and *nt*-stems (**-ēn* and **-ōn* > *-n*; **-en-* > *-in-/n-*; **-on-* > *-un-/n-*; **-n-/nt-* > *-an-*) have been distributed into three subtypes with either three ablaut variants (*n/in/an* or *n/un/an*) or four (*n/in/an/un*), e.g. nom.sg. *harsn* 'bride', gen.dat.loc. *harsin*, inst. *harsamb*, nom.pl. *harsunk'*. The gen.dat.abl.pl. *-anc'* < **-n-sKo-* and inst.sg.(pl.) *-amb(k')* < **-n-b^hi* (+ *-k'*) is common to all types, and in some old neuters like *sermn* 'seed' (old **m_ṇ[t]-stem*), pl. *sermank'*, the suffix variant *-an-* is preserved throughout the paradigm. While the historical basis of the type *gaṛn* is hysterodynamic (nom.sg. **-ēn*, loc. **-ēni*, nom.pl. **-énes*), and that of

the type *harsn* proterodynamic non-neuters (nom.sg. **-ōn*, pl. **-ones*), the type *mukn* (gen.dat.loc.sg. *-an-* < **-ŋ[t]/-*) is followed by neuters in **-ŋ(t)-* and older heteroclitics. Other sources of the *n*-stems are root nouns like *otn* ‘foot’ (< acc.sg. **pód-ŋ*), gen. *otin* (*i*-st. plural from **podisKo-* > *otic*), derivatives in **-no-* and Iranian loanwords.

An important variant of the *n*-stems consists of words in *-iwn*, including the abstracts in *-ut’iwn*, e.g. nom.acc.sg. *ankiwn* ‘corner’, gen.dat.loc. *ankean*, nom.pl. *-iwnk’*, gen.dat.abl. *-eanc’*. The main element of this type seems to be “Hoffmann formations” in **-i-h₃on(h₂)-*, **-i-h₃n(h₂)-os* (**ih₃* > **ia* > *ea* by “laryngeal breaking”), cf. the Italic pattern **-tiōn-*, **-tīn-* (Lat. *natiō*, *-iōnis* vs. Umbr. abl. *natine*). The substantival variant of the “Hoffmann formations” originally seems to have denoted ‘a load of’, or in general ‘a collective of that which is expressed by the initial stem’, e.g. Lat. *legiō* ‘a body of soldiers, a legion’ from ‘mass, group of men who have been assembled’, cf. Olsen (2010: 140).

Some *n*-stems have a more or less deviating inflection. Thus, e.g., *tun* ‘house’ (**dōm*) and *šun* ‘dog’ (**kūōn*) have the oblique forms *tan* and *šan*, apparently with analogical introduction of the preconsonantal zero grades **dŋ-*, **kūŋ-*.

In a few *n*-stem heteroclitics denoting persons, e.g. *erēc’* ‘elder’, pl. *eric’unk’*, or *nu* (< **snuso-*) ‘daughter-in-law’, pl. *nuank’*, it seems natural to ascribe the *n*-stem inflection to the “individualizing” **-n*.

The *r*- and *l*-stem forms *dsterac’*, *astelac’*, *dsterawk’*, *astelawk’* (beside *astelbk’*) are analogically introduced from the *a*-stems.

A small group of *r*-stem kinship terms, *hayr* (< **p₂tér*) ‘father’, *mayr* (< **mātér*) ‘mother’ and *elbayr* (< **b^hrātēr*) ‘brother’, have preserved an archaic ablaut pattern, thus nom.acc.sg. *hayr* < **p₂tér*, gen.dat.loc. *hawr* < **p₂tr-ós/-ej/-i*, abl. *hawrē*, inst. *harb* < **p₂t_r-b^hi*, nom.pl. *hark’* < **p₂tér-es*, acc.loc. *hars*, gen.dat.abl. *harc’* < **p₂t_r-sKo-*, inst. *harbk’*.

K’oyr < **súésōr* ‘sister’ has gen.dat.loc.sg. *keř* < **súesr-V-*, nom.pl. *k’ork’* < **súésōres* and gen.dat.abl. *k’erc’* for **k’arc’* < **k’earc’* < **súes_r-sKo-*.

The original type underlying *alewr* ‘flour’, *albewr* ‘fountain’, *eljewr* ‘horn’, and *al-tewr* ‘swamp, pool’ (textually later *-iwr*) with gen.dat.loc.sg. *-er* (< **-eueros*, *-ej*, *-i*) was a heteroclitic **-u_r/-uen*-stem, cf., e.g., Gk. *φρέαρ*, gen. *φρέατος*. However, if we take the earliest textual evidence seriously, nom.acc.sg. *albewr* (not *albiwr*) cannot be an identical match with the corresponding Greek form (**b^hrēw_r*) (cf. Olsen 1999: 156 f. for a discussion of the phonetic details).

2.4. Irregular nouns include: *akn* < **h₃k^w-mŋt-* (= Gk. *ὄμμα?*), gen.dat.loc. *akan* ‘eye’ with the plural *ač’k’* (gen.dat.abl. *ač’ac’*), whether directly from a dual **h₃k^wih₁* (cf. OChSl. *oči*, Lith. *aki*) or **h₃k^w-* + the masculine dual ending **-e*; the Greek form *ὄσσε* is similarly ambiguous, reflecting either **h₃k^wih₁* pure and simple or **h₃k^wih₁* + *-e*. In the secondary meanings ‘spring’ and ‘gem’, the nom.pl. has been normalized to *akunk’* and *akank’*, respectively.

Unkn ‘ear’, gen.dat.loc. *unkan*, is perhaps an original *nt*-stem comparable to Gk. gen. *οὔατος* (Lindeman 1980: 61) with secondary influence from *akn* ‘eye’. The plural *akanjk’* is more enigmatic.

Ayr < **h₂nér* (Gk. *ἀνήρ*) ‘man, husband’ with gen.dat.loc.sg. *ařn*, metathesized from **h₂nrós*, etc. (= Gk. *ἀνδρός*), inst. *aramb*, also with metathesis from **h₂n_rb^hi* → **h₂n_r*-

b^{hi}. The stem *ar(a)-* is generalized in the plural: nom. *ark'*, gen.dat.abl. *aranc'*. *Ayr* also survives in the archaic compound *tēr* 'lord' < **ti-ayr* (*ti-* < **dems-* 'of the house'), gen. *teařn* (cf. also *ti-kin* 'lady').

Kin < **g^wenə₂* 'woman, wife'. The peculiar inflection with gen.dat.loc. *knoj*, abl. *knojē*, is shared by the numeral *mi* (< **smih₂*), *mioj* (beside gen. *mioy*, dat.loc. *mium*), *miojē* (beside *mioy*). The ending *-oj* cannot be immediately compared with *-woj* (*-io-d^{hi}i*) of the *ea*-stems, which is essentially a locative marker, so perhaps we are dealing with a particular phonetic development of sequences containing a vocalic nasal + **-j*: (**g^wh₂ih₂os* >) **kñioh* > **køj* > **koj* → *knoj* (after *kin*); **smjah* > *(*h*)*oj* > *(*h*)*oj* → *mioj* (after *mi*) or the like. The plural *kanayk'* is perhaps historically identical with and not just superficially similar to Gk. *γυναῖκες*, since the addition of an element **-ik-* (or **-i-*) to an *a*-stem is equally unparalleled in Armenian and Greek.

3. Adjectives

3.1. After the disappearance of all gender distinctions, the borderline between substantives and adjectives became somewhat vague, though of course some derivational suffixes are particularly associated with adjectives. A few points of interest are: 1. the sporadically attested archaic contrast between *o*-stem nouns and *i*-stem compound adjectives (original *bahuvrīhis*), e.g. *gorc* (*o*-st.) 'work' vs. *angorc* (*i*-st.) 'without work' (cf., e.g., Lat. *arma* 'weapon': *inermis* 'unarmed') and 2. the isolated compound *amul* 'barren' (of women), probably < **ñ-putlo-* (= Skt. *aputra-*, Av. *apuθra-*) which is inflected as an *o*-stem (gen. *amloy*), thus pointing to the same system of "Adjektiva zweier Endungen" as is known from Greek.

A remarkable heteroclititic pattern where only the plural follows the *n*-stems is connected with old *u*-stem adjectives like *manr* (< **mñ[h]u-*) '(very) small', gen.dat.loc.sg. *manu*, nom.pl. *manunk'*, gen.dat.abl. *manunc'* or the slightly deviating *barjr* (< **b^hrġ^hu-*) 'high', *barju*, *barjunk'*, *barjanc'*. The *n*-stem forms are probably introduced from the participial *nt*-pattern (cf. in particular Skt. *brhānt-/brhat-*: Arm. *barjun-/barjan-*), while the origin of the final *-r* in the nom.acc.sg. is more obscure, insofar as neither the idea of analogical influence from old heteroclitics nor contamination with the adjective suffix **-ro-* (cf. Toch.A. *pärkär*, B. *pärkare* 'long' vs. Hitt. *parkus* 'high') is fully satisfactory. The least complicated scenario is a direct development of word-final **-u* to *-r* (cf. also the original *u*-stem neuters *artawsr* < **draku* 'tear', pl. *artasuk'*; *cunr* < **ġonu* 'knee', pl. *cungk'*; *melr* 'honey' vs. Skt. *mādhu*, etc., and contrast the *r*-less masculines *erēc'* 'elder', *zgest* 'garment'), as suggested by Alexis Manaster-Ramer (p.c.), basing himself on Pedersen (1905: 231). The latter, however, considered the formulation of a phonological rule for such a change impossible.

3.3. Armenian has not preserved the old comparative/superlative suffixes except perhaps in a few synchronically opaque relics (e.g. *ner-k'in* 'inner, inward' < **h₁en-ter-*). Instead, the comparative is indicated by the positive followed by *k'an* or *k'an z-* 'than' or by the originally intensive formation in *-agoyn*, e.g. *canragoyn* 'very heavy, too heavy' and 'heavier' with a suffix borrowed from Iranian *gaona-* 'color, kind, manner, appearance', cf. especially compounds like Sogdian *wrōγwn* ~ Arm. *vardagoyn* 'rosy'. A superlative

or relative function may be expressed by a compound in *amena-* (cf. *amenayn* ‘all’), e.g. *amenasurb* ‘holiest; very holy’, by the addition of *y-amenesin* or *amenec’un* ‘of all’, or by reduplication, e.g. *mecamec* ‘largest, very large’, *barjraberj* ‘very high’.

4. Numerals

4.1. With the exception of a couple of Iranian loanwords for the higher numbers, the Indo-European system of numerals is quite well preserved in Armenian.

4.2. Cardinals, 1–10: 1: *mi* < **smih₂* (the old feminine); the compositional form **mia-* > *me-* is preserved in *metasan* ‘eleven’. For the inflection, see 2.4. 2: *erku* < **duō*, originally a dual; **duo-* in *erkotasan* ‘twelve’, and perhaps compositional **dui-* in compounds like *erkban* ‘deceitful’ (if not simply from **erku-*); apart from the nom., *erku* is inflected as an *u*-stem plural, acc. *erkus*, etc. 3: *erek’* < **trejes*. 4: *č’ork’* < **k^wetores* (dissimilated from **k^wetūores*). 5: *hing* < **penk^we*: the final *-e* preserved in *hngetasan* ‘15’, *hngerord* ‘5th’. 6: *vec’*, probably from a Lindeman variant **suueks* (Klingenschmitt 1982: 61). 7: *ewt’n* < **septm̄* with the more obscure variant *ewt’n*. 8: *ut’*, presumably < **optō* for **oktō* with influence from ‘seven’. 9: *inn*, perhaps from a metathesized **enun* (Eichner 1978: 152). 10: *tasn* ‘ten’ < **dek̄m̄* with somewhat unclear root vocalism (distant assimilation **tesan-* > *tasan-*? Szemerényi 1960: 21). The numbers ‘three’, ‘four’, ‘five’, ‘six’ and ‘eight’ are inflected as plural *i*-stems under the influence of *erek’*, acc. *eris* < **trins* (= Goth. *þrins*), gen. *eric’*, inst. *eriwk’* < **tri-b^hi-* (Skt. *tribhiḥ*, Lat. *tribus*), ‘seven’ and ‘ten’ as *n*-stems (gen. *ewt’anc’*, *tasanc’*; also *tasin*), and ‘nine’ as either an *i-* or an *n*-stem (gen. *innic’* or *inunc’*).

11–19: The numbers 11–16 are *i*-stem compounds: *metasan*, *erkotasan*, *erek’tasan*, *č’orek’tasan*, *hngetasan*, *veštasan* (with a special development of the heavy consonant cluster **-ks-d-*) where *-tasan* should perhaps be derived from **-dek̄m̄ti-* (Winter 1992: 351). 17–19 are expressed by juxtapositions, *ewt’n ew tasn*, etc.

20–90: The decades, apart from *k’san* ‘20’ < **uik̄m̄tih₁* (dual), are characterized by the suffix *-sun* < **-k̄omt̄₂*: *eresun* ‘30’ (*ere-* < **tria-*), *k’aḫasun* ‘40’ (**k^wtūr-dk̄omt̄₂* > **k^wtūrkont̄₂* with **-f̄-* > *-aḫa-*; Szemerényi 1960: 135f, Klingenschmitt 1982: 68), *yisun* ‘50’ (the expected protoform is **hingisun* < **penk^wēk̄omt̄₂* < **penk^we-dk̄omt̄₂*, cf. Gk. *πεντήκοντα*, though the details of the phonetic development remain obscure), *vat’sun* ‘60’ ([t^hss] > [t^hs] beside *vec’* and *veštasan*), *ewt’anasun* ‘70’, *ut’sun* ‘80’, *innsun* ‘90’.

Hariwr ‘100’ has no generally accepted etymology, and *hazar* ‘1000’ and *bewr* ‘10,000’ are Iranian loanwords.

4.3. Ordinals: Apart from *aḫajin* < **pḫ₃ujō-* + **i-no-* ‘first’, the ordinals are characterized by a synchronic suffix *-rord* or *-erord* (diachronic **-ord*): *erkrord* ‘second’ < **duis-* + *-ord*, *errord* ‘third’ < **tris-* + *-ord*, *č’orrord* ‘fourth’ < analogical *č’orir-* (taken from *erkir*, *erir*) + *-ord*, *hingerord* ‘fifth’ < **penk^we-* + analogical *-rord*; hence *vec’erord*, etc. The suffix **-ord* may be derived from **k^wórtos* (cf. Lith. *kařtas*, OChSl. *kratb* ‘time’, Winter 1992: 356) or perhaps better the compositional **-k^wrt* = Skt. *(sa)kṛt*, Av. *(ha)kərət* ‘once’ (Pisani 1944: 174) with **-k^w-* > *-Ø-* before *-o-* and regular development **k^wrtC* > **k^worC*.

4.4. The numeral adverbs *erkir* ‘second(ly); in the second place’, *erir* ‘third(ly), in the third place’, may be derived from **d̥uis*, **tris* with ruki-development of the final **-s* and exact cognates in Skt. *dvīḥ* ‘twice’, Gk. *δίς* ‘twice; doubly’, Lat. *bis* ‘twice; in two ways’; Skt. *trīḥ*, Gk. *τρίς*, Lat. *ter* ‘three times’ (cf. Olsen 1989: 5–15 with references. [ed. For a more restrained assessment of the operation of ruki in Classical Armenian, with no allowance for a shift to *r*, see Macak, this handbook, 9.3.8.]).

The suffixes *-(e)k’ean* and *-(e)k’in* are used for the formation of collective numbers, e.g. *erkok’ean/erkok’in* ‘both’, *erek’ean/erek’in* ‘all three’, *ewt’anekean/ewt’anek’in* ‘all seven’. The clue to the peculiar inflection (acc. *-sean/-sin* gen.loc. *-c’unc’*, inst. *-k’ umbk’*), common to both types, may be found in the paradigm of *erkok’ean/-k’in*, *erkosean/-sin*, *erkoc’unc’*, *erkok’ umbk’* ‘both’, which was probably influenced by pronouns like pl. *nok’in*, *nosin*, *noc’unc’*, *nok’umbk’* ‘they, those, the same’, serving as the emphatic third person personal pronoun in the plural as opposed to *erkok’ean/-k’in* which might be similarly described as the third person personal pronoun in the dual.

Multiplicatives with the meaning ‘n-fold, n times’ are characterized by the suffix *-kin*, e.g. *erek’kin* ‘threefold’, *č’orek’kin* ‘fourfold’. The starting point must have been *krkin* ‘double’, a reduplicated variant of **d̥uis* > *erkir* supplied with the additional suffix **(i)no-*, i.e. **d̥ui-d̥uis-(i)no-* or the like > **(V)rkirkino-*, haplogogized to *krkin* (Olsen 1989: 7 f.). When the multiplicative par excellence *krkin* was synchronically analyzed as *erkir-* + *-kin* ‘double’, the way was open for formations like *erek’kin* ‘triple, threefold’ (cf. Dan. *tredobbelt*) and even *hingkrkin* ‘fivefold’. The same meaning is also expressed by the suffix *-patik*, apparently an Iranian loan suffix, e.g. *hariwrapatik* ‘hundredfold’.

Iteratives in *-ic’s* are attested with the meanings ‘n times’ and ‘the nth time’, e.g. *erkic’s* ‘twice’, *eric’s* ‘three times’, *vec’ic’s* ‘six times’. The key forms must be *erkic’s* < **d̥uitjo-* (rather than the traditionally assumed, functionally less satisfactory **d̥uis-ko-* as in OS *twisk*, OHG *zwisk[i]* ‘double’) and especially *eric’s* < **tritjo-* + adverbial *-s* (originally the acc.pl. denoting extension in time), cf. Skt. *dvitīya-*, Av. *daibitiia-*, *bitiia-* ‘second’, Skt. *trītiya-*, Lat. *tertius*, Lith. *trėčias* ‘third’. Apparently neuter substantivizations of these ordinals were used adverbially already by the time of the protolanguage, cf., e.g., OP *duvitiyam* ‘for the second time’, Umbr. *duti* ‘iterum’, *tertim* ‘for the third time’.

5. Pronouns

5.1. Apart from the personal pronoun with its notorious intrinsic difficulties, the Armenian pronominal system may be organized into the following categories: 1. the reflexive pronoun, the pronoun *ink’n* ‘(-)self’, the reciprocal, and the possessive pronouns, 2. interrogative, relative, and indefinite pronouns and 3. demonstrative pronouns.

5.2. The reflexive pronoun has a common form for the gen.dat.loc., *iwr*, traditionally derived from **seuero-* (or **seuetero-*, cf. Gk. *ἡμέτερος* ‘our[s]'). If we may be allowed to hypothesize an alternative **setro-* (albeit a form without precise parallel elsewhere), this would yield **ewr*, which might have gotten its *i* from the reflexive adjective, originally **ewr/iwroy* with **ew* becoming *iw* in unstressed position (de Lamberterie 2014) and later leveling to *iwr*. The plural acc. *iwreans*, etc. probably contains the adjective suffix *-ean-* < **-ih₃n(h₂)o-* (cf. Ved. *mākīna-* ‘my?’).

Ink'n '(-)self', also used as a reflexive, seems to be a conglomerate of the pronoun **im-* (acc.) + **-sue/o-*. The paradigm is based on the stem *ink'nean(-)*, again from **-ih₃n(h₂)o-*.

The reciprocal pronouns *mimeans* (acc.) and *irears* (acc.) appear to be derived from reduplicated stems. The first most likely represents **smih₂-smi(a)h₂-*, but the second has no such obvious etymological basis. Rasmussen (1999: 124–127) has suggested **etero-etero-* (cf. OCS *eterŭ ... eterŭ* 'one ... the other', Gk. *ἕτερος ... ἕτερος* 'id.' [with adaptation of the initial to the reflexive pronoun], and structurally Lat. *alter ... alterum* 'id.').

Finally the possessive pronouns *im* 'my', *mer* 'our', *jer* 'your' are identical with the genitive of the personal pronoun (cf. 6), while *k'oy* 'your' ('thy') is really the genitive of *k'o*. They are all inflected as *o*-stems with pronominal dat. *imum* and abl. *immē*, etc. *Nora* (< **no-(te)ro-* + *-ay?*), the genitive of the anaphoric pronoun (cf. 5.4), secondarily inflected as a pronominal *o*-stem (plural also *i*-st.), is used as a third person non-reflexive possessive.

5.3. When used substantively the interrogative pronoun distinguishes between an animate 'who?', nom.acc. *o(v)*, gen. *oyr* (**k^wosjo-* + *-r*, cf. OAv. gen.masc. *kahiiā*), dat.loc. *um* (**k^wo-sm-*, cf. Skt. dat. *kāsmāi*, loc. *kāsmīn*), inst. *orov*, and a neuter 'what?', nom. acc. *z-i* (with the accusative marker *z-*), gen. *ēr* (← **k^wesjo-* + *-r*, cf. OAv. gen.neut. *cahiiā*; Meillet 1936: 87), dat.loc. (*h*)*im* (← **k^we-sm-*), inst. *iw*. The stem *o-* goes back to **k^wo-* with regular loss of initial **k^w-* before the rounded vowel, while *i-* for expected **č'i-* < **k^wi-* seems to have lost its initial consonant by analogy with *o-*. The adjectival variant *or*, probably < **k^wo(te)ro-* (Schmitt 1981: 123), is inflected as a pronominal *o*-stem, i.e. with dat.loc. *orum*, abl. *ormē*. The interrogative pronoun is also used as a relative pronoun with the same distinction between substantival and adjectival forms.

The indefinite pronouns *ok'* 'someone' and *-ik'* 'something' (only *č'-ik'* 'there is not') are created by the addition of **-k^we* to the interrogative stems *o-* and *i-*, cf., e.g., Skt. *kás-ca* 'someone', Lat. *quis-que* 'each, everyone'. According to Klingenschmitt (1982: 100), *omn* 'someone' derives from **k^wos men* 'irgendeiner irgendetwie' or the like.

5.4. The dominant feature of the demonstrative pronouns is the consistent distinction between a first person deictic characterized by *s* (< **k^wo-* 'hic'), a second person with *d* (< **to-* 'iste'), and a third person with *n* (< **no-* 'ille').

The enclitic forms *-s*, *-d*, *-n* are used as definite articles. With a preceding *a(y)-*, we get the demonstrative pronouns *ays*, *ayd*, *ayn* and e.g. the local adverbs *ast*, *aydr*, *and* (ubi); *aysr*, *aydr*, *andr* (quo); *asti*, *ayti*, *anti* (unde). There are three pronouns expressing identity, *soyn*, *doyn*, and *noyn* 'this/that same', and three parallel anaphoric pronouns, *sa*, *da*, and *na* (< **say*, etc. with enclitic *-ay*), thus sg.nom.acc. *sa*, gen. *sora*, probably < **k^wo(te)ro-* + *-ay*, dat.loc. *sma*, abl. *smanē* (*-m-* < **-sm-*), inst. *sovaw*; pl. *sok'a*, *sosa*, *soc'a*, *soc'anē*, *sok'awk'*.

6. Personal pronouns

The paradigms of the personal pronouns are as follows:

	1. person	2. person
sg.nom.	es	du
acc.	is	k'ez
gen.	im	k'o
dat.	inj	k'ez
loc.	is	k'ez
abl.	inēn, injēn	k'ēn, k'ezēn
inst.	inew	k'ew
pl.nom.	mek'	duk'
acc.	mez	jez
gen.	mer	jer
dat.loc.	mez	jez
abl.	mēnj, mezēn	jēnj, jezēn
inst.	mewk'	jewk'

1.sg. The nominative *es* is probably a sandhi variant for expected **ec* < **eġō* with a similar word-final development as in the verb *asem* 'say, speak' (root *ac-* → *as-*, cf. Gk. ἄ < **h₂eġ-t*). The genitive *im* < **eme/o-*, originally a possessive pronoun, also provides the stem of the accusative *is* (< **ims*), analogically extended to the locative based on the plural personal pronouns. The dative *inj* consists of the same stem + **-ġ^{hi}* (cf. Lat. *mihī*) (for an alternative analysis, see Klein 2007: 1065–1066), while the elements of the ablative *inēn* may be **in-* from the dative + ablative suffix *-ē-* < **-eti-* + a particle *-n* (or, more daringly, *-ēn* < **-ed^hen* = Gk. *-εθεν*; cf. Klingenschmitt 1982: 19. Such a phonetic development, while not corroborated by parallels, is also not contradicted by counterexamples.). Here as well as in the other personal pronouns a by-form shows direct derivation from the dative. The instrumental *inew* goes back to *in-e-* (analogical for **eme-* > *ime-*) + **-b^{hi}*.

2.sg. Nom. *du* for expected **t'u* < **tū*, perhaps with enclitic development of the stop as also in the demonstrative *-d* < **-to-*, etc. Gen. *k'o* < **tuo-* is originally the possessive pronoun, cf. Gk. *σός*. From the stem **tue-*, acc.dat.loc. *k'ez* < **tue-ġ^{hi}*, where the full form originally comes from the dative as a substitution for **tue-b^{hi}* (cf. Lat. *tibī*) under the influence of the 1.sg. **me-ġ^{hi}*. Abl. *k'ēn* (*k'ezēn*, with *j* > *z* intervocalically) and inst. *k'ew* are parallel to *inēn* (*injēn*), *inew*.

1.pl. *mek'* is likely to form an exact counterpart to Lith. *mēs* (cf. also OCS *my*) as opposed to Skt. *vayám*, Goth. *weis*, etc. Acc. and loc. *mez* have *z-* from the dative, where it has been analogically introduced from the singular. Gen. *mer* is probably from *me-* + **(t)ero-* (cf. Gk. *ἡμέτερος*) and inst. *mewk'* from **me-* + *b^{hi}-s*, while abl. *mēnj* is unclear.

2.pl. *duk'* for **juk'* (?) < **iūs* with influence from 2.sg. *du*. Acc. (and loc.) *jez* with stem *je-*, where the vocalism may have been taken over from the 1.pl. and *-z* from the dative. Gen. *jer* < *je-* + **(t)ero-* (?), and abl. and inst. similar to 1.pl.

7. Verbs

7.1. The Armenian verbal system has undergone a series of fundamental changes since the IE protolanguage, and the end result is a fairly simple and regular system. The finite verb distinguishes between three persons, two numbers (singular and plural), three moods (indicative, subjunctive, and imperative), two verbal voices (active and mediopassive), two tenses (present and preterite), and two aspects (imperfective and aorist). As is also the case with the noun, the dual is extinct, while the optative and the perfect only survive in synchronically opaque relic forms.

The main formal distinction of an Armenian verb is between the two aspectual stems: 1. the imperfective or present stem, from which is formed the present and imperfect indicative, the present subjunctive and imperative, and the infinitive (with secondary derivatives); apart from present indicative actives in *-em* connected with mediopassives in *-im*, the imperfective has no distinction of verbal voice; 2. the aorist stem, serving as the basis of the aorist indicative, subjunctive and imperative, and the participle; finite forms of the aorist stem distinguish between active and mediopassive inflections.

7.2. The present indicative shows four parallel vocalic types, *e-*, *i-*, *a-*, and *u-*stems, of which the *i-*stems (from statives in **-eh₁-*) function as mediopassives or intransitives. Thus from *berem* 'I carry', *berim* 'I am carried', *lam* 'I weep', *hehum* 'I pour':

	e-st.	i-st.	a-st.	u-st.
sg.1.	berem	berim	lam	hehum
2.	beres	beris	las	hehus
3.	berē	beri	lay	hefu
pl.1.	beremk'	berimk'	lamk'	hehumk'
2.	berēk'	berik'	layk'	hehuk'
3.	beren	berin	lan	hefun

A single *o-*stem, *gom* 'I am, exist' (originally perfect of the root **h₂ues-*), is almost exclusively used in the 3.sg. *goy*.

This inflection represents a continuation of the Indo-European active presents where the *e-*stems constitute a compromise between the synchronically regular verb *em* 'I am' (*em*, *es*, *ē*, *emk'*, *ēk'*, *en*) and the thematic present with generalized thematic vowel **-e-*: thus 1.sg. *em* for **im* with analogical stem vowel < **h₁esmi*, whence *berem* (for **ber* < **b^herō*); 2.sg. **h₁essi* > *es* → *beres* (for **berē* < **b^heresi*); 3.sg. *b^hereti* > *berey* > *berē* → *ē* (for **est* [?] < **h₁esti*) and similarly **beriy* > *beri*, **hehuy* > *hehu*; 1.pl. **b^heromes* → *beremk'* with analogical stem vowel; hence also *emk'* (for **am(k')* < **h₁smes*); 2.pl.

**b^heretes* > **bereyk'* > *berēk'* → *ēk'* (for **ast* [?] < **h₁ste*) and similarly **beriyk'* > *berik'*, **heluyk'* > *heluk'*); 3.pl. *en* for **an* < **h₁sentī* and *beren* for **berun* (< **b^heronti*). The expected outcome of the weak forms depends on the development of **h₁-* as a prothetic vowel.

7.3. The formation of the imperfect is less transparent than that of the present. Thus, the joint paradigm of *berem* ‘I carry’ and *berim* ‘I am carried’:

	sg.	pl.
1.	berēi	berēak'
2.	berēir	berēik'
3.	berēr < *-eyr	berēin

Similarly, from *lam* ‘I weep’, *layi*, etc.; and from *helum* ‘I pour’, *helui*, etc. (3.sg. *heloyr*). As is also the case with the present, the imperfect of the verb ‘to be’, *ei*, *eir*, *ēr*, *eak'*, *eik'*, *ein*, is identical with the synchronic endings of the *e*-verbs.

The historical background of the endings is disputed: Winter (1975) assumes influence from the old optative, while Jasanoff (1979) and Klingenschmitt (1982) propose to derive the 1.sg. *-i* from the imperfect **e-h₁es-η* > **ēsη* > **ia(n)* (cf. already Meillet 1936: 126, adducing a postposed “perfect” identical with Gk. *ἦα*, Skt. *āsa*). Since the inherited imperfect paradigm with secondary endings would be extremely vulnerable, yielding zero endings throughout the singular, it seems quite plausible that the full forms of the copula were secondarily added for clarification, but while the 1.sg. would seem to have been added to the stem, the 2.sg. is open to a morphologically more straightforward analysis: *(*e*)*b^heres-* + *(*e*-*h₁es-s* >) **ēs* ‘you carried, it was you’ > *berē-ir* (with early juxtaposition and ruki-development); the outcome of a similarly structured 3.sg. *(*e*)*b^heret-ēs* ‘he carried, it was he’ is less clear, but if we assume an auslaut development **-st* > **-ss* > **-s* > *-r* (after *-i/-y-*), *berēr* could be regular, and the impact of the 2. and 3.sg. would probably be enough to establish *-i-* as an imperfect marker, causing the eventual merger between active and mediopassive in this tense. [*ed.* On these 2. and 3.sg. developments, see *ed.* note 4.4 above as well as the references that precede it.] The 1.pl. *-eak'* is possibly an analogical substitution for *-eamk'* (**ēsηmes*?) under the influence of the aorist subjunctive *-uk'* (Jasanoff 1979: 140). Finally, the post-classical 3.sg. mediopassive *beriw* is clearly a secondary form created in response to a perceived need. The most likely solution is that the speakers took the mediopassive stem *beri-* and added the unambiguously mediopassive 3.sg. aorist ending *-w* followed by the 3.sg. imperfect ending *-r*.

7.4. The present imperative, used only in prohibitive function in connection with the particle *mi* < **mē* (Skt. *mā*, Gk. *μη*), has the endings *-r* in the 2.sg. (*berer*, *berir*, *lar*, *helur*), *-yk'* in the 2.pl. (**bereyk'* > *berēk'*, **beriyk'* > *berik'*, *layk'*, **heluyk'* > *heluk'*). The singular ending may have been generalized from aorist injunctives like *dir* < **d^heh₁-s* and *tur* < **doh₃-s* with the ruki-development of **-s* > *-r* after the preserved *-i* and *-u* of monosyllabics. [*cf.* again *ed.* note, 4.4 above]. Alternatively, we might be dealing here with the change of the 2 sg. act. imperative ending **-d^hi* to *-r* (cf. Macak, this handbook, 9.4.8 with references). As expected on comparative grounds, the plural is identical with the present indicative.

7.5. The aorist is found in two variants, a “strong” or “root”-aorist and a “weak” aorist characterized by the complex morpheme *-eac’-* (> *-ec’-* in unstressed position; more rarely simply *-c’-*, e.g. *lc’i* from *lnum* ‘fill’, root **pleh₁-*), thus from *berem* ‘I carry’ and *sirem* ‘I love’:

Active

sg.1.	beri	sirec’i
2.	berer	sirec’er
3.	e-ber	sireac’
pl.1.	berak’	sirec’ak’
2.	berēk’, -ik’	sirec’ēk’, -ik’
3.	berin	sirec’in

Medio-passive

sg.1.	beray	sirec’ay
2.	berar	sirec’ar
3.	beraw	sirec’aw
pl.1.	berak’	sirec’ak’
2.	berayk’	sirec’ayk’
3.	beran	sirec’an

The active secondary endings are reminiscent of those of the imperfect, but with an endingless 3.sg. < **-et*, where the augment is preserved in forms that would otherwise become monosyllabic. The mediopassive endings, apparently added to an extended stem in *-a-*, are explained on the basis of 3.pl. **-ŋto* > *-an*, supported by a 3.sg. *-aw* which is regular in *seṭ* roots, e.g. *cnaw* ‘begot, gave birth’ < **(e)-ġenə₁to*.

Neither the aorist marker *-c’-* nor the preceding *-ea-* is fully understood, though the isolated form *ənkēc’* ‘threw’ suggests that the *-a-* must somehow be secondary (cf. Klein 2007: 1074 f.). The traditional derivation of *-c’-* from **-sĕ-* is formally objectionable and functionally somewhat shaky; and if Klingenschmitt (1982: 286 f.) is right in assuming a strengthened **-ss-* > *-c’-* based on the old *s*-aorist, it is difficult to imagine a model explaining the analogy postulated for *es* < **essi* ‘you are’ (cf. also *k’os* ‘scab’ < **kos-so-*, Lith. *kasýti* ‘scratch’). However, with a slight modification the main idea of an *s*-aorist is perhaps still the best option: As noted by Klingenschmitt himself, most archaic-looking *c’*-aorists occur with stems ending in a laryngeal: *lc’i* ‘I filled’ < **pleh₁-s-* (Gk. *πλη-σ(α)-*), *stac’ay* ‘I acquired’ < **stah₂-s-*, *bac’i* ‘I opened’ < **bah₂-s-*, *tac’i* ‘I dipped’ < **tah₂-s-*, and aorists in *-ac’i/-ac’ay* connected with denominatives in *-a-* and *-ana-* (cf. the Gk. type *τιμη-σ[α]-*), and if we accept the idea of “laryngeal hardening” before final **s* (as has been proposed for Lat. **senah₂-s* > **senex* ‘old person’), the development of **-h-s(-)* > **-ks(-)* > *-c’(-)* might have been extended from 2. and 3. sg. in **-h-s-s*, **-h-s-t* and perhaps the imperative in **-h-s*. If further the type in **-ah₂-s-* > *-ac’-* became

sufficiently successful, it may have been added to the thematic stems, thus yielding the productive type in 3.sg. *-eac'*.

7.6. The subjunctive is characterized by a morpheme *-ic'*. Thus, the active of the “strong” aorist *beric'*:

	sg.	pl.
1.	beric'	berc'uk'
2.	berc'es	berjik'
3.	berc'ē	berc'en

From the “weak” aorist *sirec'i* (prs. *sirem*) ‘love’ the corresponding forms are *sirec'ic'*, *siresc'es*, *siresc'ē*, *siresc'uk'*, *siresjik'*, *siresc'en* with dissimilation **-c'c'- > -sc'-*, **-c'j- > -sj-*. The aorist subjunctive serves to express the future tense.

The present subjunctive is of a more recent character: Here the new subjunctive marker *-ic'* is simply added to the stem and followed by the existing present endings: *beric'em*, *beric'im*, *layc'em*, *heluc'um* (for **heluc'em*).

Obviously, a simple continuation of the IE optative would have been inconvenient since only the root syllable would be left of a 3.sg. **b^heroj_ht*. The old, unextended subjunctive would be equally unsuitable, as it would have become identical with the mediopassive *i*-inflection (**-ē- > -i-*). However, the traditional derivation of *-c'* from an **s^hk-* suffix (e.g. Schmitt 1981: 143) whereby the preceding *-i-* is variously explained, is at least phonetically implausible (internal **-s^hk- > -č-* before front vowels, cf. Olsen 1987. For the alternative view that this development is securely established only for the original labiovelar **k^w*, see Macak, this handbook, 9.2.17–19.)

As an alternative historical background of the subjunctive paradigm (Olsen 1988), it may be assumed that the 3.sg., the unmarked form or “non-person” of the inherited subjunctive with generalized lengthened thematic vowel **-ē-*, served as a derivational basis of the paradigm, 1.sg. **b^herētj-* + *-ō-*, 2.sg. *-essi*, 1.pl. *-o-mes*, 3.pl. *-enti*. After the establishment of a stem *beric'*, the usual ending *-ē* was added to the opaque form of the 3.sg., *beric' + -ē > berc'ē*. The 2.pl. *berjik'* represents an old optative **-ji-iek^h < *-jeh₁-tes* (Klingenschmitt 1992: 40). The suggested scenario would be reminiscent of cases like the Old Indic precative 3.sg. *bhūyās* ⇒ 1.sg. *bhūyāsam*, etc. or Modern Persian 3.sg. *hast* ‘he is’ ⇒ 1.sg. *hastam*, etc. (For a different view of the subjunctive, see Klein 2007: 1072 f.)

The medio-passive aorist subjunctive has the following forms:

	sg.	pl.
1.	berayc'	berc'uk'
2.	berc'is	berjik'
3.	berc'i < *-iy	berc'in

From the “weak” aorist correspondingly *sirec'ayc'*, *siresc'is*, *siresc'i*, *siresc'uk'*, *siresjik'*, *siresc'in* with the same dissimilatory development as in the active forms.

Here the mediopassive *i*-endings correspond rather neatly to those of the present (*-i* < stative **-eh₁-*). 1.pl. *-uk'* may be derived from **-omes* with surviving thematic vowel **-o-*.

7.7. In opposition to the prohibitive present imperative, the aorist imperative is used only in positive orders. In the active, the 2.sg. of the “strong” aorist equals the stem, e.g. *ber* < **b^here*, while the 2.pl. has the ending *-ēk'* (< **-etes*) or *-ik'*, i.e. *berēk'/berik'*. The corresponding forms of the “weak” aorist may be exemplified by 2.sg. *sirea*, 2.pl. *sirec'ēk'/sirec'ik'*. The mediopassive forms are 2.sg. *ber/berir*, 2.pl. *beraruk'* and 2.sg. *sireac'*, 2.pl. *sirec'aruk'*. Perhaps the origin of the 2.pl. ending is **(a)-d^huue* (Jasanoff 1979: 144 f.; cf., e.g., Skt. middle ipv. *bharadhvam*), if *-r-* is accepted as a possible conditioned continuation of **-d^h-* beside *-z-* (and *-ž-?*) (cf. 7.4 above and Macak, this handbook, 9.3.10).

7.8. The Armenian infinitive, which is indifferent with respect to verbal voice, is characterized by the suffix *-l* (*o*-st.) < **-lo-* added to the imperfective stem, i.e. *berel*, *lal*, *hehul* from *berem/berim*, *lam*, *hehum* respectively.

The infinitive stem in turn is the basis of an indeclinable future participle in *-oc'*, e.g. *bereloc'* ‘to be carried’, apparently a frozen case form of a stem in **-sKo-*, and a “gerund” in *-eli* (*ea*-st.), e.g. *sireli* ‘lovable’ (cf., e.g., Toch.B. *mā yokalle* ‘one should not drink’).

The only genuine participle is formed by adding *-eal* (gen. *-eloy*) to the aorist stem, e.g. *beréal*.

The old participles in **-ont-* (active) and **-o-mə₁no-* (middle) have lost their paradigmatic connection with the verb, apparently merging in the suffix *-un* (*o*-st.), e.g. *gitun* ‘knowing’, i.e. ‘wise man’ (active) vs. *sirun* ‘beloved, dear’ (passive).

7.9. The stem formation of the Armenian imperfect stems reflects a number of inherited types more or less directly.

Thus verbs in *-em* represent a merger of several IE types: Simple thematic presents, e.g. *berem* ‘I carry’ < **b^here/o-*, *acem* ‘I conduct’ < **h₂aǵe/o-*; *i*-presents (with palatalization of the root-final consonant), e.g. *šč'em* ‘I hiss’, *mrnmjēm* ‘I mumble’; *sk̄*-verbs *čanač'em* ‘I know’, for **canač'em* < **ǵ^hh₃-sk̄e/o-*, *alač'em* ‘implore’ < **s^lh₂-sk̄e/o-* (Gk. *ἰλάσκεσθαι*), *amač'em* ‘I am ashamed’ < **^hh₃-sk̄e/o-* (against Meillet 1936: 109, followed by Macak, this handbook, 9.2.17–19, who would require **-sk̄je/o-* in each of these forms); denominatives like *lusaworem* ‘I bring light’ (‘am a *lusawor*’), *gorcem* ‘I work’ (‘I make work’, *gorc*) < **uorǵéje/o-*, *anuanem* ‘I call’ (*anun* ‘name’), *srbem* ‘I cleanse’ (‘I make clean’, *surb*); and old perfects such as *gitem* ‘I know’ < **uojd-*.

Verbs in *-anem* (mediopassive and intransitive *-anim*) are connected with old nasal presents, e.g. *lk'anem* ‘I leave’: Skt. *riṇákti*, Lat. *linquō*, *gtanem* ‘I find’, *dizanem* ‘I heap’, *awcanem* ‘I anoint’ (corresponding “root” aorists 3.sg. *elik'*, *egit*, *edēz*, *awc*). This pattern includes the complex subtypes in **-sk̄-* + *-anem*, e.g. *harc'anem* ‘ask’ (**pr̄k̄-sk̄-* + *-anem*) and **-j-* + *-anem/-anim*, e.g. *hecanim* ‘ride’ < **sed-je/o-*.

An important productive subtype is constituted by the causatives and factitives in *-uc'anem*, aor. *-uc'i* (3.sg. *-oyc'*), synchronically derived from the aorist stem, e.g. *usanim*, aor. *usay* ‘learn’ → *usuc'anem*, aor. *usuc'i* ‘teach’. Traditionally the causatives are explained by the ubiquitous **-sk̄-* suffix, but perhaps they should preferably be considered denominatives of nouns in **-eh₁uti-* > *-oyt'* (later generally extended to *-ut'awn*), cf., e.g.,

yarut'iwn 'resurrection' vs. caus. *yaruc'anem* 'I raise' ('I make a rising'), aor. *yaruc'i*, as if < *-*eut-je/o-*.

The background of the mediopassives in *-im* is to be found in the IE statives in *-*eh₁-*, cf., e.g., *nstim* 'I sit' < **ni-* + *s(e)deh₁-* (Lat. *sedēre*). As for the intransitives in *-č'im*, the inherited pattern appears to be stative inchoatives in *-*eh₁-skē/o-*, thus e.g. *hangč'im* 'I rest' < **sm-k^wieh₁-skē/o-* (aor. *hangeay*), cf. Lat. *con-quiēscō* (see above on *čanač'em*, etc.).

The type in *-am* is represented by old root presents like *bam* 'I say' (Lat. *fāri*), reduplicated presents such as *tam* 'I give' < **di-də₃-*, and deponents like *gt'am* 'I feel compassion' with corresponding mediopassive aorists.

Presents in *-C-nam* have the function of factitives: *bařnam* 'I raise' (< **barjnam*, adj. *barjr* 'high'), *dařnam* 'I turn' (aor. *darjay*), *spařnam* 'I threaten' (cf. Lat. *spernō*, ON *sporna*, Skt. aor. *aspharīti*), while the type in *-anam* is richly represented by denominatives with the meaning 'become/be X', e.g. *hiwandanam* 'I get ill, am ill' (*hiwand*), *lusanam* 'I become light' (*loys*), *ceranam* 'I grow old' (*cer*), *ylenam* 'I get pregnant' (*yli*).

Finally the type in *-num* contains verbs based on roots accidentally ending in *-u-*, e.g. *gelum* 'I turn': Lat. *volvō*, Skt. *vr̥ṇóti* (older *ūr̥ṇóti*), and old or more recent formations in *-*nu-*, e.g. *z-genum* 'I dress' < **ues-nu-mi* = Gk. *ἔννυμι*, *ařnum* 'I take' (Gk. *ἄρνωμαι*), including substitutions for older *-*neh-/nə-*presents such as *lnum* 'I fill' < **plēnu-* vs. Skt. *pr̥ṇāti*.

8. References

- Eichner, Heiner
1978 Die urindogermanische Wurzel **Hreu* 'hell machen'. *Die Sprache* 24: 144–162.
- Godel, Robert
1975 *An Introduction to the Study of Classical Armenian*. Wiesbaden: Reichert.
- Jasanoff, Jay
1979 Notes on the Armenian Personal Endings. *Zeitschrift für vergleichende Sprachforschung* 93: 133–149.
- Klein, Jared S.
2007 Classical Armenian Morphology. In: Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, Vol. 2. Winona Lake, IN: Eisenbrauns, 1051–1086.
- Klingenschmitt, Gert
1982 *Das altarmenische Verbum*. Wiesbaden: Reichert.
- Kortlandt, Frederik
1984 Proto-Armenian Case Endings. In: Gevorg B. Ĵahukyan and Andrzej Pisowicz (eds.), *International Symposium on Armenian Linguistics: Erevan, September 21–25, 1982: Reports*. Erevan: Armenian Academy Press, 97–106.
- de Lamberterie, Charles
1989 Introduction à l'arménien classique. *LALIES* 10 [1988–1989]: 233–289.
- de Lamberterie, Charles
2014 Arménien *իւր* iwr et *հիւր* hiwr, Grec *ἑταῖρος* : Le "soi" et le "sien", "l'hôte" et le "compagnon". In: Aram Mardrossian, Agnès Ouzounian, and Constantin Zuckermann (eds.), *Mélanges Jean-Pierre Mahé*. Paris: Collège de France – CNRS, 429–438.

- Lindeman, Fredrik Otto
1980 Indo-European and Classical Armenian: A Phonological Note. In: John. A. C. Greppin (ed.), *Proceedings of the First International Conference on Armenian Linguistics*. Delmar, NY: Caravan, 59–66.
- Matzinger, Joachim
1992 *Der armenische Reflex von idg. *-u-*. Unpublished ms.
- Matzinger, Joachim
1995 *Untersuchungen zur altarmenischen Deklination*. Unpublished Thesis, Vienna.
- Matzinger, Joachim
2005 *Untersuchungen zum altarmenischen Nomen: Die Flexion des Substantivs*. (Münchener Studien zur Sprachwissenschaft, Beiheft 22). Dettelbach: Röhl.
- Meillet, Antoine
1913 *Altarmenisches Elementarbuch*. Heidelberg: Winter.
- Meillet, Antoine
1936 *Esquisse d'une grammaire comparée de l'arménien classique*. Vienna: Mekhitharistes.
- Melchert, H. Craig
1994 *Anatolian Historical Phonology*. Amsterdam: Rodopi.
- Olsen, Birgit Anette
1987 On the origins of Arm. *c'*. *Arbejdspapirer udsendt af Institut for Lingvistik, Københavns Universitet* 6: 67–87.
- Olsen, Birgit Anette
1988 The Formation of the Subjunctive Paradigm in Classical Armenian. *Annual of Armenian Linguistics* 9: 7–14.
- Olsen, Birgit Anette
1989 Three notes on Armenian historical phonology. *Annual of Armenian Linguistics* 10: 5–25.
- Olsen, Birgit Anette
1999 *The Noun in Biblical Armenian: Origin and Word-formation: with Special Emphasis on the Indo-European Heritage*. Berlin: Mouton de Gruyter.
- Olsen, Birgit Anette
2010 *Derivation and Composition. Two studies in Indo-European word formation*. Innsbruck: Institut für Sprache und Literatur der Universität.
- Pedersen, Holger
1905 *Les pronoms démonstratifs de l'ancien arménien*. Copenhagen: Bianco Luno.
- Pedersen, Holger
1905 Zur armenischen sprachgeschichte. *Zeitschrift für vergleichende Sprachforschung* 38: 194–240.
- Pisani, Vittore
1944 Armenische Studien. *Zeitschrift für vergleichende Sprachforschung* 68: 157–177.
- Rasmussen, Jens Elmegård
1999 Miscellaneous Problems in Indo-European Languages I–II. 8. Armenian *irears* 'each other'. *Selected Papers in Indo-European Linguistics*. Copenhagen: Museum Tusulanum, 124–127. [First published 1985 [1987] in *Lingua Posnaniensis* 28.]
- Schmitt, Rüdiger
1981 *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Seldeslachts, Herman
1991 *Varia Etymologica*. In: Lambert Isebaert (ed.), *Studia Etymologica Indoeuropaea Memoriae A. J. van Windekens Dicata*. Leuven: Peeters, 247–261.
- Solta, Georg Rhenatus
1963 Die armenische Sprache. In: Bertold Spuler (ed.), *Handbuch der Orient Orientalistik. I/7: Armenisch und kaukasische Sprachen*. Leiden: Brill, 80–128.

Szemerényi, Oswald

1960 *Studies in the Indo-European system of numerals*. Heidelberg: Winter.

Weitenberg, Jos J. S.

1989 The inflection of proper names in Armenian. *Annual of Armenian Linguistics* 10: 57–72.

Winter, Werner

1975 Die Personalendungen des Imperfekts und Aorists im Armenischen. *Zeitschrift für vergleichende Sprachforschung* 89: 110–121.

Winter, Werner

1992 Armenian. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: De Gruyter, 347–359.

Birgit Anette Olsen, Copenhagen (Denmark)

63. The syntax of Classical Armenian

- | | |
|---|---|
| 1. Introduction | 7. Word order |
| 2. The noun phrase | 8. Discourse syntax: conjunction and deixis |
| 3. Complementation and adjunction | 9. Varia |
| 4. Clause types | 10. References |
| 5. Tense, mood, and aspect | |
| 6. Passivization, diathesis, causativity, and <i>genera verbi</i> | |

1. Introduction

1.1. The present survey is based primarily on my own analysis of Chapters 1–15 of the Gospel of Matthew, to which I have added the non-translational selections at the back of Thomson ([1975] 1989). The edition of the Classical Armenian Gospels I have utilized is Künzle (1984), who has produced a critical edition of manuscripts E (A.D. 989 and better preserved) and M (A.D. 887), presenting the text of E with variant readings of M indicated in the notes. I have followed him in this regard, with variations in M represented in brackets ([]) only when relevant to the issue at hand. Because our main corpus is taken entirely from Matthew, Gospel passages cited will be identified only by chapter and verse number, except for those relatively few cases where I have had to go outside this subcorpus to illustrate a point. I have reduced the punctuation of both Künzle's and Thomson's texts to two points: one on the line (with spacing to either side) and one about halfway up relative to the height of the letters, as well as a question mark to indicate the word that receives rising intonation in questions. I have maintained Künzle's capitalization, but do not place a line over holy words abbreviated in the text (YS 'Jesus', K'S 'Christ', TR/TN/TĒ different case forms of 'Lord', etc.). Univerbated items like *z-* and *-n* will be separated by a hyphen from their hosts only in those sections where they are discussed as features of grammar. I identify passages from Thomson according to his system (e.g. Eznik II, 3), adding the page numbers from his book before the citation.

<https://doi.org/10.1515/9783110523874-018>

Szemerényi, Oswald

1960 *Studies in the Indo-European system of numerals*. Heidelberg: Winter.

Weitenberg, Jos J. S.

1989 The inflection of proper names in Armenian. *Annual of Armenian Linguistics* 10: 57–72.

Winter, Werner

1975 Die Personalendungen des Imperfekts und Aorists im Armenischen. *Zeitschrift für vergleichende Sprachforschung* 89: 110–121.

Winter, Werner

1992 Armenian. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: De Gruyter, 347–359.

Birgit Anette Olsen, Copenhagen (Denmark)

63. The syntax of Classical Armenian

- | | |
|---|---|
| 1. Introduction | 7. Word order |
| 2. The noun phrase | 8. Discourse syntax: conjunction and deixis |
| 3. Complementation and adjunction | 9. Varia |
| 4. Clause types | 10. References |
| 5. Tense, mood, and aspect | |
| 6. Passivization, diathesis, causativity, and <i>genera verbi</i> | |

1. Introduction

1.1. The present survey is based primarily on my own analysis of Chapters 1–15 of the Gospel of Matthew, to which I have added the non-translational selections at the back of Thomson ([1975] 1989). The edition of the Classical Armenian Gospels I have utilized is Künzle (1984), who has produced a critical edition of manuscripts E (A.D. 989 and better preserved) and M (A.D. 887), presenting the text of E with variant readings of M indicated in the notes. I have followed him in this regard, with variations in M represented in brackets ([]) only when relevant to the issue at hand. Because our main corpus is taken entirely from Matthew, Gospel passages cited will be identified only by chapter and verse number, except for those relatively few cases where I have had to go outside this subcorpus to illustrate a point. I have reduced the punctuation of both Künzle's and Thomson's texts to two points: one on the line (with spacing to either side) and one about halfway up relative to the height of the letters, as well as a question mark to indicate the word that receives rising intonation in questions. I have maintained Künzle's capitalization, but do not place a line over holy words abbreviated in the text (YS 'Jesus', K'S 'Christ', TR/TN/TĒ different case forms of 'Lord', etc.). Univerbated items like *z-* and *-n* will be separated by a hyphen from their hosts only in those sections where they are discussed as features of grammar. I identify passages from Thomson according to his system (e.g. Eznik II, 3), adding the page numbers from his book before the citation.

<https://doi.org/10.1515/9783110523874-018>

1.2. A preliminary word of caution is warranted. One will not find here the type of configurational analysis (pure syntax) that is today increasingly being demanded of ancient or archaic languages. For examples of this kind of analysis, see the contributions of Hale and Keydana in this handbook. In fact, even traditional syntax (i.e. morphosyntax) has not been extensively pursued for Classical Armenian. The most exhaustive overall study in the common languages of scholarship is still Jensen (1959: esp. 133–224). Other studies, largely dealing with particular topics (a favorite is the use of the demonstrative articles), include Jungmann (1964–1965); de Lamberterie (1997); Lyonnet (1933); Meillet (1897–1898 [repr. 1962]); Minassian (1988); Ouzounian (1992, 1996); as well as my own investigations (Klein 1996a [synopsis 1996b], 1997, 2011).

2. The noun phrase

2.1. Of the eight cases traditionally reconstructed for PIE, Classical Armenian lacks only the vocative. These possess their normal values found, for example, in Sanskrit, meaning that there are no large-scale semantic case mergers, as in Greek, Latin, Germanic, and Celtic. Nevertheless, Armenian shows some highly original features. First, the preposition *z-* (always unverbated with its object) serves as a *nota accusativi* preceding the definite accusative direct object. Second, the ablative is always preceded by the preposition *i* (before a vowel, *y-* with univibration); and this is frequently true of the locative as well: 1.2 *Abraham cnaw z-Isahak* ‘Abraham begat **Isaac**’ and 3.6 *mkrtein i nmanē i Yordanan* ‘they were baptized **by him in the Jordan**’. This latter passage illustrates the use of *i* in both case roles as well as the employment of *i* + ablative to signal the agent with a passive verb. The employment of *z-* illustrated in 1.2 allows a differentiation, in the case of singular definite direct objects, of nominative and accusative, which are morphologically distinguished only in personal pronouns.

2.2. Classical Armenian lacks the category of grammatical gender even in its anaphoric and deictic pronouns, and consequently possesses a highly structured pronominal reference system based on spatial deixis. Demonstrative pronouns show three series coded with *s*, *d*, and *n*, respectively (*ays/ayd/ayn* ‘this/that’, *sa/da/na* ‘this one/that one’, *-s/-d/-n* ‘this/that/the’). The first member in each set has proximal and the third distal reference. The second is intermediate between the two and frequently signals that a referent is located in the sphere of the addressee. Consequently, the system may be conceptualized as one of personal deixis: first person *s*, second person *d*, third person *n*. Of the three triple sets cited above, the last is suffixed to a nominal form as a kind of demonstrative article with fundamentally anaphoric value, and within this set the form in *-n* functions as a neutral-deictic definite article. As such, when it occurs with an accusative object, the object is normally marked by *z-*: 2.7 *Hērovdēs galt koč'eac' z-mogs-n* ‘Herod secretly called **the Magi**’. However, the rules for the occurrence of *-n* are complex and subject to much variation. This is particularly true in those instances where a noun is accompanied by a possessive pronoun or pronominal adjective, which normally follows its head. In these instances, manuscripts E and M frequently disagree in ways that are not themselves consistent. Thus, in 5.1 E has *ašakertk'-n nora* ‘his disciples’, while M reads *ašakertk' nora*. In noun plus genitive constructions, in which the genitive normally fol-

lows its head noun, the place of the demonstrative article, which normally occurs only once to determine the syntagm as a whole, is also subject to variation. The article usually follows the genitive but may follow the head noun: 9.23 *ibrew ekn YS i town išxani-n* ‘when Jesus came into the house **of the prince** ...’ but 3.10 *tapar ař armi-n cařoc’ dni* ‘an axe is placed at **the root** of the trees’. If the genitive is a proper name, it generally avoids the article, which then is suffixed to the head: 1.11 *ař gerowt’eamb-n Babel-ac’woc’* ‘at the time of **the** Babylonian **captivity**’. In a simple collocation of adjective and noun (which may show the order NA or AN), the noun bears the article: 1.18 *i hogwoy-n srboy* ‘from **the** holy **spirit**’. The article *-n* frequently commutes with a place-adverbial *andr/and/anti* ‘thither/there/thence’ in identical value: 2.16 *xabec’aw i mogowc’ anti* (M: *i mogowc’-n*) ‘he was deceived **by the Magi**’ (also *i* + loc. + *and* and *i* + acc. + *andr*). When a proper noun is followed by a definite appositive, the latter normally bears the article: 1.19 *Yovsēp’ ayr-n nora* ‘Joseph, her **husband**’. While the *-n* demonstrative achieves article status because of its neutral-deictic anaphoric value, the *-s* and *-d* suffixes are employed in syntagms having clear proximal or second person value, respectively: 15.15 *meknea mez z-ařak-s z-ays* ‘explain to us **this parable**’ and 2.13 *ari ař z-manowk-d ew z-mayr iwr* ‘arise, take **thy child** and his mother’. In the first of these instances, the postposed proximal demonstrative *ays* (which takes the *nota accusativi* together with its head noun *ařak*) co-occurs with the corresponding demonstrative suffix *-s*; in the second, *-d* approximates in function a possessive adjective, as the angel, addressing Joseph, refers to Joseph’s child Jesus. For the discourse-level implementation of this system, see 8.5.

2.3. The demonstrative article may serve as a determiner of a definite relative or other subordinate clause, where it typically follows the first stressed word after the subordinator, irrespective of its part of speech: 1.22 *zi lc’ci or asac’aw-n i TĒ* ‘in order that **what was said** by the Lord should be fulfilled’ and 7.11 *et’e dowk’ or č’ark-d ěk’ gitĕk’ pargews baris tal ...* ‘if you **who are evil** know how to give good gifts ...’. In the first of these instances, *-n* is attached to the verb form *asac’aw* ‘was said’, and in the second, *-d* is attached to *č’ark’* ‘evil’ in a heavily marked second person context (*dowk’* ‘you [pl.], *ěk’, gitĕk’* [both 2 pl. verb forms]). Finally, the demonstrative article may possess a quotative function: 5.37 *elic’i jer ban · ayo-n . ayo · ew oč’-n . oč’* ‘Let your word be **“aye, aye”** and **“nay, nay”**’, 12.7 *et’e giteik’ zinč’? ē zolormowt’iwn-n kamim ew oč’ zzoh* ‘if you knew what is (the meaning of) **the (passage)**, **“I desire mercy and not sacrifice”**’.

2.4. The employment of the *nota accusativi z-* and its relationship to the demonstrative article is also complex. When the noun in question can, pragmatically speaking, only have a definite reading, it may take *z-* without an article: 5.5 *zi nok’a žařangesc’en z-erkir-Ø* ‘for they shall inherit **the earth**’, where *erkir* ‘earth’ is inherently definite. In a noun + genitive syntagm or a simple adjective + noun or noun + adjective collocation both the head noun and its dependent element may be marked with *z-* whether or not an article is present: 12.4 *eker z-hac’ z-ařajaworowt’ean-n* ‘he ate **the shew bread** [lit. the bread, that of presentation]’; 2.4 *žoloveal z-amenayn-Ø k’ahanayapets* [M: *-n*] ‘gathering **all the high priests**’; 1.25 *minč’ew cnaw z-ordi-n iwr z-andranik* ‘until she begat her **first born son**’. Fluctuation is also seen in the case of appositives: 4.18 *etes erkows elbars · z-Simovn z-koč’ec’cal* [M: *-n*] *Petros · ew z-Andreas Ø-elbayr nora* ‘he saw

two brothers: **Simon** [M: **the one**] **called** Peter and **Andreas** his **brother**', where the first appositional phrase is marked by *z-* and the second is not. Finally, predicate-like 'as'-adjuncts do not take *z-* even when construed with a *z-*-marked primary object: 3.9 *ownimk' mek' hayr z-Abraham* 'we have Abraham **as (our) father**' and 15.9 *owsowc'an-en vardapetowt'iwns z-[mardkan patowireals]* 'they teach **as doctrines** the things commanded by men', where, as indicated by our brackets, *z-* marks a definite phrase consisting of genitive (*mardkan*) plus head noun (*patowireals*: a substantivized participle) without an article. The order genitive plus head noun seen in this passage is unusual.

2.5. An adjective is inflected when it follows a head noun but not when it precedes. Thus, 1.18 *i hogw-oy-n srb-oy* 'by the **holy** spirit' but 7.15 *zgojš lerowk' i sowt-Ø margarē-ic'* 'be wary of **false** prophets'. But a predicate adjective is normally not inflected even if it follows a subject: 9.37 *hownj-k' bazowm-Ø en · ew mšak-k' sakaw-Ø* 'the harvest is **great** and the workers **few**'. In numeral constructions only '1' through '4' are regularly declined whether they precede or follow a noun: 14.17 *oč' inč' ownimk' ast · bayc' hing-Ø nkanak ew erkow-s jkown-s* 'we have nothing here but **five** loaves of bread and **two** fish' (*nkanak* 'loaf of bread' is a singular collective); and nouns following higher uninflected numerals are typically themselves uninflected for number: 4.2 *paheal z-k'ařasown tiw-Ø ew z-k'ařasown giřer-Ø* 'fasting for **forty days** and **forty nights**'.

2.6. The final noun-phrase constructions we shall consider are comparatives and superlatives. Where the standard of the comparison is a noun (phrase) Classical Armenian employs *k'an + z-* (+ acc.) as the pivot, while the comparative form may represent either the simple positive, the positive followed by the suffix *-agoyn*, or an inherently comparative word like *aweli* 'more'. Examples are 11.11 *č'ē yarowc'eal i cnownds kanc' mec k'an z-Yovhannēs mkrtič'* 'there has not arisen in the generations of women one **greater than John** the Baptist'; 11.22 *diwragoyn lic'i erkri-n Tiwrosi ew Sidovni yawowr-n datastani k'an jez* 'it will be **easier** for the land of Tyre and Sidon on the day of judgment **than for you**' (here **z-jez* does not occur because *jez* is dative and *z-* is licensed only by accusatives); 5.20 *e'te oč' ařawelowc'ow ardarowt'iw'n jer aweli k'an z-dprac'-n* [M: *dprac'-n*] *ew z-P'arisec'woc'* [M: *p'arisec'woc'-n*] 'if your justice will not exceed [lit. be increased **more than**] **that of the scribes and Pharisees** ...'. Here *z-* precedes the genitives *dprac'-n* and *P'arisec'woc'* in E because they have been substantivized: 'that of the scribes and Pharisees'. A less common type of comparative involves *ews* 'still, further' plus a positive adjective: 9.16 *ews čar patařowmn lini* 'a **worse** (*čar* 'bad') tear arises'. The single example of a superlative in our corpus shows an augmentative form of a positive adjective (*p'ok'rik* 'very small': *p'ok'r* 'small') followed by the demonstrative article: 11.11 *p'ok'rik-n yark'ayowt'ean erknic' · mec ē k'an z-na* '**the smallest** in the kingdom of heaven is greater than he'.

3. Complementation and adjunction

3.1. We begin with prepositional phrases. Classical Armenian is notable for its paucity of true prepositions. These number only six: *ař*, *and*, *ast*, *z-*, *i*, and *c'-*; and of these, *ast* and *c'-* are somewhat restricted. This means that the remaining four must each assume a very broad semantic range through construction with various cases. The widest ranging

of these is *i*, which takes the accusative in the value ‘into’, the locative in the sense ‘in’, and the ablative (where it is obligatory, cf. 2.1) in the value ‘from’, including partitive and (cf. Matthew 3.6 above) agentive values: 9.1 *ekn i k’alak’ iwr* ‘he came **into** his own city’; 10.27 *zor asem jez i xawari . asac’ēk’ i loys* ‘what I say to you **in darkness**, say **in light**’; 2.15 *yEgiptosē koč’ec’ic’ [M: koč’ec’i] zordi im* ‘**from Egypt** I shall call [M: I have called] my son’; 3.7 *teseal zbazowms i Sadowkec’woc’n* ‘seeing many **of the Sadducees**’. A peculiar feature of *i* is that when its object is followed by the longer, more emphatic form of the demonstrative adjective, it is repeated before this: 10.19 *i žamown y-aynmik* ‘at that time’ (lit. ‘at the time, at that’) but 11.25 *y-aynm žamanaki* ‘id.’ with the shorter form of the demonstrative preceding its head and no repetition of the preposition. Similarly, *z-*, in addition to possessing a purely grammatical value as a *nota accusativi*, indicates extent of time (acc.), ‘concerning’ (abl.), and ‘around’ (instr.): 12.40 *z-eris tiws ew z-eris gišers* ‘**for three days and for three nights**’; 10.32 *xostovanesc’ic’ ew es z-nmanē* ‘I also shall confess **concerning him**’; 8.18 *teseal YI žołowowrds bazowms z-iwrew* ‘Jesus, seeing great crowds **around him** ...’. *ař* may range as widely as ‘to(ward)’ (acc.), ‘beside, chez’ (loc.), ‘at the time of’ (instr.), and exceptionally, ‘out of [cause or basis]’ (gen.): 14.26 *ař ahin alalakec’in* ‘**out of fear** they cried out’. *and*, when construed with the accusative means ‘along, through(out)’, with the genitive ‘in exchange for’, with the locative ‘with’, and with the instrumental ‘under’; and *ast* is found only 3 times in our corpus in the values ‘according to’ (dat. or loc.) and ‘after [indicating sequence]’ (abl.): 6.11M *awr ast awrē* ‘day **after day**’. *c’-* (which, like *z-*, is univerted with a following nominal) is used primarily as a complement of verbs of speaking, especially *asel* ‘say’, to indicate the addressee: 12.11 *na asē c’ nosa* ‘he said (lit. says) to them’. Beside these true prepositions, a large number of improper prepositions govern the genitive, thereby revealing their nominal origins; e.g. *i veray* ‘above, upon’: 7.24 *i veray vimi* ‘upon a stone’. The most peculiar of these is *handerj* ‘with’, lit. ‘garment, clothing’, which appears with the instrumental and may either precede or follow its object: 2.11 *handerj Mariamaw marbn iwrov* ‘with Mary, its [sc. the child’s] mother’, 9.19 *ařkertawk’n handerj* ‘with the disciples’.

3.2. Participles occupy an important role in Classical Armenian syntax, both as arguments and as adjuncts, as well as in periphrastic verbal formations. There is only one participle in the strict sense, the aorist participle in *-eal* (gen. *-eloy*), which may have either active or passive value. The adjunct employment of the *-eal* participle is found in backgrounded temporal phrases preceding a main clause having continuity of subject. Normally, but not invariably the subject of the participle is in the nominative when the participle is intransitive and in the genitive when it is transitive: 8.14 *ekal YS i town Petrosi . etes zi zok’anč’n nora ankeal dnēr tapac’eal* ‘**Jesus, coming** into the house of Peter, saw that his mother-in-law lay ill with fever’; 9.2 *teseal YI* (gen.) *zhawats noc’a · asē c’andamaloyn* ‘**Jesus, seeing** their faith, said to the paralytic ...’. But cf. also 15.29 *gnac’eal anti YI* (gen.) *ekn i covezrn Galileac’woc’* ‘**Jesus, going from there**, came unto the seashore of Galilee’ with genitive subject, despite the fact that *gnac’eal* ‘going’ is intransitive. A peculiar impersonal usage of the participle is seen in 1.16: *Yakovb cnaw zYovsēp’ zayrn Mariama [M: -y] · orowm xawsec’eal zMariam* ‘Jacob begat Joseph, the husband of Mary, **to whom Mary was betrothed**.’ The second part of this clause follows a variant reading of the Greek text, which we have translated literally for the sake of comprehensibility; but the Armenian says, in effect, ‘to whom (one) had betrothed Mary’

(on periphrastic [“perfect”] verbal formations involving the participial and a copula [here lacking], see 5.4).

3.3. An infinitival phrase can also serve as an adverbial sentential adjunct, playing the role of an absolute construction, which Classical Armenian does not possess. Consequently, there is no continuity of subject. The infinitive is in the locative case followed by the subject of the infinitive in the genitive. This construction typically translates the Greek genitive absolute: 2.1 *i cnaneln YI i Bet'leem ... mogk' yerewelic' ekin yĒM* ‘**when Jesus was born** (lit. in the being born of Jesus) in Bethlehem, ... Magi from the East came to Jerusalem’.

3.4. Classical Armenian shows a number of patterns of complementation. Sentential complementation is normally signaled by the forms (*e*)*t'e* and *zi*: 2.16 *ibrew etes Hērov-dēs* [M: *e*] *t'e xabec'aw i mogowc' anti* ‘when Herod saw **that he had been deceived by the Magi** ...’; 3.9 *asem jez · zi karol ē AC i k'aranc' yaysc'anē yarowc'anel ordis Abrahamow* ‘I say to you **that God is able** from these stones to raise up sons to Abraham’. A common complement construction, especially following *tesanel* ‘see’ plus its object, is *zi* followed by an imperfect whose subject is the object of the higher clause. This is translatable in English as ‘He saw someone V-ing’: 9.9 *etes ayr mi zi nstēr i mak'saworowt'ean* ‘**he saw a man sitting** (lit. he saw a man that he was sitting) at a tax collector’s station’. An alternative way of achieving the same meaning is with a participle: 4.18 *etes erkows elbars ... arkeal owrkan i cov* ‘**he saw two brothers ... casting** a net into the sea’. Infinitive complementizers are frequent after *karel* ‘be able’, as well as verbs of speaking, ordering, desiring, and motion: 2.2 *ekak' erkir paganel nma* ‘we have come **to worship** (lit. to kiss the earth for) him’.

3.5. Classical Armenian employs complementizers before both direct and indirect quotations and questions. The most common of these is (*e*)*t'e*: 5.21 *lowarowk' zi asac'aw ařajnoc'n* [M: *e*] *t'e mi spananer* ‘you have heard that it was said to those of old **that** “thou shalt not kill”’, where the complementizer introduces a direct quotation, as evidenced by the fact that the verb *spananer* is an imperative. An alternative treatment involving no complementizer is 5.27 *lowarowk'? zi asac'aw mi řnar* ‘have you heard that it was said, “Do not commit adultery”?’ Examples of *et'e* preceding direct and indirect questions are Eznik II, 1 (Thomson 1989: 131) *harc'anēr · et'e ov? es du* ‘He asked (**that**), “who are you”?’ and 2.4 *harc'anēr i noc'anē t'e ovr cnanic'i K'Sn* ‘he asked of them (**that**) where the messiah would be born’.

4. Clause types

4.1. We begin with subordinate clauses, distinguishing temporal, purpose, result, causal, and conditional types together with their subordinating exponents. Temporal clauses may be subdivided into those of subsequent circumstance (when/after), concomitant circumstance (while), prior circumstance (before), and delimitative circumstance (until). The most important exponent of subsequent circumstance is *ibrew*, which typically occurs with the aorist: 2.3 *ibrew lowaw ark'ay Hērodēs · řřovec'aw* ‘**when** King Herod **heard**,

he trembled'. A second exponent of this sort is *yoržam* (*y-or-žam* 'in[to] which time'): 2.8 *yoržam gtanic'ēk'* [M: *zna*] *azd arasjik' inj* 'when you find (M: him) report (it) to me'. A specific exponent of clauses of concomitant circumstance is *minč'deř*: 1.20 *minč'deř na zays acēr zmtaw . aha hreřtak TN i teslean erewec'aw nma* 'while he was pondering this, just then an angel of the Lord appeared to him in a dream'. Clauses of prior circumstance are signaled by *minč'čew* (*minč'* 'until'+ *čew* 'not yet'): 1.18 *minč'čew ekeal ař mimeans . gtaw ylac'eal i hogwoyn srboy* 'before they had come (lit. before coming) unto each other (in sexual union), she was found (to be) pregnant from/by the holy spirit'. Clauses of delimitative circumstance are signaled by *minč'ew*: 1.25 *oč' gitēr zna . minč'ew cnaw zordin iwr zandranik* 'he (viz. Joseph) did not know her until she beget her first born son'. In a single instance in our corpus, this notion is signaled by *c'oržam* (*c'-or-žam* 'to which time'): 2.13 *and linijir c'oržam asac'ic' k'ez* 'remain there until I tell you'.

4.2. Purpose clauses are signaled by *zi*: 1.22 *ays amenayn elew zi le'c'ii or asac'awn i TĒ* 'all this happened in order that what was said by the Lord should be fulfilled'. Where the purpose is expressed as an infinitival phrase, this is signaled by a preceding *ař i*, and the infinitive is in the ablative: 5.28 *amenayn or hayi i kin mard ař i c'ankanal-oy nma . andēn řnac'aw and nma i srti iwrowm* 'any man who looks at a woman for the purpose of desiring her has right there committed adultery with her in his heart'. And of course, after many verbs, particularly verbs of motion, a simple infinitive may be used to express purpose: 5.17 *mi hamarik' et'e eki lowcanel zawrēns kam zmargarēs* 'do not think that I have come to abrogate the Law or the prophets.'

4.3. Result clauses are signaled by *orpēs t'e/zi* (*orpēs* 'in which manner, how'): 6.1 *zgojš lerowk' olormowt'ean jerowm . mi ařnel ařaji mardkan . orpēs t'e i c'oyc' inč' noc'a* 'be careful in your almsgiving not to do (it) before men, so that it is any kind of public display to them'; 8.28 *orpēs zi č'ēr hnar anc'anel owmek' and ayn čanaparh* 'so that it was not possible for anyone to pass along that path'. A second structure that signals a result is *minč'ew* + infinitive: 13.2 *žolovec'an ař na žolowowrdk' bazowmk' minč'ew mtanel nma i nawn ew nstel* 'large crowds gathered unto him, so that he entered the ship and sat down.' In this last passage the value 'until, to the point that' normally associated with *minč'ew* is still clear.

4.4. Causal clauses are signaled by either *k'anzi* or *zi* 'because': 1.20 *mi erkneř'ir ařnowl ař k'ez zMariam kin k'o . k'anzi or i nmayn cneal ē . i hogwoy srboy ē* 'do not fear to take unto you Mary, your wife; for the one who has been conceived in her is from the holy spirit'; 1.21 *cnē'i ordi ew koč'esces zanown nora YS . zi na p'rkesc'ē žžolowowrd iwr i melac' iwreanc'* 'she will beget a son, and you shall call his name Jesus. For he will save his people from their sins.'

4.5. Conditional clauses are signaled by *et'e* 'if': 5.29 *et'e akn k'o ař gayt'aglec'owc'anē zk'ez . xlea zna* 'if your right eye leads you into sin, pluck it out'.

4.6. A type of negative purpose clause is signaled by *gowc'ē* 'lest' (lit. 'it may be'), a relexicalized subjunctive of *goy* 'il y a, there is': 5.25 *ler irawaxorh and awsoxi k'owm ...*

minč'deř ic'es and nma i čanaparhi · gowc'ē matnic'ē zk'ez ... dataworin 'become reconciled with your opponent ... while you are with him on the road, **lest he hand you over ... to the judge**'.

4.7. An 'as'-clause is attested in our corpus, signaled by *orpēs*: 1.24 *Yovsēp' ... arar orpēs hramayeač' nma hreštakn TN* 'Joseph ... did as **the angel of the Lord had commanded him.**'

4.8. The most frequent subordinate clause type is the relative clause. The relative pronoun of Classical Armenian is *or*, and the relative clause is normally embedded: 2.23 *bnakec'aw i k'alak'in or koč'ēr Nazaret'* 'he dwelled **in the city which was called Nazareth**'. However, correlative constructions do occur: Eznik II, 1 (Thomson 1989: 131) *or ok' i noc'anē val ař is hasc'ē · zna řagawor araric'* '**whichever** of them comes first to me, **him** will I make king'. On the determination of a definite relative clause by *-n* see 2.2.

4.9. Interrogative clauses are signaled by the pronoun *o(v)?* (inanimate *zi[nč']?*), which normally undergoes WH-fronting. In addition, however, numerous interrogative adverbials signal a wide range of questions targeting some adverbial notion (why?, how?, of what sort?, etc.). We distinguish in addition wh-questions from yes/no-questions, direct from indirect questions, and pragmatic from rhetorical questions. Examples of pragmatic wh-questions are the following: 3.7 *o?* [M: e] *c'oyc' jez p'axč'el i barkowt'enēn or galoc'n ē* '**who** warned you to flee from the wrath which will come?'; 8.26 *andēr?* *vatasirtk' ēk' sakawahawatk'* '**why** are you fearful, O you of little faith?'; 12.34 *ziard?* *karič'ēk' baris xawsel or č'ark'd ēk'* '**how** will you who are evil be able to speak good things?'; 8.27 *orpisi?* *ok' ic'ē sa · zi ew hołmk' ew cov hnazandin sma* '**what sort** of person might this one be, that even winds and sea obey him?'. A pragmatic indirect wh-question is 2.4 *harc'anēr i noc'anē t'e ovr cnanic'i K'Sn* 'he asked of them **where** the messiah would be born'; and a pragmatic yes/no-question is 11.3 *dow? es or galoc'n es* '**are you** the one who will come?'. Rhetorical questions are 12.12 *orč'ap?* *ews ařawel ē mard k'an zoč'xar* '**how much** greater still is a man than a sheep?' and 3.14 *inj pito* [M: -y] *ē i k'ēn mkrtel · ew dow ař is? gas* 'I need to be baptized by you; **and you come to me?**'. The position of the question mark in all of these clauses reproduces a mark found in the manuscripts which regularly appears over wh-words and in yes/no-questions selects a particular item, thereby most likely giving us information about intonation peaks within the sentence. Finally, an interrogative adverb may signal an exclamation: 7.13 *k'ani andarjak ē dowřnn ew hamarjak ē čanaparhn or tani i korowst* '**how** wide is the door and roomy is the path that leads unto destruction!'.

4.10. We will discuss now some further types of clauses. Impersonal constructions are relatively frequent, e.g. *pitoy ē* 'it is necessary': 3.14 *inj pito* [M: -y] *ē i k'ēn mkrtel* '**I need** (lit. it is necessary for me) to be baptized by you'. Another type of impersonal construction is of the sort *on sait/man weiss*, for which Classical Armenian employs a third plural verb form: 5.15 *oč' lowc'anen črag ew dnen and growanaw* '**they do not light** a lamp **and place** (it) under a bushel'.

For the peculiar impersonal construction of Matthew 1.16, see 3.2 above.

4.11. In addition to employing the verb *ownel* ‘have’ to signal possession, Classical Armenian shows a construction involving ‘be’ + genitive in the same value: 12.11 *ov? ē i jēnj mard oroy ic’ē oč’xar mi* ‘who among you is a man **who would have** (lit. of whom there would be) a sheep ...?’

4.12. The verb *goy* signals existence and therefore never serves as a copula. Its original value as a verb of dwelling (PIE **h₂wes-* ‘spend the night, reside’) is seen in a passage such as 11.18: *dew goy i nma* ‘a demon **exists** in him’. The negation of *goy* is *č’ik’* ‘il n’y a pas, there is not’, as in 13.57 *č’ik’ margarē anarg . et’e oč’ yiwrowm gawaři* ‘**there is no prophet** dishonored if not in his own district.’

4.13. Finally, in both biblical and non-biblical literature we find serial verb constructions: the employment of consecutive or double verbs to signal asyndetically either the same or separate but closely related actions forming a unified whole: 4.8 *ařnow acē zna satanay i leařn mi barjr yoyž* ‘Satan **took (and) led** him onto a very high mountain’; 13.44 *ert’ay vačarē zamenayn inč’ zor owni* ‘**he went (and) sold** everything which he owned’; P’awstos Biwzandac’i IV, 23 (Thomson 1989: 154) *ařnowin kotorēin zamenayn bnakič’sn veřnagawařac’n Hayoc’* ‘**they took (and) destroyed** all the inhabitants of the upper districts of Armenia.’

5. Tense, mood, and aspect

5.1. Classical Armenian makes frequent use of the historical present, as seen in 4.8 and 13.44 just cited. This may occur with any verb but most frequently involves *asel* ‘say’: 1.20 *hreřtak TN i teslean erewec’aw nma ew asē* ‘an angel of the Lord appeared to him in a dream and **said** (lit. says) ...’. In this passage, *asē* follows the aorist *erewec’aw* ‘appeared’.

5.2. The imperfect of Classical Armenian is an imperfective past tense built to the present stem. However, it frequently signals an irrealis value, particularly that of a conditional, a verbal category which the language lacks: 12.7 *et’e giteik’ zinč’? ē zolor-mowt’iwnn kamim ew oč’ zroh . apa oč’ dataparteik’ dowk’ zampartsn* ‘**if you knew** what is (the meaning of) the (passage), “I desire mercy and not sacrifice”, then **you would not condemn** the guiltless’. In this passage, the interrogativity perceived by the Armenian translator cannot be rendered into English. Similarly, the aorist is a perfective past tense. In many instances, however, it functions as a pluperfect indicating action completed already at a particular moment in the past: 11.20 *yaynžam (= y-ayn-žam ‘at that time’) sksaw naxatel zk’alak’sn yors elen bazowm zawrowt’iwnk’ nora* ‘then he began to upbraid the cities **in which** his many miracles **had occurred**’, where *elen* indicates events which had occurred before Jesus’ censure was brought upon the cities.

5.3. Classical Armenian has no special future tense. Rather, the subjunctive mood is utilized as a future: 10.22 *or hamberic’ē i spař na kec’c’ē* ‘the one who **will last** unto the end, he **will live**’. Here the present subjunctive *hamberic’ē* is imperfective, indicating

a persistence over time, whereas the aorist subjunctive *kec'c'ē* is perfective, referring to the simple notion of living.

5.4. In addition to possessing the simple tenses present and aorist, Classical Armenian also shows compound tenses consisting of participle + the verb ‘to be’. These are sometimes called “perfect” and distinguish a present, past, and future, depending on whether the auxiliary is in the present, imperfect, or subjunctive: 1.20 *k'anzi or i nmayn cneal ē . i hogwoy srboy ē* ‘for the one who **has been conceived** in her is from the holy spirit’ (present perfect); 11.23 *et'e i Sidom eleal ein zawrowt'iwnk'n or elen i k'ez . apak'ēn kayin ewš minč'ew c'aysawr* ‘if the miracles that occurred in you **had occurred** in Sidon, then they (viz. Tyre and Sidon) would be standing unto this day’ (past perfect); 6.8 *zi gitē hayrn jer zinč' pitoy ē jez minč'č'ew jer xndreal inč' ic'ē i nmanē* ‘for your father knows what you need before **you (will) have asked** anything of him’ (future perfect).

5.5. Having discussed the tenses of Classical Armenian, we now move on to their sequential usage. The critical distinction in narration is between the imperfect and the aorist. The former must be considered marked vis-à-vis the latter, because the simple narration of linear events in sequence is carried out via the aorist, while the imperfect indicates some additional nuance relating to the nature of the action or is not properly narrative at all but rather descriptive, indicating backgrounded information: 1.19 *Yovsēp' ayrn nora k'anzi ardar ēr . ew oč' kamēr ařakel zna . xorhec'aw řrelayn arjakel zna* ‘Joseph, her husband, **because he was just and did not wish** to expose her to censure, **planned** secretly to divorce her’. Here the imperfects *ēr* ‘was’ (purely descriptive) and *kamēr* ‘wished’ together constitute the background explaining Joseph’s decision, which is signaled by the aorist *xorhec'aw* ‘he planned’. Inasmuch as the imperfect represents an imperfective past tense, it is not surprising that it is used with *minč'deř* ‘while’ to indicate a backgrounded action followed by a narrative-advancing aorist: 1.20 *minč'deř na zays acēr zmtaw . aha hreřtak TN i teslean erewec'aw nma* ‘**while he was pondering this**, right then an angel of the Lord **appeared** to him in a dream’. In other instances, the imperfect may signal an ongoing action, as opposed to the punctual aorist: 3.16 *etes zhogin AY zi řjanēr ... ew gayr i veray nora* ‘**he saw** (aor.) the spirit of God **descending** [lit. that it was descending] (impfct.) ... and **coming** [lit. was coming] (impfct.) over him’; or the aorist can signal a resultative state beside the ongoing action of the imperfect: 2.20 *zi meřan₁ or xndrein₂ zanjn mankand* ‘for those who **were seeking₂** (impfct.) the life of thy child **have died₁** (aor.)’. Sometimes the action signaled by the imperfect is iterative: 4.23–25 *řřjēr YS and amenayn kolmn Galiteac'woc' . owsowc'anēr ... ew k'arozēr ... ew břřkēr ew el lowrn nora* ‘Jesus **kept going around** throughout the bounds of Galilee; **he kept teaching** ... and **preaching** ... and **healing** ...; and the report about him went out’. Here the iterative imperfects contrast with the aorist *el* ‘went out’, which signals a single action in the past.

5.6. We have already discussed the employment of the subjunctive as a future. However, this represents just a portion of its usage. In two circumstances, the subjunctive approximates or is equivalent to an imperative. The first of these is the so-called *Gesetzessprache*: passages rendering Biblical commandments. In these instances, the aorist subjunctive is employed: 4.7 *oč' p'orjesc'es zTR AC k'o* ‘**thou shalt not try** the Lord thy God’; 5.43 *asac'aw řiresc'es znker k'o* ‘it has been said, “**thou shalt love** thy fellow man”’. As

is seen in the first of these passages, this type of structure takes the indicative negation *oč'* rather than the modal negation *mi*. The second circumstance is the third person, where Classical Armenian lacks an imperative form per se. Here again, the aorist subjunctive is used, but the negation is *mi*: 10.13 *et'e ic'ē townn aržani . ekec'ē [M: ekesc'ē] olfoyn jer i veray nora* 'if the house be worthy, **let your greeting come** upon it'; 6.3 *mi gitasc'ē jax k'o zinč' gorcē aj k'o* '**let your left (hand) not know** what your right (hand) is doing'. The possibility of a commutation of aorist subjunctive and imperative in the second person is suggested by the following passage, where E shows the latter and M the former in a conjoined structure whose second member is an imperative: 5.12 *c'ncac'ēk' [M: c'ncasc'ēk'] ew owrax lerowk'* 'rejoice and be happy'. An example showing the atypical structure *mi* + second person plural present subjunctive in imperative value is 3.9 *mi hamaric'ik' [M: hamarik'] asel yanjins [M: jer] [M: e] t'e ownimk' mek' hayr zAbraham* '**do not think** to say within yourselves that "we have Abraham as (our) father"'. Here M has the expected present imperative. Perhaps the most general usage of the subjunctive is to indicate a circumstance that stops short of being actual. This is seen especially in relative clauses where a situation is typified or imagined rather than confirmed as actualized: 3.10 *amenayn cař or oč' ařnic'ē zptowl bari . hatani* 'every tree **which does not produce good fruit** is cut down' (pres. subj.); 10.14 *or oč' ankalc'i zjez · ew oč' lowic'ē banic' jeroč'* '**whoever does not receive you and does not hearken** to your words ...' (aor. subj.). The subjunctive is also employed in deliberative questions: 11.3 *dow? es or galoc'n es . et'e ayłowm akn kalc'owk'* 'are you the one who is to come, **or should we expect another?**' The remaining usages of the subjunctive are found in a range of subordinate clause types of the sort we have discussed and illustrated in 4.1–4.5 above. These include purpose clauses (Matthew 1.22), temporal clauses of various sorts (Matthew 2.8, 2.13, 5.25), and conditional clauses: 5.23–24 *et'e matowc'anic'es zpatarag k'o i veray selanoy · ew and yišic'es [M: e] t'e elbayr k'o ownic'i inč' xēt' zk'ēn ... ert' nax hašteac' and elbawr k'owm* '**if you bring** your offering upon the altar, and **you remember** there that your brother may have some resentment toward you ... go first (and) become reconciled with your brother'.

5.7. Classical Armenian possesses imperative forms of both the present and aorist stems, but these are limited to the second person singular and plural. The present imperative is used only in negative commands (negation: *mi*), the aorist imperative in positive ones. We have already seen that the missing third person imperative forms are provided by the subjunctive; and the same is true of the first person: Eznik II, 3 (Thomson 1989: 133): *Ard mer t'oleal znosa . harc'c'uk' zsoša* 'now we, leaving those things, **let us ask** these things ...'. An example of the juxtaposition of negative present imperative and positive aorist imperative in a single passage is 10.28 *mi erkneč'ik' yaync'anē or spananen zmarmin ... ayl erkerowk' dowk' ařawel yaynmanē or karoln ē zogi ew zmarmin korows-anel* '**do not fear** those who kill the body ... but **fear** ye more that one who is able to destroy the spirit and the body'. In addition to pure imperatives, one also finds forms with endings in *-(i)jir /-(i)jik'* used in exhortative value. These have affinities with the subjunctive; and indeed they are occasionally employed in future value (though not in our corpus). They may serve independently as quasi-imperatives, as in 10.17 *zgojš lini-jik' i mardkanē · zi matnesc'en zjez yateans* '**be wary** of men, for they will hand you over to the councils'; or they may occur in sequence with true imperatives: 2.13 *ari ař zmanowkd ... ew p'axir yEgiptos · ew and linijir* '**Arise, take** thy child ... and **flee** to

Egypt; and **remain** there'. In each of these passages *linjĭk'* and *linjĭr*, despite being built to the present stem, occur in positive commands, which would be precluded for their corresponding true imperatives *linik'* and *linir*, respectively.

5.8. We shall close this section with a few final remarks about the occurrence of the modal negation *mi*. We have seen that this form occurs regularly with the present imperative. It also occurs, as we have seen, with third person (rarely, second person) subjunctives employed as imperatives. Modal negation may also be employed with both present and aorist subjunctives in a variety of clause types other than imperative, including complement clauses, purpose clauses, and even as futures: 5.29 *law ē k'ez et'e mi yandamoc' k'oc' koric'ē · ew mi amenayn marmind k'o ankanic'i i gehen* 'it is better for you that one of your limbs should be destroyed and your entire body **should not fall** into Gehenna'; 4.6 *i veray jeřac' barjc'en zk'ez · zi mi erbek' harc'es zk'ari zotn k'o* 'they shall raise you up upon their hands so that **you never strike** your foot against a stone'; 11.23 *dow kap'ařnawowm mi minč'ew yerkins barjrasc'is . ayl minč'ew i džoxs ijč'es* 'you, Kaparnaum, **will not be raised** unto the heavens but will descend into hell'. Modal negation also occurs where a command is not expressed by a finite verb but in an infinitive: 5.34 *ayl es asem jez · amenewin mi erdnowl* 'but I say to you **not to swear** at all'.

6. Passivization, diathesis, causativity, and *genera verbi*

I have so far said nothing about passivization, because it is unremarkable. The typical passive sentence shows a subject and a passive verb form. If an agent is present, it is in the ablative case (cf. 3.6, cited in 2.1). For a discussion of the morphology of active and (medio)passive verb forms, see Olsen, this handbook, 7.2, 7.5. What is remarkable about Classical Armenian, however, is its skewed, asymmetrical system of voice opposition, particularly in the present indicative. Of the four basic present-tense verb classes (-e-, -i-, -a-, -u-), only the first and second show a direct oppositional relationship of active and passive (e.g. *sirem* 'I love': *sirim* 'I am loved'), but even here the opposition does not extend to the imperfect. The third and fourth classes stand completely outside this system in the indicative. Oddly, the -a- class allows an oppositional passive in the subjunctive (although many verbs of this class are inherently mediopassive in meaning and therefore show no such opposition); and the -u- class does not even allow that. In the aorist system, oppositions are possible in all moods; however, Classical Armenian possesses an unusually large stock of *media tantum* in both present and aorist. There is no morphological distinction between mediopassive and passive forms.

Notable as well is a class of verbs in -uc'anem which signal causativity (e.g. *p'axč'im* 'I flee': *p'axuc'anem* 'I make flee'). These show the same additive syntactic effects on verbal valence and argument structure that may be observed with causatives elsewhere in Indo-European.

Finally, from an Indo-European perspective, Classical Armenian shows a paucity of verbal lexemes consisting of preverb + verb; the relatively few instances of this sort that do exist (e.g. *nstim* 'I sit', *ant'eřnum* 'I read', *zgenum* 'I wear' with *ni-, and-, and z-, respectively) have generally been lexicalized and are no longer synchronically analyzable

in such a way. In none of the examples just given do collateral verbs without the preverb exist, and **ni-* in particular has no synchronic existence in Classical Armenian.

7. Word order

7.1. Word order is not amenable to rigid description in terms of S(ubject), V(erb), and O(bject). Rather, word order is determined by pragmatic or discourse factors. Thus, we find very close to each other the sentences 7.17 *Ayspēs amenayn caṙ bari ptowl bari aṙnē* ‘thus every good tree **creates good fruit**’ (OV) and 7.20 *Apa i ptloy noc’a canajik’ znosa* ‘Accordingly, by their fruit **know them**’ (VO). In the first instance, the focus is on the good fruit produced by the good tree; in the second, the Armenian translator was certainly avoiding the awkward juxtaposition *noc’a znosa* ‘their them’), while at the same time emphasizing the idea of judging by results.

7.2. We have already seen that nouns precede dependent genitives and that adjectives may precede or follow head nouns. There is a strong tendency for the verb to appear immediately after a subordinator: 1.24 *arar orpēs hramayeač’ nma hreštakn TN* ‘he did **as** the angel of the Lord **commanded** him’; 2.15 *zi lc’c’i asac’ealn i TĒ* ‘**in order that** the thing said by the Lord **should be fulfilled**’; 5.18 *minč’ew anc’c’en erkink’ ew erkir* ‘**until** heaven and earth **shall pass away** ...’; 8.4 *mato zpataragn zor hramayeač’ Movsēs* ‘offer the sacrifice **which** Moses **commanded**’. The same can be said for sentence-initial adverbials generally: 1.12 *Yet gerowt’eann Babelac’woc’ cnaw Yek’onia zSalat’iēl* ‘**after the Babylonian captivity** Jechonia **begat** Salathiel’. Examples of focus fronting of a constituent other than a verb are 9.13 *zolormowt’iwn kamim ew oč’ zzoḥ* ‘I desire **mercy** and not sacrifice’ and Eznik II, 3 (Thomson 1989: 133) *Ayl asen · p’aṙac’ aṙnēr zyaštñ* ‘but, they say, **for glory** he performed the sacrifice’. Similarly, in the correlative diptych (*or ... na* ‘which one ... that one’), the correlative is normally fronted within its own clause (see Eznik II, 1 cited in 4.8). Finally, the negative particle *oč’* normally precedes the verb in proclisis, often being reduced to *č’-*: 8.8 *č’em bawakan et’e and yarkaw imov mtc’es* ‘**I am not** worthy that you should enter under my roof’. But where it may be interpreted as applying to a proposition as a whole, it precedes that proposition: 15.11 *oč’ or inč’ mtanē and beran [M: -n] plcē zmard* ‘**it is not the case that** whatever enters into the mouth defiles man’.

8. Discourse syntax: conjunction and deixis

8.1. We begin with conjunction. The basic additive conjunction of Classical Armenian is *ew* (**epi*), which is capable of conjoining structures at all levels, from the word or phrase to the sentence: 4.24 *matowc’in aṙ na zamenayn hiwands ... ew zdewahars · ew zlowsnots · ew zandamaloysc* ‘they brought to him all the sick ... **and** those possessed of demons, **and** lunatics, **and** paralytics’; 8.11 *bazmesc’in and Abrahamow ew and Sahakay ew and Yakovbow* ‘they will sit down to eat with Abraham **and** with Isaac **and** with Jacob’; 6.17 *awc zglowx k’o . ew lowa zeress k’o* ‘anoint your head **and** wash your face’. *ew* may also possess an adverbial value ‘also, even’ with narrow scope in both

positive and negative clauses: 10.32 *amenayn or xostovanesc'i yis ařaji mardkan . xostovanesc'ic' ew es znmanē ařaji hawr imoy or yerkins* 'everyone who will profess (faith) in me before men, **I also** will profess concerning him before my father in heaven'; 6.29 *ew oč' Solovmovn yamenayn p'ařsn iwrowm zgec'aw ibrew zmi i noc'anē* '**not even** Solomon in all his glory was clothed like one of them'. Finally, *ew* may be employed in generalizing value following an indefinite relative construction: 8.19 *ekic' ew es zkni k'o yor vavr ew ert'ic'es* 'I also will follow you **wherever** you will go'. The first occurrence of *ew* here is adverbial.

8.2. Alternative conjunction is signaled in either of two ways. In a noninterrogative sentence, *kam* signals a simple alternative: 5.17 *mi hamarik' et'e eki lowcanel zawrēns kam zmargarēs* 'do not think that I have come to abrogate the law **or** the prophets'. Where *kam* occurs twice in a conjoined set, it conveys the value 'either ... or': 12.33 *kam ararēk' zcařn bari ... kam ararēk' zcařn č'ar* '**either** make the tree good ... **or** make the tree bad'. In the case of questions, however, *et'e* is employed to express an alternative: 9.5 *zinč'? diwrin ē · asel t'oleal lic'in k'ez melk' k'o · et'e asel ari ew šrjeac'* 'which is easier: to say, "may your sins be forgiven you" **or** to say, "arise and move about"?'.

8.3. Adversative conjunction is signaled primarily by *ayl/l* (also 'other') and *bayc'*: 4.4 *oč' hac'iw miayn kec'c'ē mard ayl amenayn baniw or elanē i beranoy AY* 'not by bread alone shall man live, **but** by every word which emanates from the mouth of God' and 3.11 *es mkrtē zjez jrov ... bayc' or zknin im gay hzawragoy ē k'an zis* 'I baptize you with water ... **but** the one who comes after me is stronger than I'.

8.4. A final level of conjunction involves solely intersentential structures and may be labeled "discourse continuity". The main exponents of this in our corpus are *ard*, *isk*, *apa*, and *ayl/l*. The first of these means basically 'now' but may also convey a conclusion 'therefore'. In the former sense, its value may be merely discourse continuative, as in English 'now then': 1.17 *Ard amenayn azgk'n yAbrahamē minč'ew i Dawit' · azgk' č'orektasank'* '**now then**, all the generations from Abraham to David (total) fourteen generations'. A consequential value is reflected in 5.48 *ard elerowk' dowk' kataarealk' · orpēs ew hayrn jer erkawor katareal ē* '**therefore** be ye perfect, just as also your heavenly father is perfect.' *isk* 'truly', a generally asseverative particle, may convey weak adversative or occasionally inferential value, and frequently renders Greek *dé*. It may enter the functional sphere of *ard*, so that in the following passage the E text reads *isk*, while M shows *ard*: 6.23 *isk [M: ard] et'e loysd or i k'ez ē xawar ē · xawarn orč'ap'? ews* '**but** [M: now] if the light which is in you is darkness, how much more is the darkness?' *Apa* possesses a decidedly consequential value but, like English 'then', may also be temporal: 7.20 *apa i ptloy noc'a canijik' znosa* 'know them, **then**, by their fruit'; 4.11 *apa et'ol zna satana [M: -y]* '**then** Satan left him'. Finally, *ayl* appears in weak discourse continuative value at 1.22 *ayl ays amenayn elew zi lc'ci or asac'awn i TĒ* '**but** all this happened in order that what was said by the Lord should be fulfilled.'

8.5. An important feature of discourse syntax is the implementation of deixis. Our discussion of this complex topic will focus on deixis as it relates to discourse continuity and discourse perspective. In a language without gender distinctions, even among pro-

nouns, another means of effecting reference is required; and Classical Armenian's response to this challenge has been to index referents by position rather than by gender. The most surcharged of the three deictic settings (cf. 2.1) is *-n-*, which must function both as a distal deictic and a neutral anaphoric. In narrative, it is normally the latter that comes into play. We expect narrative reference to be effectuated overwhelmingly via *n*-deixis, particularly by *na* 'he/she/it'; and this expectation is confirmed: 4.24–25 *Ew el lowrn nora and amenayn erkir Asorwoc' · ew matowc'in aṙ na zamenayn hiwands ... Ew ertayin zhet nora žolovowrdk' 'And his fame went out throughout all the land of Syria, and they brought unto him all the sick ... And the crowds followed him.'* However, a number of factors can lead to a referential setting of either *-d-* or *-s-* (the latter being instantiated by such factors as empathy or heightened interest); and when this setting is adopted, it is continued, in accordance with a principle of deictic consistency, unless the discourse perspective changes. In order to illustrate this, we must move outside our corpus: Mk. 4.41 *o? ok' ardewk' ic'ē sa . zi ew hołm . ew cov hnazandin sma* 'Who, indeed, might **this one** be? For both wind and sea obey **him**.' In this passage, the initial deictic setting is evoked by the heightened interest of the disciples in the identity of the man who has by his word alone saved them from a tempest; but the second occurrence of *sa* (dative *sma*) is the result of deictic consistency.

The use of deixis to indicate shift of discourse perspective is illustrated, again outside our corpus, by the following passage: Lk. 10.40–42 *Ew Mart'a ... ekn ekac' aṙ nma ew asē · TR · oč' inč' ē p'oyt' k'ez zi k'oyr-d im miayn et'oł zis i spasow · ard asa dma zi awgnesc'ē inj.* (41) *Patasxani et nma YS ew asē ...* (42) *... Mariam masn bari əntreac' or oč' barjc'i i smanē* 'And Martha ... came (and) stood by him (viz. Jesus), and said, "Lord, is it of no consequence to thee that **my sister (there by thee)** has left me alone in service? Now say **to her** that she should help me ..." (41) Jesus answered her and said, "... (42) ... Mary has chosen a good part, which will not be taken away **from her**." At first, this passage appears to be an exception to the principle of deictic consistency just noted. However, it must be understood in terms of deictic shift. Jesus has visited the house of Martha, whose sister Mary sits at his feet, while Martha is left to tend house by herself. Martha complains of this to Jesus, referring to Mary with second person *d*-forms (*k'oyr-d im ... asa dma*), thereby designating her as being in Jesus' proximity. When Jesus replies, the deictic perspective is shifted, his use of the *s*-form *smanē* 'from her (here)' corroborating the position of Mary relative to Jesus from the latter's perspective. Finally, this passage also illustrates another important implementation of deictic settings: whereas *-n-* is overwhelmingly associated with narrative discourse (cf. the employment of dat.-loc. *nma* twice in narrative stretches of the passage just cited), *-s-* and *-d-* are primarily associated with conversation.

9. Varia

9.1. One of the more remarkable features of Classical Armenian is its system of indefinite pronouns and the relationship of these to polarity. Two sets of indefinites occur: one in *-mn* (*omn/imn*) and one in *-k'* (*ok'/inč'* [the expected *ik'* occurs only in the collocation *č'ik'* 'il n'y a pas, there is not']). The forms with *o-* are animate, those with *i-* inanimate. The difference between the two sets can be described in terms of polarity: *-k'* forms are

negative polarity items and *-mn* forms are positive polarity items. Classical contexts for negative polarity are negative, relative, interrogative, and conditional clauses, as well as clauses of prior circumstance: 6.24 *oč' ok' karē erkowc' teranc' cařayel* 'nobody can serve two masters'; 15.11 *or inč' elanē i beranoy . ayn plcē zward* 'whatever emanates from the mouth, that defiles man'; 12.29 *ziard? karē ok' mtanel i town hzawri* 'how can anyone enter the house of a mighty man ...?'; 5.39 *et'e ok' acic'ē aptak yař cnawt k'o . darjo nma ew zmiwsn* 'if anybody should strike (you) on your right cheek, turn also the other to him'; 6.8 *zi gitē hayrn jer zinč' pitoy ē jez minč'č'ew jer xndreal inč' ic'ē i nmanē* 'for your father knows what you need before you (will) have asked anything of him'. Another classical negative polarity context is the complement of a comparative construction. For this we must go beyond our corpus to Lk 17.2 *law ēr nma et'e vēm erkanak'ar kaxēr zparanoc'ē nora . ew ankanēr i cov . k'an t'e gayt'aktec'owc'anic'ē zmi ok' i p'ok'rkanc's yaysc'anē* 'it were better for him that a millstone were hung about his neck and he were cast into the sea than that he should cause any one of these little ones to stumble'. Moreover, it is not merely the generalizing head of a relative clause that licenses negative polarity, as in 15.11 above, but any other argument within the clause may do so as well. To see this, we must again go beyond our corpus to Lk 6.30 *zor inč' hanē ok' i k'ēn . mi pahanjer* 'Whatever anyone takes from thee, don't demand back'. In all these instances, English uses its negative polarity item *anybody/anyone* rather than positive *somebody/someone*. In positive statements, one finds rather *-mn* forms in Classical Armenian: 8.21 *mi omn yařakertac'n nora asē c'na* '(some/a certain) one of his disciples said to him ...'; 12.38 *patasxani etown nma omank' i dprac'n* 'some/certain of the scribes answered him'. The *-k'/-mn* opposition extends also to adverbials: *erbek'/erbemn* 'at some time/at any time' and *owrek'/owremn* 'somewhere/anywhere' For both of these, we must once again go outside our corpus: Mk. 2.25 *oč' erbek' ic'ē ant'erc'eal jer* 'Have ye never read?', Lk. 22.32 *ew dow erbemn darjc'is* 'and you will some day return'; Mk. 9.8 *hayec'eal aysr andr. oč' ews zok' owrek' tesin* 'Looking this way and that, they no longer saw anybody anywhere' (with three negative polarity items), Lk. 18.2 *Datawor mi ēr i k'alak'i owremn* 'There was a judge in a city somewhere'. This last passage introduces a parable. It also illustrates the fact that Classical Armenian possesses an indefinite article: the numeral *mi* '1' suffixed to a head noun. In addition to possessing negative polarity features, the *-k'* forms also signal minimal value scalar readings when following low-value quantifiers: 5.19 *or ok' lowcc'ē mi inč' i patowiranač' [M: -s] yaysc'anē i p'ok'ownc'* 'whoever abrogates even/so much as one of these smallest commandments ...'. Finally, the fact that indefinite pronouns share with interrogative pronouns the unique feature that they alone among all nominal items in Classical Armenian are capable of distinguishing animacy has led to the usage of *inč'* in particular as a marker of inanimacy with certain pronominal adjectives: Jh 16.12 *Ews bazowm inč' ownim asel jez* 'I have still much to say to you'.

9.2. Reflexivization in Classical Armenian is signaled by one of three words: *iwr*, *ink'n*, or *anjn*, lit. 'person'. The first of these has a defective paradigm which does not allow it to be employed in the nominative or accusative singular or the nominative plural, so it is particularly in these cases where one finds *ink'n*. Each of these terms may refer to all three persons, but *anjn* (and sometimes *ink'n*) may be followed by a personal possessive pronoun, especially in the second and third person: 14.15 *zi ert'ic'en šowrj i šēnsn [M: ew] gnesc'en iwreanc' kerakowrs* 'so that they may go around in the villages (and

they may buy food **for themselves**’; 8.24 *ink’n nnjēr* ‘**he himself** was sleeping’; 15.10 *koč’ec’eal aṙ ink’n zžolovowrdsn asē c’nosa* ‘calling **unto himself** the crowds, he said to them ...’; 8.4 *ert’ c’oyc’ zanjn k’o k’ahanayin* ‘go show **yourself** (lit. your person) to the priest’; 12.25 *amenayn t’agaworowt’iwn bažaneal yanjn iwr aweri* ‘every kingdom divided **against itself** is destroyed.’ This last passage illustrates the strengthening of *anjn* by the reflexive possessive pronominal adjective *iwr*, which possesses a full paradigm. Finally, a simple personal pronoun may be employed in reflexive value: 11.29 *aṙēk’ zlowc im i jez* ‘Take my yoke **upon yourselves**’.

9.3. Although Classical Armenian is basically a pro-drop language, subject pronouns do in fact frequently occur, particularly in cases of emphasis or contrast: 3.11 *es mkrtem zjez jrov ... bayc’ or zknin im gay hzawragoyñ ē k’an zis* ‘**I baptize** you with water ... but the one who comes after me is stronger than I’; 3.9 *ownimk’ mek’ hayr zAbraham* ‘**we have** Abraham as (our) father’ (i.e. we Jews but not non-Jews); 6.26 *oč’ apak’ēñ dowk’ aṙawel ēk’ k’an znosa* ‘**are you not** greater than they?’; 3.14 *inj pito [M: -y] ē i k’ēñ mkrtel · ew dow aṙ- is? gas* ‘I need to be baptized by you; and **you come** to me?’.

9.4. Finally, we shall discuss a number of idiomatic features which lie on the borderline between syntax and lexicon. First, Classical Armenian makes substantial use of periphrasis in the creation of verbal lexemes. Some examples drawn from our corpus are *patasxani tal* (lit. ‘give response’) ‘answer’, *erkir paganel* (lit. ‘kiss the earth’) + dat. ‘worship’, *t’oyl tal* (lit. ‘give leave’) ‘allow’, *kal yalawt’s* (lit. ‘stand in prayer’) ‘pray’, and *akn ownel* (lit. ‘have an eye’) ‘expect, wait for’. A particularly large number of these involve the verb *aṙnel* ‘make, do’: *p’aṙawor aṙnel* (lit. ‘make glorious’) ‘glorify’, *c’aṙ aṙnel* (lit. ‘make laughter’) ‘ridicule’, *kin aṙnel* (lit. ‘make a woman’) ‘marry’. In addition, *tal* + infinitive (lit. ‘give to X’) produces a periphrastic causative: 5.32 *amenayn or arjakē zkin iwr aṙanc’ bani poṙnkowt’ean · na tay nma šnal* ‘Whoever repudiates his wife without a cause of prostitution, he **makes** her **commit adultery**’. Some verbs may take unpredictable object marking: *gohanal z* + abl. ‘praise’ (*z* + abl. ‘concerning’), *anc’anel z* + instr. ‘transgress’ (*z* + instr. ‘around’). Other interesting instances of governance or case usage include the use of a nominative in place of a genitive in certain locutions which denote quantity: 10.42 *bažak mi jowr c’owrt* ‘a cup (of) cold **water**’ (cf. German *ein Trunk Wasser, ein Stück Brot*). An agentive construction that may ultimately go back to Semitic is the use of *i jeṙn* + gen. (lit. ‘at the hand of’), especially associated with the pronouncements of a prophet: 2.5 *zi ayspēs greal ē i jeṙn margarēin* ‘for thus has it been written **by** the prophet’ (Gk. *dià tou prophētou*; but cf. Hebrew ‘*al yad*’ ‘by the hand of, through’). A lexical construction worth noting is *t’ol z* + acc. (lit. ‘leave [out]’) ‘except for’: 14.21 *ork’ kerann ein ibrew hing hazar · t’ol zkanays ew zmankti* ‘those who ate were about five thousand, **except for** women and children’. The employment here of *ibrew* ‘like’ in an approximative sense (‘about’) is conceivably to be traced to Semitic.

9.5. Four remaining points we shall make are, first, that Classical Armenian, in referring to places, often names them by a designation of their inhabitants rather than directly: 10.15 *diwragoyñ lic’i erkrin Sodomac’woc’ ew Gomorac’woc’ yawowrn datastani · k’an k’atak’in aynmik* ‘it will be easier **for the land of Sodomites and Gomorrans** (i.e. for Sodom and Gomorra) on the day of judgment than for that city’. Second, many nouns

are *pluralia tantum*: *hawatk* ‘faith’, *p’aṙk* ‘glory’, *xawsk* ‘speech’, *dowrk* ‘door’, etc. These are construed with plural verbs: 15.28 *ov kin dow · mec en hawatk k’o* ‘O thou woman, great is (lit. are) thy **faith**’. In passing, we note here the use of the nominative case for the missing vocative. For clarity (and no doubt a degree of affect), the nominative in this value may be followed by the second person demonstrative *-d*: Mk 15.28 *olj er ark’ay-d hrēic* ‘Hail, **O (thou) king** of the Jews’. Finally, nominalization cum lexicalization of participles is particularly frequent, the most notable example being *aṙak’ealk* ‘disciples’, lit. ‘those sent’.

10. References

- Jensen, Hans
1959 *Altarmenische Grammatik*. Heidelberg: Winter.
- Jungmann, Paul
1964–1965 L’emploi de l’article défini avec le substantif en arménien classique. *Revue des Études Arméniennes* 1: 47–99; 2: 43–116.
- Klein, Jared S.
1996a *On Personal Deixis in Classical Armenian. A Study of the Syntax and Semantics of the n-, s-, and d-Demonstratives in Manuscripts E and M of the Old Armenian Gospels*. (Münchener Studien zur Sprachwissenschaft, Beiheft 17, Neue Folge). Dettelbach: Röhl.
- Klein, Jared S.
1996b Remarks on the Personal Demonstratives of Classical Armenian. In: Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan. (synopsis of Klein 1996a), 21–34.
- Klein, Jared S.
1997 Indefinite Pronouns, Polarity, and Related Phenomena in Classical Armenian. A Study Based on the Old Armenian Gospels. *Transactions of the Philological Society* 95: 189–245.
- Klein, Jared S.
2011 Negation and Polarity in the Greek, Gothic, Classical Armenian, and Old Church Slavic Gospels: A Preliminary Study. In: Eirik Welø (ed.), *Indo-European syntax and pragmatics: contrastive approaches*. (Oslo Studies in Language 3[3]). Oslo: University of Oslo Press, 131–154.
- Künzle, Beda
1984 *Das altarmenische Evangelium/L’Évangile arménien ancien*. Teil I/I^{ère} partie: Edition, Teil II/II^e partie: *Lexikon/lexique*. (Europäische Hochschulschriften, Reihe XXI, Linguistik und Indogermanistik 33). Bern: Peter Lang.
- de Lamberterie, Charles
1997 L’article dans la relative en arménien classique. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft Madrid, 21–24 de septiembre de 1994*. Wiesbaden: Reichert, 311–326.
- Lyonnet, Stanislas
1933 *Le Parfait en Arménien Classique. Principalement dans la Traduction des Évangiles et chez Eznik*. (Collection Linguistique Publiée par la Société de Linguistique de Paris 37). Paris: Champion.

Meillet, Antoine

1897–1898 Recherches sur la syntaxe comparée de l'arménien, I. les démonstatifs. *Mémoires de la Société de Linguistique* 10: 241–271. [Reprinted in Meillet 1962, 5–35.]

Meillet, Antoine

1962 *Études de linguistique et de philologie arméniennes*, I. Lisbon: Imprensa Nacional.

Minassian, Martiros

1988 L'article du premier terme non substantivé d'une subordonnée en arménien classique. *Le Muséon* 101: 97–158.

Ouzounian, Agnès

1992 *Le discours rapporté en arménien classique*. (Bibliothèque des Cahiers de l'Institut de Linguistique de Louvain 63). Louvain-La-Neuve: Peeters.

Ouzounian, Agnès

1996 À propos des relatives en arménien classique: le fonctionnement de l'article et le cas des relatives à antécédent démonstratif d'après les évangiles et les actes des apôtres. In: Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan, 75–92.

Sakayan, Dora (ed.)

1996 *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan.

Thomson, Robert W.

1989 [1975] *An Introduction to Classical Armenian*. 2nd edn. Delmar, NY: Caravan.

Jared S. Klein, Athens, GA (USA)

64. The lexicon of Armenian

- | | |
|--|-------------------|
| 1. Inherited vocabulary | 4. Word-formation |
| 2. Loanwords | 5. References |
| 3. Technical and scientific vocabulary | |

1. Inherited vocabulary

1.1. In 1875, Heinrich Hübschmann conclusively showed that Armenian constituted a separate branch of Indo-European (Hübschmann 1875). In the decades before Hübschmann, Indo-Europeanists had judged the language to belong to the Iranian subgroup (see Schmitt 1975). The reason for this mistaken analysis of the origin of the Armenian language lies in the nature of the lexicon. Armenian has a smaller proportion of inherited vocabulary than other Indo-European languages of comparable age, and it contains a large proportion of Iranian loanwords, which are not limited to specific semantic fields but extend into the basic vocabulary, as will be shown in section 3. Moreover, since Armenian has undergone a complex series of sound changes in its development from Indo-European, many words which can be shown to be inherited from Indo-European

Meillet, Antoine

1897–1898 Recherches sur la syntaxe comparée de l'arménien, I. les démonstatifs. *Mémoires de la Société de Linguistique* 10: 241–271. [Reprinted in Meillet 1962, 5–35.]

Meillet, Antoine

1962 *Études de linguistique et de philologie arméniennes*, I. Lisbon: Imprensa Nacional.

Minassian, Martiros

1988 L'article du premier terme non substantivé d'une subordonnée en arménien classique. *Le Muséon* 101: 97–158.

Ouzounian, Agnès

1992 *Le discours rapporté en arménien classique*. (Bibliothèque des Cahiers de l'Institut de Linguistique de Louvain 63). Louvain-La-Neuve: Peeters.

Ouzounian, Agnès

1996 À propos des relatives en arménien classique: le fonctionnement de l'article et le cas des relatives à antécédent démonstratif d'après les évangiles et les actes des apôtres. In: Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan, 75–92.

Sakayan, Dora (ed.)

1996 *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan.

Thomson, Robert W.

1989 [1975] *An Introduction to Classical Armenian*. 2nd edn. Delmar, NY: Caravan.

Jared S. Klein, Athens, GA (USA)

64. The lexicon of Armenian

- | | |
|--|-------------------|
| 1. Inherited vocabulary | 4. Word-formation |
| 2. Loanwords | 5. References |
| 3. Technical and scientific vocabulary | |

1. Inherited vocabulary

1.1. In 1875, Heinrich Hübschmann conclusively showed that Armenian constituted a separate branch of Indo-European (Hübschmann 1875). In the decades before Hübschmann, Indo-Europeanists had judged the language to belong to the Iranian subgroup (see Schmitt 1975). The reason for this mistaken analysis of the origin of the Armenian language lies in the nature of the lexicon. Armenian has a smaller proportion of inherited vocabulary than other Indo-European languages of comparable age, and it contains a large proportion of Iranian loanwords, which are not limited to specific semantic fields but extend into the basic vocabulary, as will be shown in section 3. Moreover, since Armenian has undergone a complex series of sound changes in its development from Indo-European, many words which can be shown to be inherited from Indo-European

are not immediately recognizable as such, for example, *elbayr* ‘brother’, cognate with Sanskrit *bhrātar-*, Latin *frater*, and English *brother*; and *k’oyr* ‘sister’, cognate with Sanskrit *svasar-*, Latin *soror*, and English *sister*. Only after Hübschmann worked to show systematic correspondences between the inherited vocabulary and that of the other branches of Indo-European could the true nature of the language be revealed.

1.2. The comparatively small number of inherited items in Armenian can be shown by comparing the volume of Armenian entries in the standard etymological dictionaries of Indo-European. In the index to Pokorny’s *Indogermanisches Etymologisches Wörterbuch* (Pokorny 1959), Armenian entries occupy seven and a half pages. In comparison, the index for Old Irish is nineteen pages long, for Gothic ten and a third pages, for Lithuanian twenty-eight pages, and Old Church Slavonic sixteen pages. Of course the number of entries in an etymological dictionary is only a very rough guide to the proportion of inherited vocabulary in a language. It is possible that the proportion of entries in an etymological dictionary corresponds to the amount of attention scholars have paid to a language rather than the real rate of survival among the vocabulary. Untangling the sound changes which have operated in the pre-history of Armenian is not always easy and many of the essential research materials, such as the Venice dictionary (Awetik’ean, Siwrmēlean, and Awgorean 1836–1837), or Ačařyan’s excellent etymological dictionary (Ačařyan 1971–1979), are not easily accessible to western scholars since they are written in Armenian. It would therefore be no surprise if the comparative paucity of Armenian data in Pokorny’s dictionary reflected its comparative neglect by Indo-Europeanists. However, if we consider the position of Armenian in the more recent *Lexikon der indogermanischen Verben* (Rix et al. 2001), we see that the situation is still largely similar to what was found in Pokorny: there are 159 entries for Armenian, compared with 253 for Old Irish, 277 for Gothic, 606 for Lithuanian, and 388 for Old Church Slavonic. Even Tocharian B, which had only had two and a half pages in Pokorny’s index now surpasses Armenian in Rix et al. (2001) with 179 entries. It is no longer possible to argue that the Armenian vocabulary has been neglected, since a sizeable monograph on the Armenian verb (Klingenschmitt 1982) included a thorough account of the attested Armenian verbal roots and their possible etymologies and unearthed several connections that had previously been overlooked. It now seems unlikely that there are many other verbal stems in the Classical language with Indo-European etymologies which have been escaped notice. The reduced number of inherited words in Armenian can also be demonstrated by considering Indo-European roots which are widely attested in other branches, but which are not found in Armenian. These include basic vocabulary items such as **h₁ei-* ‘go’ (retained in Anatolian, Indo-Iranian, Greek, Italic, Germanic, Baltic, Slavic, and possibly Albanian), **sek^w-* ‘follow’ (retained in Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, Slavic, and possibly Albanian), **pek^w-* ‘cook’ (retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Baltic, Slavic, Albanian, and possibly Anatolian), **h₁reud^h-* ‘red’ (derivatives retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, and Slavic), **k^hmtóm* ‘hundred’ (retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, and Slavic), and **ǵ^hd^hom-* ‘earth’ and derivatives (retained in Anatolian, Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, Slavic, and Albanian). Some of this material has been replaced by other items inherited from Indo-European: the standard word meaning ‘go’ *ertam* (aorist *čogay*) is derived from Klingenschmitt (1982: 103 f.) as a compound of the original IE word for ‘stand’; the

word for ‘follow’ *hetewem* is formed on the inherited noun *het* ‘footstep’. Others are replaced by words only found in a minority of the IE languages, such as the word for ‘cook’ *ep'em*, which has a cognate only in Greek *ἔψω* (see Clackson 1994: 172). Iranian loanwords have supplanted others; thus, the word for ‘red’ in Armenian is a loanword, *karmir*, as are most of the Armenian color terms. However, much of the Armenian vocabulary is of unknown or doubtful origin, as the words for ‘hundred’ *hariwr* and ‘earth’ *erkir* (see further below). Indeed, around 50% of the Armenian vocabulary is of unknown origin (following the count of the material in Ačařyan 1971–1979 given by Solta 1990: 13), neither clearly Indo-European nor borrowed from any other language of which we have records. This “unknown” element of the vocabulary includes many other basic vocabulary items, including, for example, *mort* ‘skin’, *masn* ‘part’, *xawsem* ‘speak’, *c'owrt* ‘cold’.

Absolute numbers of Armenian roots inherited from Indo-European are difficult to know precisely, since many Armenian etymologies remain contested. Hübschmann’s *Armenische Grammatik* (Hübschmann [1895/1897] 1992), considers 438 words which previous to him had been proposed as inherited, but many of these are uncertain and several are dismissed out of hand by Hübschmann. In the time since Hübschmann, many new etymologies have been proposed, and there have been attempts to resurrect etymologies which were explicitly rejected by him, so that the number of words now thought to be inherited has increased. Solta (1960) lists 453 inherited items, Ačařyan (1971–1979) gives 713 (figure from Solta 1990: 13), Džaukjan (1987) has around 1400, and Martirosyan (2010) has 961. The figure given in Solta (1960) is likely to be too low: we have seen that Rix et al. (2001) include 159 entries for Armenian verbs, but Solta includes only 114 verbs. Conversely, Džaukjan’s figure is probably too high, since a number of the forms he includes are doubtful and not accepted by all scholars. For example, an earlier connection of the word for ‘cub, young animal’ *korıwn*, with Greek *βρέφος* and Slavic *žrěbę* was already rejected by Hübschmann (1992: 461) since the expected Armenian cognate would be *korb-*. Pedersen (1911) argued for the retention of the etymology, however, and he is followed tentatively by Džaukjan (1987) and with more certainty by Olsen (1999: 492). The etymology is not included in Solta (1960) nor endorsed by Ačařyan (1971–1979). With such a small proportion of inherited words in Armenian, the weight accorded to these controversial etymologies becomes of greater importance when assessing questions such as the relationship of the Armenian vocabulary to that of other languages.

1.3. Despite the general loss of inherited vocabulary in Armenian, there are several examples of words that can be assumed to derive from archaic Indo-European lexical items. The Armenian word for ‘hand’ *jeřn*, has cognates in Anatolian, Tocharian, Greek and (with a different suffix) Indo-Iranian, and appears to have been replaced in the other branches of Indo-European by neologisms. The word for ‘blood’ *ariwn*, has cognates in Anatolian, Sanskrit, Greek, archaic Latin, and Baltic, but the root is elsewhere lost (as it is in later Greek and Classical Latin). The adjective *barjr* ‘high’ continues a primary adjective elsewhere limited to Hittite, Tocharian and Indo-Iranian, with the *u*-stem form only found in Armenian and Hittite. The pair *han* ‘grandmother’ and *haw* ‘grandfather’ have their closest cognates in Hittite *hanna-* ‘grandmother’ and *huhha-* ‘grandfather’, and no other single language outside the Anatolian branch preserves both members of the pair with the same meanings. Indeed, Armenian has retained most of the reconstructi-

ble Indo-European kinship terms, including the terms for the wife's relations by marriage: *skesowr* 'husband's mother', *skesayr* 'husband's father', *taygr* 'husband's brother', *tal* 'husband's sister', and *nēr* 'husband's brother's wife' (see Szemerényi 1977 on these terms and, for *nēr*, most recently Olsen 1999: 190 f. and Kortlandt 1997). Some other Armenian lexical items have been seen to continue words of high or "poetic" registers of the parent language (Porzig 1954: 482; Schmidt 1985: 215; Ritter 2006): *arcowi* or *arciw* 'eagle', plausibly connected to Sanskrit *ṛjipyá-*, an epithet of eagles in the Ṛg-Veda (despite the doubts of de Lamberterie 1978: 251–262); *ji* 'horse', cognate with the poetic Sanskrit word *haya-* 'steed' (note that the Armenian word for 'donkey', *žš*, corresponds to the normal word for horse elsewhere, Latin *equus* etc.; see de Lamberterie 2006); *erg* 'song' cognate with Sanskrit *arc-* 'sing, praise'; *arew* and *aregagn* 'sun' cognate with Sanskrit *ravi-* (see Eichner 1978 for the history of this pair and Hittite cognates). It is also possible that the Armenian words for 'heaven' and 'earth' *erkin* and *erkir*, which share the same initial sequence as Armenian *erkow* 'two', continue an inherited conception of the heaven and earth as the two halves of the universe, found also in Vedic hymns and early Greek cosmologies (Meillet 1937; subsequent scholars have admired Meillet's suggestion without wishing to endorse it fully; see de Lamberterie 1988: 225 f.).

Some words appear to be related to forms in other IE languages but show peculiar deviations in their phonological development or morphological behavior. A number of these have been explained through blending or "contamination" of two different inherited terms. The idea that a particular lexical item can arise from the merger of two separate terms is of course a useful tool for the desperate etymologist, but Armenian does provide a larger number of plausible candidates for this type of development than many other languages. A few examples will suffice to illustrate. The word for 'honey', *melr*, appears to be cognate with the family of Hittite *melit*, Greek μέλι, Latin *mel*, etc. However, the Armenian word is a *u*-stem and this has led scholars to think that there has been some influence from a different etymon, **med^hu* 'mead' (Sanskrit *mádhu*, Greek μέθυ, Lithuanian *medūs* etc.; see further Martirosyan 2010: 462). The word for 'son' in opposition to 'daughter', *owstr*, appears to continue the same root as found elsewhere to mean 'son', Greek υἱός, Sanskrit *sūnú-* etc., but with the sequence *-str* taken from *dowstr* 'daughter' (Szemerényi 1977: 19; Olsen 1999: 149). The word for 'nail' or 'claw' *elowngn* is generally taken to be cognate with the set Greek ὄνυξ, Latin *unguis*, Old Irish *ingen*, but the explanation for the initial *el-* is not known; it was seen by Osthoff (1901: 248) as contamination from the word for 'horn' *eljewr*. The adjective *k'alcr* 'sweet' has been explained by Ačařyan (1971–1979 IV: 546), Hamp (1983: 38–42) and de Lamberterie (1990: 502) as the result of some sort of combination of the stem **sweh₂du-* 'sweet' (Greek ἡδύς, Latin *suāuis* etc.), **dluku-* 'sweet' (Greek γλυκύς, Latin *dulcis* etc.) and perhaps **saldu-* (Lithuanian *saldūs*, Old Church Slavonic *sladükū* 'sweet'). Klingenschmitt (1982: 227) proposed that the verb meaning 'kill', *spananem*, was a blending of two different roots, one of which survives in Greek σφάζω, the other in Greek κτείνω, and this suggestion is endorsed by Watkins (1995: 521), who however suggests that the second root involved is that of Greek θείνω, Sanskrit *han-*.

1.4. The inherited vocabulary of Armenian has been examined several times for the light it can bring to bear on the question of the position of Armenian within the Indo-European family. The first systematic attempt to use the vocabulary to assess the proxim-

ity of Armenian to other branches of Indo-European was made by Pedersen, who gave figures for lexical agreements between Armenian and other IE branches in an encyclopedia article (Pedersen 1924: 224). By Pedersen's count, Armenian had 95 special agreements with Greek, 53 with Indo-Iranian, 34 with Balto-Slavic, 24 with Germanic, 15 with Italic, and only 3 with Celtic. (Pedersen also gave figures for lexical agreements working only from the material in Hübschmann 1992: Greek 59, Indo-Iranian 37, Balto-Slavic 14, Germanic 10, Italic 10, and none for Celtic). Unfortunately, Pedersen never gave the full list of lexical material he had worked on, and it is not quite clear what he means by his "special agreements". In 1960, Solta undertook a much more thorough survey of the Armenian vocabulary and again found the lexical agreements between Armenian and Greek to be the most numerous and important. In his summary (1960: 459–485), he lists 53 noteworthy correspondences between Armenian and Greek (that is, both cognate pairs which are exclusive to the two languages and ones where there is a unique morphological or semantic development between the languages), but only 19 between Armenian and Indo-Iranian, 19 with Germanic, 17 with Latin, 13 with Baltic, 10 with Slavic, 4 with Celtic, 4 with Tocharian, and 3 with Albanian. The Greek-Armenian material was re-examined by Arutjunjan (1983) and Clackson (1994). In Clackson (1994), I argued that the number of innovative lexical agreements was much lower than had previously been thought, although there were undoubted cases where the two languages had made shared innovation. For example, I dismissed (1994: 158) the etymological connection of the initial *o-* of Armenian (*o*)č' 'not' with Greek οὐ from PIE **h₂oyu* 'life' proposed by Cowgill (1960), given the lack of parallels for the phonological development in Armenian (in Clackson 2005, I gave an alternative explanation of the origin of the *o-* of Armenian *o*č', through consideration of the frequent collocation *o*č' *ok*' 'no one'). However, it is likely that the short form, č', of the Armenian negative particle can be connected to the final element of the epic form of the Greek negative οὐ-κι, through grammaticalization of a particle **k^wid* 'something'. Note that Kortlandt (1986) and Klingenschmitt (1994: 245) have drawn attention to the lexical agreements between Armenian and Albanian. In more recent years, several scholarly teams have undertaken mass comparison of basic vocabulary lists to unearth the relationships between the lexicons of different Indo-European branches. Gray and Atkinson, who produced a tree for the Indo-European language family (2003) using basic vocabulary lists compiled by Isidore Dyen and described in Dyen, Kruskal, and Black (1992), note that the vocabulary list for Armenian was based on a modern Western Armenian variety, rather than the Classical language. In Gray and Atkinson's tree, the Armenian branch forms a node with the Greek branch. Armenian is also grouped with Greek on the phylogenetic tree generated by a team of mathematicians and linguists including Don Ringe (see for example Taylor, Warnow, and Ringe 2000 and Ringe, Warnow, and Taylor 2002). Ringe, Warnow, and Taylor (2002: 102) list six Armenian-Greek lexical isoglosses that support a subgroup, three of which (the words for 'day', 'not', and 'wind') are seen as likely candidates for shared innovations. It is worth repeating the data here:

Armenian *awr* 'day', Greek ἡμᾶρ, derived respectively from **āmōr* and **ām̄r*; Latin and Vedic words for 'day' derive from **dyew-*.

Armenian *ayr* 'man, husband', Greek ἀνὴρ 'man, husband'; both share the meaning 'husband' for which Tocharian and Indo-Iranian use a derivative of **poti-*.

Armenian (*o*)č' 'not', Greek οὐ(κ) 'not' (see above); most other IE languages show a reflex of **ne*.

Armenian *holm* ‘wind’, Greek ἄνεμος ‘wind’ derived, respectively, from **h₂onh₁mo-* and **h₂enh₁mo-* (see the doubts of Beekes 1972: 129 on this equation); other IE languages form the word for ‘wind’ as a derivative of the root **h₂weh₁-*.

Armenian *alam* ‘I grind’, Greek ἀλέω ‘I grind’; other language branches (Hittite, Italo-Celtic, Germanic, and Balto-Slavic) form words for ‘to grind’ from the root **melh₂-*.

Armenian *nor* ‘new, fresh, young’, Greek νέος ‘new, young’; Greek and Armenian extend the inherited term for ‘new’ to mean ‘young’, whereas the stem **h₂yu-Hen-* is employed elsewhere.

2. Loanwords

2.1. The Iranian languages have contributed a large number of loanwords to the Armenian lexicon. Hübschmann listed 686 Iranian words in Armenian (Hübschmann 1992), and since then the number of identified loans has increased substantially with the improved knowledge of Middle Iranian languages. There is no comprehensive modern collection of the Iranian loanwords (although a project to produce one was announced in Considine 1979), but much of the material is gathered in Bailey (1986), Schmitt (1983), Schmitt (1986), and Olsen (1999: 857–920). Much work has been done on locating the source of Armenian loanwords within the Iranian language group. Armenia and Armenians are first mentioned by name in the Old Persian Bisutun inscription dated to the end of the sixth century BCE, and thereafter in the work of the Greek historian Xenophon. Since most of those named in these sources have Iranian names (for example Dādarši- in the Bisutun inscription [DB II 29], Tigranes in Xenophon), it seems plausible that some of the Iranian loanwords stretch back to Achaemenid times; and this is very likely for the name of the Aryans, *Arik’* (genitive plural *Areac’*), which exactly reflects the stem *ariya-* seen in Old Persian *ariya-*, Avestan *airiia-* (Schmitt 1983: 77). Indeed, it has been argued that some Iranian loans took place so early that they participated in the Armenian consonant shift, and show voiceless stops for Iranian voiced stops, and Armenian aspirates for Iranian voiceless stops. The most likely example for a pre-sound shift loan is the word for ‘garden’ *partēz*, beside Avestan *pairidaēza-*, New Persian *palēz*, which was explained in this way already by Meillet (1911: 250). But other scholars have added further possible examples: de Lamberterie (1988: 245–262) proposed *arcat’* ‘silver’ (cf. Old Persian *ardata-*, Avestan *arəzata*) and *arcowi* ‘eagle’ (but this word may also be inherited, see 1.3) and Olsen (1999: 857) added *ciran* ‘apricot’, *cirani* ‘purple’ (both from an Iranian cognate of Avestan *zaraniia-* ‘golden’), and *parc* ‘proud’ from an Iranian word cognate with Avestan *bərəz-* ‘high’.

2.2. However, the vast majority of Iranian loans in Armenian appear to have entered the language at a later date, during the period of the Arsacid dynasty’s rule over Armenia from the 1st to the 3rd century CE. The language of the Arsacids was the north-west Iranian dialect Parthian, and many of the loanwords in Armenian can now be directly compared to forms found in surviving Parthian texts, and many more with putative Parthian forms reconstructed through attested words in other Iranian languages. The Arsacid loans permeate the Armenian lexicon, referring not just to aspects of administra-

tion, religion, and military service (such as *azat* ‘free’, Parthian *āzāt*; *hraman* ‘command’, Parthian *framān*; *marz* ‘province’, Parthian *marz*; *den* ‘religion’, Parthian *dēn*; *dēw* ‘devil’, Parthian *dēw*; *zoh* ‘offering’, Middle Persian *zōhr*; *nizak* ‘spear’, Parthian *nēzag*; *zēn* ‘weapon’, Parthian *zēn*); but also numerous aspects of everyday life including names of colors (*seaw* ‘black’, Parthian *syāw*; *karmir* ‘red’, Middle Persian *karmir*), domestic objects (*črag* ‘lamp’, Modern Persian *čarāy*; *šapik* ‘shirt’, Middle Persian *ša-bīg*; *xan* ‘basket’, Parthian *xawān*), people and relationships (*harazat* ‘blood(-brother), (-son)’, Avestan *hadō-zāta-*; *tohm* ‘family’, Parthian *tōxm*; *pařaw* ‘old woman’, Modern Persian *pārāw*), abstracts and emotions (*oyž* ‘strength’, Middle Persian *ōz*; *ah* ‘fear’, Middle Persian *āhal*), adverbs (*yavēt* ‘always’, Middle Persian *yāwēt*; *-pēs* ‘-how’, Avestan *-paēsa-*) and numerals (*hazar* ‘1,000’, Parthian *hazār*; *biwr* ‘10,000’, Parthian *bē-war*). The large influx of loans during the Arsacid period was most probably assimilated over a long time period, and there are some linguistic grounds for dividing the loanwords up into different chronological strata. The more recent Iranian loans in Armenian show the equivalence of Armenian *e* for Iranian *ē*, and Armenian *o* for Iranian *ō*, for example *den* ‘religion’ beside *dēn* and *zoh* ‘offering’ beside Persian *zōhr*. But other words, which may reflect earlier borrowings, show *ē* where the Iranian source had *ē* and *oy* where the source had *ō*, as *zēn* ‘weapon’ beside Parthian *zēn*, and *oyž* ‘strength’ beside Middle Persian *ōz*. Furthermore, some loans appear to have undergone an inner Armenian development of a prothetic vowel before initial *r-*, whereas in later loans initial *r-* is preserved: thus *erasan* ‘rein’ derives from an Iranian stem attested in Middle Persian as *rasan*, and *řočik* ‘provisions’ derives from the same stem as Middle Persian *rōzīg* (note that this word also shows the “late” development of the Iranian vowel *ō*).

It is also possible to isolate a yet later stratum of Iranian loanwords in Armenian, those that stem from the Sasanian period, when Armenia had moved away from the Persian cultural and religious sphere. Loans from this period are generally restricted to technical military or religious terms, they are restricted in their textual occurrences, and they tend not to show the same range of derivative formations as the loans of the Arsacid period. These loans also reflect the phonology of the south-west Iranian language of the Sasanian Persians. Thus *řatagov* ‘proponent, advocate’ (Middle Persian *řatagōv*), *řahan-řah* ‘king of kings’ (Middle Persian *řāhān řāh*) and *payik* ‘guard’ (Middle Persian *paik*) show characteristic Middle Persian developments of initial *y-* to *ř-* (*řatagov*), *xš-* to *ř-* and *-θr-* to *-h-* (*řahanřah*), and medial *-đ-* to *-y-* (*payik*, see Bolognesi 1960: 56, 21, and 43). None of these words appears in the Armenian Bible translation, and from *payik* and *řahanřah* there are no further derivatives formed.

While the majority of Iranian loans can be sorted into their appropriate chronological and dialectal layer, whether early or late, and from a north-west or south-west variety, there remain a number of words for which the origin is identifiably Iranian, but the details are unclear or unparalleled. Thus the word for ‘road’, *čanaparh*, was taken by Bailey (1930: 61) as an Iranian compound, the second half of which represented *-přθ-* or *-parθ-* ‘way’, but the origin of the initial element *čana-* remains disputed; the best explanation is probably still that of Nyberg (1931 s.v. *puhl*) that it is a dissimilation from *čarana-parθ-*, with the first element from the Iranian root *čar-* ‘go’. A particularly noteworthy pair of words are those which have their closest cognates with forms in eastern Iranian dialects: *kari* ‘very’ and *margarē* ‘prophet’; compare Sogdian *k’đy* and *m’rkr’y*. Henning (1958: 93) provided an explanation for these words that is now widely accepted: they are elements of the speech of the Parnians, the eastern Iranian tribe which

migrated westward as overlords of the Parthians. For many other words the exact Iranian source remains disputed: thus the word for ‘treasure’, *ganj*, has cognates in many Iranian languages, including Parthian *gazn*, Middle Persian *ganz* and Modern Persian *ganj*; Henning took Armenian *ganj* to have originated in a Median word (1963: 197), but Olsen (1999: 861) argues for it as another Parnian loan.

2.3. Classical Armenian has also borrowed a substantial number of words from Greek and Syriac. The majority of Syriac loans entered the language when the Armenians adopted Christianity as the state religion (the traditional date for this is 306 CE), and the most common words are in some way connected with the Christian religion, for example, *k'ahanay* ‘priest’ (Syriac *kāhanā*), *šabat* ‘week, sabbath’ (Syriac *šabbāthā*), and *tlay* ‘infant, child’ (Syriac *ṭalyā*), used in the Armenian Bible translation particularly to denote the innocent and receptive child as opposed to the wise and knowing sceptic. Some Greek loans are also connected with the Christian religion, for example, *salmos* ‘psalm’ (ψαλμός), *episkopos* (ἐπίσκοπος), *pentakostē* ‘pentecost’ (πεντηκοστή), and *martivros* (μάρτυρος). However, these reveal themselves as recent specialized borrowings, and Meillet (1926) showed that the majority of the Armenian words used for Christian concepts are either borrowed from Syriac or Iranian (Iranian loans include *ašakert* ‘disciple’, *hrestak* ‘angel’, *margarē* ‘prophet’, *awetaran* ‘gospel’) or are native Armenian terms, some of which lack any straightforward or known etymology (for example *erēc* ‘priest’ which can be connected to the family of Greek *πρεσβύς*, Latin *priscus*, and, with unknown etymologies *Astowac* ‘God’ [see Olsen 1999: 545 f. for a survey of some proposed etymologies], *ařak'em* ‘I send’ [from which *ařak'ič* ‘apostle’ is derived], and *mkrt'em* ‘I baptize’). Accordingly, Meillet proposed that the Armenian word for ‘church’ *ekelec'i*, was not a direct loan from Greek ἐκκλησία but entered the language via Syriac or possibly an Iranian source, just as was the case for other Greek loans such as *p'ilisop'ay* ‘philosopher’, *ak'sorank* ‘exile’, *katapar* ‘model’ and *lambar* ‘lamp’ (respectively Greek φιλόσοφος [via Syriac], ἔξορία [via Syriac], καλοπόδιον [via Parthian], and λαμπάδα, accusative of λαμπάς [via Parthian]). The possibility of a Syriac or Iranian intermediary unfortunately complicates Hübschmann’s theory (1906: 472–477) that the Armenian word for the antichrist, *neřn*, was a very early loan of the Greek name for the Emperor Nero, Νέρων, which had undergone the Armenian loss of final syllables. Armenian did however borrow a number of words from Greek to denote items that reflected Greek technology, thought or cultural life: for example *pnak* ‘dish’ (πίναξ), *balistr* ‘catapult’ (βαλίστρα), *hiwl* ‘matter’ (ῥήλη), *hřetor* ‘orator’ (ῥήτωρ), and *t'atr* ‘theatre’ (θέατρον). Thumb (1900) argued that the majority of these words were “eye-borrowings” rather than “ear-borrowings” and represented the learned pronunciation of the Greek words, since most of the Armenian representations of the Greek sounds were in accord with the learned pronunciation of Greek, rather than the spoken Greek of the middle of the first millennium CE. In support of this theory, he noted that Armenian represented, for example, the Greek consonants written φ θ χ and β as aspirates rather than fricatives. However, more recent evidence has suggested that these loanwords need not be of a particularly elevated register; note in particular the discovery of a papyrus fragment, dated to around the 6th century, with Greek words written in Armenian script. The text appears to have been taken down by dictation, and it shows much the same equivalences for the Greek vowels and consonants as are found in the Greek loanwords (see Clackson 2000 and 2003). Many Armenian words of scientific and technological vocabulary, from

the sixth century to the present day are also formed through a process of calquing of Greek compound and suffixed forms, and these will be discussed in more detail in the next section.

2.4. Several Classical Armenian words for flora and fauna appear to be borrowed from Mediterranean or Near Eastern languages which are now irrecoverable, since the same words appear as loanwords in other ancient languages: for example, *ewl* (post-Classical *iwl*) ‘oil, unguent’ appears to be related to the word borrowed into Greek as ἔλαιον ‘olive oil’; *tōwz* ‘fig’ appears to be in origin the same word as Greek σῦκον and Latin *ficus*; *xstor* ‘garlic’ links with Greek σκόροδος, Albanian *hurdhë*; *inc* ‘leopard’ with Sanskrit *simha-* ‘lion’, Tibetan *señ-ge*; and *owlt* ‘camel’ with Assyrian *utru*, Avestan *uštra-*, and Urartian *ultu*. Urartian is the most likely source of the Armenian word for ‘camel’, and further plausible loans from Urartian include Armenian *sowr* ‘sword’ (Urartian *šuri* ‘weapon’), *san* ‘cauldron’ (Urartian *šani* ‘vessel’). Since the Urartian corpus is very small, the related language Hurrian sometimes provides supporting evidence for early Urartian loans into Armenian. One possible case is the Armenian word for ‘apple’ *xnjor*, which is an exact match for Hurrian *hinzuri* ‘apple’, although the word is also found in Aramaic *ħazzūrā*. Diakonoff (1985), Greppin and Diakonoff (1991), and Greppin (2010) list other possible Hurro-Urartian borrowings into Armenian, but many are either unattested in our surviving Hurrian and Urartian texts or are semantically or phonologically problematic. The Armenian word for ‘wine’, *gini*, is most likely another example of a Mediterranean *Wanderwort*; compare Mycenaean Greek *wo-no*, Latin *uīnum*, Arabic *wain*, Hebrew *yayin*, Hittite *wiyana-*, Hieroglyphic Luwian *wayana-*. The Georgian word for wine, *γvino*, is of particular interest, as it shares the presence of a velar consonant at the beginning of the word with the Armenian word. Hübschmann (1992: 397) noted the correspondence alongside other words which appear to be shared by Armenian and South Caucasian languages, but left open the question of whether the loan was from, or to Armenian. Many scholars have subsequently taken the Armenian as the original form in the two languages, and the Georgian initial sequence *γv-* to reflect an intermediary stage in the development of Armenian *g-* from original **w-* (see Greppin 1998). Indeed, there are almost no loanwords from South Caucasian languages which are widespread and long established in the Armenian lexicon; most of the loans are restricted to Armenian dialects which have long been in contact with neighboring South Caucasian varieties (see the survey in Greppin 2000).

2.5. In the time since the earliest Armenian texts, speakers have continued to adopt loanwords from the languages with which they came in contact: Persian and Iranian dialects, Arabic, South Caucasian, Greek, Turkish, French, Russian, and English. The earliest French loans date from the 12th century, when the Cilician Armenians came into contact with crusaders, but have since spread throughout the language: for example *paron* (in Modern Western Armenian *baron*) is borrowed from French *baron* and has become the standard equivalent to the title ‘Mr.’, or the address ‘sir’. However, since the formation of the modern standard literary languages in the nineteenth century, there has been a reluctance to incorporate foreign loans into the standard. Thus Standard Modern Eastern Armenian avoids the virtually pan-European terms *music*, *problem*, and *coffee* using instead the “native” terms *eražštowt’iwn* (in fact an early Iranian borrowing), *xndir* (of unknown origin), and *sowrč* (plausibly explained as a metathesis of *sew jowr*

‘black water’, the first element of which is an Iranian loan, the second an inherited word).

3. Technical and scientific vocabulary

3.1. The Armenian technical and scientific vocabulary is formed to a large extent through an extensive system of calquing of compound or suffixed words in Greek, and to a lesser extent Latin. Many of the words used to denote items of modern technology are formed from calques in this way. For example, the word for ‘photograph’, *lowsankar*, is a compound incorporating the lexical stems of *loys* ‘light’ and *nkar* ‘picture’; the word for ‘television’, *heřowstatesayowt’iwn*, is built from *heřowst* ‘from a distance’ and the root of the verb meaning ‘see’, *tesanem*, with the abstract noun suffix *-owt’iwn*. This calquing technique was first inaugurated in the translations of Greek grammatical and philosophical works during the late 5th and 6th centuries of the Christian era, a period of Armenian literature known as “the Hellenizing School”. The works of the Hellenizing school marked a departure from the earlier translation practice of the Armenian Bible translators and writers of the “golden age”. For example, the fifth century Armenian author Eznik used the noun *dprowt’iwn* ‘scribery’ (formed from *dpir* ‘scribe’) to translate Greek γραμματική, whereas the translators of the Hellenizing School used the new artificial term *k’erakanowt’iwn*, formed from the verb *k’erem* ‘I scrape’, which was taken as the equivalent to the base meaning of Greek γράφω, presumably following the discussion of the original meaning of γραμματα given in the grammatical work attributed to Dionysius Thrax. The term *dprowt’iwn* seems to have been rejected as a satisfactory translation of γραμματική because its range of meanings was too wide to allow it to be used in the specific sense of ‘grammar’. In the Armenian Bible translation, *dprowt’iwn* can mean ‘learning’, ‘letters’, and even ‘book’. The new term *k’erakanowt’iwn* could stand in technical works without leading to any ambiguity.

3.2. A distinctive feature of the calques of the Hellenizing School is their reliance on a completely artificial system of composition. In Classical Armenian (that is, the Armenian of the Bible translation), nominal compounds are not uncommon, but composite verbs are only regularly formed from three prefixes, *ař-*, *y-*, and *and-*, none of which has an exact correspondence with any of the Greek prepositions used in composition. The Hellenizing School vastly increased the number of Armenian prefixes by creating a range of “artificial prefixes”, which they used to translate Greek prepositions in composition. The following gives some of the Armenian equivalents to Greek prefixes that are used (see further Mowradyan 1971: 136–152 and Mercier 1978–1979: 64–67).

<i>ver-</i>	անá
<i>der-</i>	անτί
<i>bac’-</i>	áπό
<i>tram-, hastat-</i>	διά
<i>art-</i>	έκ

<i>ner-</i>	ἐν
<i>mak-, ver-</i>	ἐπί
<i>ənd-</i>	μετά
<i>yar-, mat-, tar-</i>	παρά
<i>par-, bak-</i>	περί
<i>nax-, (y)araj̄-</i>	πρό
<i>aṙ-</i>	πρός
<i>bał-, šar-, šal-, jok-, p'al-, par-</i>	σύν
<i>ger-</i>	ὑπερ
<i>stor-, ent-, p'al-</i>	ὑπό

Some of these prepositions originate from earlier Armenian prepositions, such as *aṙ-*, used to translate Greek πρός, but most of them are newly formed from nouns or verbs in the language. Thus, for example, the prefixes *šal-* and *p'al-*, which are both used to translate Greek σύν, are derived, respectively, from the verbs *šalem* 'I encircle' and *p'alem* 'I join'. It is immediately noticeable from this table that the correspondence between the Armenian prefixes and the Greek is not one-to-one, since some Greek prefixes are rendered by more than one Armenian prefix. However, the same Armenian prefix (with the exceptions of *ver-* and *p'al-*) does not translate different Greek prefixes and nearly every Greek prefix has an Armenian equivalent. The equivalence of several different Armenian prefixes to a single Greek prefix does not seem to have arisen from a desire to translate different nuances of the Greek prefixes, but rather it seems to reflect the difficulties the translators had in finding native equivalents to the Greek. This rich array of newly-formed prefixes also helped compensate for the relative poverty of the Armenian suffixal system in comparison with Greek. The Armenian translators were not so bold or so rigorous in their treatment of the different suffixes used to form technical terms. Rather than attempt to create new artificial suffixes, they relied upon those already existing in the Classical language, and the replacement of a Greek suffix with its Armenian "equivalent" is not entirely consistent. However, the following correspondences between Greek and Armenian suffixes are used with some degree of regularity:

<i>-akan</i>	-ικός
<i>-eli</i>	-τέον
<i>-ac</i>	-μα
<i>-ot̄</i>	-τηρ, -τωρ
<i>-ayin</i>	-ιος
<i>-eal</i>	-μενος
<i>-pēs-, -bar</i>	-ως

Note also that the Armenian abstract noun suffix *-owt'iwñ* is frequently used to translate a variety of different Greek abstract noun suffixes, such as -ία, -σις, -ισμός etc. Armenian

also uses a far greater number of unsuffixed nouns than Greek, and these are widely used to render both Greek agent and action nouns.

4. Word-formation

4.1. Word-formation processes in Armenian largely correspond closely to those found in other branches of Indo-European. There is widespread use of a large number of nominal suffixes, a more restricted set of verbal suffixes, and compounding. As in other IE languages, the basic unit of word formation in Armenian is the lexical root. In some verbs, the aorist-stem is identical with the lexical root: for example, root *tes-* ‘see’, 3rd singular aorist *e-tes* ‘s/he saw’, but this pattern is of limited productivity, and most present and aorist verbal stems are formed through suffixation, as is the case for the present stem *tesanem* ‘I see’. Most noun and adjective stems are also suffixed forms of the root, although in many cases the original root shape may be obscured by phonological developments. Often the nominative singular is coincidental in shape with the root, and the suffix is only apparent in oblique cases. For example, the root *gorc-* means ‘work’ and among its synchronic derivatives are the noun *gorc* ‘work’, the verb *gorcem* ‘I work’ and a compound adjective *angorc* ‘lazy’. The noun *gorc* and adjective *angorc* belong to different declensions: *gorc* is an *o*-stem, genitive plural *gorcoc*, whereas *angorc* is an *i*-stem, genitive plural *angorcic*; accordingly, *gorc* and *angorc* are best seen synchronically as suffixed forms. Note that the relationship between the simplex *o*-stem *gorc* and the *i*-stem compound *angorc* can be compared to similar processes in other IE languages, for example, Latin *o*-stem *arma* ‘weapons’, *i*-stem compound *inermis* ‘unarmed’, and is taken to be an inheritance from the parent language. Although most derivation in Armenian is effected through lexical stems, Armenian is unusual among Indo-European languages in that nouns and verbs may also be derived directly from an inflected nominal form, or from a complete syntagm. For example: *kanambi* ‘having a wife’ is derived from the instrumental singular, *kanamb*, of *kin* ‘woman, wife’; a common word for ‘night’ *c’ayg* (*o*-stem) derives from the prepositional phrase *c’-ayg* where *c’-* means ‘until’ and *ayg* (normally *ow*-stem) is the word for ‘dawn’; the adjective *č’k’mel* ‘innocent, free from sin’ is formed from a complete sentence, *č’ik’ mel* ‘there is no sin’.

4.2. Unlike the older Indo-European languages, Classical Armenian does not exhibit vowel alternation (ablaut) within a lexical root as a productive derivational marker. Ablaut alternations are still found in some of the inherited vocabulary items, for example, *barjr* ‘high’ and compounds *barjraberj* ‘very high’ and *erknaberj* ‘sky-high’, where the second member of the compound shows the *e*-grade as opposed to original zero-grade in the simplex adjective (see de Lamberterie 1986 on compounds of this type); note also *snanim* ‘I nourish’ and *san* ‘nursling’ (see Klingenschmitt 1982: 226). Inherited vowel alternations can also be found preserved in isolated suffixed formations, as in the well-known example of *anjn* ‘person, self’ nominative plural *anjnk*, compound *mianjn* ‘monk’, nominative plural *mianjownk*, reflecting an Indo-European distinction between *e*- and *o*-grade, which was first identified by Meillet (1901). However, the only semi-productive use of vocalic alternations in the root is in some reduplicated compounds. In Classical Armenian total, or partial, reduplication of a lexical root or stem is frequently

employed in order to create words with an intensive or distributive sense: for example, *mecamec* ‘very big’ (*mec* ‘big’; note that this is one way to express a superlative adjective in Armenian, which has no comparative or superlative suffix); *dasadas* ‘in divisions’ (*das* ‘division’); *atxamatx* ‘diverse goods for sale’ (*atx* ‘box, baggage’). It is not unusual to find reduplication of this type with alternation of the root vowel, or even consonants: for example, *kerakowr* ‘food’ (suppletive aorist *ker-ay* ‘I ate’); *sarsowr* ‘cold’ (*sařn* ‘ice’); *xařamowř* ‘vulgar, of the rabble’ (*xař* ‘rude’; further examples and discussion given in Leroy 1986).

4.3. The nominal suffixes of Classical Armenian are discussed in extensive detail by Olsen (1999). Some examples of Armenian patterns of nominal suffixation can be shown from the following derivatives (excluding compounds) of *gorc* ‘work, action, manufacture’ found in 5th and 6th century Armenian: *gorcawor* ‘workman, anyone who works’ formed with a suffix *-awor* which derives historically from a verbal compounding element **bhoros*, and forms nouns denoting occupation or profession (see Olsen 1999: 358–368); *gorcaworowt’iwn* ‘work (in the abstract), labor’ shows the addition of the extremely common abstract noun suffix *-owt’iwn* (Olsen 1999: 546–584) to the previous word; *gorci* ‘tool’ (instrumental *gorceaw*) is formed with a suffix *-i* which is sometimes used, as here, to denote an instrument (Olsen 1999: 440); a derivative of *gorci* is formed with the adjectival suffix *-akan*, borrowed from Iranian, *gorciakan* ‘instrumental’; *gorcakic* ‘fellow-worker’ employs a suffix which is used to form words denoting companions or participants; *gorcac* and *gorcowac*, both ‘work’, are each derived from the verb *gorcem* using frequent suffixes for action nouns (see Olsen 1999: 231–239 and 543–545). Armenian derivative verbal suffixes are described in detail by Klingenschmitt (1982). The suffix used to form causative verbs has the form present *-owc’ane-*, aorist *-owc’-* (3rd singular *-oyc’*) and is widespread and productive in the Classical language. Causative verbs are formed through the addition of the suffix to the verbal aorist stem: for example, *darřnam* ‘I turn (intransitive)’, aorist *darřay*, causative *darřowc’anem* ‘I turn (transitive)’.

4.4. The examples of *angorc* ‘lazy’, *erknaberj* ‘sky-high’, and *mecamec* ‘very big’ given above show that compounding is a productive process of word formation in Classical Armenian; indeed several of the derivative suffixes of Armenian, such as *-awor* mentioned above, derive from generalized compound forms. For all compounds, the head occurs as the second member. The first member of a compound, if a noun or adjective, normally stands in the stem form which, for most items, is identical with the nominative singular. When the second element of a compound does not begin with a vowel, the productive pattern is to insert a liaison vowel *-a-* between the two members of the compound. However, a number of compounds are formed without the liaison vowel *-a-*, and in derivatives of compounds the liaison vowel is often dropped.

The principal productive types of compounding found in Armenian are as follows (see Olsen 1999: 657–759). First, exocentric compounds of the type modifier + head noun, for example *mecatown* ‘rich’ from *mec* ‘big’ and *town* ‘house’; *anmit* ‘mad, senseless’ from *an-* ‘without-’ and *mit(k’)* ‘mind’. Exocentric compounds frequently follow the same declension class as their head noun, but many are declined as *i*-stems: in the Bible translation, *anmit* is found declined both as an *i*-stem and as an *a*-stem, the declension class of the simplex *mit(k’)*. Second, endocentric compounds of the type modifier +

head noun: for example, *aysawr* ‘today’ from *ays* ‘this’ and *awr* ‘day’; *k’alak’orm* ‘city-wall’ from *k’alak’* ‘city’ and *orm* ‘wall’. Third, governing compounds, with a verbal element as the second member, are highly productive: for example, *jknors* ‘fisherman’ from *jowkn* ‘fish’ and the stem of the verb *orsam* ‘I hunt’ (note the reduction of the vowel *ow* when not in the final accented syllable); *andamaloyc* ‘paralytic, one who has had his limbs loosened’ from *andam* ‘limb’ and the stem of the verb *lowcanem* ‘I loose’ (3rd singular aorist *eloyc* ‘s/he loosed’). This second example shows a compound that appears to be exocentric with the first element as its head: ‘having loosened limbs’. There are also several copulative compounds in Armenian. Sometimes these show the conjunction *ew* between the two elements, as in *ert’ewek* ‘coming and going’ derived from the stems of *ert’am* ‘I go’ and *eki*, suppletive aorist of *gam* ‘I come’; but there are also examples without the conjunction, as *lrtes* ‘spy’ which combines the aorist imperatives *lowr* ‘listen’ from the verb *lsem* (aorist *loway*), and *tes* ‘see’ from the verb *tesanem*. Other compound types of Armenian, reduplicated compounds, and compounds which are calqued on Greek models are discussed above in 4.2 and 3.2.

5. References

- Ačařyan, Hraçya
1971–1979 *Hayeren armatakan bařaran* [Armenian etymological dictionary]. 2nd edition. Erevan: Erevan University Press.
- Arutjunjan, Covinar Rafikovna
1983 Armjano-grećeskie leksićeskie izoglossy [Armenian-Greek lexical isoglosses]. In: Gevorg Džaukjan (ed.), *Oćerki po sravnitel’noj leksikologii armjanskogo jazyka* [Essays on the comparative lexicology of the Armenian language]. Erevan: Armenian Academy of Sciences, 221–355.
- Awetik’ean, Gabriel, Xačatur Siwrmēlean, and Mkrtič Awgorean
1836–1837 *Nor bařgirk’ haykazean lezowi* [New dictionary of the Armenian language]. Venice: Tparani Srboyn Ghazaru.
- Bailey, Harold W.
1930 To the Žāmāsp Nāmā. *Bulletin of the School of Oriental Studies* 6: 55–85.
- Bailey, Harold W.
1986 Armenia and Iran iv. Iranian Influences in Armenian, 2. Iranian Loanwords in Armenian. In: Ehsan Yarshater (ed.), *Encyclopaedia Iranica* II 4–5, 459–465.
- Beekes, Robert S. P.
1972 H₂O. *Die Sprache* 18: 117–131.
- Bolognesi, Giancarlo
1960 *Le fonti dialettali degli imprestiti iranici in armeno*. Milan: Società editrice Vita e pensiero.
- Clackson, James
1994 *The Linguistic Relationship between Armenian and Greek*. (Publications of the Philological Society 30). Oxford: Blackwell.
- Clackson, James
2000 A Greek Papyrus in Armenian Script. *Zeitschrift für Papyrologie and Epigraphik* 129: 223–258.
- Clackson, James
2003 Vox Armeniaca: New Evidence for the pronunciation of Classical Armenian. In: Anaïd Donabédian and Agnes Ouzounian (eds.), *Actes du Sixième Colloque international de*

Linguistique arménienne. INALCO- Académie des Inscriptions et Belles-Lettres 5–9 juillet 1999. Paris: Langues Orientalistes, 43–52 (= *Slovo* 26–27, 2001–2002).

Clackson, James

- 2005 Review of Frederik Kortlandt. 2003. *Armeniaca: Comparative Notes with an Appendix on the historical phonology of Classical Armenian by Robert S. P. Beekes.* *Annual of Armenian Linguistics* 24–25: 153–158.

Considine, Patrick

- 1979 A semantic approach to the identification of Iranian loan-words in Armenian. In: Bela Brogyanyi (ed.), *Studies in diachronic, synchronic and typological linguistics: Festschrift for Oswald Szemerényi on the occasion of his 65th Birthday.* Amsterdam: Benjamins, 213–228.

Cowgill, Warren

- 1960 Greek *ὄν* and Armenian *oč*. *Language* 36: 347–350.

Diakonoff, Igor M.

- 1985 Hurro-Urartian Borrowings in Old Armenian. *Journal of the American Oriental Society* 105: 597–603.

Dyen, Isidore, Joseph B. Kruskal, and Paul Black

- 1992 An Indo-European Classification: A Lexicostatistical Experiment. *Transactions of the American Philosophical Society* 82: Part 5.

Džaukjan, Geworg. B. (ed.)

- 1983 *Očerki po sravnitel'noj leksikologii armjanskogo jazyka* [Essays on the comparative lexicology of the Armenian language]. Erevan: Armenian Academy of Sciences.

Džaukjan, Geworg. B. [Jahowkjan, Geworg B.]

- 1987 *Hayoc' lezvi patmowt'iown: naxagrayin žamanakašrjan* [History of the Armenian language: the pre-literary period]. Erevan: Haykakan SSH GA Hratarakč'ut'yun.

Eichner, Heiner

- 1978 Die urindogermanische Wurzel **H₂reu* 'hell machen'. *Die Sprache* 24: 144–162.

Gray, Russell D. and Quentin D. Atkinson

- 2003 Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 435–439.

Greppin, John A. C.

- 1998 Armenian *gini*, Georgian *γvino* 'wine'. *Annual of Armenian Linguistics* 19: 65–69.

Greppin, John A. C.

- 2000 Armenian and Kartvelian. In: Yoel Arbeitman (ed.), *The Asia Minor Connection: Studies on the Pre-Greek Languages in Memory of Charles Carter.* (Orbis Supplementa 13). Leuven: Peeters, 85–91.

Greppin, John A. C.

- 2010 Hurro-Urartian-North-East-Caucasian-Armenian Linguistic Connections. *Aramazd, Armenian Journal of Near Eastern Studies* V.1: 116–120.

Greppin, John A. C. and Igor M. Diakonoff

- 1991 Some effects of the Hurro-Urartian people and their languages upon the earliest Armenians. *Journal of the American Oriental Society* 111: 720–730.

Hamp, Eric

- 1983 Mi k'ani kisašp'akanneri masin hayereni patmowt'yan meġ [On some affricates in the history of Armenian]. *Istoriko-Filologičeskij Žurnal* 1: 38–42.

Henning, W. B.

- 1958 Mitteliranisch. In: Bertold Spuler (ed.), *Handbuch der Orientalistik*, 1. IV: *Iranistik*, 1. *Linguistik*. Leiden: Brill, 20–130.

Henning, W. B.

- 1963 Coriander. *Asia Minor* 11: 195–199.

- Hübschmann, Heinrich
1877 Ueber die stellung des armenischen im kreise der indogermanischen sprachen. *Zeitschrift für vergleichende Sprachforschung* 23: 5–49.
- Hübschmann, Heinrich
1992 [1895/1897] *Armenische Grammatik. I Theil: Armenische Etymologie*. Hildesheim: Olms. [Leipzig: Breitkopf & Härtel.]
- Hübschmann, Heinrich
1906 *Armeniaca. Indogermanische Forschungen* 19: 457–480.
- Klingenschmitt, Gert
1982 *Das Altarmenische Verbum*. Wiesbaden: Reichert.
- Klingenschmitt, Gert
1994 Die Verwandtschaftsverhältnisse der indogermanischen Sprachen. In: Jens Elmegård Rasmussen (ed.), *In honorem Holger Pedersen: Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Wiesbaden: Reichert, 235–251.
- Kortlandt, Frederik
1986 Armenian and Albanian. In: Leroy and Mawet (eds.), 38–47.
- Kortlandt, Frederik
1997 Arm. *nēr* ‘sister-in-law’. *Annual of Armenian Linguistics* 18: 7–9.
- de Lamberterie, Charles
1978 *Armeniaca I–VIII: études lexicales*. *Bulletin de la Société de Linguistique* 73: 243–285.
- de Lamberterie, Charles
1986 Arménien *barjrberj* ‘très haut’. In: Dickran Kouymjian (ed.), *Armenian Studies/Études arméniennes in memoriam Haïg Berbérian*. Lisbon: Calouste Gulbenkian Foundation, 469–476.
- de Lamberterie, Charles
1988 Meillet et l’arménien. *Histoire/Épistémologie/Langage* 10: 217–234.
- de Lamberterie, Charles
1990 *Les adjectifs grecs en -ος*, 2 vols. Leuven: Peeters.
- de Lamberterie, Charles
2006 Traces de la langue poétique indo-européenne dans la langue arménienne. In: Pinault and Petit (eds.), 213–234.
- Leroy, Maurice
1986 Le redoublement comme procédé de formation nominale en arménien classique. In: Leroy and Mawet (eds.), 62–75.
- Leroy, Maurice and Francine Mawet (eds.)
1986 *La place de l’arménien dans les langues indo-européennes*. Leuven: Peeters.
- Martirosyan, Hrach K.
2010 *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.
- Meillet, Antoine
1901 De quelques archaïsmes remarquables de la déclinaison arménienne. *Zeitschrift für armenische Philologie* 1: 139–148.
- Meillet, Antoine
1911 Sur les mots iraniens empruntés par l’arménien. *Mémoires de la Société de Linguistique* 17: 242–250.
- Meillet, Antoine
1926 Le mot *ekelec'i*. *Revue des Études Arméniennes* 9: 131–136.
- Meillet, Antoine
1937 Arménien *erkin* et *erkir*. *Annuaire de l’Institut de Philologie et d’Histoires Orientales et Slaves* 5: 1–2.
- Mercier, Charles
1978–1979 L’école hellénistique dans la littérature arménienne. *Revue des Études Arméniennes* n.s. 13: 59–75.

- Mowradyan, Arowsyak Nersesi
 1971 *Hownaban dproc'ə ev nra derə hayereni kerakanakan terminabanowt'iwn stelcman gorcwm* [The philhellenic school and its role in the establishment of Armenian grammatical terminology]. Erevan: Haykakan SSR GA Hratarakč'owt yown.
- Nyberg, Henrik Samuel
 1931 *Hilfsbuch des Pehlevi. 2. Glossar*. Uppsala: Almqvist & Wiksell.
- Olsen, Birgit Anette
 1999 *The Noun in Biblical Armenian: Origin and Word-formation: with Special Emphasis on the Indo-European Heritage*. Berlin: Mouton de Gruyter.
- Osthoff, Hermann
 1901 *Etymologische Parerga*. Leipzig: Hirzel.
- Pedersen, Holger
 1911 Arm. *Koriun*. In: *Huschardzan. Festschrift aus Anlass des 100 jährigen Bestandes der Mechitaristen-Kongregation in Wien (1811–1911) und des 25. Jahrganges der philologischen Monatsschrift "Handes amsorya" (1887–1911)*. Vienna: Mechitaristen, 287–288.
- Pedersen, Holger
 1924 Armenier. In: Max Ebert (ed.), *Reallexikon der Vorgeschichte*, Band I. Berlin: De Gruyter, 219–226.
- Pinault, Georges-Jean and Daniel Petit (eds.)
 2006 *La langue poétique indo-européenne: Actes du colloque de travail de la Société des Études Indo-Européennes (Indogermanische Gesellschaft/Society for Indo-European Studies), Paris, 22–24 octobre 2003*. Leuven: Peeters.
- Pokorny, Julius
 1959 *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.
- Porzig, Walter
 1954 *Die Gliederung des indogermanischen Sprachgebiets*. Heidelberg: Winter.
- Ringe, Don, Tandy Warnow, and Ann Taylor
 2002 Indo-European and Computational Cladistics. *Transactions of the Philological Society* 100: 59–129.
- Ritter, Ralf-Peter
 2006 Zur Frage der Reste indogermanischer Dichtersprache im Armenischen. In: Pinault and Petit (eds.), 213–234.
- Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
 2001 *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstambildungen*. 2nd, corrected edn. Wiesbaden: Reichert.
- Schmidt, Karl-Horst
 1985 Die indogermanischen Grundlagen des altarmenischen Verbuns. *Zeitschrift für vergleichende Sprachforschung* 98: 214–237.
- Schmitt, Rüdiger
 1975 Von Bopp bis Hübschmann: Das Armenische als indogermanische Sprache. *Zeitschrift für vergleichende Sprachforschung* 89: 3–30.
- Schmitt, Rüdiger
 1983 Iranisches Lehngut im Armenischen. *Revue des Études Arméniennes* n.s. 17: 73–112.
- Schmitt, Rüdiger
 1986 Armenia and Iran. In: Ehsan Yarshater (ed.), *Encyclopaedia Iranica* II 4–5: 417–459.
- Solta, Georg Renuat
 1960 *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*. Vienna: Mechitaristen.
- Solta, Georg Renuat
 1990 Die Stellung des Armenischen im Kreise der indogermanischen Sprachen. Ein Überblick. In: E. M. Ruprechtsberger (ed.), *Armenien: Frühzeit bis 1. Jahrtausend; Sprache, Kunst und Geschichte*. Linz: Stadtmuseum, 7–18.

- Szemerényi, Oswald
 1977 *Studies in the kinship terminology of the Indo-European languages*. Leiden: Brill.
- Taylor, Ann, Tandy Warnow, and Don Ringe
 2000 Character-Based Reconstruction of a Linguistic Cladogram. In: John Charles Smith and Delia Bentley (eds.), *Historical linguistics 1995: selected papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995. Volume 1: general Issues and Non-Germanic Languages*. Amsterdam: Benjamins, 393–408.
- Thumb, Albert
 1900 Die griechischen Lehnwörter im Armenischen. *Byzantinische Zeitschrift* 9: 388–452.
- Watkins, Calvert
 1995 *How to Kill a Dragon. Aspects of Indo-European Poetics*. New York: Oxford University Press.

James Clackson, Cambridge (United Kingdom)

65. The dialectology of Armenian

1. General description
2. Armenian dialects and PIE
3. The early isoglosses
4. References

1. General description

1.1. Armenian dialects: number, source, geography

With the term “Armenian dialects”, linguists usually designate the dialects of the Modern Armenian language in their geographical distribution from the 19th century onwards up to the beginning of the 20th century (for the extinct Armenian dialects of the Ottoman empire) or the present time. The number of Armenian dialects obviously depends on the classifications applied. In his comprehensive overview of Armenian dialectology, Ĵahukyan (1972: 132–136) distinguishes two main branches (East and West) in Modern Armenian and 11 dialect groups comprising 44 individual dialects. An enumeration in English is given by Ĵahukyan (1986; in this work on p. 22 insert between No. 23 and 24 the line “VI. The Mush-Tigranakert or South-Central intergroup”). Dialects identified after the appearance of Ĵahukyan (1986) or not accounted for in this work are Bolu, a Karabagh dialect in Western Turkey (Samuelian n. d.), Stanoz, Yozgat (Mkrtč‘yan 2006), and Jerusalem (Vaux 2002).

The sources of our knowledge of individual modern dialects are the currently spoken dialects (including the dialects of the Anatolian area that are still used by survivors of the 1915 massacres and their descendants), texts collected in ethnographic studies (for a

Meillet, Antoine

1897–1898 Recherches sur la syntaxe comparée de l'arménien, I. les démonstatifs. *Mémoires de la Société de Linguistique* 10: 241–271. [Reprinted in Meillet 1962, 5–35.]

Meillet, Antoine

1962 *Études de linguistique et de philologie arméniennes*, I. Lisbon: Imprensa Nacional.

Minassian, Martiros

1988 L'article du premier terme non substantivé d'une subordonnée en arménien classique. *Le Muséon* 101: 97–158.

Ouzounian, Agnès

1992 *Le discours rapporté en arménien classique*. (Bibliothèque des Cahiers de l'Institut de Linguistique de Louvain 63). Louvain-La-Neuve: Peeters.

Ouzounian, Agnès

1996 À propos des relatives en arménien classique: le fonctionnement de l'article et le cas des relatives à antécédent démonstratif d'après les évangiles et les actes des apôtres. In: Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan, 75–92.

Sakayan, Dora (ed.)

1996 *Proceedings of the Fifth International Conference on Armenian Linguistics, McGill University, Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan.

Thomson, Robert W.

1989 [1975] *An Introduction to Classical Armenian*. 2nd edn. Delmar, NY: Caravan.

Jared S. Klein, Athens, GA (USA)

64. The lexicon of Armenian

- | | |
|--|-------------------|
| 1. Inherited vocabulary | 4. Word-formation |
| 2. Loanwords | 5. References |
| 3. Technical and scientific vocabulary | |

1. Inherited vocabulary

1.1. In 1875, Heinrich Hübschmann conclusively showed that Armenian constituted a separate branch of Indo-European (Hübschmann 1875). In the decades before Hübschmann, Indo-Europeanists had judged the language to belong to the Iranian subgroup (see Schmitt 1975). The reason for this mistaken analysis of the origin of the Armenian language lies in the nature of the lexicon. Armenian has a smaller proportion of inherited vocabulary than other Indo-European languages of comparable age, and it contains a large proportion of Iranian loanwords, which are not limited to specific semantic fields but extend into the basic vocabulary, as will be shown in section 3. Moreover, since Armenian has undergone a complex series of sound changes in its development from Indo-European, many words which can be shown to be inherited from Indo-European

are not immediately recognizable as such, for example, *elbayr* ‘brother’, cognate with Sanskrit *bhrātar-*, Latin *frater*, and English *brother*; and *k’oyr* ‘sister’, cognate with Sanskrit *svasar-*, Latin *soror*, and English *sister*. Only after Hübschmann worked to show systematic correspondences between the inherited vocabulary and that of the other branches of Indo-European could the true nature of the language be revealed.

1.2. The comparatively small number of inherited items in Armenian can be shown by comparing the volume of Armenian entries in the standard etymological dictionaries of Indo-European. In the index to Pokorny’s *Indogermanisches Etymologisches Wörterbuch* (Pokorny 1959), Armenian entries occupy seven and a half pages. In comparison, the index for Old Irish is nineteen pages long, for Gothic ten and a third pages, for Lithuanian twenty-eight pages, and Old Church Slavonic sixteen pages. Of course the number of entries in an etymological dictionary is only a very rough guide to the proportion of inherited vocabulary in a language. It is possible that the proportion of entries in an etymological dictionary corresponds to the amount of attention scholars have paid to a language rather than the real rate of survival among the vocabulary. Untangling the sound changes which have operated in the pre-history of Armenian is not always easy and many of the essential research materials, such as the Venice dictionary (Awetik’ean, Siwrmēlean, and Awgorean 1836–1837), or Ačařyan’s excellent etymological dictionary (Ačařyan 1971–1979), are not easily accessible to western scholars since they are written in Armenian. It would therefore be no surprise if the comparative paucity of Armenian data in Pokorny’s dictionary reflected its comparative neglect by Indo-Europeanists. However, if we consider the position of Armenian in the more recent *Lexikon der indogermanischen Verben* (Rix et al. 2001), we see that the situation is still largely similar to what was found in Pokorny: there are 159 entries for Armenian, compared with 253 for Old Irish, 277 for Gothic, 606 for Lithuanian, and 388 for Old Church Slavonic. Even Tocharian B, which had only had two and a half pages in Pokorny’s index now surpasses Armenian in Rix et al. (2001) with 179 entries. It is no longer possible to argue that the Armenian vocabulary has been neglected, since a sizeable monograph on the Armenian verb (Klingenschmitt 1982) included a thorough account of the attested Armenian verbal roots and their possible etymologies and unearthed several connections that had previously been overlooked. It now seems unlikely that there are many other verbal stems in the Classical language with Indo-European etymologies which have been escaped notice. The reduced number of inherited words in Armenian can also be demonstrated by considering Indo-European roots which are widely attested in other branches, but which are not found in Armenian. These include basic vocabulary items such as **h₁ei-* ‘go’ (retained in Anatolian, Indo-Iranian, Greek, Italic, Germanic, Baltic, Slavic, and possibly Albanian), **sek^w-* ‘follow’ (retained in Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, Slavic, and possibly Albanian), **pek^w-* ‘cook’ (retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Baltic, Slavic, Albanian, and possibly Anatolian), **h₁reud^h-* ‘red’ (derivatives retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, and Slavic), **k^hmtóm* ‘hundred’ (retained in Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, and Slavic), and **ǵ^hd^hom-* ‘earth’ and derivatives (retained in Anatolian, Tocharian, Indo-Iranian, Greek, Italic, Celtic, Germanic, Baltic, Slavic, and Albanian). Some of this material has been replaced by other items inherited from Indo-European: the standard word meaning ‘go’ *ertam* (aorist *čogay*) is derived by Klingenschmitt (1982: 103 f.) as a compound of the original IE word for ‘stand’; the

word for ‘follow’ *hetewem* is formed on the inherited noun *het* ‘footstep’. Others are replaced by words only found in a minority of the IE languages, such as the word for ‘cook’ *ep'em*, which has a cognate only in Greek $\epsilon\psi\omega$ (see Clackson 1994: 172). Iranian loanwords have supplanted others; thus, the word for ‘red’ in Armenian is a loanword, *karmir*, as are most of the Armenian color terms. However, much of the Armenian vocabulary is of unknown or doubtful origin, as the words for ‘hundred’ *hariwr* and ‘earth’ *erkir* (see further below). Indeed, around 50% of the Armenian vocabulary is of unknown origin (following the count of the material in Ačařyan 1971–1979 given by Solta 1990: 13), neither clearly Indo-European nor borrowed from any other language of which we have records. This “unknown” element of the vocabulary includes many other basic vocabulary items, including, for example, *mort* ‘skin’, *masn* ‘part’, *xawsem* ‘speak’, *c'owrt* ‘cold’.

Absolute numbers of Armenian roots inherited from Indo-European are difficult to know precisely, since many Armenian etymologies remain contested. Hübschmann’s *Armenische Grammatik* (Hübschmann [1895/1897] 1992), considers 438 words which previous to him had been proposed as inherited, but many of these are uncertain and several are dismissed out of hand by Hübschmann. In the time since Hübschmann, many new etymologies have been proposed, and there have been attempts to resurrect etymologies which were explicitly rejected by him, so that the number of words now thought to be inherited has increased. Solta (1960) lists 453 inherited items, Ačařyan (1971–1979) gives 713 (figure from Solta 1990: 13), Džaukjan (1987) has around 1400, and Martirosyan (2010) has 961. The figure given in Solta (1960) is likely to be too low: we have seen that Rix et al. (2001) include 159 entries for Armenian verbs, but Solta includes only 114 verbs. Conversely, Džaukjan’s figure is probably too high, since a number of the forms he includes are doubtful and not accepted by all scholars. For example, an earlier connection of the word for ‘cub, young animal’ *korıwn*, with Greek $\beta\rho\epsilon\phi\omicron\varsigma$ and Slavic *žrěbę* was already rejected by Hübschmann (1992: 461) since the expected Armenian cognate would be *korb-*. Pedersen (1911) argued for the retention of the etymology, however, and he is followed tentatively by Džaukjan (1987) and with more certainty by Olsen (1999: 492). The etymology is not included in Solta (1960) nor endorsed by Ačařyan (1971–1979). With such a small proportion of inherited words in Armenian, the weight accorded to these controversial etymologies becomes of greater importance when assessing questions such as the relationship of the Armenian vocabulary to that of other languages.

1.3. Despite the general loss of inherited vocabulary in Armenian, there are several examples of words that can be assumed to derive from archaic Indo-European lexical items. The Armenian word for ‘hand’ *jeřn*, has cognates in Anatolian, Tocharian, Greek and (with a different suffix) Indo-Iranian, and appears to have been replaced in the other branches of Indo-European by neologisms. The word for ‘blood’ *ariwn*, has cognates in Anatolian, Sanskrit, Greek, archaic Latin, and Baltic, but the root is elsewhere lost (as it is in later Greek and Classical Latin). The adjective *barjr* ‘high’ continues a primary adjective elsewhere limited to Hittite, Tocharian and Indo-Iranian, with the *u*-stem form only found in Armenian and Hittite. The pair *han* ‘grandmother’ and *haw* ‘grandfather’ have their closest cognates in Hittite *hanna-* ‘grandmother’ and *huhha-* ‘grandfather’, and no other single language outside the Anatolian branch preserves both members of the pair with the same meanings. Indeed, Armenian has retained most of the reconstructi-

ble Indo-European kinship terms, including the terms for the wife's relations by marriage: *skesowr* 'husband's mother', *skesayr* 'husband's father', *taygr* 'husband's brother', *tal* 'husband's sister', and *nēr* 'husband's brother's wife' (see Szemerényi 1977 on these terms and, for *nēr*, most recently Olsen 1999: 190 f. and Kortlandt 1997). Some other Armenian lexical items have been seen to continue words of high or "poetic" registers of the parent language (Porzig 1954: 482; Schmidt 1985: 215; Ritter 2006): *arcowi* or *arciw* 'eagle', plausibly connected to Sanskrit *ṛjipyá-*, an epithet of eagles in the Ṛg-Veda (despite the doubts of de Lamberterie 1978: 251–262); *ji* 'horse', cognate with the poetic Sanskrit word *haya-* 'steed' (note that the Armenian word for 'donkey', *žš*, corresponds to the normal word for horse elsewhere, Latin *equus* etc.; see de Lamberterie 2006); *erg* 'song' cognate with Sanskrit *arc-* 'sing, praise'; *arew* and *aregagn* 'sun' cognate with Sanskrit *ravi-* (see Eichner 1978 for the history of this pair and Hittite cognates). It is also possible that the Armenian words for 'heaven' and 'earth' *erkin* and *erkir*, which share the same initial sequence as Armenian *erkow* 'two', continue an inherited conception of the heaven and earth as the two halves of the universe, found also in Vedic hymns and early Greek cosmologies (Meillet 1937; subsequent scholars have admired Meillet's suggestion without wishing to endorse it fully; see de Lamberterie 1988: 225 f.).

Some words appear to be related to forms in other IE languages but show peculiar deviations in their phonological development or morphological behavior. A number of these have been explained through blending or "contamination" of two different inherited terms. The idea that a particular lexical item can arise from the merger of two separate terms is of course a useful tool for the desperate etymologist, but Armenian does provide a larger number of plausible candidates for this type of development than many other languages. A few examples will suffice to illustrate. The word for 'honey', *melr*, appears to be cognate with the family of Hittite *melit*, Greek μέλι, Latin *mel*, etc. However, the Armenian word is a *u*-stem and this has led scholars to think that there has been some influence from a different etymon, **med^hu* 'mead' (Sanskrit *mádhu*, Greek μέθυ, Lithuanian *medūs* etc.; see further Martirosyan 2010: 462). The word for 'son' in opposition to 'daughter', *owstr*, appears to continue the same root as found elsewhere to mean 'son', Greek υἱός, Sanskrit *sūnú-* etc., but with the sequence *-str* taken from *dowstr* 'daughter' (Szemerényi 1977: 19; Olsen 1999: 149). The word for 'nail' or 'claw' *elowngn* is generally taken to be cognate with the set Greek ὄνυξ, Latin *unguis*, Old Irish *ingen*, but the explanation for the initial *el-* is not known; it was seen by Osthoff (1901: 248) as contamination from the word for 'horn' *eljewr*. The adjective *k'alcr* 'sweet' has been explained by Ačařyan (1971–1979 IV: 546), Hamp (1983: 38–42) and de Lamberterie (1990: 502) as the result of some sort of combination of the stem **sweh₂du-* 'sweet' (Greek ἡδύς, Latin *suāuis* etc.), **dluku-* 'sweet' (Greek γλυκός, Latin *dulcis* etc.) and perhaps **saldu-* (Lithuanian *saldūs*, Old Church Slavonic *sladükū* 'sweet'). Klingenschmitt (1982: 227) proposed that the verb meaning 'kill', *spananem*, was a blending of two different roots, one of which survives in Greek σφάζω, the other in Greek κτείνω, and this suggestion is endorsed by Watkins (1995: 521), who however suggests that the second root involved is that of Greek θείνω, Sanskrit *han-*.

1.4. The inherited vocabulary of Armenian has been examined several times for the light it can bring to bear on the question of the position of Armenian within the Indo-European family. The first systematic attempt to use the vocabulary to assess the proxim-

ity of Armenian to other branches of Indo-European was made by Pedersen, who gave figures for lexical agreements between Armenian and other IE branches in an encyclopedia article (Pedersen 1924: 224). By Pedersen's count, Armenian had 95 special agreements with Greek, 53 with Indo-Iranian, 34 with Balto-Slavic, 24 with Germanic, 15 with Italic, and only 3 with Celtic. (Pedersen also gave figures for lexical agreements working only from the material in Hübschmann 1992: Greek 59, Indo-Iranian 37, Balto-Slavic 14, Germanic 10, Italic 10, and none for Celtic). Unfortunately, Pedersen never gave the full list of lexical material he had worked on, and it is not quite clear what he means by his "special agreements". In 1960, Solta undertook a much more thorough survey of the Armenian vocabulary and again found the lexical agreements between Armenian and Greek to be the most numerous and important. In his summary (1960: 459–485), he lists 53 noteworthy correspondences between Armenian and Greek (that is, both cognate pairs which are exclusive to the two languages and ones where there is a unique morphological or semantic development between the languages), but only 19 between Armenian and Indo-Iranian, 19 with Germanic, 17 with Latin, 13 with Baltic, 10 with Slavic, 4 with Celtic, 4 with Tocharian, and 3 with Albanian. The Greek-Armenian material was re-examined by Arutjunjan (1983) and Clackson (1994). In Clackson (1994), I argued that the number of innovative lexical agreements was much lower than had previously been thought, although there were undoubted cases where the two languages had made shared innovation. For example, I dismissed (1994: 158) the etymological connection of the initial *o-* of Armenian (*o*)č' 'not' with Greek οὐ from PIE **h₂oyu* 'life' proposed by Cowgill (1960), given the lack of parallels for the phonological development in Armenian (in Clackson 2005, I gave an alternative explanation of the origin of the *o-* of Armenian *o*č', through consideration of the frequent collocation *o*č' *ok*' 'no one'). However, it is likely that the short form, č', of the Armenian negative particle can be connected to the final element of the epic form of the Greek negative οὐ-κι, through grammaticalization of a particle **k^wid* 'something'. Note that Kortlandt (1986) and Klingenschmitt (1994: 245) have drawn attention to the lexical agreements between Armenian and Albanian. In more recent years, several scholarly teams have undertaken mass comparison of basic vocabulary lists to unearth the relationships between the lexicons of different Indo-European branches. Gray and Atkinson, who produced a tree for the Indo-European language family (2003) using basic vocabulary lists compiled by Isidore Dyen and described in Dyen, Kruskal, and Black (1992), note that the vocabulary list for Armenian was based on a modern Western Armenian variety, rather than the Classical language. In Gray and Atkinson's tree, the Armenian branch forms a node with the Greek branch. Armenian is also grouped with Greek on the phylogenetic tree generated by a team of mathematicians and linguists including Don Ringe (see for example Taylor, Warnow, and Ringe 2000 and Ringe, Warnow, and Taylor 2002). Ringe, Warnow, and Taylor (2002: 102) list six Armenian-Greek lexical isoglosses that support a subgroup, three of which (the words for 'day', 'not', and 'wind') are seen as likely candidates for shared innovations. It is worth repeating the data here:

Armenian *awr* 'day', Greek ἡμᾶρ, derived respectively from **āmōr* and **ām̄r*; Latin and Vedic words for 'day' derive from **dyew-*.

Armenian *ayr* 'man, husband', Greek ἀνὴρ 'man, husband'; both share the meaning 'husband' for which Tocharian and Indo-Iranian use a derivative of **poti-*.

Armenian (*o*)č' 'not', Greek οὐ(κ) 'not' (see above); most other IE languages show a reflex of **ne*.

Armenian *holm* ‘wind’, Greek ἄνεμος ‘wind’ derived, respectively, from **h₂onh₁mo-* and **h₂enh₁mo-* (see the doubts of Beekes 1972: 129 on this equation); other IE languages form the word for ‘wind’ as a derivative of the root **h₂weh₁-*.

Armenian *alam* ‘I grind’, Greek ἀλέω ‘I grind’; other language branches (Hittite, Italo-Celtic, Germanic, and Balto-Slavic) form words for ‘to grind’ from the root **melh₂-*.

Armenian *nor* ‘new, fresh, young’, Greek νέος ‘new, young’; Greek and Armenian extend the inherited term for ‘new’ to mean ‘young’, whereas the stem **h₂yu-Hen-* is employed elsewhere.

2. Loanwords

2.1. The Iranian languages have contributed a large number of loanwords to the Armenian lexicon. Hübschmann listed 686 Iranian words in Armenian (Hübschmann 1992), and since then the number of identified loans has increased substantially with the improved knowledge of Middle Iranian languages. There is no comprehensive modern collection of the Iranian loanwords (although a project to produce one was announced in Considine 1979), but much of the material is gathered in Bailey (1986), Schmitt (1983), Schmitt (1986), and Olsen (1999: 857–920). Much work has been done on locating the source of Armenian loanwords within the Iranian language group. Armenia and Armenians are first mentioned by name in the Old Persian Bisutun inscription dated to the end of the sixth century BCE, and thereafter in the work of the Greek historian Xenophon. Since most of those named in these sources have Iranian names (for example Dādarši- in the Bisutun inscription [DB II 29], Tigranes in Xenophon), it seems plausible that some of the Iranian loanwords stretch back to Achaemenid times; and this is very likely for the name of the Aryans, *Arik’* (genitive plural *Areac’*), which exactly reflects the stem *ariya-* seen in Old Persian *ariya-*, Avestan *airiia-* (Schmitt 1983: 77). Indeed, it has been argued that some Iranian loans took place so early that they participated in the Armenian consonant shift, and show voiceless stops for Iranian voiced stops, and Armenian aspirates for Iranian voiceless stops. The most likely example for a pre-sound shift loan is the word for ‘garden’ *partēz*, beside Avestan *pairidaēza-*, New Persian *palēz*, which was explained in this way already by Meillet (1911: 250). But other scholars have added further possible examples: de Lamberterie (1988: 245–262) proposed *arcāt* ‘silver’ (cf. Old Persian *ardata-*, Avestan *arəzata*) and *arcowī* ‘eagle’ (but this word may also be inherited, see 1.3) and Olsen (1999: 857) added *ciran* ‘apricot’, *cirani* ‘purple’ (both from an Iranian cognate of Avestan *zaraniia-* ‘golden’), and *parc* ‘proud’ from an Iranian word cognate with Avestan *bərəz-* ‘high’.

2.2. However, the vast majority of Iranian loans in Armenian appear to have entered the language at a later date, during the period of the Arsacid dynasty’s rule over Armenia from the 1st to the 3rd century CE. The language of the Arsacids was the north-west Iranian dialect Parthian, and many of the loanwords in Armenian can now be directly compared to forms found in surviving Parthian texts, and many more with putative Parthian forms reconstructed through attested words in other Iranian languages. The Arsacid loans permeate the Armenian lexicon, referring not just to aspects of administra-

tion, religion, and military service (such as *azat* ‘free’, Parthian *āzāt*; *hraman* ‘command’, Parthian *framān*; *marz* ‘province’, Parthian *marz*; *den* ‘religion’, Parthian *dēn*; *dēw* ‘devil’, Parthian *dēw*; *zoh* ‘offering’, Middle Persian *zōhr*; *nizak* ‘spear’, Parthian *nēzag*; *zēn* ‘weapon’, Parthian *zēn*); but also numerous aspects of everyday life including names of colors (*seaw* ‘black’, Parthian *syāw*; *karmir* ‘red’, Middle Persian *karmir*), domestic objects (*črag* ‘lamp’, Modern Persian *čarāy*; *šapik* ‘shirt’, Middle Persian *ša-bīg*; *xan* ‘basket’, Parthian *xawān*), people and relationships (*harazat* ‘blood(-brother), (-son)’, Avestan *hadō-zāta-*; *tohm* ‘family’, Parthian *tōxm*; *pařaw* ‘old woman’, Modern Persian *pārāw*), abstracts and emotions (*oyž* ‘strength’, Middle Persian *ōz*; *ah* ‘fear’, Middle Persian *āhal*), adverbs (*yavēt* ‘always’, Middle Persian *yāwēt*; *-pēs* ‘-how’, Avestan *-paēsa-*) and numerals (*hazar* ‘1,000’, Parthian *hazār*; *biwr* ‘10,000’, Parthian *bē-war*). The large influx of loans during the Arsacid period was most probably assimilated over a long time period, and there are some linguistic grounds for dividing the loanwords up into different chronological strata. The more recent Iranian loans in Armenian show the equivalence of Armenian *e* for Iranian *ē*, and Armenian *o* for Iranian *ō*, for example *den* ‘religion’ beside *dēn* and *zoh* ‘offering’ beside Persian *zōhr*. But other words, which may reflect earlier borrowings, show *ē* where the Iranian source had *ē* and *oy* where the source had *ō*, as *zēn* ‘weapon’ beside Parthian *zēn*, and *oyž* ‘strength’ beside Middle Persian *ōz*. Furthermore, some loans appear to have undergone an inner Armenian development of a prothetic vowel before initial *r-*, whereas in later loans initial *r-* is preserved: thus *erasan* ‘rein’ derives from an Iranian stem attested in Middle Persian as *rasan*, and *řočik* ‘provisions’ derives from the same stem as Middle Persian *rōzīg* (note that this word also shows the “late” development of the Iranian vowel *ō*).

It is also possible to isolate a yet later stratum of Iranian loanwords in Armenian, those that stem from the Sasanian period, when Armenia had moved away from the Persian cultural and religious sphere. Loans from this period are generally restricted to technical military or religious terms, they are restricted in their textual occurrences, and they tend not to show the same range of derivative formations as the loans of the Arsacid period. These loans also reflect the phonology of the south-west Iranian language of the Sasanian Persians. Thus *řatagov* ‘proponent, advocate’ (Middle Persian *řatagōv*), *řahan-řah* ‘king of kings’ (Middle Persian *řāhān řāh*) and *payik* ‘guard’ (Middle Persian *paik*) show characteristic Middle Persian developments of initial *y-* to *ř-* (*řatagov*), *xš-* to *ř-* and *-θr-* to *-h-* (*řahanřah*), and medial *-đ-* to *-y-* (*payik*, see Bolognesi 1960: 56, 21, and 43). None of these words appears in the Armenian Bible translation, and from *payik* and *řahanřah* there are no further derivatives formed.

While the majority of Iranian loans can be sorted into their appropriate chronological and dialectal layer, whether early or late, and from a north-west or south-west variety, there remain a number of words for which the origin is identifiably Iranian, but the details are unclear or unparalleled. Thus the word for ‘road’, *čanaparh*, was taken by Bailey (1930: 61) as an Iranian compound, the second half of which represented *-přθ-* or *-parθ-* ‘way’, but the origin of the initial element *čana-* remains disputed; the best explanation is probably still that of Nyberg (1931 s.v. *puhl*) that it is a dissimilation from *čarana-parθ-*, with the first element from the Iranian root *čar-* ‘go’. A particularly noteworthy pair of words are those which have their closest cognates with forms in eastern Iranian dialects: *kari* ‘very’ and *margarē* ‘prophet’; compare Sogdian *k’đy* and *m’rkr’y*. Henning (1958: 93) provided an explanation for these words that is now widely accepted: they are elements of the speech of the Parnians, the eastern Iranian tribe which

migrated westward as overlords of the Parthians. For many other words the exact Iranian source remains disputed: thus the word for ‘treasure’, *ganj*, has cognates in many Iranian languages, including Parthian *gazn*, Middle Persian *ganz* and Modern Persian *ganj*; Henning took Armenian *ganj* to have originated in a Median word (1963: 197), but Olsen (1999: 861) argues for it as another Parnian loan.

2.3. Classical Armenian has also borrowed a substantial number of words from Greek and Syriac. The majority of Syriac loans entered the language when the Armenians adopted Christianity as the state religion (the traditional date for this is 306 CE), and the most common words are in some way connected with the Christian religion, for example, *k'ahanay* ‘priest’ (Syriac *kāhanā*), *šabat* ‘week, sabbath’ (Syriac *šabbāthā*), and *tlay* ‘infant, child’ (Syriac *ṭalyā*), used in the Armenian Bible translation particularly to denote the innocent and receptive child as opposed to the wise and knowing sceptic. Some Greek loans are also connected with the Christian religion, for example, *salmos* ‘psalm’ (ψαλμός), *episkopos* (ἐπίσκοπος), *pentakostē* ‘pentecost’ (πεντηκοστή), and *martivros* (μάρτυρος). However, these reveal themselves as recent specialized borrowings, and Meillet (1926) showed that the majority of the Armenian words used for Christian concepts are either borrowed from Syriac or Iranian (Iranian loans include *ašakert* ‘disciple’, *hrestak* ‘angel’, *margarē* ‘prophet’, *awetaran* ‘gospel’) or are native Armenian terms, some of which lack any straightforward or known etymology (for example *erēc* ‘priest’ which can be connected to the family of Greek *πρεσβύς*, Latin *priscus*, and, with unknown etymologies *Astowac* ‘God’ [see Olsen 1999: 545 f. for a survey of some proposed etymologies], *ařak'em* ‘I send’ [from which *ařak'ič* ‘apostle’ is derived], and *mkrt'em* ‘I baptize’). Accordingly, Meillet proposed that the Armenian word for ‘church’ *ekelec'i*, was not a direct loan from Greek ἐκκλησία but entered the language via Syriac or possibly an Iranian source, just as was the case for other Greek loans such as *p'ilisop'ay* ‘philosopher’, *ak'sorank* ‘exile’, *katapar* ‘model’ and *lambar* ‘lamp’ (respectively Greek φιλόσοφος [via Syriac], ἔξορία [via Syriac], καλοπόδιον [via Parthian], and λαμπάδα, accusative of λαμπάς [via Parthian]). The possibility of a Syriac or Iranian intermediary unfortunately complicates Hübschmann’s theory (1906: 472–477) that the Armenian word for the antichrist, *neřn*, was a very early loan of the Greek name for the Emperor Nero, Νέρων, which had undergone the Armenian loss of final syllables. Armenian did however borrow a number of words from Greek to denote items that reflected Greek technology, thought or cultural life: for example *pnak* ‘dish’ (πίναξ), *balistr* ‘catalpult’ (βαλίστρα), *hiwl* ‘matter’ (ῥήλη), *hřetor* ‘orator’ (ῥήτωρ), and *t'atr* ‘theatre’ (θέατρον). Thumb (1900) argued that the majority of these words were “eye-borrowings” rather than “ear-borrowings” and represented the learned pronunciation of the Greek words, since most of the Armenian representations of the Greek sounds were in accord with the learned pronunciation of Greek, rather than the spoken Greek of the middle of the first millennium CE. In support of this theory, he noted that Armenian represented, for example, the Greek consonants written φ θ χ and β as aspirates rather than fricatives. However, more recent evidence has suggested that these loanwords need not be of a particularly elevated register; note in particular the discovery of a papyrus fragment, dated to around the 6th century, with Greek words written in Armenian script. The text appears to have been taken down by dictation, and it shows much the same equivalences for the Greek vowels and consonants as are found in the Greek loanwords (see Clackson 2000 and 2003). Many Armenian words of scientific and technological vocabulary, from

the sixth century to the present day are also formed through a process of calquing of Greek compound and suffixed forms, and these will be discussed in more detail in the next section.

2.4. Several Classical Armenian words for flora and fauna appear to be borrowed from Mediterranean or Near Eastern languages which are now irrecoverable, since the same words appear as loanwords in other ancient languages: for example, *ewl* (post-Classical *iwl*) ‘oil, unguent’ appears to be related to the word borrowed into Greek as ἔλαιον ‘olive oil’; *tōwz* ‘fig’ appears to be in origin the same word as Greek σῦκον and Latin *figus*; *xstor* ‘garlic’ links with Greek σκόροδος, Albanian *hurdhë*; *inc* ‘leopard’ with Sanskrit *simha-* ‘lion’, Tibetan *señ-ge*; and *owlt* ‘camel’ with Assyrian *utru*, Avestan *uštra-*, and Urartian *ultu*. Urartian is the most likely source of the Armenian word for ‘camel’, and further plausible loans from Urartian include Armenian *sowr* ‘sword’ (Urartian *šuri* ‘weapon’), *san* ‘cauldron’ (Urartian *šani* ‘vessel’). Since the Urartian corpus is very small, the related language Hurrian sometimes provides supporting evidence for early Urartian loans into Armenian. One possible case is the Armenian word for ‘apple’ *xnjor*, which is an exact match for Hurrian *hinzuri* ‘apple’, although the word is also found in Aramaic *ħazzūrā*. Diakonoff (1985), Greppin and Diakonoff (1991), and Greppin (2010) list other possible Hurro-Urartian borrowings into Armenian, but many are either unattested in our surviving Hurrian and Urartian texts or are semantically or phonologically problematic. The Armenian word for ‘wine’, *gini*, is most likely another example of a Mediterranean *Wanderwort*; compare Mycenaean Greek *wo-no*, Latin *uīnum*, Arabic *wain*, Hebrew *yayin*, Hittite *wiyana-*, Hieroglyphic Luwian *wayana-*. The Georgian word for wine, *γvino*, is of particular interest, as it shares the presence of a velar consonant at the beginning of the word with the Armenian word. Hübschmann (1992: 397) noted the correspondence alongside other words which appear to be shared by Armenian and South Caucasian languages, but left open the question of whether the loan was from, or to Armenian. Many scholars have subsequently taken the Armenian as the original form in the two languages, and the Georgian initial sequence *γv-* to reflect an intermediary stage in the development of Armenian *g-* from original **w-* (see Greppin 1998). Indeed, there are almost no loanwords from South Caucasian languages which are widespread and long established in the Armenian lexicon; most of the loans are restricted to Armenian dialects which have long been in contact with neighboring South Caucasian varieties (see the survey in Greppin 2000).

2.5. In the time since the earliest Armenian texts, speakers have continued to adopt loanwords from the languages with which they came in contact: Persian and Iranian dialects, Arabic, South Caucasian, Greek, Turkish, French, Russian, and English. The earliest French loans date from the 12th century, when the Cilician Armenians came into contact with crusaders, but have since spread throughout the language: for example *paron* (in Modern Western Armenian *baron*) is borrowed from French *baron* and has become the standard equivalent to the title ‘Mr.’, or the address ‘sir’. However, since the formation of the modern standard literary languages in the nineteenth century, there has been a reluctance to incorporate foreign loans into the standard. Thus Standard Modern Eastern Armenian avoids the virtually pan-European terms *music*, *problem*, and *coffee* using instead the “native” terms *eražštowt’iwn* (in fact an early Iranian borrowing), *xndir* (of unknown origin), and *sowrč* (plausibly explained as a metathesis of *sew jowr*

‘black water’, the first element of which is an Iranian loan, the second an inherited word).

3. Technical and scientific vocabulary

3.1. The Armenian technical and scientific vocabulary is formed to a large extent through an extensive system of calquing of compound or suffixed words in Greek, and to a lesser extent Latin. Many of the words used to denote items of modern technology are formed from calques in this way. For example, the word for ‘photograph’, *lowsankar*, is a compound incorporating the lexical stems of *loys* ‘light’ and *nkar* ‘picture’; the word for ‘television’, *heřowstatesayowt’iwn*, is built from *heřowst* ‘from a distance’ and the root of the verb meaning ‘see’, *tesanem*, with the abstract noun suffix *-owt’iwn*. This calquing technique was first inaugurated in the translations of Greek grammatical and philosophical works during the late 5th and 6th centuries of the Christian era, a period of Armenian literature known as “the Hellenizing School”. The works of the Hellenizing school marked a departure from the earlier translation practice of the Armenian Bible translators and writers of the “golden age”. For example, the fifth century Armenian author Eznik used the noun *dprowt’iwn* ‘scribery’ (formed from *dpir* ‘scribe’) to translate Greek γραμματική, whereas the translators of the Hellenizing School used the new artificial term *k’erakanowt’iwn*, formed from the verb *k’erem* ‘I scrape’, which was taken as the equivalent to the base meaning of Greek γράφω, presumably following the discussion of the original meaning of γραμματα given in the grammatical work attributed to Dionysius Thrax. The term *dprowt’iwn* seems to have been rejected as a satisfactory translation of γραμματική because its range of meanings was too wide to allow it to be used in the specific sense of ‘grammar’. In the Armenian Bible translation, *dprowt’iwn* can mean ‘learning’, ‘letters’, and even ‘book’. The new term *k’erakanowt’iwn* could stand in technical works without leading to any ambiguity.

3.2. A distinctive feature of the calques of the Hellenizing School is their reliance on a completely artificial system of composition. In Classical Armenian (that is, the Armenian of the Bible translation), nominal compounds are not uncommon, but composite verbs are only regularly formed from three prefixes, *ař-*, *y-*, and *and-*, none of which has an exact correspondence with any of the Greek prepositions used in composition. The Hellenizing School vastly increased the number of Armenian prefixes by creating a range of “artificial prefixes”, which they used to translate Greek prepositions in composition. The following gives some of the Armenian equivalents to Greek prefixes that are used (see further Mowradyan 1971: 136–152 and Mercier 1978–1979: 64–67).

<i>ver-</i>	անά
<i>der-</i>	άντί
<i>bac’-</i>	άπό
<i>tram-, hastat-</i>	διά
<i>art-</i>	έκ

<i>ner-</i>	ἐν
<i>mak-, ver-</i>	ἐπί
<i>ənd-</i>	μετά
<i>yar-, mat-, tar-</i>	παρά
<i>par-, bak-</i>	περί
<i>nax-, (y)araj̄-</i>	πρό
<i>aṙ-</i>	πρός
<i>bał-, šar-, šal-, jok-, p'al-, par-</i>	σύν
<i>ger-</i>	ὑπερ
<i>stor-, ent-, p'al-</i>	ὑπό

Some of these prepositions originate from earlier Armenian prepositions, such as *aṙ-*, used to translate Greek πρός, but most of them are newly formed from nouns or verbs in the language. Thus, for example, the prefixes *šal-* and *p'al-*, which are both used to translate Greek σύν, are derived, respectively, from the verbs *šalem* 'I encircle' and *p'alem* 'I join'. It is immediately noticeable from this table that the correspondence between the Armenian prefixes and the Greek is not one-to-one, since some Greek prefixes are rendered by more than one Armenian prefix. However, the same Armenian prefix (with the exceptions of *ver-* and *p'al-*) does not translate different Greek prefixes and nearly every Greek prefix has an Armenian equivalent. The equivalence of several different Armenian prefixes to a single Greek prefix does not seem to have arisen from a desire to translate different nuances of the Greek prefixes, but rather it seems to reflect the difficulties the translators had in finding native equivalents to the Greek. This rich array of newly-formed prefixes also helped compensate for the relative poverty of the Armenian suffixal system in comparison with Greek. The Armenian translators were not so bold or so rigorous in their treatment of the different suffixes used to form technical terms. Rather than attempt to create new artificial suffixes, they relied upon those already existing in the Classical language, and the replacement of a Greek suffix with its Armenian "equivalent" is not entirely consistent. However, the following correspondences between Greek and Armenian suffixes are used with some degree of regularity:

<i>-akan</i>	-ικός
<i>-eli</i>	-τέον
<i>-ac</i>	-μα
<i>-ot̄</i>	-τηρ, -τωρ
<i>-ayin</i>	-ιος
<i>-eal</i>	-μενος
<i>-pēs-, -bar</i>	-ως

Note also that the Armenian abstract noun suffix *-owt'iwñ* is frequently used to translate a variety of different Greek abstract noun suffixes, such as -ία, -σις, -ισμός etc. Armenian

also uses a far greater number of unsuffixed nouns than Greek, and these are widely used to render both Greek agent and action nouns.

4. Word-formation

4.1. Word-formation processes in Armenian largely correspond closely to those found in other branches of Indo-European. There is widespread use of a large number of nominal suffixes, a more restricted set of verbal suffixes, and compounding. As in other IE languages, the basic unit of word formation in Armenian is the lexical root. In some verbs, the aorist-stem is identical with the lexical root: for example, root *tes-* ‘see’, 3rd singular aorist *e-tes* ‘s/he saw’, but this pattern is of limited productivity, and most present and aorist verbal stems are formed through suffixation, as is the case for the present stem *tesanem* ‘I see’. Most noun and adjective stems are also suffixed forms of the root, although in many cases the original root shape may be obscured by phonological developments. Often the nominative singular is coincidental in shape with the root, and the suffix is only apparent in oblique cases. For example, the root *gorc-* means ‘work’ and among its synchronic derivatives are the noun *gorc* ‘work’, the verb *gorcem* ‘I work’ and a compound adjective *angorc* ‘lazy’. The noun *gorc* and adjective *angorc* belong to different declensions: *gorc* is an *o*-stem, genitive plural *gorcoc*, whereas *angorc* is an *i*-stem, genitive plural *angorcic*; accordingly, *gorc* and *angorc* are best seen synchronically as suffixed forms. Note that the relationship between the simplex *o*-stem *gorc* and the *i*-stem compound *angorc* can be compared to similar processes in other IE languages, for example, Latin *o*-stem *arma* ‘weapons’, *i*-stem compound *inermis* ‘unarmed’, and is taken to be an inheritance from the parent language. Although most derivation in Armenian is effected through lexical stems, Armenian is unusual among Indo-European languages in that nouns and verbs may also be derived directly from an inflected nominal form, or from a complete syntagm. For example: *kanambi* ‘having a wife’ is derived from the instrumental singular, *kanamb*, of *kin* ‘woman, wife’; a common word for ‘night’ *c’ayg* (*o*-stem) derives from the prepositional phrase *c’-ayg* where *c’-* means ‘until’ and *ayg* (normally *ow*-stem) is the word for ‘dawn’; the adjective *č’k’mel* ‘innocent, free from sin’ is formed from a complete sentence, *č’ik’ mel* ‘there is no sin’.

4.2. Unlike the older Indo-European languages, Classical Armenian does not exhibit vowel alternation (ablaut) within a lexical root as a productive derivational marker. Ablaut alternations are still found in some of the inherited vocabulary items, for example, *barjr* ‘high’ and compounds *barjraberj* ‘very high’ and *erknaberj* ‘sky-high’, where the second member of the compound shows the *e*-grade as opposed to original zero-grade in the simplex adjective (see de Lamberterie 1986 on compounds of this type); note also *snanim* ‘I nourish’ and *san* ‘nursling’ (see Klingenschmitt 1982: 226). Inherited vowel alternations can also be found preserved in isolated suffixed formations, as in the well-known example of *anjn* ‘person, self’ nominative plural *anjnk*, compound *mianjn* ‘monk’, nominative plural *mianjownk*, reflecting an Indo-European distinction between *e*- and *o*-grade, which was first identified by Meillet (1901). However, the only semi-productive use of vocalic alternations in the root is in some reduplicated compounds. In Classical Armenian total, or partial, reduplication of a lexical root or stem is frequently

employed in order to create words with an intensive or distributive sense: for example, *mecamec* ‘very big’ (*mec* ‘big’; note that this is one way to express a superlative adjective in Armenian, which has no comparative or superlative suffix); *dasadas* ‘in divisions’ (*das* ‘division’); *atxamatx* ‘diverse goods for sale’ (*atx* ‘box, baggage’). It is not unusual to find reduplication of this type with alternation of the root vowel, or even consonants: for example, *kerakowr* ‘food’ (suppletive aorist *ker-ay* ‘I ate’); *sarsowr̄* ‘cold’ (*sařn* ‘ice’); *xařamowř* ‘vulgar, of the rabble’ (*xař* ‘rude’; further examples and discussion given in Leroy 1986).

4.3. The nominal suffixes of Classical Armenian are discussed in extensive detail by Olsen (1999). Some examples of Armenian patterns of nominal suffixation can be shown from the following derivatives (excluding compounds) of *gorc* ‘work, action, manufacture’ found in 5th and 6th century Armenian: *gorcawor* ‘workman, anyone who works’ formed with a suffix *-awor* which derives historically from a verbal compounding element **bhoros*, and forms nouns denoting occupation or profession (see Olsen 1999: 358–368); *gorcaworowt’iwn* ‘work (in the abstract), labor’ shows the addition of the extremely common abstract noun suffix *-owt’iwn* (Olsen 1999: 546–584) to the previous word; *gorci* ‘tool’ (instrumental *gorceaw*) is formed with a suffix *-i* which is sometimes used, as here, to denote an instrument (Olsen 1999: 440); a derivative of *gorci* is formed with the adjectival suffix *-akan*, borrowed from Iranian, *gorciakan* ‘instrumental’; *gorcakič* ‘fellow-worker’ employs a suffix which is used to form words denoting companions or participants; *gorcac* and *gorcowac*, both ‘work’, are each derived from the verb *gorcem* using frequent suffixes for action nouns (see Olsen 1999: 231–239 and 543–545). Armenian derivative verbal suffixes are described in detail by Klingenschmitt (1982). The suffix used to form causative verbs has the form present *-owc’ane-*, aorist *-owc’-* (3rd singular *-oyc’*) and is widespread and productive in the Classical language. Causative verbs are formed through the addition of the suffix to the verbal aorist stem: for example, *dařnam* ‘I turn (intransitive)’, aorist *darjay*, causative *darjowc’anem* ‘I turn (transitive)’.

4.4. The examples of *angorc* ‘lazy’, *erknaberj* ‘sky-high’, and *mecamec* ‘very big’ given above show that compounding is a productive process of word formation in Classical Armenian; indeed several of the derivative suffixes of Armenian, such as *-awor* mentioned above, derive from generalized compound forms. For all compounds, the head occurs as the second member. The first member of a compound, if a noun or adjective, normally stands in the stem form which, for most items, is identical with the nominative singular. When the second element of a compound does not begin with a vowel, the productive pattern is to insert a liaison vowel *-a-* between the two members of the compound. However, a number of compounds are formed without the liaison vowel *-a-*, and in derivatives of compounds the liaison vowel is often dropped.

The principal productive types of compounding found in Armenian are as follows (see Olsen 1999: 657–759). First, exocentric compounds of the type modifier + head noun, for example *mecatown* ‘rich’ from *mec* ‘big’ and *town* ‘house’; *anmit* ‘mad, senseless’ from *an-* ‘without-’ and *mit(k’)* ‘mind’. Exocentric compounds frequently follow the same declension class as their head noun, but many are declined as *i*-stems: in the Bible translation, *anmit* is found declined both as an *i*-stem and as an *a*-stem, the declension class of the simplex *mit(k’)*. Second, endocentric compounds of the type modifier +

head noun: for example, *aysawr* ‘today’ from *ays* ‘this’ and *awr* ‘day’; *k’alak’orm* ‘city-wall’ from *k’alak’* ‘city’ and *orm* ‘wall’. Third, governing compounds, with a verbal element as the second member, are highly productive: for example, *jknors* ‘fisherman’ from *jowkn* ‘fish’ and the stem of the verb *orsam* ‘I hunt’ (note the reduction of the vowel *ow* when not in the final accented syllable); *andamaloyc* ‘paralytic, one who has had his limbs loosened’ from *andam* ‘limb’ and the stem of the verb *lowcanem* ‘I loose’ (3rd singular aorist *eloyc* ‘s/he loosed’). This second example shows a compound that appears to be exocentric with the first element as its head: ‘having loosened limbs’. There are also several copulative compounds in Armenian. Sometimes these show the conjunction *ew* between the two elements, as in *ert’ewek* ‘coming and going’ derived from the stems of *ert’am* ‘I go’ and *eki*, suppletive aorist of *gam* ‘I come’; but there are also examples without the conjunction, as *lrtes* ‘spy’ which combines the aorist imperatives *lowr* ‘listen’ from the verb *lsem* (aorist *loway*), and *tes* ‘see’ from the verb *tesanem*. Other compound types of Armenian, reduplicated compounds, and compounds which are calqued on Greek models are discussed above in 4.2 and 3.2.

5. References

- Ačařyan, Hračya
1971–1979 *Hayeren armatakan bařaran* [Armenian etymological dictionary]. 2nd edition. Erevan: Erevan University Press.
- Arutjunjan, Covinar Rafikovna
1983 Armjano-grečeskie leksičeskie izoglossy [Armenian-Greek lexical isoglosses]. In: Gevorg Džaukjan (ed.), *Očerki po sravnitel’noj leksikologii armjanskogo jazyka* [Essays on the comparative lexicology of the Armenian language]. Erevan: Armenian Academy of Sciences, 221–355.
- Awetik’ean, Gabriel, Xačatur Siwrmēlean, and Mkrtič Awgerean
1836–1837 *Nor bařgirk’ haykazean lezowi* [New dictionary of the Armenian language]. Venice: Tparani Srboyn Ghazaru.
- Bailey, Harold W.
1930 To the Žāmāsp Nāmāk. *Bulletin of the School of Oriental Studies* 6: 55–85.
- Bailey, Harold W.
1986 Armenia and Iran iv. Iranian Influences in Armenian, 2. Iranian Loanwords in Armenian. In: Ehsan Yarshater (ed.), *Encyclopaedia Iranica* II 4–5, 459–465.
- Beekes, Robert S. P.
1972 H₂O. *Die Sprache* 18: 117–131.
- Bolognesi, Giancarlo
1960 *Le fonti dialettali degli imprestiti iranici in armeno*. Milan: Società editrice Vita e pensiero.
- Clackson, James
1994 *The Linguistic Relationship between Armenian and Greek*. (Publications of the Philological Society 30). Oxford: Blackwell.
- Clackson, James
2000 A Greek Papyrus in Armenian Script. *Zeitschrift für Papyrologie and Epigraphik* 129: 223–258.
- Clackson, James
2003 Vox Armeniaca: New Evidence for the pronunciation of Classical Armenian. In: Anaïd Donabédian and Agnes Ouzounian (eds.), *Actes du Sixième Colloque international de*

Linguistique arménienne. INALCO- Académie des Inscriptions et Belles-Lettres 5–9 juillet 1999. Paris: Langues Orientalistes, 43–52 (= *Slovo* 26–27, 2001–2002).

Clackson, James

- 2005 Review of Frederik Kortlandt. 2003. *Armeniaca: Comparative Notes with an Appendix on the historical phonology of Classical Armenian by Robert S. P. Beekes.* *Annual of Armenian Linguistics* 24–25: 153–158.

Considine, Patrick

- 1979 A semantic approach to the identification of Iranian loan-words in Armenian. In: Bela Brogyanyi (ed.), *Studies in diachronic, synchronic and typological linguistics: Festschrift for Oswald Szemerényi on the occasion of his 65th Birthday.* Amsterdam: Benjamins, 213–228.

Cowgill, Warren

- 1960 Greek *ὄν* and Armenian *oč*. *Language* 36: 347–350.

Diakonoff, Igor M.

- 1985 Hurro-Urartian Borrowings in Old Armenian. *Journal of the American Oriental Society* 105: 597–603.

Dyen, Isidore, Joseph B. Kruskal, and Paul Black

- 1992 An Indo-European Classification: A Lexicostatistical Experiment. *Transactions of the American Philosophical Society* 82: Part 5.

Džaukjan, Geworg. B. (ed.)

- 1983 *Očerki po sravnitel'noj leksikologii armjanskogo jazyka* [Essays on the comparative lexicology of the Armenian language]. Erevan: Armenian Academy of Sciences.

Džaukjan, Geworg. B. [Jahowkjan, Geworg B.]

- 1987 *Hayoc' lezvi patmowt'iown: naxagrayin žamanakašrjan* [History of the Armenian language: the pre-literary period]. Erevan: Haykakan SSH GA Hratarakč'ut'yun.

Eichner, Heiner

- 1978 Die urindogermanische Wurzel **H₂reu* 'hell machen'. *Die Sprache* 24: 144–162.

Gray, Russell D. and Quentin D. Atkinson

- 2003 Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 435–439.

Greppin, John A. C.

- 1998 Armenian *gini*, Georgian *γvino* 'wine'. *Annual of Armenian Linguistics* 19: 65–69.

Greppin, John A. C.

- 2000 Armenian and Kartvelian. In: Yoel Arbeitman (ed.), *The Asia Minor Connection: Studies on the Pre-Greek Languages in Memory of Charles Carter.* (Orbis Supplementa 13). Leuven: Peeters, 85–91.

Greppin, John A. C.

- 2010 Hurro-Urartian-North-East-Caucasian-Armenian Linguistic Connections. *Aramazd, Armenian Journal of Near Eastern Studies* V.1: 116–120.

Greppin, John A. C. and Igor M. Diakonoff

- 1991 Some effects of the Hurro-Urartian people and their languages upon the earliest Armenians. *Journal of the American Oriental Society* 111: 720–730.

Hamp, Eric

- 1983 Mi k'ani kisašp'akanneri masin hayereni patmowt'yan meġ [On some affricates in the history of Armenian]. *Istoriko-Filologičeskij Žurnal* 1: 38–42.

Henning, W. B.

- 1958 Mitteliranisch. In: Bertold Spuler (ed.), *Handbuch der Orientalistik*, 1. IV: *Iranistik*, 1. *Linguistik*. Leiden: Brill, 20–130.

Henning, W. B.

- 1963 Coriander. *Asia Minor* 11: 195–199.

- Hübschmann, Heinrich
1877 Ueber die stellung des armenischen im kreise der indogermanischen sprachen. *Zeitschrift für vergleichende Sprachforschung* 23: 5–49.
- Hübschmann, Heinrich
1992 [1895/1897] *Armenische Grammatik. I Theil: Armenische Etymologie*. Hildesheim: Olms. [Leipzig: Breitkopf & Härtel.]
- Hübschmann, Heinrich
1906 *Armeniaca. Indogermanische Forschungen* 19: 457–480.
- Klingenschmitt, Gert
1982 *Das Altarmenische Verbum*. Wiesbaden: Reichert.
- Klingenschmitt, Gert
1994 Die Verwandtschaftsverhältnisse der indogermanischen Sprachen. In: Jens Elmegård Rasmussen (ed.), *In honorem Holger Pedersen: Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Wiesbaden: Reichert, 235–251.
- Kortlandt, Frederik
1986 Armenian and Albanian. In: Leroy and Mawet (eds.), 38–47.
- Kortlandt, Frederik
1997 Arm. *nēr* ‘sister-in-law’. *Annual of Armenian Linguistics* 18: 7–9.
- de Lamberterie, Charles
1978 *Armeniaca I–VIII: études lexicales*. *Bulletin de la Société de Linguistique* 73: 243–285.
- de Lamberterie, Charles
1986 Arménien *barjrberj* ‘très haut’. In: Dickran Kouymjian (ed.), *Armenian Studies/Études arméniennes in memoriam Haïg Berbérian*. Lisbon: Calouste Gulbenkian Foundation, 469–476.
- de Lamberterie, Charles
1988 Meillet et l’arménien. *Histoire/Épistémologie/Langage* 10: 217–234.
- de Lamberterie, Charles
1990 *Les adjectifs grecs en -ος*, 2 vols. Leuven: Peeters.
- de Lamberterie, Charles
2006 Traces de la langue poétique indo-européenne dans la langue arménienne. In: Pinault and Petit (eds.), 213–234.
- Leroy, Maurice
1986 Le redoublement comme procédé de formation nominale en arménien classique. In: Leroy and Mawet (eds.), 62–75.
- Leroy, Maurice and Francine Mawet (eds.)
1986 *La place de l’arménien dans les langues indo-européennes*. Leuven: Peeters.
- Martirosyan, Hrach K.
2010 *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden: Brill.
- Meillet, Antoine
1901 De quelques archaïsmes remarquables de la déclinaison arménienne. *Zeitschrift für armenische Philologie* 1: 139–148.
- Meillet, Antoine
1911 Sur les mots iraniens empruntés par l’arménien. *Mémoires de la Société de Linguistique* 17: 242–250.
- Meillet, Antoine
1926 Le mot *ekelec'i*. *Revue des Études Arméniennes* 9: 131–136.
- Meillet, Antoine
1937 Arménien *erkin* et *erkir*. *Annuaire de l’Institut de Philologie et d’Histoires Orientales et Slaves* 5: 1–2.
- Mercier, Charles
1978–1979 L’école hellénistique dans la littérature arménienne. *Revue des Études Arméniennes* n.s. 13: 59–75.

- Mowradyan, Arowsyak Nersesi
 1971 *Hownaban dproc'ə ev nra derə hayereni k'erakanakan terminabanowt'iwn stelcman gorcwm* [The philhellenic school and its role in the establishment of Armenian grammatical terminology]. Erevan: Haykakan SSR GA Hratarakč'owt yown.
- Nyberg, Henrik Samuel
 1931 *Hilfsbuch des Pehlevi. 2. Glossar*. Uppsala: Almqvist & Wiksell.
- Olsen, Birgit Anette
 1999 *The Noun in Biblical Armenian: Origin and Word-formation: with Special Emphasis on the Indo-European Heritage*. Berlin: Mouton de Gruyter.
- Osthoff, Hermann
 1901 *Etymologische Parerga*. Leipzig: Hirzel.
- Pedersen, Holger
 1911 Arm. *Koriun*. In: *Huschardzan. Festschrift aus Anlass des 100 jährigen Bestandes der Mechitaristen-Kongregation in Wien (1811–1911) und des 25. Jahrganges der philologischen Monatsschrift "Handes amsorya" (1887–1911)*. Vienna: Mechitaristen, 287–288.
- Pedersen, Holger
 1924 Armenier. In: Max Ebert (ed.), *Reallexikon der Vorgeschichte*, Band I. Berlin: De Gruyter, 219–226.
- Pinault, Georges-Jean and Daniel Petit (eds.)
 2006 *La langue poétique indo-européenne: Actes du colloque de travail de la Société des Études Indo-Européennes (Indogermanische Gesellschaft/Society for Indo-European Studies), Paris, 22–24 octobre 2003*. Leuven: Peeters.
- Pokorny, Julius
 1959 *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.
- Porzig, Walter
 1954 *Die Gliederung des indogermanischen Sprachgebiets*. Heidelberg: Winter.
- Ringe, Don, Tandy Warnow, and Ann Taylor
 2002 Indo-European and Computational Cladistics. *Transactions of the Philological Society* 100: 59–129.
- Ritter, Ralf-Peter
 2006 Zur Frage der Reste indogermanischer Dichtersprache im Armenischen. In: Pinault and Petit (eds.), 213–234.
- Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
 2001 *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstambildungen*. 2nd, corrected edn. Wiesbaden: Reichert.
- Schmidt, Karl-Horst
 1985 Die indogermanischen Grundlagen des altarmenischen Verbuns. *Zeitschrift für vergleichende Sprachforschung* 98: 214–237.
- Schmitt, Rüdiger
 1975 Von Bopp bis Hübschmann: Das Armenische als indogermanische Sprache. *Zeitschrift für vergleichende Sprachforschung* 89: 3–30.
- Schmitt, Rüdiger
 1983 Iranisches Lehngut im Armenischen. *Revue des Études Arméniennes* n.s. 17: 73–112.
- Schmitt, Rüdiger
 1986 Armenia and Iran. In: Ehsan Yarshater (ed.), *Encyclopaedia Iranica* II 4–5: 417–459.
- Solta, Georg Rhenatus
 1960 *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*. Vienna: Mechitaristen.
- Solta, Georg Rhenatus
 1990 Die Stellung des Armenischen im Kreise der indogermanischen Sprachen. Ein Überblick. In: E. M. Ruprechtsberger (ed.), *Armenien: Frühzeit bis 1. Jahrtausend; Sprache, Kunst und Geschichte*. Linz: Stadtmuseum, 7–18.

- Szemerényi, Oswald
 1977 *Studies in the kinship terminology of the Indo-European languages*. Leiden: Brill.
- Taylor, Ann, Tandy Warnow, and Don Ringe
 2000 Character-Based Reconstruction of a Linguistic Cladogram. In: John Charles Smith and Delia Bentley (eds.), *Historical linguistics 1995: selected papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995. Volume 1: general Issues and Non-Germanic Languages*. Amsterdam: Benjamins, 393–408.
- Thumb, Albert
 1900 Die griechischen Lehnwörter im Armenischen. *Byzantinische Zeitschrift* 9: 388–452.
- Watkins, Calvert
 1995 *How to Kill a Dragon. Aspects of Indo-European Poetics*. New York: Oxford University Press.

James Clackson, Cambridge (United Kingdom)

65. The dialectology of Armenian

1. General description
2. Armenian dialects and PIE
3. The early isoglosses
4. References

1. General description

1.1. Armenian dialects: number, source, geography

With the term “Armenian dialects”, linguists usually designate the dialects of the Modern Armenian language in their geographical distribution from the 19th century onwards up to the beginning of the 20th century (for the extinct Armenian dialects of the Ottoman empire) or the present time. The number of Armenian dialects obviously depends on the classifications applied. In his comprehensive overview of Armenian dialectology, Ĵahukyan (1972: 132–136) distinguishes two main branches (East and West) in Modern Armenian and 11 dialect groups comprising 44 individual dialects. An enumeration in English is given by Ĵahukyan (1986; in this work on p. 22 insert between No. 23 and 24 the line “VI. The Mush-Tigranakert or South-Central intergroup”). Dialects identified after the appearance of Ĵahukyan (1986) or not accounted for in this work are Bolu, a Karabagh dialect in Western Turkey (Samuelian n. d.), Stanoz, Yozgat (Mkrtč’yan 2006), and Jerusalem (Vaux 2002).

The sources of our knowledge of individual modern dialects are the currently spoken dialects (including the dialects of the Anatolian area that are still used by survivors of the 1915 massacres and their descendants), texts collected in ethnographic studies (for a

- Weitenberg, Jos J. S.
1977c The prepositional group *i y-* and the orthography of Gospel manuscript M (Matenadaran 6200). *Annual of Armenian Linguistics* 18: 39–50.
- Weitenberg, Jos J. S.
1999–2000 On the early development of Armenian dialects. II. The monophthongization of *ay*. *Annual of Armenian Linguistics* 20: 1–26.
- Weitenberg, Jos J. S.
2001 Thoughts on Armenian accentuation. *Annual of Armenian Linguistics* 21: 65–73.
- Weitenberg, Jos J. S.
2002 Aspects of Armenian dialectology. In: Jan Berns and Jaap Van Marle (eds.), *Present-day Dialectology. Problems and Findings*. (Trends in Linguistics. Studies and Monographs 137). Berlin: Mouton de Gruyter, 141–157.
- Weitenberg, Jos J. S.
2004 Armenian *barwok* ‘good, well’. In: Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson, and Thomas Olander (eds.), *Per Aspera ad Asteriscos. Studia Indogermanica in Honorem Jens Elmegaard Rasmussen*. Innsbruck: Institut für Sprachwissenschaft der Universität, 627–632.
- Weitenberg, Jos J. S.
2006 Aspects of Classical Armenian orthography: Armenian *e*, and the Greek names in the Gospels. In: Anna Krasnuolska, Kinga Maciuszak, and Barbara Mękarska (eds.), *In the Orient where the Gracious Light ... Satura Orientalis in Honorem Andrzej Pisowicz*. Cracow: Księgarnia Akademicka, 215–228.
- Weitenberg, Jos J. S.
2008 Diphthongization of initial *E-* and the development of initial *Y-* in Armenian. In: Alexander Lubotsky, Jos Schaeken, and Jeroen Wiedenhof (eds.), *Evidence and Counter-Evidence. Essays in Honour of Frederik Kortlandt. Volume I. Balto-Slavic and Indo-European Studies*. Amsterdam: Rodopi, 609–616.

† Jos J. S. Weitenberg

66. The evolution of Armenian

- | | |
|--------------------------|---------------|
| 1. Varieties of Armenian | 5. Syntax |
| 2. Foreign influences | 6. Lexicon |
| 3. Phonology | 7. Texts |
| 4. Morphology | 8. References |

(In order to accommodate Armenians, armenologists, and linguists alike we render all linguistic forms in both Armenian script and the International Phonetic Alphabet. Bibliographic references, Classical Armenian forms, and names of authors and dialects, on the other hand, are rendered in the ALA-LC system (<http://goo.gl/z0rs0m>) so as to facilitate bibliographic research.)

<https://doi.org/10.1515/9783110523874-021>

1. Varieties of Armenian

This chapter outlines the historical development of the dialects making up the Armenian branch of the Indo-European language family. We follow Kortlandt (1985) in referring to the common ancestor of the modern dialects as “Common Armenian”; the language is not attested in any written sources at this stage, and so Common Armenian must be reconstructed based on comparison of the modern dialects and comparative evidence from other branches of Indo-European. The language is first known to have been written down with the introduction of Christianity in around the year 405, when a cleric named Mesrop Mashtots created a new alphabet for writing Armenian translations of Christian texts (Russell 1999). There have been four main literary varieties of Armenian, at different times and places in history:

Classical Armenian was the standard literary language in the fifth century, and is the earliest attested form of the language. No literature from before the Christianization of Armenia survives; most of the early literature in the classical period consists of translations of Christian texts from Greek or Syriac (Thomson 1989).

We cannot identify Common Armenian with Classical Armenian, since some modern dialects preserve archaisms reconstructible from Indo-European that are lost in Classical Armenian, so they must descend from an unwritten earlier language also preserving these archaisms. One example is the form for ‘milk’, Classical Armenian **կաթն** *kat^hn*, which in comparison with Greek γάλα, γάλακτος and Latin *lac*, *lactis* suggests a Mediterranean proto-form (if not Indo-European root) **g₁gt-* (Martirosyan 2010). Both the syllabic **ɟ* and the second **g* have been lost in the classical form, while the remaining two stops have undergone the expected shift to *k* and *t^h*, respectively (Karst 1901). In the dialect of Agulis, however, we have **կախց** *kaxts^h*, with a *-χ-* not found in the classical form (but found in certain other dialects; cf. Hawarik **կախս** *kaxs*, Acharyan 1973: 481). Acharyan (1901) takes this *-χ-* to be the result of a proto-form **kałts^h*, where the dark **ɫ* is a reflex of the original **l* seen in Latin and Greek. Since these dialects preserve an archaic form not found in the classical language, it seems reasonable to propose that these dialects split off from Common Armenian before Classical Armenian innovated the loss of the **ɫ*, and so Common and Classical Armenian are not the same stage of the language. (See Martirosyan 2010 for extensive discussion of further archaisms found in the modern dialects.)

Middle Armenian is attested from the 11th to the 15th centuries (Karst 1901: 1). Most Armenian varieties of this period have only fragmentary attestation; Cilician Armenian happens to have survived because it was the official literary language of Cilicia, a kingdom founded by Armenian refugees in southern Anatolia. This Cilician variety is ancestral to modern dialects of Cilicia (notably those of Hadjin, Marash, and Zeytun). As Karst (1901) extensively documents, many of the morphological and syntactic features of Modern Armenian already appear in Middle Armenian, such as the agglutinative number-case system in nouns and the use of particles such as **կու** *ku/gu* in the verbal system.

Many scholars, including Parnassian (1985) and Lassiter (2016), believe that there was a further major stage in the evolution of Modern Armenian from Middle Armenian, which following Kostandnupölsets‘i (1674: 3) and Schröder (1711) they call *Civil Armenian* (**քաղաքական հայերէն** *k^havak^hakan hajeren*, or for Schröder *lingua civilis*). According to Motalová (*apud* Zgusta 1971: 192–193), this emerged in the seventeenth

century and contained elements of Grabar (the classical literary language) and various non-standard dialects, and was predominantly used in letters, administration, the courts, scientific literature, and newspapers. (Zekiyan 1997: 338 states that Civil Armenian first emerged in the fourteenth century and was initially called *րամկորէն ramkoren* [‘common’ or ‘vulgar’ language].) According to Zekiyan, Civil Armenian predated the split of modern literary Armenian into Eastern and Western varieties in the middle of the nineteenth century; it had “almost a unitary character as much as unity at a formation stage of the language is allowed” and was “a common means of literary expression for all Armenians. Hence most probably it was also, aside from the various dialects, a common means of oral communication, especially used and developed by the travelling merchants” (Zekiyan 1997: 338).

Parnassian (1985) argues that 17th-century Civil Armenian texts such as Zak‘aria Agulets ‘i’s diary (1647–1664; cf. Ter-Avetisyan 1938) show a mix of Western and Eastern features; for example, they tend to construct the present tense with forms of *կու ku*, the perfect tense with forms of the *-էր -er* participle, and the ablative case with *-է -e* (all Western); but they tend to employ the locative *-ում -um* (Eastern) and the genitive plural in *-ի -i* (E; Western *-ու -u*). Nichanian (1989: 273–277) adds that Kostandnupōlsets ‘i uses *պիտի piti* as a future (W) rather than an obligatory (E) marker in his version of Civil Armenian.

These generalizations do not hold for all Civil Armenian documents; for example, the version of Civil Armenian described by Schröder (1711), which appears to be based on the speech of his informant Lucas Nurigianides (cf. Schröder 1711: 1.6), uses *կու ku* for the future tense (E) and both W/Classical *-է -e* and E *-ից -its^h* for the ablative (Lassiter 2016).

Zekiyan (1997: 338) asserts that the power and influence of the Armenian merchant elites in Constantinople, Tbilisi, and Erevan most prominently associated with Civil Armenian began to wane in the mid-eighteenth century, precipitating a decline in the use of this form of the language and ultimately leading to the rise of the Modern Eastern and Western literary languages in the mid-nineteenth century. Zekiyan (1997: 338) echoes the general belief among Armenians and armenologists that these were based on the dialects of Erevan and Constantinople respectively, but dialectologically informed examination of the phonological, morphological, and lexical features of the literary varieties suggests that the situation is more complicated than this.

Consider first *Standard Western Armenian* (SWA), which is normally taken to be based on the dialect of Istanbul (IA). Adjarian (1906) outlines this view of the history of SWA in a series of steps. When the Turks invaded in the 11th century, the survivors spoke Middle Armenian as a lingua franca based on spoken Armenian. This began to fall apart for two reasons: firstly, the Cilician Armenians dispersed in the 15th century, after the kingdom of Cilicia dissolved following the invasion of the Mamluks in the 14th century. Secondly, the Turco-Persian wars of the 15th century onwards were fought mainly in Armenia, scattering the Armenian language as refugees established communities elsewhere, including in Constantinople. With the late 18th century came the establishment of the first Armenian schools in Constantinople; shortly afterwards, the new dialect began to be written in Constantinople, Smyrna, and Venice, and disseminated in print through newspapers, journals, and by missionaries. The 1848 revolutions spurred movements to replace the classical language with this new dialect as a standard, and since

Standard Eastern Armenian (SEA) was banned by the Ottoman government at that point, SWA subsequently developed with minimal influence from eastern varieties.

If we compare the traditional Istanbul dialect described in 1941 by Achařean, himself a speaker of the dialect, to SWA as described by Bardakjian and Thomson (1977), a somewhat different picture emerges. The outcomes of the stop series differ, for example: IA belongs to Group 3 and preserves original word-initial voiced stops unchanged (e.g. բան ‘thing’ → [ban]), whereas SWA, like all Group 5 dialects, devoices and aspirates them ([p^han]). Word-initial mid vowels diphthongize in SWA (երկու [jergu] ‘two’, ոսկի [vɔsgɪ] ‘gold’), but not in IA ([ergu], [ɔsgɪ]). Morphologically, IA differs from SWA in negating the present tense with the old preposition *ի* ‘in’, as in չեմ *ի* գար [tʃ^hem i gar] ‘I’m not coming’ vs. SWA չեմ գար [tʃ^hem k^har]. Yes-no questions employ the Turkish clitic *-մի -mi* in IA but not SWA; compare IA կը սիրէ՞ մի [gə sire mi] ‘does he love?’ with SWA կը սիրէ [gə sire]. Numerous verbs belong to the *ա-* *a*-conjugation in IA but the *ե-* *ε-* or *ի-* *i*-conjugation in SWA, such as IA տեսնալ [desnal] ‘see’ vs. SWA տեսնել [desnel], IA խօսալ [χɔsəl] ‘speak’ vs. SWA խօսիլ [χɔsil].

The lexical differences between Istanbul and SWA are equally striking and pervasive: IA բայրինկուն [p^hajringun] ‘good evening’ vs. SWA բարի իրիկուն [p^hari irigun], IA օյրորդ [ɔjɔr^h] ‘lady’ vs. SWA օրիորդ [ɔriɔr^h], IA վով [vɔv] ‘who’ vs. SWA ով [ɔv], etc.

It is similarly difficult to maintain the common belief that *Standard Eastern Armenian* is based on the Erevan dialect (EA) once one examines the relationship between the two. The pronunciation of the vowels is significantly different in the Erevan dialect than in SEA, as was noted by Gharibyan (1948: 75) and remains true at the time of writing this chapter. EA stresses the penultimate syllable, whereas SEA resembles French in stressing the rightmost full vowel in a word; EA also undergoes extensive reduction of unstressed vowels, unlike SEA: contrast SEA գրում եմ [gərum em] ‘I write’ with EA գըրում էմ [g^hərəm em] (loc. cit.). This example shows moreover that EA differs from SEA in having a fourth stop series, voiced aspirates, corresponding to SEA (and Classical Armenian) plain voiced stops in word-initial position. EA, but not SEA, also voices original plain voiceless stops in medial and final position, as in SEA փայտ [p^hajt] ‘wood’ : EA փէտ [p^hed], SEA կատու [katu] ‘cat’ : EA կադու [kadu] (Gharibyan 1948: 76). Morphologically, EA differs from SEA in forming the plural of many polysyllabic vowel-final nouns with *-ք* [-k^h], whereas SEA employs the standard polysyllabic plural suffix *-ներ* [-ner]: contrast SEA այգիներ [ajginer] ‘vineyards’ with EA իքիք [ik^hik^h] (Gharibyan 1948: 77). The obligatory in EA can optionally be formed in a manner not found in SEA, wherein the obligatory marker is placed after the infinitive and conjugated, as in EA կարժալ պրդէմ [kart^hal pədem] ‘I must read’ : SEA պիտի կարդամ [piti kardam] (Gharibyan 1948: 83; see also Asatryan 1980).

In sum, the literary dialects appear to have arisen not from the dialects of Constantinople and Erevan, but from something like Civil Armenian combined with elements of various Western or Eastern dialects, including but not limited to the varieties spoken in the respective capitals. To take just two SWA examples, the Classical verb ուղարկել [uɔarkel] ‘send’ surfaces in current SWA as դրկել [ɔrgel]; this does not accord with the Istanbul form խրկել [χərgel], but rather is what we find in the dialects of Akhaltskha and Sivas among others. For the verb ‘swell’, Classical Armenian attests ուռնուլ [urnul] and ուռչիլ [urtʃ^hil]; corresponding to this we find ուռնալ [urnal] (HAB [=Achařyan 1973] 607) in Istanbul (as well as Rodosto and Sivas). The SWA form however is ուռիլ

[uril] (Guyumchean 1970: 640; Sargsyan 1991: 293), as we find not in Istanbul but in many other Western dialects including Mush, Suczawa, Van, Axalcxa, Erzerum, Moks, Tigranakert, Nor Naxichevan, and Hamshen (HAB 607). The fact that the standard literary varieties draw on dialects beyond those of Constantinople and Erevan is perhaps not surprising, given the diverse regional origins of the speakers and writers of the language from its formative period to the present day.

2. Foreign influences

The evolution of Armenian has been affected not just by the endogenous dialectal forces just described, but also by an array of exogenous influences. One of the hallmarks of Armenian is the extent to which it has adapted elements from other languages, a process which has continued from the pre-historic and Classical periods (cf. Clackson, this handbook) to the present day. While all forms of Armenian have been extensively influenced by, among others, Middle Iranian languages, Turkish and Azeri, and Arabic, over the past two centuries, the modern literary varieties have been differentially affected by the ambient languages of their sociopolitical milieu: Russian in the case of SEA, Persian in the case of Iranian Armenian, and Turkish in the case of SWA (though most of the modern Turkish lexical elements in the latter were excised following the Genocide of 1915–1924).

In the lexical domain, exogenous influences can take the form of direct lexical borrowing, calquing, and semantic differentiation. SEA for instance contains a host of Russian borrowings not found in SWA, such as *մայկա* [majka] and *ֆուտբոլկա* [futbolkɑ] ‘(t-)shirt’, from Russian *майка* and *футболка*, respectively. The SWA equivalent is *շապիկ* [ʃabig], itself a loan from Middle Iranian *šabīg*, which Mackenzie (1971) glosses as ‘Mazdean ritual undershirt’, though the derivation from Iranian *xšap-ika- ‘night’ suggests that it originally meant something like ‘night-shirt’. Similarly on a hot day in the Republic of Armenia or Moscow one might ask for *մարոճնի* [marɔʒni] ‘ice cream’, from Russian *морозное* [mɐˈrozɨnəjə]; compare SWA *պարպարակ* [bɑʁbɑʁɑg], a reduplicated derivative of native *պար* [bɑʁ] ‘cold’. (Some SEA speakers now use their equivalent, [րաբրաբɑk].)

Calquing from Russian surfaces in words such as *ինքնաթիռ* [inkʰnatʰir] ‘airplane’, based on Russian *самолёт* [səmɐˈlʲɵt], both of which have the morphological structure ‘self-fly’; compare SWA (and Iranian Armenian) *օդանավ* [ɔtʰɑnɑv], based on archaic English *airship*. The SEA verb *զանգել* [zɑngɛl] ‘call on the telephone’ is a calque on Russian *звонить* [zvɐˈnʲitʲ], both literally meaning ‘to ring (a bell)’; SWA *հեռաձայնել* [heratsʰajnel] on the other hand is a calque on French *téléphoner*, both being verbal derivatives of ‘sound-from-a-distance’. With respect to cars, SEA *ավտոմոբիլ* [ɑvtomɔbil] is a direct borrowing from Russian *автомобиль* [ɐftəmɐˈbʲilʲ], whereas SWA *ինքնաշարժ* [inkʰnɑʃɑʁʒ] is a calque on French *automobile*, literally ‘self-moving’.

When borrowed words come into competition with pre-existing native forms, semantic differentiation (change in the meaning of either the original or the incoming synonym) often results. An SEA example involves the edible tuber of the *Solanum tuberosum*

plant, or ‘potato’. Originally cultivated in the Andes, this vegetable made its way to Europe following the Spanish conquest in the mid-16th century, and shows up in Armenian as *գետնախնձոր* *getnaxnjor* (literally ‘earth-apple’) soon thereafter. (Awetik‘ean et al. 1836: 540 cites the two earliest examples as being from an unspecified medical text [*Բժշկարան* *Bžškaran*], but this must be later than the most famous Bžškaran by Amirdovlat‘ Amasiats‘i, which dates to the late fifteenth century, before the introduction of the potato to the Old World. It is not clear to us how to square the linguistic evidence from Armenian with the proposal that the potato was introduced to Iran [and hence presumably Armenia] by John Malcolm in the early 19th century [Reader 2008: 246].) The Armenian form *գետնախնձոր* *getnaxnjor* appears to be a calque on French *pomme de terre* (perhaps via Persian *sebi zamīnī*), and remains the form for ‘potato’ in SWA and Iranian Armenian. Modern SEA has now imported Russian картофель [*kər‘tofʲilʲ*] as *կարտոֆիլ* [*kartɔfil*], and this serves as the word for potato for many speakers, whereas [*gɛtnɑxəndzɔɾ*] has shifted in meaning to ‘yam’ or ‘Jerusalem artichoke’.

3. Phonology

Turning to the phonological evolution of Armenian in the historical period, one significant way in which the varieties of the language vary is in their treatment of the three stop series inherited from Proto-Indo-European. The outcomes of these series vary mostly with respect to Voice Onset Time (and indeed, Adjarian first came up with the concept of Voice Onset Time while studying laryngeal contrasts across Armenian dialects; cf. Adjarian 1889 and Braun 2013). A summary of the outcomes of word-initial stops is given in Table 66.1, together with the traditional classification and representative dialects from each group.

Tab. 66.1: Armenian dialect stop series

group	*D	*D ^h	*T	example dialects
1	D	D ^h	T ^h	Sivas
2	T	D ^h	D	Erevan, New Julfa
3	D	D	T ^h	Istanbul
4	D	T	T ^h	Sasun, MidA, Kesab
5	D	T ^h	T ^h	Malatya, SWA
6	T	D	T ^h	Classical, SEA
7	T	T	T ^h	Van

Here D stands for voiced stops, D^h for voiced aspirated stops, T for voiceless stops, and T^h for voiceless aspirated stops. Note that in some dialects, such as Van, we see a merger of two of the series. The behavior of loanwords suggests that these changes occurred across the dialects between the 6th and 13th centuries (Weitenberg 2002: 148). By way

of example, consider the PIE words for ‘ten’, ‘I carry’, and ‘eight’, and their reflexes in a sample of dialects in Table 66.2:

Tab. 66.2: Reflexes of the PIE stop series

group	*D	*D ^h	*T	variety
	*deḱmt	*b ^h er-	*oktō:	PIE
	‘ten’	‘I carry’	‘eight’	
1	dasə	b ^h εrɛm	ut ^h ə	Sivas
2	tasə	b ^h iɛ.riɛm	ut ^h	New Julfa
3	dasə	bεrɛm	ut ^h u	Istanbul
4	das	pεrɛm	ut ^h	Sasun
5	dasə	p ^h εrɛm	ut ^h ə	SWA
6	tasn	bεrɛm	ut ^h	Classical
7	tas	piɛrɛm	ut ^h	Van

Under the assumption that Group 6 represents the Proto-Armenian situation (see Pisowicz 1976), as Classical Armenian is the oldest attested variety, an ordering paradox arises in deriving the Group 4 series. The problem is that we seem to have an inversion of the *D and *D^h series: Classical *teḷi* ‘place’ corresponds to Kesab *diεk*, while Classical *deḷ* ‘drug’ to Kesab *tiεk*. But if either the d > t or the t > d change happened prior to the other, we would see a merger, as the second change would in each case restore the original sound, extending it to both original series.

One solution to this problem would be to take another series to be original; taking Group 6 to be original produces, as we have seen, an ordering paradox with respect to 4; and 3, 5, and 7 all involve mergers and so cannot be original; therefore, it appears that either 1 or 2 must be original. If so, this has the consequence that the voiced aspirates in groups 1 and 2 descend directly from the original Proto-Indo-European voiced aspirates, and so also avoids the complication of having to propose any sound changes in between (Garrett 1998).

The behavior of Greek and Iranian loanwords poses problems for this analysis of the voiced aspirates as original, however. If Group 1 is original, the incoming voiced stops should be assigned to the *D series, which would then come out as plain voiceless in Group 6; a /b/ borrowed into Common Armenian should come out as a /p/ in Classical Armenian, for example. But we see loanwords present in multiple dialects (and so reconstructible for Common Armenian) that do not show this behavior: Greek *bēma* ‘stage’ appears as Classical *բէմ* *bēm*, and not *pēm. By the same reasoning, the change w > g ought to give *k in Group 6, if Group 1 were original, whereas from the word ‘wine’ we see Classical *գինի* *gini*, and not *kini, as we would expect if it were first borrowed into an ancestral Group 1 dialect as *gini*.

We therefore need to treat Group 6 as original after all, and somehow derive the Group 4 situation. Taking inspiration from our last hypothesis, we can propose that Group 4 arose via an intermediate Group 1 stage, even though this stage was not part

of Common Armenian. We need to propose a sound change $*D > *D^h$ in Groups 1 and 2, whereby the original Group 6 voiced series (which in turn came from the Proto-Indo-European voiced aspirate series) became a voiced aspirate series. Pisowicz (1976) proposes that this sound change happened in Group 4 as well; Group 4 can then be derived from Group 1 by deaspirating (and subsequently devoicing) the voiced aspirate series, to reach the attested facts.

An interesting change affecting vowels in some modern dialects is Adjarian's Law, in which back vowels undergo certain changes after voiced obstruents (Vaux 1992). Adjarian first noticed this change in the modern dialect of Van, in which back vowels are consistently fronted after historically voiced obstruents (which are synchronically voiceless in Van) as shown in Table 66.3:

Tab. 66.3: Examples of Adjarian's Law

Classical form	Modern form	Gloss
bah	pæχ	'spade'
danak	tænæk	'knife'
gair̄n	k ^h æɾ	'sheep'
bołk	pøχk	'radish'
dzu	tsy	'egg'

Some dialects exhibit a similar rule affecting vowel quality without fronting, giving a clue as to the origin of this change. Consider the data from Malatya Armenian in Table 66.4 (Danielyan 1967); these IPA values are the authors' interpretation of Danielyan's phonetic descriptions.

Tab. 66.4: Vowel differentiation in Malatya Armenian

Earlier form	Malatya	Gloss
t ^h as	t ^h as	'cup'
p ^h ot ^h	p ^h ot ^h	'plait'
p ^h uk ^h	p ^h ok ^h	'breath'
das	t ^h es	'lesson'
boyt ^h	p ^h ot ^h	'thumb'
buk ^h	p ^h uk ^h	'snowstorm'

The Malatya minimal pairs in Table 66.4 suggest that the historical contrast in voicing on consonants has turned into a different contrast in vowels, involving [atr] (advanced tongue root); /e o u/ can be analysed as [+atr], and /a ɔ ʊ/ as [-atr]. If voiced consonants are specified with the feature [+atr] (Vaux 1998), this change is simply represented as the spreading of [atr] from consonants to their following vowels. In Malatya, voicing on stops has then neutralized, phonologizing this change.

This can be seen as similar to an earlier stage of the rule in dialects such as Van, where the rule specifies fronting rather than a change in [atr]. In the relevant dialects, the first change was that the feature [atr] spread from consonants to their following vowels. Dialects such as Van then innovated a further rule, whereby [ɑatr] → [-ɑback]; the contrast in [atr] on vowels was mapped by this second rule into a contrast in [back], the end result of which is equivalent to back vowels being fronted after historical voiced obstruents. Then, in Van, Malatya, and many other dialects, the voicing contrast in stops was neutralized.

A significant number of modern Armenian dialects have innovated a system of vowel harmony that was not present in Classical Armenian; these systems are typologically very common, and their presence in Armenian likely reflects areal influence from other neighbouring vowel harmony systems like that of Turkish (Vaux 1998).

The dialect of Aresh, for example, was spoken until 1918 in the region around the southeast corner of the Mingachevir reservoir, near Yevlax and Mingachevir (Lusents' 1982). We may interpret the description of the surface vowel inventory in Aresh provided by Lusents' in the following way:

Tab. 66.5: Aresh vowel inventory

i	y	əi		u
ε	œ	ə	ə̇	ɔ
æ		ɑ		

Of these vowels, Lusents' designates [ɑ ə ə̇ əi u ɔ] as “heavy”, [i æ y œ] as “light”, and [ε] as “neutral”.

Each affix in Aresh comes in two forms: a “heavy” (back) one and a “light” (front) one, determined by the vowel in the root of the word. For the Classical derivational suffix *-akan*, for example, we have allomorphs *-akan* and *-ækæɛn*; Aresh *talakan* ‘debt’ < *tal* ‘give’, while *g'ə̇lækæɛn* ‘future’ (with root vowel fronted by Adjarian’s Law) < *gal* ‘come’. The same holds for inflectional affixes, such as the genitive ending *-i* > *-a* ~ *r* -æ. For *k^har* ‘stone’ we have *k^har*, gen. *k^hara*; but for *amis* ‘month’ we have *æmis*, gen. *æmisæ*.

4. Morphology

The nominal morphology of Armenian has been radically slimmed down over its history from Classical Armenian, and even more so from Proto-Indo-European. Classical Armenian has seven nominal declensions showing limited ablaut, a change from the Indo-European system in which the vowels in roots, suffixes, and endings underwent extensive ablaut in accordance with the position of the mobile accent (Beekes 2011). Classical Armenian had already lost the inherited gender distinction of masculine, feminine, and neuter, so nominals were inflected only for number and case; this was entirely done by fusional markers, where for the most part a single morpheme marked each number-case combination (Meillet 1936; Olsen 1999).

Modern Armenian, by contrast, has only one productive declension, with separate number and case morphemes and no fusion (Halle and Vaux 1998). We can see the contrast between Classical Armenian’s multiple fusional genitive plural affixes and SEA’s separation of number and case and levelling of declensions in the forms for ‘water’ and ‘hair’ (Table 66.6).

Tab. 66.6: Development of agglutination in the SEA genitive plural

Classical	ջրոց	յ(u)r	oc'	
	հերաց	her	ac'	
	gen. pl.	ROOT	PL	GEN
SEA	ջ(ու)րեբրի	dʒ(u)r	ɛr	i
	հերեբրի	hɛr	ɛr	i

The verbal system has undergone a significant realignment since the Classical Armenian period, in which each morphological form shifts in meaning to become the marker for a different verbal category, and new grammaticalized forms are created for categories to which no previously existing morphological form has shifted. The overall picture is something resembling a chain shift, although it should be noted that this shift is not a single entity, since the individual changes did not take place in the same period. There is also considerable dialectal variation with regard to which shifts took place, and which new grammaticalized forms were created.

The Classical Armenian verb has a present, an aorist, and an imperfect; of these, the present and aorist have indicative and subjunctive forms. There is no future tense inherited from Indo-European, but the aorist subjunctive endings already come to have future meaning by the classical period, as well as conveying intention or desire (Vaux 1995). Also during the classical period, the present subjunctive form disappears, and it is preserved in none of the modern dialects (Weitenberg 1993).

In Cilician Middle Armenian, the function of the present subjunctive comes to be filled by the original present indicative of Classical Armenian, inherited from Indo-European; in almost all of the modern dialects, this form continues to have subjunctive rather than indicative force (see Vaux 2013 for one exception – the Khodorjur dialect – which preserves it as an indicative in certain contexts).

In turn, the present indicative function is filled by a periphrastic formation that arose late in the classical period. This uses the collocation *կայ եւ* *kay ew*, literally ‘there exists and’, followed by a conjugated form of the classical present (Karst 1901). This formation may originally have had a progressive meaning, evidenced by certain grammatical restrictions in the modern dialects; in SWA, for example, the descendant *կու* /gu/ of *kay ew* cannot normally be used with stative verbs, resembling the behavior of progressive forms in other languages such as English. In Cilician Middle Armenian, the reduced form /gu/ had become the standard marker of the present tense. In SWA and many other western dialects, this marker continues the present tense function, but in SEA and other eastern dialects, it has become a future tense marker.

The late Classical Armenian future tense, expressed by the original aorist subjunctive, disappears in Middle Armenian and the modern dialects, to be replaced by various periphrastic formations. In Middle Armenian, the future tense is expressed by the conjugated verb *կամիմ gamim* ‘want’ followed by the infinitive; none of the modern dialects retain this formation. As noted above, most eastern dialects have the future function performed by the Middle Armenian present tense, from the original progressive. This use of ‘want’ to form the future, as well as being cross-linguistically common, is a characteristic of the Balkan Sprachbund (Joseph 1983); this is geographically linked to the Byzantine and Ottoman empires, which also happen to be the most plausible sources of influence on Cilician Armenian.

In most modern western dialects, the future is instead expressed by an old periphrastic form that came from an obligatory mood formation in Middle Armenian: this was an invariant form *պիտի piti* ‘it is necessary’ followed by a conjugated form of the classical present. This was not present as a separate mood as such in Classical Armenian, but the formation developed from an impersonal construction *պիտի որ piti or* ‘it is necessary that’. In SEA and many other eastern dialects, this obligatory formation survives with its original force (Dum-Tragut 2009).

An interesting new verbal formation has developed in some Western dialects. The Balkan Sprachbund was coextensive with the Byzantine domain (Sandfeld 1930) and the Ottoman domain, and Armenia was under the control of both of these for around one thousand years, so we might expect there to be some Balkan influence on the morphology of modern Armenian. One potential such feature is the evidential, or mediative (Donabédian 1996, 2001).

The normal shape of the perfect in Standard Western Armenian is periphrastic: it consists of the aorist stem marked with a perfect participle suffix, followed by a form of the auxiliary ‘be’ inflected for tense, person, and number. But many varieties of spoken Western Armenian contain two different perfect participle suffixes which impart two slightly different meanings, as in the following two constructions:

- (1) a. *պարկած են*
 barg-adz ε-n
 lie.down-PPL be-3.PL
 ‘they are lying down’
- b. *պարկեր են*
 barg-εr ε-n
 lie.down-MED be-3.PL
 ‘they are supposedly/probably/unfortunately lying down’

The former of these is the unmarked perfect, using the perfect participle suffix *-ած -adz*. The latter is the marked, “mediative” perfect, using *-եր -er*. Donabédian (1996, 2001) reports that there are significant differences in the interpretation of the evidential participle that distinguish it from the unmarked perfect participle in *-adz*; these differences have to do with aspect, modality, and discourse conditions.

The evidential is appropriate when there is some contribution from the speaker over and above the assertion that the proposition P expressed by the sentence is true. This contribution can take the following forms, hence the variable gloss ‘supposedly/probably/unfortunately’ above:

- The speaker did not witness the event reported in P; P is asserted on the basis of hearsay or inference.
- P is contrary to the speaker’s expectation or accidental.
- The speaker does not approve of what is related in P.
- The speaker does not commit herself to the veracity of P.

Taking the ‘lying down’ example in (1), Donabédian (2001) observes that if the mediative perfect *barger en* is uttered by a mother-in-law about her daughters-in-law, a likely interpretation would be that the former does not approve of or is surprised by the latter’s behavior. If, on the other hand, the unmarked perfect participle *bargadz* is used, the utterance has a more matter-of-fact flavor, in which the speaker does not include any implicit commentary on how she feels about the fact that her daughters-in-law are resting.

One can also use a fixed 3rd singular past subjunctive form of ‘be’, *Եղբր յեբեր*, as an evidential particle with predicates of any form, mediative or otherwise; this encodes the same four properties already seen. Aytėnean (1883) attributes this use of the form to influence from Turkish *-miş*.

5. Syntax

With respect to a specific cluster of syntactic features, the varieties of Armenian fall into two typological categories, summarized in Table 66.7:

Tab. 66.7: Typological categorization of Armenian varieties

	Type	Varieties of Armenian	Features
i.	head-initial	Classical Armenian Middle Armenian	unmarked SVO word order prepositions adjectives can follow head noun Indo-European (fusional) inflection
ii.	head-final	Modern Armenian	unmarked SOV word order postpositions preposed modifiers agglutinative inflection

To illustrate the differences between the two categories in Table 66.7, consider the following representative noun phrases and relative clauses in Classical and Modern Armenian (data modified from Acharyan 1911: 24):

(2) noun phrase: ‘my neighbor’s son Leon’s books’ pages’

a. Classical

<i>Թերթք</i>	<i>գրոց</i>	<i>Լեւոնի</i>	<i>որդույ</i>	<i>դրացույ</i>	<i>իմոյ</i>
t ^h ert ^h -k ^h	gr-ɔts ^h	Lewɔn-i	ɔrdw-ɔj	drats ^{hw} -ɔj	im-ɔj
page-PL	book-GEN.PL	Leon-GEN	son-GEN	neighbor-GEN	my-GEN

b. MWA

դրացիիս t ^h ərats ^h i-i-s neighbor-GEN-1.POSS	տղուն dəɣ-u-n son-GEN-DEF	Լեոնին ləvɔn-i-n Leon-GEN-DEF	գիրքերուն k ^h irk ^h -ɛr-u-n book-PL-GEN-DEF
Թերթերը t ^h ɛrt ^h -ɛr-ə page-PL-DEF			

(3) relative clause: ‘I saw the bird that was singing in the tree’

a. Classical

տեսի tes-i see-AOR.1.SG	զթռչունն z-t ^h rtʃ ^h un-n SPEC-bird-DEF	որ երգէր ɔr ɛrg-ɛ:r REL sing-3.SG.IMF	ի վերայ ծառոյն i vɛraj tsar-ɔj-n in on tree-GEN-DEF
--------------------------------------	--	--	--

b. MWA

ծառին dzar-i-n tree-GEN-DEF	վրայ երգող vəɾɑ jɛrk ^h -ɔɣ on sing-SUBJ.PPL	թռչունը t ^h ərtʃ ^h un-ə bird-DEF	տեսայ des-ɑ see-AOR.1.SG
--	---	---	---------------------------------------

As the facts in (2) and (3) demonstrate, a significant syntactic realignment occurred at some point between the Middle and Modern Armenian periods (according to Karst 1901: 407, Middle Armenian preserved “pure [Classical] Armenian syntax”). (One should not infer from this that all of the changes that now distinguish Modern Armenian from Classical Armenian happened after the Middle Armenian period. In fact, many of the characteristics of Modern Armenian first appear in Middle Armenian, such as periphrastic verb formations [e.g. Middle Armenian and MWA *gu dam* ‘I give’, *bidi dam* ‘I will/must give’ vs. Classical *tam*, *tac^h*, respectively] and the Modern *-[n]er* plural morpheme [cf. Classical *-k^h*].) This realignment is traditionally linked to the significant influence of Turkish in the Armenian-speaking world following the invasion of Asia Minor by various Turkic tribes beginning in the eleventh century. In fact, it is often observed anecdotally that Modern Armenian is simply Armenian phonology and morphology with Turkish syntax (cf. Pedersen 1906: 472; Adjarian 1909: 8). The syntactic similarities between Modern Armenian and Turkish can be seen by comparing the Modern Armenian structures in (2b) and (3b) to their Turkish equivalents in (4a) and (4b) respectively; the primary difference is in the relative order of the genitive and possessive markers – GEN-POSS in Armenian but POSS-GEN in Turkish.

(4) Comparison of Modern Armenian and Turkish noun phrases and relative clauses

a. ‘my neighbor’s son Leon’s books’ pages’

t ^h ərats ^h i-i-s neighbor-GEN-1.POSS	dəɣ-u-n son-GEN-DEF	ləvɔn-i-n Leon-GEN-DEF	k ^h irk ^h -ɛr-u-n book-PL-GEN-DEF
komşu-m-un neighbor-1.POSS-GEN	oğl-u son-GEN	Levon-ın Leon-GEN	kitab-lar-i-nın book-PL-3-GEN
t ^h ɛrt ^h -ɛr-ə page-PL-DEF			
yaprak-lar-ı page-PL-3			

- b. 'I saw the bird that was singing in the tree'

dzar-i-n	vəra	jɛrk ^h -əɤ	t ^h ərtʃ ^h un-ə	dəs-a
tree-GEN-DEF	on	sing-SUBJ.PPL	bird-DEF	see-AOR.1.SG
ağac-ın	üstün-de	öt-en	kuş-u	gör-dü-m
tree-GEN	on-LOC	sing-SUBJ.PPL	bird-DEF.ACC	see-AOR-1.SG

Other morphosyntactic innovations in Modern Armenian that have been attributed to Turkish influence include the development of agglutinative nominal morphology (discussed above, 4), the appearance of a fixed position for nonspecific objects (immediately before the verb; cf. Comrie 1984), the creation of a special construction for yes-no questions, the development of periphrastic passive formations with ablative agents, and the declension of adpositions (Acharyan 1952: 198, section 7.5).

Armenian has had extensive contact with other languages as well; in fact most speakers of Armenian are bilingual, typically also speaking one or more of the languages Russian, Turkish, Georgian, Arabic, English, and French. However, these languages do not (with the exception of Turkish) appear to have had a significant influence on Armenian syntax. A notable exception is Persian, which has noticeably influenced the syntax of several Iranian Armenian varieties. Acharyan (1911: 284) noted that the Maragha dialect appears to have borrowed from Persian the ability to attach direct object clitics to verbs; in Maragha these clitics are homophonous with the possessive clitics, as in (5).

- (5) direct object clitics in Maragha (Acharyan 1911: 284)

- a. մէ ձի պռնէնքս
 mə dzi pərn-ε-nk^h-ət
 a horse take-THEME.V-1PL-2SG.CLITIC
 'let's take a horse for you'

SEA equivalent:

մի ձի բռնենք	քեզ	համար
mi dzi bərn-ε-nk ^h	k ^h εz	hamar
a horse take-THEME.V-1PL	you.SG.ACC	for

- b. փստտէս էրդ
 p^hətət-ε-s εɪ-əd
 seek-THEME.V-PPL PAST-2SG.CLITIC
 'he/she/it was looking for you'

SEA equivalent:

քեզ	կը փստտէր
k ^h εz	kə p ^h əntɾ-εɾ
you.SG.ACC	IMF seek-3SG.IMF

The same construction is found in Teheran Armenian, as in (6) (from Karine Megerdoo-mian (p.c.); cf. also Muradyan et al. 1977, feature 675).

- (6) կխփեմս
 kə-xəp^h-ε-m-ət
 FUT-hit-THEME.V-1SG-2SG.CLITIC
 'I will hit you'

Syntactic influence from Persian arguably also surfaces in the formation of relative clauses with a resumptive pronoun in Teheran Armenian, as with *իրանից* *ւան-its^h* in (7) (data from Karine Megerdooian, p.c.).

- (7) *էն կինը որ ես իրանից էս գիրքը վերցրած*
 en kin-ə vɔɪ jɛs ʷan-its^h ɛs gɪrk^h-ə vɛɪts^həɪ-ɑ-m
 that woman-DEF which I her-ABL this book-DEF buy-THEME.V-1SG
 ‘the woman I bought this book from’

SEA equivalent:

- այն կինը վորից այս գիրքը վերցրեցի*
 aɟn kin-ə vɔɪ-its^h aɟs gɪrk^h-ə vɛɪts^həɪ-ɛ-ts^h-i
 that woman-DEF which-ABL this book-DEF buy-THEME-AOR-1SG

One syntactic similarity between Armenian dialects (past and present) and European languages such as French, German, and Dutch is in the formation of the perfect using an auxiliary ‘have’ or ‘be’. In Benveniste’s (1952) analysis, he draws an analogy between the so-called “transitive perfect” – which he describes as being used with transitive verbs – and a possessive construction:

- (8) “transitive perfect”
նորա է գործեալ
 nora ē gorc-eal
 3SG.GEN be.3SG do-PPL
 ‘(s)he/it has done/accomplished’

- (9) possession
նորա է հանդերձ
 nora ē handerj
 3SG.GEN be.3SG garment
 ‘(s)he has a garment’

In both constructions, we have a copula whose “subject” is in the genitive. This is supported by the fact that the lexical verb *unim* ‘have’ is sometimes used to express the perfect in Middle Armenian, as in *գբաղաքն առած ունէին շէտակն արաւն* ‘they had taken the city’, *ի ներքսէ պահած ունի i nerk’sē pahac uni* ‘he has kept [it] inside’ (Aytēnean 1866.2: 96–97).

This can also be seen in the development of some modern dialects. In the Hamshen subdialect of Köprücü, for example, we see by comparing the forms in (10a) with those in (10b) that the perfect is expressed using a marker *-ui*.

- (10) a. *dzidzaɣ-adz a* ‘he laughed’
 mɛɪ-adz a ‘he died’
 ɛg-adz a ‘he came’
 b. *dzidzaɣ-adz ui* ‘I (have) laughed’
 **mɛɪ-adz ui* ‘I (have) died’
 **ɛg-adz ui* ‘I (have) come’

In the second person singular we have *-ues*, in the third person singular *-ua*. Given its use in the past tense, deriving from the original perfect, we can take this *-ui*, *-ues*, *-ua* marker to come from a grammaticalization of the original lexical verb *unim*.

Benveniste analyzes the Classical Armenian data in (8) as part of a split ergative system, in which transitives behave differently from intransitives in taking this kind of perfect construction; split ergativity is common to other languages of the region. Indeed, the similarity between this system and that of western European languages had already been noticed by earlier Armenian writers; Aytĕnean (1866, 2: 96–97) commented: “The modern European languages seem to parallel our 12th century language, using *have* with transitive verbs and *be* with intransitives and middles.”

In fact, after Benveniste’s time (Burzio 1986), it was realized that the division in the western European perfect is not transitive-intransitive, but unaccusative-non-unaccusative. Unaccusative verbs (passives, raising verbs [*seem*, *appear*], *come*, *go*, *arrive*, *fall* ...) take only an internal argument, prototypically when this argument is a patient or undergoes a change of state. The rest of intransitive verbs, unergative verbs (*work*, *sing*, *dance*, *cough*, *laugh* ...) take an external argument, which prototypically plays an agentive thematic role.

6. Lexicon

As well as changes to the lexicon in the form of loanwords, there have been various idiosyncratic semantic changes to individual native Armenian words. We quickly outline two examples here to give a flavor of historical Armenian-internal lexicology.

- The word Հաւ *haw* is often glossed for Classical Armenian as ‘bird’, which is the reconstructed meaning of its PIE etymon *h₂éwis (Martirosyan 2010), cf. Latin *avis*, Greek αἰετός ‘eagle’. Strohmeier (1983) concludes from a philological investigation that the word has a somewhat narrower meaning in Armenian than its etymon, primarily referring to “birds which are useful to men” (Martirosyan 2010). In most modern dialects (with the exception of Van), the word’s range has narrowed further to mean only ‘chicken’; the frozen plural *haw-k^h* has come to mean the more generic ‘bird’.
- The Classical Armenian verb քունեմ *k’unem* ‘sleep’ has two reflexes in the modern dialects. The original sense of ‘sleep’ is continued in the irregularly altered form *k^hənel*, perhaps generalized from the oblique stem *k’n-* of the associated noun *k’un* ‘sleep’. The regular outcome *k^hunel* has come to mean ‘futuere’ throughout the informal registers of the modern dialects. Petrosyan (2007) suggests that this meaning was influenced by a reflex of PIE *keh₂- ‘love’ (cf. Sanskrit *kā-* ‘desire’, English *whore*), but Martirosyan (2010) contests that the two roots are formally too distant, and argues that the shift ‘sleep’ > ‘futuere’ is semantically plausible without interference from other lexemes.

7. Texts

To illustrate the evolution of Armenian in the historical period, we give the text of the Lord’s Prayer in several varieties of Armenian dialects.

- (11) Classical Armenian (Ējmiacin ms. 229, 989 A.D.)

Հայր մեր որ յերկինս, սուրբ եղիցի անուն քո, եկեցէ արքայութիւն քո, եղիցին կամք քո որպէս յերկինս եւ յերկրի. զհաց մեր հանապազորդ տուր մեզ այսաւր, եւ թող մեզ զպարտիս մեր, որպէս եւ մեք թողումք մերոց պարտապանաց, եւ մի տանիր զմեզ ի փորձութիւն, այլ փրկեա զմեզ ի չարէ. զի քո է արքայութիւն եւ զաւրութիւն եւ փառք յաւիտեանս ամէն.

hayr mer or yerkins. surb elic'i anun k'o. ekec'ē ark'ayut'iwñ k'o. elic'in kamk' k'o orpēs yerkins ew yerki. zhac' mer hanapazord tur mez aysawr. ew to'ł mez zpartis mer. orpēs ew mek' t'olumk' meroc' partapanac'. ew mi tanir zmez i p'ordzut'iwñ. ayl p'rkea zmez i č'arē. zi k'o ē ark'ayut'iwñ ew zawrut'iwñ ew p'ark' yawiteans amēñ.

- (12) Standard Eastern Armenian

Հայր մեր, վոր յերկնքում ես, սուրբ թող լինի քո անունը. քո թագավորությունը թող գա. քո կամքը թող լինի յերկրի վրա, ինչպես վոր յերկնքում է. մեր հանապազորյա հացը տուր մեզ այսօր, յեվ թող մեզ մեր պարտքերը ինչպես յեվ մենք ենք թողնում մեր պարտականներին. յեվ մի տար մեզ փորձության, այլ փրկիր մեզ չարից. վորովհետեւ քօնն է թագավորությունը յեվ գորությունը յեվ փառքը հավիտյանս. ամեն.

hajr mer, vor jerkənk^hum es. surp^h t^həw lini k^hə anunə. k^hə t^hagavərut^hjunə t^həw ga. k^hə kamk^hə t^həw lini jerki vərə, intf^hpes vor jerkənk^hum e. mer hanapazordja hats^hə tur mez ajsor. jev t^həw mez mer partk^herə intf^hpes jev menk^h enk^h t^həwnum mer partakaner. jev mi tar mez p^hordzut^hjan, ajl p^harkir mez t^harits^h. vorəvhetev k^həw: e t^hagavərut^hjunə jev zərut^hjunə jev p^hark^hə havitjanəs. amēn.

- (13) Standard Western Armenian

Ով հայր մեր որ երկինքն ես, քու անունդ սուրբ ըլլայ. քու թագաւորութիւնդ գայ. քու կամքդ ըլլայ ինչպէս երկինքը՝ նոյնպէս երկրի վրայ. մեր ամէնօրուան հացը այսօր ալ մեզի տուր, մեզի ներէ մեր պարտքերը ինչպէս մենք ալ կը ներենք մեր պարտականներուն. ու մեզ փորձութեան մի տանիր, հապա չարէն մեզ ազատէ. քանզի քուկդ է թագաւորութիւնը եւ գորութիւնը ու փառքը յաւիտեանս: Ամէն.

ov hajr mer vor jergink^hn es, k^hu anunəth surp^h əl:a. k^hu t^hak^havərut^hjunəth k^ha. k^hu gamk^həth əl:a intf^hbes jergink^hə nujnbes jergri vərə. mer amenərvan hats^hə ajsor al mezi dur, mezi nerē mer bardk^herə intf^hbes menk^h al gə nerenk^h mer bardagan:erun. u mez p^harts^hut^hjan mi danir, haba t^harēn mez azade. k^hanzi k^hugəth e t^hak^havərut^hjunə jev zərut^hjunə u p^hark^hə havidjanəs: amēn.

- (14) Zok (Vaux 2008)

Անուն հօր եւ օրթօ եւ օքույն սրբօ ամմէն: Միր ափի օր երկնքումնըս, սուրբ նէան քու անունու. քու թագավորությունը միր վէարին մնօ միչտ, համման ըզհհհհաճըթ կատարվի, ուրտի երկնքումն, անտի էալ երկրի վէարին: Միր օրվա հօըը հըսանի միզ, եւ թուդ միչտ միզ պարտական, օր միք էալ պօրտք չըմընօնք միրուց, եւ միչտ միզ հէառի պէհհի չարից, մէանէակ փրկիս չարությունից. քանի օր ըչսարքամըս թագավորությունու, ուժու, նէան պատիվը, քունն ա միչտ: Ամմէն:

ánun hór jev ərthə jev ək^hújñ sərp^hə am:en. mir ap^hi ər jerkənk^húmnəs, surp^h næn k^hu ánunu. k^hu t^hak^havərut^hjunə mir vərín mənə mijt, ham:an əzəhətsət^h

katarvi, úrti jærkænk^húmni, ánti æl jærkrí værín: mir ɔrva hɔts^hə həsáni miz, jev t^huæ miʃt miz pærtakan, ɔr mik^h æl pɔrtk^h tʃ^həmənɔnk^h miruts^h, jev miʃt miz hæri pæhis tʃ^harits^h, mænæk p^hærkis tʃ^harut^hjúnits^h. k^hani ɔr əʃxárk^haməs t^hak^havɔrut^hjúnu, úzu, næn pátivə, k^hunn a miʃt: am:ən.

- (15) Marash (<http://hyeforum.com/index.php?showtopic=15940>)

Միր պոպը քի իրկինքն իս Քինիստ ատունիդ սուրբ թօննօ Քինիդ թէգէվիւրիւթիւնդ թող գօ Քինիստ կումքիդ թօննօ չոց քի իրկինքի ինտէն է կիտենան իվրէն Միր ամէնավիւր հոցը էօտօր միդ սուր հէմ միր պորտէքը միդ պաղըլէ չոց քի մինք է միր պարտականոցը կը պաղըլինք: Հէմ միդ փորձըթան մի տանա հապր չորէն միդ փրկէ: Ինչու քի Քինիստ է թէգէվիւրիւթիւնը գորիւթիւնը փառքը յաւիտեանս յաւիտենից Ամէն:

mir bəbə k^hi irgink^hn is k^hinid adunit surp t^hɔn:ɔ k^hinid t^hɛkɛvyryt^hynət t^hɔk kɔ k^hinid gɔmk^hit t^hɔn:ɔ tʃ^hɔts^h k^hi irgink^hi indɛn ɛ gidɛnan ivrɛn mir amɛnavyr hɔts^hə ɛsɛər miz dur hɛm mir bɔrdɛk^hə miz bəkəʃɛ tʃ^hɔts^h k^hi mink^h ɛ mir bɑrdɑɡɑnɔts^hə gə bəkəʃɛnk^h: hɛm miz p^hɔrtsət^hɑn mi dɑnɑ hɑbɔ tʃ^hɔrɛn miz p^hərgɛ: intʃ^hu k^hi k^hinid ɛ t^hɛkɛvyryt^hynə zɔryt^hynə p^hɑrk^hə hɑvidjɑnəs hɑvidɛnits^h ɑmɛn.

- (16) Zeytun (<http://hyeforum.com/index.php?showtopic=15940>)

Ով մեյ պօպը ոյ իյկինքն իս Քու անունդ սուր թողնա Քու թէգէվիւրիւթիւնդ թող գօ Քու կօմքդ թող լա ինչպէս իյկինքը ինտեն էլ իյկեյին վիս Միյ ամէնէօյվէն հօցը եսէօյ միդ սուր: Եւ միդ նեյէ միժ պօյտքը չոց որ մինք էլ կը նեյինք միյ պօյտքի տէյեյուն Եւ միդ փոյձութան մի տանեյ հապր չօյէն միդ ազատէ ոյեհետեւ Քինն է թէգէվիւրիւթիւնը Եւ գոյութիւնը ու փառքը Յաւիտեանս յաւիտենից Ամէն:

ɔv mej bəbə ɔj iʃgink^hn is k^hu ɑnɔnət sujɔ t^hɔkɑnɑ. k^hu t^hɛkɛvyryt^hynət t^huæ kɔ. k^hu gɔmk^hət t^huæ lɑ, intʃ^hbɛs iʃgink^hə indɛn ɛl ijgɛjin vijɔ: mij amɛnɛɛjvɛn hɔts^hə ɛsɛj miz duj: yɛv miz nɛjɛ mits bɔjɔdk^hə, tʃ^hɔts^h ɔr mink^h ɛl gə nɛjɛnk^h mij bɔjɔdk^hi dɛjɛjun. jɛv miz p^hɔjtsut^hɑn mi dɑnɛj, hɑbɔ tʃ^hɔjɛn miz ɑzɑdɛ. vɔjɛvhɛdɛv k^hinn ɛ t^hɛkɛvyryt^hynə jɛv zɔjɔt^hynə u p^hɑrk^hə. hɑvidjɑnəs hɑvidɛnits^h ɑmɛn:

- (17) Kesab (Adjarian 1911)

էօվ մեր բիւբը, սուրփի էդնի քէ անուն, քէ թէքէվիւրութիւնը թող գ'օ. քէ իրադէթդ ըննօ, չիւց ըր քի իրգայանքը, թըրգէն էլ ի գ'էդինը. մեր ամէնէվիւր հօցըը դուր մեզ էս էվիւր էլ, մեր բօրդքը մեզի բաղըլլամուշ ըրօ, չիւց ըր քի մենք գինօնք մերօցը, վէ զըզմեզ փօրցիւթյան մի դանօ. հաբօ խալլսօ ի չարէն, չիւնքի քէ է թէքէվիւրութիւնը, շէրէֆը, դուվէթը, հավիդիյինս հավիդօնից ամօն:

ɛv mɛr bybə, surp^h ɛbn k^hɛ ɛnɔn, k^hɛ t^hɛk^hɛvyryt^hynə t^hɔk g^hɔ. k^hɛ iradɛt^hɔd ɔn:ɔ, tʃ^hyts^h ɔr k^hi irgjank^hə, t^hərzɛn ɛl i g^hɛdinə. mɛr amɛnevyr hɔts^hə dur mɛz ɛs ɛvyr ɛl, mɛr bɔrdk^hə mɛzi bəkəʃlɑmuʃ ɔrɔ, tʃ^hyts^h ɔr k^hi mɛnk^h ɡinɔnk^h mɛrɔnts^hə, vɛ zɔzɛmɛz p^hɔrts^hiut^hjɑn mi dɑnɔ. hɑbɔ xɑlɔsɔ i tʃ^hɑrɛn, tʃ^hynk^hi k^hɛ ɛ t^hɛk^hɛvyryt^hynə, ʃɛrɛfə, kɔvɛt^hə, hɑvidijins hɑvidɔnits^h ɑmɔn.

8. References

- Adjarian, Hrachea [= Acharyan, Hrach'ya]
 1889 Les explosives de l'ancien arménien étudiées dans les dialectes modernes. *La Parole ou Revue internationale de Rhinologie, Otologie, Laryngologie et Phonétique expérimentale* 1: 119–127.
- Acharyan, Hrach'ya
 1901 Barbarakhōsut'ean ōgutnerē ew matenakhōsakan 'Ewdokioy gawarabarbar'-in [Useful spoken and literary expressions of the regional dialect of Eudokia]. (Review of Gazanchean 1899.) *Banasēr* (Paris) 3: 76–82.
- Acharyan, Hrach'ya
 1906 *Patmut'awn Hayots' Nor Grakanut'ean* [History of Modern Armenian Literature]. Vaghharshapat. [Reprinted 1939. Beirut: Yoys Tparan.]
- Adjarian, Hrachea [= Acharyan, Hrach'ya]
 1909 *Classification des dialectes arméniens*. Paris: Champion.
- Acharyan, Hrach'ya
 1911 *Hay barbaragitut'awn* [Armenian Dialectology]. Ēminean Azgagrakan Zhoghovatsu 8. Moscow: Lazarean Chemaran.
- Acharyan, Hrach'ya
 1941 *K'nnut'yun polysahay barbari* [A study of the Istanbul dialect]. *Gitakan ashkhatut'yunner* [Scientific works]. Vol. 19. Erevan: Petakan Hamalsarani Hratarakch'ut'yun, 19–250.
- Acharyan, Hrach'ya
 1952 *K'nnut'yunyun Vani barbari* [A study of the Van dialect]. Erevan: Erevani Petakan Hamalsarani Hratarakch'ut'yun.
- Acharyan, Hrach'ya
 1973 *Hayeren armatakan bararan* [Armenian Etymological Dictionary]. Hator 2: E–K. Erevan: Erevani Hamalsarani Hratarakch'ut'yun.
- Asatryan, Manvel
 1980 Arevelahay grakan lezvi ev araratyan barbari dzevabanakan hamakargeri p'okkharaberut'yunnerē [The relationships between the morphological systems of the Eastern Armenian literary language and the Ararat dialect]. *Banber Erevani Hamalsarani* 3 (42): 81–95.
- Awetik'ean, Gabriēl, Khach'atur Siwrmēlean, and Mkrtich' Awgerean
 1836 *Nor bargirk' haykazean lezui* [New dictionary of Armenian]. Venice: Tparan i Srboyn Ghazaru.
- Aytēnean, Arsēn
 1866 *K'nnakan k'erakanut'awn ashkharhabar kam ardi hayerēn lezui* [A critical grammar of the colloquial or modern Armenian language]. Vienna: Mkhit'areants' Tparan.
- Aytēnean, Arsēn
 1883 *K'nnakan k'erakanut'awn ashkharhabar kam ardi hayerēn lezui* [A critical grammar of the colloquial or modern Armenian language], 2nd edn. Vienna: Mkhit'areants' Tparan.
- Bardakjian, Kevork and Robert Thomson
 1977 *A textbook of Modern Western Armenian*. Delmar, NY: Caravan.
- Beekes, Robert S. P.
 2011 *Comparative Indo-European linguistics: An introduction*. Amsterdam: Benjamins.
- Benveniste, Émile
 1952 La construction passive du parfait transitif. *Bulletin de la Société de Linguistique de Paris* 48: 52–62.
- Braun, Angelika
 2013 An early case of “VOT”. *Interspeech* 14: 119–122.
- Burzio, Luigi
 1986 *Italian syntax: A government-binding approach*. Dordrecht: Reidel.

- Comrie, Bernard
1984 Some formal properties of focus in Modern Eastern Armenian. *Annual of Armenian Linguistics* 5: 1–21.
- Danielyan, T'ewan
1967 *Malat'iayi barbarë* [The dialect of Malatya]. Erevan: Haykakan SSH GA Hratarakch'ut'yun.
- Donabédian, Anaïd
1996 Perfect and Mediative in Modern Western Armenian. In: Dora Sakayan (ed.), *Proceedings of the Fifth International Conference on Armenian linguistics: McGill University Montreal, Quebec, Canada May 1–5, 1995*. Delmar, NY: Caravan, 149–166.
- Donabédian, Anaïd
2001 Towards a semasiological account of evidentials: An enunciative approach of *-er* in Modern Western Armenian. *Journal of Pragmatics* 33: 421–442.
- Dum-Tragut, Jasmine
2009 *Armenian: Modern Eastern Armenian*. Amsterdam: Benjamins.
- Fortson, Benjamin W. IV.
2004 *Indo-European language and culture: an introduction*. Oxford: Blackwell.
- Garrett, Andrew
1998 Adjarian's law, the glottalic theory, and the position of Armenian. In: Benjamin K. Bergen, Madelaine C. Plauché, and Ashlee C. Bailey (eds.), *Proceedings of the 24th annual meeting of the Berkeley Linguistics Society (February 14–16, 1998): Special session on Indo-European subgrouping and internal relations (February 14, 1998)*. Berkeley: Berkeley Linguistics Society, 12–23.
- Gharibyan, Ararat
1948 *Hay barbaragitut'yun* [Armenian dialectology]. Erevan: Haykakan SSR Petakan Heraka Mankavarzhakan Institut.
- Gyumchean, Mesrop
1970 *Ēndardzak baṙaran hayerēnē anglierēn* [Large Armenian-English dictionary]. Beirut: Tparan Atlas.
- HAB
1973 See Acharyan.
- Halle, Morris and Bert Vaux
1998 Theoretical Aspects of Indo-European Nominal Morphology: The Nominal Declensions of Latin and Armenian. In: Jay Jasanoff, H. Craig Melchert, and Lisi Oliver (eds.), *Mír Curad: Studies in Honor of Calvert Watkins*. Innsbruck: Institut für Sprachwissenschaft der Universität, 223–240.
- Joseph, Brian
1983 *The synchrony and diachrony of the Balkan infinitive*. Cambridge: Cambridge University Press.
- Karst, Josef
1901 *Historische grammatik des kilikisch-armenischen*. Strassburg: Trübner.
- Kortlandt, Frederik
1985 Proto-Indo-European glottalic stops: The comparative evidence. *Folia linguistica historica* 6: 183–202.
- Kostandnupōlsets'i, Yovhannēs
1674 *Pvritas lingvae armenicæ = Ztowtiwn haykabanowtēan, kam kërakanowtiwn haykakan* [Purity of the Armenian language or Armenian grammar]. Rome: Ex Typographia Sacrae Congreg. de Propaganda Fide.
- Lassiter, Daniel
2016 The Dialectal Position of Civil Armenian as Recorded in Schröder's *Thesaurus Linguae Armenicæ*. Manuscript, Stanford University.

- Lusents', Ashot
1982 *Areshi Barbarē* [The dialect of Aresh]. Erevan: Erevani Petakan Hamalsarani Hratarakch'ut'yun.
- Mackenzie, D. Neil
1971 *A concise Pahlavi dictionary*. London: Oxford University Press.
- Martirosyan, Hrach
2010 *Etymological dictionary of the Armenian inherited lexicon*. Leiden: Brill.
- Meillet, Antoine
1936 *Esquisse d'une grammaire comparée de l'arménien classique*. 2nd edn. Vienna: Mekhitaristes.
- Muradyan, Hovhannes, Dariko Kostandyan, Anahit Haneyan, Mat'evos Muradyan, and Aharon Grigoryan
1977 *Hayereni barbaragitakan atlası nyut'eri havak'man dzragir* [Program for the collection of materials for an Armenian dialectological atlas]. Erevan: Haykakan SSH Gitut'yunneri Akademiayi hratarakch'ut'yun.
- Nichanian, Marc
1989 *Agēs et usages de la langue arménienne*. Paris: Entente.
- Olsen, Birgit Anette
1999 *The noun in biblical Armenian: with Special Emphasis on the Indo-European Heritage*. Berlin: Mouton de Gruyter.
- Parnassian, Nevard
1985 On the Formation of Ashkharhabar. *Annual of Armenian Linguistics* 6: 67–73.
- Pedersen, Holger
1906 Armenisch und die Nachbarsprachen. *Zeitschrift für vergleichende Sprachforschung* 39: 334–484.
- Petrosyan, Armen
2007 Stugabanut'iwinner [Etymologies]. *Handēs Amsōreay* 121: 1–24.
- Pisowicz, Andrzej
1976 *Le développement du consonantisme arménien*. (Polska Akademia Nauk. Oddział w Krakowie. Prace Komisji Językoznawstwa 43). Warsaw: Polska akademia nauk.
- Reader, John
2008 *The untold history of the potato*. London: Vintage Books.
- Russell, James
1999 Alphabets. In: Glen Warren Bowersock, Peter Robert Lamont Brown, and Oleg Grabar, (eds.), *Late Antiquity: A Guide to the Postclassical World*. Cambridge, MA: Harvard University Press, 289.
- Sandfeld, Kristian
1930 *Linguistique balkanique: Problèmes et résultats*. Paris: Klincksieck.
- Sargsyan, Artem
1991 *Arewmtahayereni bararan* [Western Armenian dictionary]. Erevan: Arewik.
- Schröder, Johann Joachim
1711 *Thesaurus Linguae Armenicae, antiquae et hodiernae ...* [A thesaurus of the Armenian language, old and modern ...]. Amsterdam: Amstelodanum.
- Strohmeyer, Virgil B.
1983 The meaning of *haw* and *t'řč'un* in Classical Armenian. *Revue des Études Arméniennes* 13: 43–48.
- Ter-Avetisyan, Smbat
1938 *Zak'aria Agulets'u Ōragrut'iwñē* [The diary of Zak'aria Agulets']. SSSH GA Haykakan Filiali Hratarakut'yun.
- Thomson, Robert
1989 [1975] *An Introduction to Classical Armenian*. 2nd edn. Delmar, NY: Caravan.

Vaux, Bert

- 1992 Adjarian's Law and consonantal ATR in Armenian. In: John A. C. Greppin (ed.), *Proceedings of the Fourth International Conference of Armenian Linguistics*. Delmar, NY: Caravan, 271–293.

Vaux, Bert

- 1995 A problem in diachronic Armenian verbal morphology. In: Jos Weitenberg (ed.), *New approaches to the medieval Armenian language and literature*. Amsterdam: Rodopi, 135–148.

Vaux, Bert

- 1998 *The phonology of Armenian*. Oxford: Clarendon Press.

Vaux, Bert

- 2008 Zok: The Armenian dialect of Agulis. In: Barlow Der Muğrdichian (ed.), *Between Paris and Fresno. Armenian Studies in Honor of Dickran Kouymjian*. (Armenian Studies Series 10). Costa Mesa, CA: Mazda Press, 283–301.

Vaux, Bert

- 2012 The Armenian dialect of Khodorjur. In: Aram Arkun and Victoria Rowe (eds.), *Khodorchur: Lost Paradise. Memories of a land and its people*. Monterey, CA: Mayreni Publishing, 555–572.

Weitenberg, Jos J. S.

- 1993 The use of the classical Armenian subjunctive forms in the fifth century and in middle Armenian Grabar texts: Koriwn and Grigor Aknerc'i. In: Suren Simonian and Jos Weitenberg (eds.), *Computers in Armenian Philology*. Erevan: Armenian Academy Press, 72–100.

Zekiyani, Boghos Levon

- 1997 Modern Armenian culture: some basic trends between continuity and change, specificity and universality. In: Nicholas Awde (ed.), *Armenian Perspectives. 10th Anniversary Conference of the Association Internationale des Études Arméniennes. School of Oriental and African Studies. London*. Richmond, Surrey: Curzon, 323–354.

Zgusta, Ladislav

- 1971 *Manual of Lexicography*. The Hague: Mouton.

Ollie Sayeed and Bert Vaux, Cambridge (UK)

XI. Celtic

67. The documentation of Celtic

1. Historical aspects: Ancient tradition
2. Historical aspects: Archaeology
3. The term “Celtic”
4. The genealogy of Celtic languages
5. The documentation of Celtiberian and other Hispano-Celtic languages
6. The documentation of Gaulish and other Continental *p*-Celtic languages
7. The documentation of Irish
8. The documentation of British Celtic
9. Celtic languages today
10. References

1. Historical aspects: Ancient tradition

The earliest reference to Celtic people may be ascribed to the end of the 6th century BCE in connection with the Greek settlement in *Massalia* (Hekataios of Milet, only indirectly attested in the *Ora maritima* of Festus Rufus Avienus in the 4th century CE, *GLQFM* I: 44 f.; Tomaschitz 2002: 5 f.). In the following centuries the Celts and their customs were described in the works of Herodotus (5th c. BCE), Aristotle (4th c. BCE), extensively in Polybios (2nd c. BCE), Poseidonius (1st c. BCE) and Caesar’s *De bello Gallico* (1st c. BCE). For further information and details see *GLQFM*; Tomaschitz (2002); Tierney (1959/1960); Dobesch (1991); and Freeman (1996).

2. Historical aspects: Archaeology

Archaeologists associate the material dating from Hallstatt Culture (HaC-HaD, ca. 800–475 BCE) and La Tène Culture (LtA-LtD, ca. 475–25 BCE) with Celtic tribes, beginning in Central Europe and then spreading during the La Tène Culture to Spain, France, Northern Italy, and to southern parts of Eastern Europe (Trachsel 2004; Dillon and Chadwick 1967). Accordingly, the “homeland” of Common Celtic speaking tribes is often located in an area which includes Eastern France, Northern Switzerland, and South West Germany. This view was confirmed by a hydronymic study by Busse (2007). But recently some scientists have proposed to seek the “homeland” of Common Celtic far more in the west, namely in the cultures of the Atlantic Bronze Age (Cunliffe and Koch 2010; Koch and Cunliffe 2013).

The origin of the Celtic speaking inhabitants of the Iberian peninsula is still a matter of debate. The invasion theory proposed by Bosch Gimpera (cf. e.g. Bosch Gimpera 1940) is rejected by most researchers today. Though connections in the archaeological record are visible between Spain and Central Europe during the Iron Age (see Lenerz-De Wilde 1991; Neumaier 1995; Stary 1994) there is hardly any hint of large-scale migrations of “Hallstatt” – or “La Tène” – people to Iberia. One could instead assume a sort of “cumulative celticity” as was proposed for the British Isles by Christopher

Hawkes (1973). Quite possible seems a Late Bronze Age Origin of the Iberian Celts, too (a recent overview of this topic can be found in Lorrio and Ruiz Zapatero 2005).

3. The term “Celtic”

The word “Celtic” and its derivational base “Celt” are derived from Greek *Κελτοί* and Latin *Celtae*, first used by ancient Greek and Latin authors to denote a group of peoples and tribes who – according to the point-of-view of most ancient authors – had ethnic, religious, cultural, and linguistic features in common. The region inhabited by these Celtic people was the central part of Europe (so e.g. Ephoros; cf. *GLQFM* I: 50 f.). Greek authors called this region *Κελτική*, Latin authors (e.g. Plinius Secundus, *Nat.Hist.* 4,111,7) *Celticum*. Besides the term *Celtae* Latin authors used the word *Galli*. The words *Κελτοί*, *Celtae* and *Galli* must once have been names of some neighboring Celtic tribes known to the Greeks and Romans, who took the names to refer to “the Celts” in general (see especially Birkhan 1997: 55 on this point; compare French *les Allemands* used to denote all Germans). They are etymologically quite clear, although their derivational bases or verbal roots, respectively, allow different semantic interpretations (Birkhan 1997: 47 f.; Ziegler 1996). Today the linguistic term “Celtic” is assigned to a group of languages showing common features, which could be termed “*Leitformen*”. This designation is borrowed from archaeology, where it is used for the different artefacts which are typical of an archaeological culture and equally for artefacts which are typical of a certain stage in the relative chronology of an archaeological culture, e.g. Bell Beakers, which are typical of the eponymous archaeological culture that covered large parts of Western Europe in the 3rd millennium BCE; but “*Leitformen*” could also be used for the fibulae of the Münsingen-type or the Duchcov-type which are characteristic of a horizon within the relative chronological stage Lt B (cf. for the horizon Duchcov-Münsingen Kruta 1979; for the Bell Beaker Mallory 1997; for the term “*Leitformen*” and its implication for and implementation in linguistics, Vath *apud* Ziegler 2012: 322 f.).

4. The genealogy of the Celtic languages

4.1. The genealogy of the Celtic languages is still a matter of debate. There are at least two important classifications still in use: the subdivision into *p*- and *q*-Celtic languages and the subdivision into Continental Celtic and Insular Celtic. The former classification uses the treatment of Proto-Celtic (and Indo-European) **/k^w/* yielding **/p/* in the *p*-Celtic languages and remaining */k^w/* in the *q*-Celtic languages (in the first step **/k^w/* undergoes a change to */k/* in Early Old Irish). While Goidelic and Hispano-Celtic belong to the *q*-Celtic languages, Brythonic, Gaulish, Lepontic, and Galatian are languages of *p*-Celtic character. The other classification is a primarily geographical matter with Goidelic and Brythonic belonging to the Insular Celtic branch, and Hispano-Celtic, Gaulish, Lepontic, and Galatian belonging to the Continental Celtic branch. Nevertheless, both Insular Celtic languages show some common new features probably due to an underlying substratum, to language contact (Russell 1995: 17 f.), or, according to Schrijver (1995: 465), to an “Insular Celtic linguistic unity”. A genealogical model using the subdivision

into *p*- and *q*-Celtic was developed by Karl-Horst Schmidt on the basis of the development of PIE */kʷ/ and the sonantic variations of */m/ and */n/ (cf. e.g. Schmidt 1986). It was criticized by Kim McCone and others who see a chronological-chorological correspondence for morphological-syntactical reasons (cf. for this model e.g. McCone 1996: 104). Further classifications were presented by Calvert Watkins (1999), who used the development of sibilants to elaborate his model, and David Stifter (2008) by sketching isoglosses of the development of nasal clusters.

4.2. Recently, two more elaborated models were developed by Patrick Sims-Williams, using diverse Celtic isoglosses for his classification (cf. Sims-Williams 2007), and by Jesús Alberto Arenas-Esteban and Patrizia de Bernardo Stempel, who see a development from Common Celtic to the documented Celtic languages in five strata (see e.g. Arenas-Esteban and de Bernardo Stempel 2011). The latter model owes much to the archaeological way of organizing material. Besides these more or less philological classifications mentioned above there are models using methods from other scientific disciplines. We shall confine ourselves to mentioning only two recent examples here: the phylogenetic model elaborated by Peter Forster and Alfred Toth (see Forster and Toth 2003) and a glottochronological classification presented by Václav Blažek (see Blažek 2009).

Good overviews of the Celtic languages and literatures in addition to the above-mentioned works are Prosdocimi and Solinas (1991); Green (1995); Meid (1997); Kruta (2000); Ball with Fife (2002); and Pilch (2007).

5. The documentation of Celtiberian and other Hispano-Celtic languages

5.1. After the Romans had occupied most parts of Spain in the 2nd half of the 2nd c. BCE, the Iberian and Celtic inhabitants adopted (or rather had to adopt) the Roman system of administration (cf. Arenas-Esteban 2012). From that time the Celtiberian-speaking peoples began writing in Latin and in their own language, respectively. The documents written in Celtic languages are mainly Celtiberian, a word coined by the Romans to denote the Celtic speaking inhabitants of Spain. Celtiberian inscriptions are attested in Central and Northern Spain, but there are traces of other Celtic languages or rather dialects in Spain, especially Lusitanian (collection of all hitherto known inscriptions in *MLH IV*, 723–758) and Tartessian (collection in *MLH IV*, 93–348) in the southwest of the Iberian peninsula, though their attestation is meagre and their affiliation as Celtic, Indo-European, or even non-Indo-European is controversial (cf. Broderick 2010: 302 ff.; see also the discussion in Mallory et al. 2014).

5.2.1. Inscriptions in Celtiberian languages are attested from the 2nd century BCE to the first century CE. They are written in two orthographic systems: the Iberian syllabic script, which is similar to the Northern Etruscan syllabic script, and the Latin alphabetic script.

5.2.2. Inscriptions in Iberian script are transliterated in bold (sometimes italic non-bold) minuscules, except for the occlusive graphemes, which may denote both *tenues* and

mediae (see below). The Iberian script is very inadequate to render an Indo-European language because it does not allow consonant clusters to be written directly. The script consists of monographemic signs for the vowels *a, e, i, o, u* and the continuants *r, l, m, n, s* and *z* (and/or *ð*). Digraphemic signs containing a consonant and a vowel include *Ka, Ke, Ki, Ko, Ku*. The difference between *tenues* and *mediae* cannot be expressed in the eastern variant of the Iberian script, and for this reason we write these consonant-vowel signs with a capital character for the consonant. The western variant of the Iberian script allows differentiation of the occlusive velar and dental signs according to their sonority by the use of an additional diacritic stroke (see Ferrer i Jané 2005). There are no signs for the combination of vowel + consonant or for clusters of two or more consonants. Therefore, some rules have to be given to interpret Celtiberian inscriptions written in the Iberian script (all cited examples are from Botorríta I, see below, unless otherwise noted):

- a) Signs for *tenues* may denote *mediae* as well in the eastern variant, e.g. **a-m-Pi-Ti-n-Ko-u-n-e-i** /ambi-dingounei/ ‘to build/for building’.
- b) Consonant clusters are written with a combination of signs with the same inherent vowel, e.g. **Ti-r-i-s** /trīs/ ‘three’, **e-n-Ta-r-a** /entra/ ‘within’.
- c) Sometimes nasals may be omitted in writing before consonants, e.g. **s-e-Ko-Ti-a-ð** /segontiãð/ abl.sg. ‘from Segontia’ on a coin (*MLH I*, A 77-1, see 5.3, compare **s-e-Ko-n-Ti-o-s** nom. sg. ‘inhabitant of Segontia’, Botorríta III), at other times they are written, e.g. **Ti-r-i-Ka-n-Ta-m** /trikantãm/ (toponym).
- d) Vowel quantity is not marked. Some rare examples of plene-writing like **a-l-e-Tu-u-r-e-s** (an ethnonym?) do not allow us to suggest plene-writing as an orthographical means to write vowel length.
- e) Sometimes a word-divider denotes word boundaries and is rendered in transcription with a colon : .

5.2.3. Inscriptions in Latin script are transliterated in capitals, e.g. nom. sg. **SEGOBRIS** /segobris/ ‘Segobriga’ (*MLH I*, 89-3; a toponym on a coin). As can be seen from this example the Latin script regularly denotes the difference between *tenues* and *mediae* in Celtiberian. This toponym **SEGOBRIS** is also attested on the same coin in Celtiberian script: abl. sg. **s-e-Ko-Pi-r-i-Ke-ð** /segobrigẽð/ abl. sg. ‘from Segobriga’ (*MLH I* A 89-1); this example shows the obvious differences between the two writing systems.

5.3. The Celtiberian inscriptions on stone and metal are presented by Untermann in *MLH IV*, the coins in *MLH I*, together with the dictionary in *MLH V.1* and diverse word lists in *MLH I* and *MLH IV* by Wodtko. Many of the attested coins contain only the name of a city (e.g. **s-e-Ko-Ti-a-ð** /segontiãð/ ‘from Segontia’, (*MLH I*, A 77-1) and/or the name of a leader (e.g. **Pi-u-l-a-Ko-s**; *MLH I*, A 33-13). Another category of inscriptions comprises the so-called *tesserae hospitales* created after the Roman manner (see e.g. Díaz and Jordán 2006). They are formed like figures of animals or like hands, e.g. the Bronze tessera from Monte Cildá (*MLH IV*, K.27.1) formed like two hands shaking each other and reading **TURIASICA CAR** in Latin script. **CAR** is thought to be an abbreviation of *Cib. karuo-* ‘friendship’.

The best known inscription of at least some length is Botorríta 1 (*MLH IV*, K 1.1.A, K 1.1.B), an edict of the magistrate of Contrebia Belaisca concerning the sacred land of

two gods (?) named Tocoit- and Sarnicios. This inscription consists of about 112 words followed by a list of magistrates, their titles, and their hometowns. The beginning of this inscription reads: **Ti-r-i-Ka-n-Ta-m : Pe-r-Ku-n-e-Ta-Ka-m : To-Ko-i-To-s-Ku-e : s-a-r-n-i-Ki-o {;} Ku-e : s-u-a : Ko-m-Pa-l-Ke-δ : n-e-l-i-To-m** ‘concerning the mountainous region (or: the region *Trikanta Bergunetaka*) of (the gods?) Togoit- and Sarnikios the magistrate decreed as not allowed’ (for various interpretations of this inscription see Eichner 1989; Eska 1989; Meid 1993 and 1994; Jordán Cólera 2004: 185 ff.; Ziegler 2005 [2004]; Prósper 2008; Broderick 2010: 296 ff.).

6. The documentation of Gaulish and other Continental *p*-Celtic languages

6.1. Lepontic

6.1.1. The first attested *p*-Celtic language from the European Mainland is the Lepontic language of Northern Italy and Southern Switzerland. About 150 Lepontic inscriptions are found in an area within a diameter of about 100 km around the town of Lugano in the Canton Ticino, Switzerland. The oldest inscriptions are from the 6th century BCE; the youngest date to the first century BCE. Other Celtic inscriptions of Northern Italy outside of this diameter are traditionally regarded as written in Cisalpine Gaulish. Some scholars hold the opinion that Lepontic is an early dialect of Gaulish rather than an independent branch of the Celtic languages. The Lepontic language is named after the Lepontians, who are considered to be the inhabitants of the Canton Ticino, the southern parts of the Canton Graubünden, both in Switzerland, and the area between and around Lago di Como and Lago Maggiore, in the Italian provinces of Novara and Verbano-Cusio-Ossola. The Lepontians settled there during the first millennium BCE, although their attestation is quite late (Caesar, BG IV, 10; Strabon, IV). This area matches perfectly with the extension of the archaeological Golasecca culture, which begins in the 14th Century BCE with the *Facies Canegrata*, sometimes regarded as an offshoot of north western alpine communities (see e.g. de Hoz 1992; Pauli 1971, 1992: 179 f.). It is therefore likely that immigrants from the north brought the ancestor of the Lepontic language to Northern Italy.

6.1.2. The Lepontic inscriptions are written in the Lugano alphabet, a variant of the North Etruscan alphabet. Like the Iberian script, the Etruscan alphabet does not distinguish between *tenues* and *mediae*. The direction of the script is mostly sinistroversive. The majority of the inscriptions are found on pottery, less frequently on other materials like stone. Besides these there are a few coin inscriptions. The archaeological context often makes a dating possible, so Uhlich (1999) proposed a chronological division into three stages: Early Lepontic (6th–beg. 4th century BCE), Middle Lepontic (beg. 4th–3rd century BCE) and Late Lepontic (2nd–1st century BCE).

6.1.3. There are at least three important collections of the Lepontic inscriptions: Lejeune (1971), Solinas (1995), and the collection of the *Lexicon Leponticum* (*LexLep*) by David Stifter, Martin Braun, and Michaela Vignoli. Lejeune (1971) and *LexLep* furthermore

give solid information concerning what is known about the grammar of Lepontic. The oldest Lepontic inscription of Northern Italy was found on a potsherd in grave 5 of the necropolis of Castelletto Ticino, Via Aronco. The inscription **Xosioiso** is sometimes interpreted as a gen. sg. to a thematic *o*-stem, with metathesis from ***Xosiosio** (see Gambari and Colonna 1988). This inscription is dated to Golasecca IIA (ca. 580–550 BCE). Note also that an inscription recently found on a pottery sherd from the settlement of Montmorot, dép. Jura, France, is probably written in Lepontic. The sherd belongs to an imported Golasecca vessel. The archaeological context in which this sherd was found is dated to around 600 BCE (cf. Kaenel 2000: 153 with Fig. 3; Zeidler 2003: 96). This could indicate that the tradition of writing in the Lepontic-speaking Golasecca culture had been established already in the 7th century BCE. A further example of an Early Lepontic inscription is the somewhat longer inscription of Prestino, Fondo Giulini, which must be read as: **uvamoKozis Plialeθu uvITiauiopos ariuonePos siTeś TeTu**, with **TeTu** as a rare example of an attested verb in Lepontic (for interpretations of this inscription see e.g. Solinas 1995: 344; Markey and Mees 2002: 145 ff.). A good example for the Middle Lepontic period is an inscription on a grave stone from Davesco, Distr. Lugano, Ticino: **slaniai uerKalai Pala Tisiui PiuoTialui Pala**. From Late Lepontic an inscription on a pottery vessel which was found in a grave of the cemetery of San Bernardo di Ornavasso, Italy may be mentioned. This grave is dated to Lt D1 (late 2nd c. BCE–beg. 1st c. BCE). The inscription reads **laTumarui saPsuTai Pe uinom naśom**. See also Morandi (2004).

6.2. Cisalpine Gaulish

6.2.1. As mentioned above Celtic inscriptions of Northern Italy have also been found outside of a diameter of 100 km around Lugano. These are traditionally regarded as written in Cisalpine Gaulish. One may assume that this language was brought to northern Italy by the invasion of northern alpine tribes shortly after 400 BCE.

6.2.2. Again, the Cisalpine Gaulish inscriptions mainly employ a variant of the North Etruscan alphabet.

6.2.3. Modern collections of Cisalpine Gaulish inscriptions can be found in Lejeune (1988), Solinas (1995), and in LexLep. There are three somewhat longer inscriptions that can be assigned to Cisalpine Gaulish. The first is the inscription of San Bernardino di Briona from the Province of Novara in the Piedmont, found on a stone stela in the area of an ancient (Gaulish or Lepontic) graveyard, written in a variant of the North Etruscan alphabet, and dated to the 1st century BCE (for the inscription and its interpretation see e.g. Lambert 2003: 72 ff.; Solinas 1995: 379 ff.). Another important Cisalpine Gaulish inscription, dated to 150–100 BCE, was found in Todi, in the province of Perugia in Umbria. It is a bilingual written on stone with a Gaulish part using the North Etruscan alphabet and a Latin part using the Latin alphabet (for the inscription and its interpretation see e.g. Lambert 2003: 74 ff.; Solinas 1995: 382 f.). The third inscription worthy of mention here is the bilingual inscription (Latin and Cisalpine Gaulish) of Vercelli. Again this inscription was written in Latin and North Etruscan letters (for the

inscription and its interpretation see e.g. Lambert 2003: 76 ff.; Solinas 1995: 381 f.). The direction of the Gaulish parts of these three inscriptions is dextroverse.

6.3. Gaulish

6.3.1. The Gaulish language was mainly spoken in the area of modern France, and the majority of the Gaulish inscriptions are from this region. Very few inscriptions are attested in an area even as close by as northern Switzerland. Nevertheless, one may imagine that a variety of Gaulish was spoken in the adjacent regions of south-western Germany, Austria, and Czechia (Bohemia). Further research may lead to dialectal subdivisions in such a huge area (as is indicated by de Bernardo Stempel 2005: 189 Abb. 1, and Whatmough 1970). To what extent parts of Eastern Europe and Eastern Central Europe – where one can observe the expansion of the archaeological La Tène Culture (parts of the Balkan Regions, Romania, Hungary, Slovakia, etc.) in the 4th–2nd century BCE – may have been parts of a Celtic or Gaulish *κοινή* remains unclear. Another Gaulish dialect seems to have been spoken far to the east: the Galatian language of Anatolia (see Eska 2013; Freeman 2001). For the regions outside of France we must look primarily to toponymic and hydronymic evidence and historic information or “Nebenüberlieferungen” of the classical writers. There are very few attested written documents in the Gaulish language outside of the French borders, rare examples being e.g. from Bern, Manching, and Mayence (cf. Koch 2007; Stüber 2006; Sims-Williams 2006).

6.3.2. Three writing systems were used for the Gaulish inscriptions: the North Etruscan alphabet, used for the Cisalpine Gaulish inscriptions already mentioned; the Greek alphabet, used mainly in Southern France from the 3rd century BCE until the Roman conquest of Gaul; and the Roman alphabet, which thereafter rapidly replaced the Greek script in all of France. It is possible to divide the Gaulish inscriptions of France into two or three stages. While Pierre-Yves Lambert (1997) proposes a division into two stages (*Gauloise ancien*, *Gauloise tardif*), David Stifter (*LexLep*) distinguishes three stages (Early, Middle, and Late Gaulish). Stifter’s Early Gaulish (cf. Stifter: Gaulish) dates from the 3rd to the 1st century BCE, perhaps even until the 1st century CE. It consists mainly of the Gaulish inscriptions in the Greek alphabet. Also to be reckoned here are some of the earliest inscriptions in the Roman alphabet as well as Gaulish coins. Middle Gaulish lasts from the beginning of Christian times until the 2nd or 3rd century CE. Typically, in this period, the Roman alphabet in plain and cursive variants is used. Late Gaulish covers the period from the 3rd c. CE until the extinction of the Gaulish language in the second half of the 1st millennium CE. Besides grave inscriptions and dedicatory inscriptions on stone in the Greek alphabet one finds few inscriptions that indicate ownership or manufacturing. The latter are mostly on various materials like bronze (Bern, Thormebodewald), iron (Port, Canton Bern) or ceramics (Manching, Bavaria). In addition to the inscriptions in Greek letters, one also finds grave inscriptions and votive inscriptions on stone using the letters of the Latin alphabet. Furthermore, Gallo-Latin inscriptions are found on spindle-whorls, rings, pottery, clay figures, pendants, glass jugs, bronze vessels, etc. Very special examples are the bronze fragments of two Gaulish Calendars: the famous Calendar of Coligny, dép. Ain, and the calendar from Villards-d’Héria, dép. Jura.

Furthermore, there are a few longer inscriptions on lead tablets (Hospitalet-du-Larzac, dép. Aveyron; Chamalières, dép. Puy-de-Dôme; Lezoux, dép. Puy-de-Dôme; Rom, dép. Deux-Sèvres; Bath, Somerset), on tiles (Châteaubleau, dép. Seine-et-Marne), on a plate (Lezoux, dép. Puy-de-Dôme), on a gold plate (Baudécet à Gembloux, Belgium), and on a silver tablet (Poitiers, dép. Vienne).

6.3.3. The Gaulish inscriptions are collected in the RIG-Series (Gaulish calendars: Duval/Pinault 1986; Gaulish coin legends: Colbert de Beaulieu and Fischer 1998; Gaulish inscriptions in the Greek alphabet: Lejeune 1985; Gaulish inscriptions in the Latin alphabet on stone and Cisalpine Gaulish: Lejeune 1988; Gaulish inscriptions in the Latin alphabet on implements etc.: Lambert 2002). For the graffiti of La Graufesenque see Marichal (1988). A survey of Gaulish grammar is found in Lambert (2003); an overview of Gaulish vocabulary is Delamarre (2003). An example of Early Gaulish is the dedicatory inscription to the “mothers of Glanum” of Saint-Rémy-en-Provence, dép. Bouche-du-Rhône, the ancient city of Glanum, using the common formula **BPATOY ΔEKAN TEM/N: MATPEBO ΓΛANEIKABO BPATOY ΔEKANTEM** (cf. e.g. Meid 1992, with 26 Fig. 20). Another Early Gaulish example – from outside the French borders – is the Swiss inscription from Bern, Thormebodewald which perhaps contains the ancient name of Bern **BPENOΔΩP**. The whole inscription goes: **ΔOBNOPHΔO ΓOΒANO BPENOΔΩP NANTAPΩP** (cf. Fellmann 1999; Stüber 2006). As examples for Middle Gaulish one can cite the calendar of Coligny (cf. e.g. Duval and Pinault 1986; Lambert 2003: 108 ff.) and the inscription of Alise-Sainte-Reine, dép. Côte-d’Or:

**MARTIALIS DANNOTALI
IEVRV · VCVETE · SOSIN
CELICNON ETIC
GOBEDBI · DVGIÓNTIÍO
VCUETIN
IN ... ALISIÁ**

(cf. e.g. Lejeune 1988: 147 ff.; Lambert 2003: 98 ff.). Finally, Late Gaulish is represented e.g. by the tile of Châteaubleau (see Lambert 1998–2000; Lambert 2002: 238 ff.) and the dedication on the vase of Séraucourt à Bourges, dép. Cher: **BVSCILLA SOSIO LEGASIT IN ALIXIE MAGALU** (cf. e.g. Lambert 2002: 205 ff.; 2003: 136 f.).

7. The documentation of Irish

7.1. The Irish language belongs to the *q*-Celtic languages and is attested from the 4th century CE until today, with discontinuity of transmission only in the mid-first millennium CE. Ireland was never conquered by the Romans, so it has undergone less influence from Latin than British Celtic. Latin influences increased only with christianization; and many Latin and Greek words were adopted from the British Celts (on British see Haarmann 1970, 1973; on Irish see McManus 1983; Vendryes 1902; Ziegler 2000 with lit.).

7.2. The earliest documents in Irish are the so-called *Ogam* (variantly: *Ogham*) inscriptions (Korolev 1984; McManus 1991; Ziegler 1994) written in an alphabet consisting of up to five strokes and points looking somewhat like a written Morse-alphabet. These

were used mainly for inscriptions on tombstones, although Irish lore informs us that they were also employed in cursing formulas on wood (which were not preserved due to the climate; cf. Ziegler 1994: 9 ff.). About 370 Ogam stones are known. The language of most of these is very archaic, in part because it preserves many sounds and inflections which vanished in Classical Old Irish (see McManus 1991: 83 ff.; McCone 1996; Ziegler 1994: 36 ff.), e.g. VALUVI (ca. 500–550; Ballyelan, Waterford; gen. sg. of a PN), which later became *Fáilbi* in Classical Old Irish, or the syntagm TRIA MAQA MAILAGNI ‘of the three sons of Maelán’, which should be *tre* (archaic gen.pl. besides Old Irish *trí*, see Ziegler 1994: 61) *m(m)acc* gen.pl. *M(m)aeláin* gen.sg. in Classical Old Irish.

7.3. The first directly attested literary documents are the Cambray Homily (end of the 7th / beginning of the 8th c. CE) with some alternating short Latin and Early Old Irish text passages (Stokes and Strachan 1903, II: 244 ff.) and the Early Old Irish glosses and insertions in the Computus Einsidlensis (ca. 700; Bisagni and Warntjes 2008), the Münchener Computus (ca. 719; Warntjes 2010), and the *prima manus* of the Würzburg manuscript of the Epistles of St. Paul (ca. 725, Stokes and Strachan 1903, I: 499 ff.; Kavanagh and Wodtke 2001) followed by the Classical Old Irish glosses of the *secunda* and *tertia manus* (ca. 750) of the Würzburg manuscript. Showing an archaic stage of language but copied only in younger times from older manuscripts with mistakes and some Middle Irish word forms are the *Amra Choluim Chille* by *Dallán Forgaill* (ca. 600; cf. Richter 2005: 54 f.; in the *Lebor na hUidre* from ca. 1100); the poems by *Lucreth Moccu Chiara* (early 7th cent.; Henry 1997), containing the oldest surviving reference to the Ulster Cycle (see below); and parts of the Old Irish Law tracts (see 7.4 below), dating from the 6th and 7th c. CE.

From the 8th century on we have an overwhelmingly rich literature (Welch 1996). The most important is the *Táin Bó Cúailnge* (O’Rahilly 1967, 1976; both available online at *CELT*) in prose with some inserted archaic verses. It belongs to the Ulster Cycle (see 7.4).

Most Old and Middle Irish texts are collected in the *Lebor na h-Uidre* (‘Book of the Dun Cow’, 11th/12th c. CE; Best and Bergin 1929), the *Lebor Laignech* (‘Book of Leinster’, early 12th c. CE; Best et al. 1954–1983), the Rawlinson manuscript B 502 (Meyer 1909), the *Leabhar Bhaile an Mhóta* (‘Book of Ballymote’, 1390/1391, Atkinson 1887), the *Leabhar Buidhe Leácain* (‘Yellow Book of Lecan’, early 15th cent., only partly edited and translated, contents in Abbott and Gwynn 1921), and the *Leabhar Mór Leacain* (‘Great Book of Lecan’, 1397–1418, only partly edited, facsimile on *ISOS*).

7.4. The Old and Middle Irish sagas and tales are collected in four cycles called Mythological Cycle, Ulster Cycle, Fenian Cycle, and Historical Cycle (O’Rahilly 1946; Maier 1994 s.vv.). The most important and most voluminous is the Ulster Cycle containing stories about the Ulidian ruler *Conchobor mac Nessa* and his most prominent hero, *Cú Chulainn* (Kinsella 1969; Thurneysen 1921; Koch and Carey 2003). The so-called Mythological Cycle (see d’Arbois de Jubainville 2006) comprises stories set up around some pre-Christian pagan gods, fairies, and elves, although these stories are interwoven with the other cycles so that in some cases no explicit classification of the Old and Middle Irish stories is possible. The Fenian (or Ossian) Cycle is younger than the Heroic Cycle. The most prominent heroes are *Oisín* (Eng. *Ossian*) and his son *Oscar*, the protagonists in this popular body of romance in Old and Middle Irish times (cf. Murphy and

O’Cuív 1971). The Historical Cycle (cf. Mac Eoin 1989) was composed – mainly in poetry – by medieval Irish bards, who recorded the history of the families and the genealogy of the kings they served. To these belongs the well-known story *Buile Suibhne* ‘the frenzy of Suibhne’ (Eng. *Sweeney*) about the king *Suibhne*, who was cursed and became a hybrid – half man, half bird – condemned to live out his life in the woods (English translation by Joyce 1976).

Besides the four cycles there is a huge amount of other Old and Middle Irish literature: Christian literature (e.g. *Bethada náem nÉirenn* ‘the lives of Irish Saints’; some *aislinge* ‘visions’; a nice parody of these is *Aislinge Meic Con Glinne* ‘the vision of Mac Conglinne’); translations or adaptations of Latin and Greek tales (e.g. *imtheachta Aeniasa* ‘the journeys of Aeneas’; *togail troi* ‘the destruction of Troy’), historical treatments (many *annála* ‘annals’, e.g. *Annála Connacht* ‘the annals of Connaught’), and Early Irish Law (Binchy 1978; Breatnach 2005). For detailed overviews see Kelleher and O’Leary (2006); Brady and Cleeve (1985).

Many of the above-mentioned texts can be found as facsimiles in *ISOS* or edited and/or translated in *CELT*.

Irish is also spoken in Scotland (Scottish Gaelic or Gáidhlig), where it was most likely brought by Irish settlers who founded the kingdom of the Dál Riata in the 4th or 5th c. CE (Broun 2001), although this supposition is doubted by some archaeologists on purely archaeological grounds (Campbell 2001). Beginning with the 13th c. Scottish Gaelic separated from Irish and developed as an independent language still in use today (see 9). In 1622 Scottish people began to migrate to North America, especially to Nova Scotia (Canada), but most of them came in the wake of the Jacobite Rebellions between 1688 and 1746. Some descendants of these settlers still speak the Canadian variant of Scottish Gaelic (Gáidhlig Chanada) as a native language in Cape Breton Island and in isolated areas in Nova Scotia and Prince Edward Island (<http://www.novascotia.ca/oga/aboutgaelic.asp>).

On the Isle of Man a variety of Irish called Manx was also spoken as a native language until 1765, when the United Kingdom annexed the island, and the English language superseded Manx in the following centuries. Manx became extinct in the 20th c. with the death in 1974 of Ned Maddrell, the last native speaker of Manx (Broderick 2002).

8. The documentation of British Celtic

8.1. Brythonic (Brittonic) languages

Knowledge of the Brythonic (or: Brittonic) languages comes from a variety of sources. The earliest attestations of British Celtic words and personal or place names are on coins (e.g. *Catamanus*, Old Welsh *Catman*, Modern Welsh *Cadfan*; Jackson 1953: 620 f.; cf. Bammesberger and Wollmann 1990) and some old inscriptions, both from Roman times (1st–5th c. CE), together with “Nebenüberlieferung” by classical writers (e.g. Lat. *Londonium*, MWelsh *Llundein* ‘London’, Rivet and Smith 1979: 49 ff.). From the 5th century on in the eastern parts of Wales inscriptions in Ogam script are attested, containing Irish names and a few definitely Old British names (cf. Jackson 1953 *passim* and Ziegler

1994: 90 f. on their typical British features, e.g. VOTEPORIGIS *‘king of retreat’ = Welsh *Godebri** with Old Brit. *-tepo-* ‘retreat, flight’ as in Middle Welsh *tebed* m. ‘id.’ < PCelt. **tek^w-eto-* with regular sound change from labiovelar **k^w* to labial *p*; cf. OIr. *techid* ‘flees, retreats’ with velar *ch* from labiovelar **k^w*, both from PIE **tek^w-* ‘run, flow, flee’, LIV² 620 f.).

8.2. Welsh

The first attested literary documents of Welsh (in Welsh: *Cymraeg*) are the poems of Taliesin and Aneirin in the 6th and 7th centuries CE (*canu Taliesin* ‘the songs of Taliesin’, cf. Williams and Caerwyn Williams 1968; *canu Aneirin*, cf. Williams 1938; Koch 1997); both are also known as the *cynfeirdd* ‘early bards or poets’. Their original language is Early Old Welsh but the manuscripts containing them are younger (from the 12th c.) and show some adaptations to the later language. In Nennius’ *Historia Brittonum* (ca. 800, written in Latin) Talhaearn Tad Awen, Blwchfardd, and Cian are mentioned besides Taliesin and Neirin (= Aneirin) as Welsh poets of the end of the 6th century, although we know nothing of their works. From that time on Old Welsh glosses are attested in several Latin manuscripts (cf. Falileyev 2008; Huws 2000), the oldest of these, the so-called “*surexit*”-*memorandum*, is from the beginning or the first half of the 9th c. and contains words like *tutri* ‘king of peoples’ (cf. Gaul. *Toutorix*, Falileyev 2008, 91). The famous laws of *Huwel Dda* [Howell the Good; ca. 880–950] contain the codification of traditional Welsh law and are therefore of great interest (Edwards 1973; Ellis 1926–1927).

In the Middle Welsh period from the 12th to the 14th century literacy in Wales increased significantly. The most prominent texts of this period are the *pedeir keinc y mabinogi* ‘the four branches of the mabinogi’ (Williams 1982) and the works of the early *gogynfeirdd* ‘rather early poets’ (Hughes and Williams 1910; Caerwyn Williams 1994; translation Clancy 1970), who were also called *beirdd y tywysogion* ‘bards of the princes’ because they were mainly biographers of kings (Lewis 1992: 123 ff.; Lloyd 1992: 157 ff.).

The beginning of Modern Welsh is marked by the works of Dafydd ap Gwilym (ca. 1320–1370). About 170 of his poems are believed to have survived, although many more have been attributed to him. He was one of the typical medieval novelists belonging to the so-called courtly troubadour poetry circle (see ap Gwilym and Parry 1996). For further reading see Jarman and Rees Hughes (eds. 1992–1997) and Bromwich (1974).

8.3. Cumbrian

The Cumbrian language, formerly spoken in some parts of southern Scotland and the Southern Uplands, is known only from three words in the *Leges inter Brettos et Scotos*: *galnes/galnys* ‘blood-fine’, *mercheta* ‘daughter, girl’, and *kelchyn* ‘circle’ (Jackson 1953: 9).

8.4. Cornish

Cornish (in Cornish: *Kernowek*, *Kernewek*) is one of the lesser-known Brythonic languages and is today almost extinct except in some associations devoted to reviving the Cornish language by translating modern literature into “modern” Cornish (e.g. *adro dhe’n Bys in Peswar Ugans Dëdh* ‘Around the World in Eighty Days’, an adventure story by the famous French writer Jules Verne). The oldest Cornish document is the Latin poem *Prophetiae Merlini* ‘the Prophecies of Merlin’ with some Cornish glosses and marginal notes from the 12th century. One of the earliest surviving works of Cornish literature is *Pascon agan Arluth* ‘the Passion of Our Lord’, a poem of 259 eight-line verses composed in the 2nd half of the 14th century; the longest original surviving work is *Beunans Meriasek* ‘the Life of Meriasek’ dated from 1504, but probably copied from an earlier manuscript. Cornish died out in the 19th century. For further information see Ellis (1974).

8.5. Breton

With the beginning of the Anglo-Saxon invasions in England some tribes of Brythonic speakers fled to the North-Western part of the continent, which since then has been named Bretagne (Brittany). The earliest surviving manuscript containing instances of Old Breton is the Leiden Leechbook (cf. Falileyev and Owen 2005) from the end of the 8th century. Until the 11th century only a few words and glosses in Old Breton can be found in Latin manuscripts, and their language is very similar to the Welsh of that time. The Norman invasion of England brought forth another exodus of people which marks the beginning of the Middle Breton period. The literature is mainly courtly (of the troubadour circle, whose most famous author was the French-writing Chrétien de Troyes, ca. 1140–1190) or ecclesiastical metrical literature. An example of the first is *Dialog etre Arzur ha Gwynglaff* ‘The Dialogue between Arthur and Gwynglaff’, composed ca. 1450 (Koch 2006: 585); an example of the latter is *Buhez Sante Barba* ‘The Life of St. Barbara’ (first printed 1557, cf. Ernault 1888).

9. Celtic languages today

The Celtic languages are today in retreat. Once spoken in an area that included Western Europe and the western parts of Central Europe, Celtic languages are now spoken only in the Western fringe of Great Britain and Ireland – besides a few speakers of Irish and Welsh in America. While in the Republic of Ireland there are about 1.8 million persons who are competent to speak Irish, there are only estimated to be 40,000–80,000 people who use Irish – outside of the education system – as their everyday language. The situation seems to be better in Wales, where about 15% (431,000 in total) of Welsh citizens regard themselves as fluent speakers of Welsh. Scottish Gaelic is still used by about 59,000 speakers, while the Breton language is still spoken by 206,000 people in Brittany (see for these data URL http://en.wikipedia.org/wiki/Celtic_languages).

Irish, Scottish Gaelic, and Welsh are also spoken in some enclaves in North America (on Canada see 7.4), Australia, and New Zealand, mainly by members of associations trying to keep their ancestors' languages alive (on Scottish Gaelic in Australia and New Zealand see Skilton 2004; on Irish in the USA see <http://www.iaci-usa.org/aboutus.html>, on Welsh in the USA see <http://www.cardiff.ac.uk/welsh/subsites/welshamericanstudies/index.html>, on all Celtic languages in Australia see <http://www.celticcouncil.org.au/>).

10. References

- Abbott, Thomas K. and Edward J. Gwynn
 1921 *Catalogue of the Irish Manuscripts in the Library of Trinity College, Dublin*. Dublin: Hodges, Figgis & Co., 94–110.
- d'Arbois de Jubainville, Henry
 2006 *The Irish Mythological Cycle and Celtic Mythology*. Kila, MT: Kessinger. [Reprint of the English translation by Richard Irvine Best 1903, Dublin: O'Donoghue, of *Le cycle mythologique irlandais et la mythologie celtique*, Paris: Thorin. 1884.]
- Arenas-Esteban, Jesús A.
 2012 The Celtiberian World: A Long Process of Cultural Hybridization. In: Peter Anreiter, Eszter Bánffy, László Bartosiewicz, Wolfgang Meid, and Carola Metzner-Nebelsick (eds.), *Archaeological, Cultural and Linguistic Heritage. Festschrift for Erzsébet Jerem in Honour of her 70th Birthday*. Budapest: Archaeolingua Alapítvány, 33–46.
- Arenas-Esteban, Jesús A. and Patrizia de Bernardo Stempel
 2011 Celtic Dialects and Cultural Contacts in Protohistory: the Italian and Iberian Peninsulae. *Études Celtiques* 37: 119–139.
- Atkinson, Robert (ed.)
 1887 *The Book of Ballymote*. Dublin: The Royal Irish Academy.
- Ball, Martin J. with James Fife (eds.)
 2002 *The Celtic Languages*. London: Routledge.
- Bammesberger, Alfred and Alfred Wollmann (eds.)
 1990 *Britain 400–600: Language and History*. Heidelberg: Winter.
- de Bernardo Stempel, Patrizia
 2005 Indogermanisch und keltisch 'geben': Kontinentalkelt. *Gabiae, gabi /gabas*, keltib. *gabizeti* und altir. *ro-(n)-gab* und Zugehöriges. *Historische Sprachforschung* 118: 185–200.
- Best, Richard I. and Osborn Bergin (eds.)
 1929 *Lebor na hUidre* [Book of the Dun Cow]. Dublin: The Royal Irish Academy.
- Best, Richard I, Osborn Bergin, Michael A. O'Brien, and Anne O'Sullivan
 1954–1983 *The Book of Leinster; formerly Lebar na Núachongbála*. 6 vols. Dublin: Dublin Institute for Advanced Studies.
- Binchy, D[aniel] A.
 1978 *Corpus Iuris Hibernici* [Corpus of Irish law]. 6 vols. Dublin: Dublin Institute for Advanced Studies.
- Birkhan, Helmut
 1997 *Kelten. Versuch einer Gesamtdarstellung ihrer Kultur*. Vienna: Österreichische Akademie der Wissenschaften.
- Bisagni, Jacopo and Immo Warntjes
 2008 The Early Old Irish material in the newly discovered *Computus Einsidlensis* (c. CE 700). *Ériu* 58: 77–105.

- Blažek, Václav
 2009 On the position of Gaulish within Celtic from the point of view of glottochronology. *Indogermanische Forschungen* 114: 257–299.
- Bosch Gimpera, Pedro
 1940 *Two Celtic Waves in Spain*. Proceedings of the British Academy XXVI. London: Humphrey Milford.
- Brady, Anne M. and Brian Cleeve (eds.)
 1985 *A Biographical Dictionary of Irish Writers*. Mullingar: Lilliput.
- Breatnach, Liam
 2005 *A companion to the Corpus Iuris Hibernici*. Dublin: Dublin Institute for Advanced Studies.
- Broderick, George
 2002 Manx. In: Ball with Fife (eds.), 228–285.
- Broderick, George
 2010 Die vorrömischen Sprachen auf der iberischen Halbinsel. In: Uwe Hinrichs (ed.), *Handbuch der EuroLinguistik*, Wiesbaden: Harrassowitz, 287–320.
- Bromwich, Rachel
 1974 *Dafydd ap Gwilym*. Cardiff: University of Wales Press.
- Broun, Dauvit
 2001 Dál Riata. In: Michael Lynch (ed.), *Oxford Companion to Scottish History*. Oxford: Oxford University Press, 161–162.
- Busse, Peter
 2007 Hydronymie und Urheimat: Ein neuer Ansatz zur Lokalisierung der Urheimat der Kelten. In: Helmut Birkhan (ed.), *Kelten-Einfälle an der Donau*. Vienna: Österreichische Akademie der Wissenschaften, 89–98.
- Caerwyn Williams, John E.
 1994 *The Poets of the Welsh Princes*. Cardiff: University of Wales Press.
- Campbell, Ewan
 2001 Were the Scots Irish? *Antiquity* 75: 285–292.
- CELT = *Corpus of electronic texts*.
 The Free Digital Humanities Resource for Irish history, literature and politics. URL: <http://www.ucc.ie/celt/>
- Clancy, Joseph
 1970 *The earliest Welsh Poetry*. London: Macmillan.
- Colbert de Beaulieu, Jean-Baptiste and Brigitte Fischer
 1998 *Recueil des inscriptions Gauloises, Vol. 4: Les légendes monétaires*. Paris: Centre National de la Recherche Scientifique.
- Cunliffe, Barry and John T. Koch (eds.)
 2010 *Celtic from the West. Alternative Perspectives from Archaeology, Genetics, Language and Literature*. Oxford: Oxbow.
- Delamarre, Xavier
 2003 *Dictionnaire de la langue gauloise. Une approche linguistique du vieux-celtique continental*. 2nd edn. revised and expanded. Paris: Errance.
- Díaz, María Antonia and Carlos Jordán
 2006 Dos téseras de hospitalidad procedentes de Fitero (Navarra). *Palaeohispanica* 6: 257–266.
- Dillon, Myles and Nora K. Chadwick
 1967 *The Celtic Realms*. London: Weidenfeld and Nicholson.
- Dobesch, Gerhard
 1991 Ancient Literary Sources. In: Moscati et al. (eds.), 35–50.

- Duval, Paul-Marie and Georges-Jean Pinault
 1986 *Recueil des inscriptions Gauloises, Vol. 3: Les calendriers*. Paris: Centre National de la Recherche Scientifique.
- Edwards, John G.
 1973 Hywel Dda and the Welsh Law-books. In: Dafydd Jenkins (ed.), *Celtic Law Papers*. Brussels: Libraire Encyclopédique, 135–160.
- Eichner, Heiner
 1989 Damals und heute: Probleme der Erschließung des Altkeltischen zu Zeußens Zeit und in der Gegenwart. In: Bernhard Forssman (ed.), *Erlanger Gedenkfeier für Johann Kaspar Zeuß*. Erlangen: Universitätsbund Erlangen-Nürnberg, Auslieferung, Universitätsbibliothek Erlangen, 9–56.
- Ellis, Peter B.
 1974 *The Cornish Language and its Literature*. London: Routledge.
- Ellis, Thomas P.
 1926–1927 Hywel Dda: Codifier. *Transactions of the Honourable Society of Cymmrodorion*, 1–69.
- Ernault, Émile. (ed. and tr.)
 1888 *Le mystère de Sainte Barbe. Tragédie bretonne, texte de 1557, publié avec traduction française, introduction et dictionnaire étymologique du breton moyen*. Paris: Thorin.
- Eska, Joseph F.
 1989 *Towards an interpretation of the Hispano-Celtic inscription of Botorrita*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Eska, Joseph F.
 2013 A salvage grammar of Galatian. *Zeitschrift für Celtische Philologie* 60: 51–63.
- Falileyev, Alexander
 2008 *Le Vieux-Gallois*. (en collaboration avec Hildegard L. C. Tristram pour la rédaction et Yves le Berre pour la traduction française.) Potsdam: Universitätsverlag.
- Falileyev, Alexander and Morfydd E. Owen
 2005 *The Leiden Leechbook. A Study of the Earliest Neo-Brittonic Medical Compilation*. With two appendices contributed by Helen McKee. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Fellmann, Rudolf
 1999 Das Zinktäfelchen vom Thormebodewald auf der Engehalbinsel bei Bern und seine keltische Inschrift (mit Beiträgen von Willem B. Stern, Andreas Burkhardt und Thilo Rehren). *Archäologie im Kanton Bern 4B*: 133–175.
- Ferrer i Jané, Joan
 2005 Novetats sobre el sistema dual de diferenciació gràfica de les oclusives sordes i sonores [Developments concerning the dual system of graphic differentiation of voiced and voiceless stops]. *Palaeohispanica* 5: 957–982.
- Forster, Peter and Alfred Toth
 2003 Toward a phylogenetic chronology of ancient Gaulish, Celtic, and Indo-European. *Proceedings of the National Academy of Sciences* 100: 9079–9084.
- Freeman, Philip M.
 1996 The earliest Greek sources on the Celts. *Études Celtiques* 32: 11–48.
- Freeman, Philip M.
 2001 *The Galatian Language*. Lewiston, NY: Edwin Mellon.
- Gambari, Filippo M. and Giovanni Colonna
 1988 Il bicchiere con iscrizione arcaica da Castelletto Ticino e l'Adozione della scrittura nell'Italia nord-occidentale. *Studi Etruschi* 54: 119–164.
- GLQFM = Griechische und lateinische Quellen zur Frühgeschichte Mitteleuropas bis zur Mitte des 1. Jahrtausends u. Z.*
 4 vols. Ed. Joachim Herrmann. Berlin: Akademie-Verlag, 1988–1992.

- Green, Miranda J. (ed.)
 1995 *The Celtic World*. London: Routledge.
- ap Gwilym, Dafydd and Thomas Parry (eds.)
 1996 *Gwaith Dafydd ap Gwilym*³ [The Work of Dafydd ap Gwilym]. Cardiff: University of Wales Press.
- Haarmann, Harald
 1970 *Der lateinische Lehnwortschatz im Kymrischen*. Bonn: Romanisches Seminar der Universität.
- Haarmann, Harald
 1973 *Der lateinische Lehnwortschatz im Bretonischen*. Hamburg: Buske.
- Hawkes, Christopher
 1973 “Cumulative Celticity” in Pre-Roman Britain. *Études Celtiques* 13: 607–628.
- Henry, Patrick L.
 1997 *Conailla Medb Míchuru* and the Tradition of Fiacc Son of Fergus. In: Séamus Mac Mathúna and Ailbhe Ó Corráin (eds.), *Miscellanea Celtica in memoriam Heinrich Wagner*. Uppsala: University Press, 53–70.
- de Hoz, Javier
 1992 Lepontic, Celt-Iberian, Gaulish and the archaeological evidence. *Études Celtiques* 29: 223–239.
- Hughes, Arthur and Sir Ifor Williams
 1910 *Gemau'r Gogynfeirdd* [Jewels of the Gogynfeirdd]. Pwllheli: Argraffwyd i'r Golygydd.
- Huws, Daniel
 2000 *Medieval Welsh Manuscripts*. Cardiff: University of Wales Press.
- ISOS = *Irish Script on Screen*, URL: <http://www.isos.dias.ie/english/index.html>.
- Jackson, Kenneth H.
 1953 *Language and History in Early Britain: a Chronological Survey of the Brittonic Languages First to Twelfth Century A.D.* Edinburgh: University Press.
- Jarman, Alfred O. H. and Gwilym Rees Hughes (eds.)
 1992–1997 *A guide to Welsh literature*. 2 vols. Cardiff: University of Wales Press.
- Jordán Cólera, Carlos
 2004 *Celtibérico*. Zaragoza: Universidad de Zaragoza.
- Joyce, Trevor
 1976 *The poems of Sweeny Peregrine*. Ed. by Michael Smith. Dublin: New Writer's Press.
- Kaenel, Gilbert
 2000 Les relations transalpines à l'Age du Fer: territoire “lépontien” – Suisse occidentale – Jura. In: Rafaele C. de Marinis and Simonetta Biaggio Simona (eds.), *I Leponti: tra mito e realtà*. Tomo 2. Verbania: Armando Dadò, 151–158.
- Kavanagh, Séamus and Dagmar S. Wodtko
 2001 *A lexicon of the old Irish glosses in the Würzburg manuscript of the Epistles of St. Paul*. Vienna: Österreichische Akademie der Wissenschaften.
- Kelleher, Margaret and Philip O'Leary (eds.)
 2006 *The Cambridge History of Irish Literature*. 2 vols. Cambridge: University Press.
- Kinsella, Thomas
 1969 *The Táin*. Oxford: University Press.
- Koch, John T.
 2006 *Celtic Culture. A Historical Encyclopedia*. Santa Barbara: ABC-CLIO.
- Koch, John T.
 2007 *An Atlas for Celtic Studies. Archaeology and Names in Ancient Europe and Early Medieval Ireland, Britain, and Brittany*. In collaboration with Raimund Karl, Antone Minard, and Simon Ó Faoláin. Oxford: Oxbow.
- Koch, John T. (ed.)
 1997 *The Gododdin of Aneirin*. Cardiff: University of Wales Press.

- Koch, John T. and John Carey
2003 *The Celtic Heroic Age*. Aberystwyth: Celtic Studies Publications.
- Koch, John T. and Barry Cunliffe (eds.)
2013 *Celtic from the West 2. Rethinking the Bronze Age and the Arrival of Indo-European in Atlantic Europe*. Oxford: Oxbow.
- Korolev, Andrej Alexandrovič
1984 *Drevnejšie pamjatniki irlandskogo jazyka* [The oldest monuments of the Irish language]. Moscow: Nauka.
- Kruta, Venceslas
1979 Duchcov-Münsingen: Nature et diffusion d'une phase laténienne. In: Paul-Marie Duval and Venceslas Kruta (eds.), *Les mouvements Celtiques du V^e au I^{er} siècle avant notre ère. Actes du XXVIII^e Colloque organisé à l'occasion du IX^e Congrès International des Sciences Préhistoriques et Protohistoriques*, Nice, le 19 septembre 1976. Paris: Centre National de la Recherche Scientifique, 81–115.
- Kruta, Venceslas
2000 *Les Celtes. Histoire et dictionnaire. Des origines à la romanisation et au christianisme*. Paris: Laffont.
- Lambert, Pierre-Yves
1997 Gaulois tardif et latin vulgaire. *Zeitschrift für Celtische Philologie* 49–50: 396–413.
- Lambert, Pierre-Yves
1998–2000 La tuile gauloise de Châteaubleau (Seine-et-marne). *Études Celtiques* 34: 57–115.
- Lambert, Pierre-Yves
2002 *Recueil des inscriptions gauloises, Vol. 2,2: Textes gallo-latins sur instrumentum*. Paris: Centre National de la Recherche Scientifique.
- Lambert, Pierre-Yves
2003 *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*². Paris: Errance.
- Lejeune, Michel
1971 *Lepontica*. (Monographies Linguistiques I). Paris: Les Belles Lettres.
- Lejeune, Michel
1985 *Recueil des inscriptions Gauloises, Vol. 1: Textes gallo-grecs*. Paris: Centre National de la Recherche Scientifique.
- Lejeune, Michel
1988 *Recueil des inscriptions Gauloises, Vol. 2,1: Textes gallo-étrusques, textes gallo-latins sur pierre*. Paris: Centre National de la Recherche Scientifique.
- Lenerz-De Wilde, Majolie
1991 *Iberia Celtica. Archäologische Zeugnisse keltischer Kultur auf der Pyrenäenhalbinsel*. 2 vols. Stuttgart: Steiner.
- Lewis, Ceri
1992 The court poets: their function, status and craft. In: Jarman and Rees Hughes (eds.), 123–156.
- Lexlep*: see Stifter, Braun, and Vignolie.
- LIV*: see Rix.
- Lloyd, David M.
1992 The poets of the princes. In: Jarman and Rees Hughes (eds.), 157–188.
- Lorrio, Alberto J. and Gonzalo Ruiz Zapatero
2005 The Celts in Iberia: An Overview. *E-Keltoi: Journal of Interdisciplinary Celtic Studies* 6: 167–254.
URL: http://www4.uwm.edu/celtic/ekeltoi/volumes/vol6/6_4/lorrio_zapatero_6_4.html

- Mac Eoin, Gearóid
 1989 Orality and Literacy in some Middle-Irish King-Tales. In: Stephen Tranter and Hildegard L. C. Tristram (eds.), *Mündlichkeit und Schriftlichkeit in der frühen irischen Literatur*. Tübingen: Narr, 149–183.
- Maier, Bernhard
 1994 *Lexikon der keltischen Religion und Kultur*. Stuttgart: Kröner.
- Mallory, James P.
 1997 Beaker Culture. In: James P. Mallory and Douglas Q. Adams (eds.), *Encyclopedia of Indo-European Culture*. London: Fitzroy Dearborn, 53–55.
- Mallory, James P., John T. Koch, Joseph F. Eska, Miguel Valéro, and Blanca Maria Prósper
 2014 Discussion of Tartessian as a Celtic language. *Journal of Indo-European Studies* 42/3–4: 332–524.
- Marichal, Robert
 1988 *Les graffites de la Graufesenque*. Paris: Centre National de la Recherche Scientifique.
- Markey, Thomas and Bernard Mees
 2002 Prestino, patrimony and the Plinys. *Zeitschrift für Celtische Philologie* 53: 116–167.
- McCone, Kim R.
 1996 *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change*. (Maynooth Studies in Celtic Linguistics 1). Maynooth: Department of Old Irish, St. Patrick's College.
- McManus, Damian
 1983 A Chronology of the Latin Loanwords in Early Irish, *Ériu* 34: 21–72.
- McManus, Damian
 1991 *A Guide to Ogam*. (Maynooth Monographs 4). Maynooth: An Sagart.
- Meid, Wolfgang
 1992 *Gaulish inscriptions. Their interpretation in the light of archaeological evidence and their value as a source of linguistic and sociological information*. Budapest: Archaeolingua Alapítvány.
- Meid, Wolfgang
 1993 *Die erste Botorrita-Inschrift. Interpretation eines keltiberischen Sprachdenkmals*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Meid, Wolfgang
 1994 *Celtiberian Inscriptions*. Budapest: Archaeolingua Alapítvány.
- Meid, Wolfgang
 1997 *Die keltischen Sprachen und Literaturen. Ein Überblick*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Meyer, Kuno
 1909 *Rawlinson B 502: a collection of pieces in prose and verse in the Irish language compiled during the eleventh and twelfth centuries*. Oxford: Clarendon Press.
- MLH I: Jürgen Untermann
 1975 *Monumenta Linguarum Hispanicarum I. Die Münzlegenden*. Wiesbaden: Reichert.
- MLH IV: Jürgen Untermann and Dagmar S. Wodtke
 1997 *Monumenta Linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert.
- MLH V.1: Dagmar S. Wodtke
 2000 *Monumenta Linguarum Hispanicarum V.1. Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Reichert.
- Morandi, Alessandro
 2004 *Celti d'Italia*. Vol. II. Rome: Spazio Tre.
- Moscatti, Sabatino, Otto H. Frey, Venceslas Kruta, Barry Raftery, and Miklos Szabo (eds.)
 1991 *The Celts – Catalog to the Legendary Exhibition in Venice*. Milan: Rizzoli.

- Murphy, Gerard and Brian Ó Cuív
 1971 *The Ossianic Lore and Romantic Tales of Medieval Ireland: Fianaíocht agus Rómán-saíocht*. 2nd edn. rev. by Brian Ó Cuív. Cork: Mercier.
- Neumaier, Joachim
 1995 Eine iberische Fundgruppe Südwestdeutschlands in der Späthallstattzeit. *Germania* 73: 495–503.
- O’Rahilly, Cecile
 1967 *Táin Bó Cuailnge from the Book of Leinster*. Dublin: Dublin Institute for Advanced Studies.
- O’Rahilly, Cecile
 1976 *Táin Bó Cuailnge Recension I*. Dublin: Dublin Institute for Advanced Studies.
- O’Rahilly, Thomas F.
 1946 *Early Irish History and Mythology*. Dublin: Dublin Institute for Advanced Studies.
- Pauli, Ludwig
 1971 Studien zur Golasecca-Kultur. *Mitteilungen des deutschen Archäologischen Instituts. Römische Abteilung*. (Ergänzungsheft 19). Heidelberg: Kerle.
- Pauli, Ludwig
 1992 Die historische Entwicklung im Gebiet der Golasecca-Kultur. In: Luciana Aigner-Foresti (ed.), *Etrusker nördlich von Etrurien. Etruskische Präsenz in Norditalien und nördlich der Alpen sowie ihre Einflüsse auf die einheimischen Kulturen. Akten des Symposions von Wien – Schloß Neuwaldegg 2.–5. Oktober 1989*. Vienna: Österreichische Akademie der Wissenschaften, 179–196.
- Pilch, Herbert
 2007 *Die keltischen Sprachen und Literaturen*. Heidelberg: Winter.
- Prodocimi, Aldo Luigi and Patrizia Solinas
 1991 The Language and Writing of the Early Celts. In: Moscati et al. (eds.), 51–59.
- Prósper, Blanca Maria
 2008 *El bronce Celtibérico de Botorrita – I*. Rome: Serra.
- Richter, Michael
 2005 *Medieval Ireland*. Revised Edition. Dublin: Gill and Macmillan.
- Rivet, Albert L. F. and Colin Smith
 1979 *The Place-names of Roman Britain*. London: Batsford.
- Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
 2001 *Lexikon der indogermanischen Verben: die Wurzeln und ihre Primärstambildungen*. 2nd, corrected edn. Wiesbaden: Reichert.
- Russell, Paul
 1995 *An Introduction to the Celtic Languages*. London: Longman.
- Schmidt, Karl-Horst
 1986 Zur Rekonstruktion des Keltischen. Festlandkeltisches und inselkeltisches Verbum. *Zeitschrift für Celtische Philologie* 41: 159–179.
- Schrijver, Peter
 1995 *Studies in British Celtic Historical Phonology*. Amsterdam: Rodopi.
- Sims-Williams, Patrick
 2006 *Ancient Celtic Place-Names in Europe and Asia Minor*. Oxford: Wiley-Blackwell.
- Sims-Williams, Patrick
 2007 Common Celtic, Gallo-Brittonic and Insular Celtic. In: Pierre-Yves Lambert and Georges-Jean Pinault (eds.), *Gaulois et celtique continental*. Geneva: Droz, 309–354.
- Skilton, St John
 2004 *The Survey of Scottish Gaelic in Australia and New Zealand*. Thèse de Doctorat Univ. Fribourg (Switzerland). <http://ethesis.unifr.ch/theses/downloads.php?file=SkiltonS.pdf>
- Solinas, Patrizia
 1995 Il Celtico in Italia. *Studi Etruschi* 60: 311–408.

- Stary, Peter F.
1994 *Zur eisenzeitlichen Bewaffnung und Kampfweise auf der iberischen Halbinsel*. 2 vols. Berlin: De Gruyter.
- Stifter, David
2008 www.univie.ac.at/indogermanistik/download/CE/Stifter/oldcelt2008_1_general.pdf
- Stifter, David: Gaulish = http://eprints.nuim.ie/3740/1/DS_Gaulish.pdf
- Stifter, David, Martin Braun, and Michela Vignoli with the assistance of and contributions by Anna Adaktylos, Chiara Dezi, Eva Lettner, Corinna Scheungraber, and Marcel Schwarz
Lexlep = *Lexicon Leponticum*. www.univie.ac.at/lexlep/wiki/Main_Page
- Stokes, Whitley and John Strachan
1903 *Thesaurus Palaeohibernicus*. 2 vols. Cambridge: University Press.
- Stüber, Karin
2006 Sprachliche Spuren der Kelten in der Schweiz. *Helvetica Archaeologica* 145: 2–22.
- Thurneysen, Rudolf
1921 *Die irische Helden- und Königsage bis zum siebzehnten Jahrhundert*. Halle/Saale: Niemeyer.
- Tierney, James J.
1959/1960 The Celtic Ethnography of Posidonius. *Proceedings of the Royal Irish Academy*. Section C: Archaeology, Celtic Studies, History, Linguistics, Literature 60: 189–275.
- Tomaschitz, Kurt
2002 *Die Wanderungen der Kelten in der antiken literarischen Überlieferung*. (Österreichische Akademie der Wissenschaften. Philosophisch-historische Klasse. Mitteilungen der prähistorischen Kommission vol. 47). Vienna: Österreichische Akademie der Wissenschaften.
- Trachsel, Martin
2004 *Untersuchungen zur relativen und absoluten Chronologie der Hallstattzeit*. (Universitätsforschungen zur prähistorischen Archäologie 104). 2 vols. Bonn: Habelt.
- Uhlich, Jürgen
1999 Zur sprachlichen Einordnung des Lepontischen. In: Stefan Zimmer, Rolf Ködderitzsch, and Arndt Wigger (eds.), *Akten des zweiten deutschen Keltologen-Symposiums (Bonn, 2.–4. April 1997)*. Tübingen: Niemeyer, 277–304.
- Vendryes, Joseph
1902 *De hibernicis vocabulis quae a Latina lingua originem duxerunt* [On the Irish words which originated from the Latin language]. Paris: Klincksieck.
- Warntjes, Immo
2010 *The Munich Computus: text and translation. Irish computistics between Isidore of Seville and the Venerable Bede and its reception in Carolingian times*. Stuttgart: Steiner.
- Watkins, Calvert
1999 Two Celtic Notes. In: Peter Anreiter and Erzsébet Jerem (eds.), *Studia Celtica et Indogermanica. Festschrift für Wolfgang Meid zum 70. Geburtstag*. Budapest: Archaeolingua Alapítvány, 539–543.
- Welch, Robert (ed.)
1996 *The Oxford Companion to Irish Literature*. Oxford: University Press.
- Whatmough, Joshua
1970 *The Dialects of Ancient Gaul*. Cambridge, MA: Harvard University Press.
- Williams, Sir Ifor (ed.)
1938 *Canu Aneirin* [The poems of Aneirin]. Cardiff: University of Wales Press.
- Williams, Sir Ifor (ed.)
1982 [1930] *Pedeir Keinc y Mabinogi* [The four branches of the mabinogi]. Cardiff: University of Wales Press.
- Williams, Sir Ifor and John E. Caerwyn Williams (eds.)
1968 *The Poems of Taliesin*. Dublin: Dublin Institute for Advanced Studies.

- Zeidler, Jürgen
2003 A Celtic script in the eastern La Tène Culture? *Études Celtiques* 35: 69–132.
- Ziegler, Sabine
1994 *Die Sprache der altirischen Ogam-Inschriften*. (Historische Sprachforschung Ergänzungsheft 36). Göttingen: Vandenhoeck & Ruprecht.
- Ziegler, Sabine
1996 Gallien, Namenkunde. In: Heinrich Beck, Heiko Steuer, and Dieter Timpe (eds.), *Reallexikon der Germanischen Altertumskunde* 10: 369–370.
- Ziegler, Sabine
2000 Keltische Sprachen. In: Hubert Cancik and Helmut Schneider (eds.), *Der Neue Pauly. Rezeptions- und Wissenschaftsgeschichte*. Vol. 14. Stuttgart: Metzler, 874–878.
- Ziegler, Sabine
2005 Keltiberisch *sua* „Vollversammlung“. *Historische Sprachforschung* 117 [2004]: 292–297.
- Ziegler, Sabine
2012 Die indogermanischen Sprachen nebst einer kurzen Darstellung der Ziele und Inhalte der Indogermanistik. In: Bettina Bock, Susanne Zeilfelder, and Sabine Ziegler. *Deutsche Wortfeldetymologie in europäischem Kontext* Vol. 1: *Der Mensch und sein Körper*. Wiesbaden: Reichert, 316–334.

Bernd Vath, Würzburg (Germany)
Sabine Ziegler, Jena/Leipzig (Germany)

68. The phonology of Celtic

- | | |
|------------------------------------|------------------------------------|
| 0. Preliminaries | 3. The most important developments |
| 1. From PIE to Proto-Celtic | of British in approximate order |
| 2. The most important developments | 4. References |
| of Irish in approximate order | |

0. Preliminaries

This chapter outlines the phonological development of Celtic (Clt.) from PIE to Proto-Celtic (PC; the latest reconstructable unitary ancestor of all Celtic languages) and some important Common Celtic developments (CC; after the break-up of Celtic, but wave-like affecting all Celtic languages) in detail, and sketches the most salient developments leading to the medieval Insular Celtic (IC) languages. It is not easy to distinguish in each case between a feature of Proto-Celtic and a Common Celtic development (the terms are used differently here from Sims-Williams 2007: 309–311), or whether a change affected only a section of the Celtic speaking world. The latter distinction is particularly relevant in the case of Celtiberian, which did not participate in several developments that affected all other Celtic languages and that had been considered as constitutive for Celtic before the decipherment of Celtiberian.

<https://doi.org/10.1515/9783110523874-023>

- Zeidler, Jürgen
2003 A Celtic script in the eastern La Tène Culture? *Études Celtiques* 35: 69–132.
- Ziegler, Sabine
1994 *Die Sprache der altirischen Ogam-Inschriften*. (Historische Sprachforschung Ergänzungsheft 36). Göttingen: Vandenhoeck & Ruprecht.
- Ziegler, Sabine
1996 Gallien, Namenkunde. In: Heinrich Beck, Heiko Steuer, and Dieter Timpe (eds.), *Reallexikon der Germanischen Altertumskunde* 10: 369–370.
- Ziegler, Sabine
2000 Keltische Sprachen. In: Hubert Cancik and Helmut Schneider (eds.), *Der Neue Pauly. Rezeptions- und Wissenschaftsgeschichte*. Vol. 14. Stuttgart: Metzler, 874–878.
- Ziegler, Sabine
2005 Keltiberisch *sua* „Vollversammlung“. *Historische Sprachforschung* 117 [2004]: 292–297.
- Ziegler, Sabine
2012 Die indogermanischen Sprachen nebst einer kurzen Darstellung der Ziele und Inhalte der Indogermanistik. In: Bettina Bock, Susanne Zeilfelder, and Sabine Ziegler. *Deutsche Wortfeldetymologie in europäischem Kontext* Vol. 1: *Der Mensch und sein Körper*. Wiesbaden: Reichert, 316–334.

Bernd Vath, Würzburg (Germany)
Sabine Ziegler, Jena/Leipzig (Germany)

68. The phonology of Celtic

- | | |
|------------------------------------|------------------------------------|
| 0. Preliminaries | 3. The most important developments |
| 1. From PIE to Proto-Celtic | of British in approximate order |
| 2. The most important developments | 4. References |
| of Irish in approximate order | |

0. Preliminaries

This chapter outlines the phonological development of Celtic (Clt.) from PIE to Proto-Celtic (PC; the latest reconstructable unitary ancestor of all Celtic languages) and some important Common Celtic developments (CC; after the break-up of Celtic, but wave-like affecting all Celtic languages) in detail, and sketches the most salient developments leading to the medieval Insular Celtic (IC) languages. It is not easy to distinguish in each case between a feature of Proto-Celtic and a Common Celtic development (the terms are used differently here from Sims-Williams 2007: 309–311), or whether a change affected only a section of the Celtic speaking world. The latter distinction is particularly relevant in the case of Celtiberian, which did not participate in several developments that affected all other Celtic languages and that had been considered as constitutive for Celtic before the decipherment of Celtiberian.

<https://doi.org/10.1515/9783110523874-023>

Celtic consists of four separate branches, the internal grouping of which is not sufficiently clear. The first branch consists of Celtiberian (Cib.). The second branch consists of Gaulish (G) and Lepontic (Lep.). Together, these two branches make up Continental Celtic (Cont. Clt.). The third branch is British (B), the most important stages of which are, for present purposes, Middle Welsh (MW), Middle Breton (MBr.) and Middle Cornish (MCo.). The fourth branch is Goidelic (Gd.) whose most important representative is Old Irish (OIr.). The latter two branches make up Insular Celtic (IC). The terms Continental Celtic and Insular Celtic are used in a strictly geographical or chronological sense. Because of the relatively sharp chronological dividing line (Continental Celtic being attested in antiquity from the middle of the 1st millennium BCE to the middle of the 1st millennium CE, Insular Celtic being attested from the middle of the 1st millennium CE onward), one can also speak of Old Celtic and Neo-Celtic.

The fundamental handbooks for the historical phonology of the Celtic languages are Jackson (1953), Schrijver (1995), McCone (1996), Sims-Williams (2003), as well as the relevant chapters in Ternes (2011). Pedersen (1909 [1976]) and Thurneysen (1946) are in many respects outdated, but offer insights into specific problems. Concise descriptions can be found in Schumacher (2004: 115–138) and Matasović (2009: 4–11).

1. From PIE to Proto-Celtic

Vowels and consonants went largely separate ways, affecting each other only rarely in PC sound changes. Only the development of syllabic resonants and of laryngeals introduced new vowels into the language and changed the relative distribution of vowels. This situation changed drastically in the emergence of the Neo-Celtic Languages, where vowels and consonants show heavy interaction.

1.1. Obstruents

The palatal series merged with the velar series (* \hat{K} > * K ; * \hat{kun} - > * kun - > G *cuno*-, OIr. *con*-, MW *kyn*- ‘dog’; * \hat{g}^huHtus > G *gut*-, OIr. *guth* ‘voice’).

Newly arisen sequences of velar + labial glide merged with the labiovelars (* h_1ek^uo - > * ek^u - > G *epo*-, OIr. *ech* ‘horse’).

Labiovelars were delabialized in various contexts, before * i (* K^u > * K / $_i$ PIE * nig^u i - eti > * $nigjeti$ → OIr. *nigid* ‘washes’; * d^heg^uho - > * $degio$ - > MW *de* ‘burning’), before * n (* h_2og^unos > * $ognos$ > OIr. *úan*, W *oen* ‘lamb’), and before * u (* $\hat{k}u\bar{o}(n)$ > * $k^u\bar{u}$ > * $k\bar{u}$ > OIr. *cú*, MW *ki* ‘dog’). The remaining instances of * g^u became * b (* g^uous > * $bous$ > Cib. *bou*-, OIr. *bó* ‘cow’; * $g^uHd^heh_1$ - ‘to bestow praise’ → PC **bardos* > G *bardus*, OIr. *bard*, MW *bardd* ‘bard’). Only after this change did PIE * g^uh become PC * g^u , thus filling a slot in the phonological system that had briefly been empty.

If there was ever an opposition between * T and * T^h , there is no reflex of it.

For the developments affecting the voiced obstruents posterior to * g^u > * b , two alternative scenarios of differing complexity are conceivable. The simple scenario is that all voiced aspirates lost their aspiration (* D^h > * D), and that at a later stage (but still affecting all Celtic languages) intervocalic voiced obstruents were “lenited”, i.e. they

became the corresponding voiced fricatives ($*D > *Ḍ/V_V$). Since the two classes were in complementary distribution, they were mere allophones within the system.

A possible alternative, albeit more complex scenario is the following. The voiced aspirates became corresponding fricatives ($*D^h > *Ḍ$). This change coincided partly with the developments in Italic, Venetic, and Germanic, and probably Lusitanian. As in Germanic, these fricatives were delenited in certain contexts, most notably after homorganic nasals ($*Ḍ > *Ḍ/N_$), a natural phonetic process that led to the phonetic merger with voiced stops in this position. Word-internally, the voiced fricatives occurred preponderantly between vowels. Combined with a natural tendency for “lenition” in this position, the voiced obstruents came under pressure to become fricativized in such contexts, too ($*D > *Ḍ/V_V$). At this stage, word-internal voiced fricatives and stops had merged in almost all positions and were found in complementary distribution between vowels and after nasals. In theory, the two series would have contrasted phonemically in word-initial position, but even here sandhi-effects could have obliterated the contrast. The only stable phonemic contrast existed word-internally after *l* and *r*. Eventually, this basis may have proven too small to maintain the phonetic distinction between the two series, which in consequence merged even after liquids. Regular allophonic variation in morphological processes (inflection, derivation, verbal stem formation, especially of nasal-infixed verbs) further bolstered the speakers’ awareness of the allophonic relationship between the two sets of sounds. The advantage of this more complex scenario over the simpler one is that it embeds the first stage of the Celtic development of the PIE voiced aspirate series in a wider phonetic trend of Western IE languages. In any case, after either scenario voiced stops and voiced fricatives were in allophonic distribution.

A handful of words display sporadic devoicing of initial voiced stops or voiced aspirates: $*ḡ^h eng^h-$ $> *keng-$ ‘go, step’, $*dḡḡ^h ueh_2 \rightarrow *tang^h āt-$ ‘tongue’, $*d^h reg^h/g^h-$ \rightarrow *tre/ag-* ‘run (?)’.

PIE $*p$ underwent a series of very diverse changes that in the long term eliminated the sound from the system. It suffered its first minor loss of ground word-initially by distance assimilation to $*k^u$ ($*\#p...k^u > \#k^u...k^u$; $*penk^ue > *k^u enk^ue >$ OIr. *cóic*, MW *pypm* ‘5’, cf. G *pimpetos* ‘5th’).

Where PIE $*p$ had survived the change to $*\chi$ before $*s$ and $*t$ (see below), it was first weakened to a voiceless bilabial fricative $*\varphi$. It is perhaps attested as $\langle v \rangle$ in Lep. *uvamokozis* $< *upm̄h_2o-$ ‘highest’. This sound was subject to further transformations: before liquids, it was voiced to β ($*\varphi > \beta/_\{r,l\}$; $*duej-plos > *duej-blos >$ OIr. *diabul* ‘double’; $*pipr̄h_3seti > *\varphi i\varphi r̄āseti > *\varphi i\beta r̄āseti >$ OIr. *ebraid* ‘will grant’), thereby adding to the number of “lenited” consonants in the language; after non-front vowels and before *n*, it became the bilabial glide $*u$ ($*\varphi > u/_\{a,o,u\}_n$; $*supnos > *su\varphi nos > *suunos > *sounos >$ OIr. *súan*, W *hun* ‘sleep’; $*kḡ_2pnos > *ka\varphi nos > *kaunos >$ OIr. *cúan* ‘haven’), but it was apparently lost without a trace if the vowel was *e* ($*tepnēs-$ ‘heat’ \rightarrow OIr. *teine* ‘fire’).

After *s*, $*p$ survived as a marginal phoneme (perhaps an allophone of *b*; less likely $*s\varphi$) into the historic period (G. *Bratuspantium*). This sequence eventually merged with $*su$ in Goidelic and with $*f$ in British.

In all other contexts, $*\varphi$ ultimately became \emptyset , probably passing through an intermediate stage $*h$. This is perhaps the stage at which preceding liquids were geminated ($*Rh > *RR$; $*\sqrt{telp}-$ ‘make space’ \rightarrow OIr. *tella* ‘finds room’; $*stirpāko-$ $>$ OIr. *serrach* ‘young animal’; $*serpeh_2 >$ OIr. *serr*, W *serr* ‘sickle’). It is very doubtful whether this $*h$

is actually attested as such. The <h> in Lat. *Heluetii* (< **pelu-* ‘many’) is likely to be graphical or ornamental (3rd c. BCE; Etruscans heard no *h*, to judge from the name *eluveitie*), as are various spellings with initial <h> in OIr. (pace Schrijver 1997). Perhaps the aspiration in Ἐρκύνιον ὄρος (< **perku*^o) may lay claim to reflecting a trace of PIE **p*.

**VpV* led first to hiatus sequences **V.V*. If one of the vowels was a high vowel that came into word-initial or post-vocalic position, it became the corresponding glide (e.g., **uper* > **u.er* > **uer* > G *uer-*, Cib. *uer-*, OIr. *for*, MW *gwar*, *gor* ‘over, on’; **epirom* > **e.irom* > **eïrom* > OIr. *iar*^N ‘after, along’; **upelo-* > OIr. *fel* ‘bad’). Otherwise, the hiatus was retained, e.g., **nepot-* > **ne.ot-* > OIr. gen. *niad*, B pl. *nied* ‘nephew’; **ka-pero*^o > **ka.ero*^o > OIr. gen. *cáerach* ‘sheep’, MW *caeriwrch* ‘roebuck’.

The resulting CC system was a rather imbalanced one that contained the three voiceless obstruents *t*, *k*, *kʰ*, the four voiced obstruents *b*, *d*, *g*, *gʰ* and their lenited allophones *β*, *δ*, *γ*, *γʰ* between vowels, perhaps even across word boundaries. Phonetically, the contrast between the D-series and the T-series may actually have been that between “lenis” voiceless consonants and “fortis” aspirated consonants. Some evidence for this is provided by processes that display aspiration in the further development of the T-series in the British languages (nasalization, spirantization), by the phonetics of the modern Celtic languages, and by spelling variations in Gaulish.

1.2. Developments, especially simplifications of obstruent clusters

Several independent changes conspired to reduce the number of phonotactically tolerated consonant clusters, and to reduce the average number of phonemes acceptable in clusters.

Tautosyllabic sequences of dental + velar (PIE “thorn”) were metathesized (**TK* > **KT*; e.g., **d^hg^homios* ‘earthly’ > **gdonios* > **γdonios* > G *-χtonion*, OIr. *duine*, W *dyn* ‘human’; but **h₁rud^hk₂ieh₂* ‘reddening’ > **rukkīā* > OIr. *ruccae* ‘blushing, shame’). In clusters of dentals, the sibilant that had arisen subphonemically in this position already in PIE probably became phonemicized, i.e. **-DD-* > **-DsD-* (D = dental), before it underwent further changes (see below).

All non-dental obstruents, including **φ* < **p*, became **χ* before **t* or **s* (*T*_[-dent] > *χ*/_{-{t/s}}). Perhaps a very early dialectal split is reflected by the treatment of the complex cluster **Ktl* which becomes **χtl* in Gaulish and British, but **kk*l in Goidelic, perhaps with metathesis **tkl* at an intermediate stage (PC **anegt_lom* > G *anex_{tl}o-*, OIr. *anacul* ‘protection’; **uok^utlom* > OIr. *focul* ‘word’, W *gwaethl* ‘dispute’). However, the Proto-Goidelic cluster could conceivably have arisen from an erstwhile **χtl*.

Clusters of three or more consonants with a sibilant in the middle underwent rather complex changes; not all details are sufficiently clear and some of the regular outcomes may have been obscured by subsequent analogical restructurings. Stops and probably also nasals were lost before sequences of **s* + another consonant (for dentals see below), or the stop was first assimilated to the *s* with subsequent simplification of the geminate sibilant (**TsC* > **sC*). Velars and labials had already become **χ* in a previous step, but this is not indicated in the following reconstructions: e.g., **nHds^hketi* > **nasketi* > OIr. *nascaid* ‘binds’; **syēks(d)koms* > **syeskoms* > OIr. *sesca* ‘60’; **komsk^uom* > **konsk^u-om* > **kosk^uom* > OIr. *cosc*, MW *cosp* ‘reproving, punishment’; **tep^ustus* > **testus* > OIr. *tess*, MW *tes* ‘heat’; **reǵstrom* > **reǵstrom* > MW *rwyrstr* ‘obstacle’, **segsm^uŋ* >

sesman* > **semman* > OIr. *seimm*, MW *hem* ‘rivet’. The simplification occurred before the resolution of syllabic **ʃ* and **ʎ* (cp. **dʎg^hsketi* > **dʎsketi* > **dalsketi* > **dalketi* > Brit. *dalch-* ‘hold, keep’); the case of **ŋ* is not so clear (g^hŋg^hsmŋ* >> **kangsman* > **kāmmān* (with a nasalized **a*? see 1.6.1 below) > Cib. *kamanom* ‘path’, OIr. *céimm*, MW *cam* ‘step’).

When the lost consonants were reintroduced by reanalysis or recomposition, the new sequences of **TsC* fed a subsequent rule whereby a fricative, mostly **s* but also **χ*, sandwiched in three- or four-consonant clusters, was lost. Apart from the newly arisen instances of **TsC*, this rule most prominently affected fricatives between a liquid and another consonant, mostly voiceless stops (**LsC* > **LC*, e.g., **tʃstū* > **tarstus* > **tartus* > OIr. *tart* ‘thirst’; **pr^hksketi* > **φarsketi* > **arketi* > OIr. *airc*, W *eirch* ‘asks’; **mēl(h₂)st* > **mīlt* > OIr. *mīlt* ‘ground’; perhaps **ḱ^hsnom* ‘horn’ > MW *carn* ‘hoof’ (Hill 2012); **perk^htos* > **k^herk^htos* > **k^herχtos* > **k^hertos* > W *perth* ‘bush’). An example for secondary **TsC* is **h₃rēg^hst* ‘stretched out’ > **rīst* >> **regst* > **reχt* > OIr. *recht*, MW *-reith* ‘arose’; a special case is PC **orgst* > **ort* > OIr. *ort*, MW *-orth* ‘slew’. Whether at this stage **s* was likewise lost between a stop and a voiced stop, i.e. **TsD* > **DD*, perhaps with voicing of the entire cluster at an intermediate stage, is uncertain. All examples involve the preverbs **eks* and **uts/ups*, and thus *s*-less allomorphs **ek-* and **ut-/up-* could be involved (e.g., PC **eks-bereti* > **egzbereti* > **egbereti* > **ebbereti* > OIr. *epir* ‘says’).

Instead of being parallel to the developments outlined above, the treatment was probably special when the **s* was sandwiched between a dental and a **t*, including instances of **tst* from earlier **Dt* (D = any dental stop). If **Dst* had resulted in ***st*, it would be expected to show up as such in those Celtic languages that (occasionally) retain inherited **st* (Brit. and Cib., Schrijver 1995: 399–430). Since this appears not to be the case, it must be assumed that **Dst* developed differently from **st*, probably > **t^s* (approximative). By the orthographic evidence of G, the result was not just a geminate sibilant, but a new phoneme, perhaps a dental affricate [t^s], for which the term *tau Gallicum* has come to be used. OIr. *ress* ‘was put’ < **krit^so-* < **ḱ^hrdto-* shows that the first element of this sound had obstruent quality when syllabic resonants were resolved. In Gaulish, and certainly in Irish, original **st* merged with **t^s*.

Three-consonant clusters with *s* at the beginning remained as such (**alistro-* > OIr. *ailistar*, MW *elystyr* ‘iris [flower]’).

After the above changes, surviving word-internal **s* assimilated to a following resonant except **r* (**VsR* > **VRR*; **koslos* > **kollos* > OIr. *coll*, MW *coll* ‘hazel’; **pr^hh₃sneh₂* > **φrasnā* > **φrannā* > OIr. *rann*, MW *ran* ‘part, share’; **H₂iosmōj* > **iomūj* > Cib. dative singular relative pronoun *iomui*). Before **r*, it perhaps became **ð* (perhaps passing through a stage **z*; **tisres* > G *tidres*, OIr. *teóir*, MW *teir* ‘3 [fem.]’) (Kim 2008: 160–161). Word-internally, the sequence **Ls* became **LL* (e.g., **ḱ^hrsos* > **karsos* > **karros* > Gallo-Lat. *carrus*, OIr. *carr* ‘cart’; **p^hlseh₂* > **φalsā* > **φallā* > OIr. *all* ‘cliff’).

1.3. Sibilant

Although undergoing heavy changes in clusters (see above), **s* was otherwise largely stable. Possibly pan-Celtic is the dissimilatory loss of **s* at the beginning of a second

syllable if the first syllable also began with *s (*s > Ø/#s(u)V_(R)V; *suesōr > *sue-ūr > G *suioerebe*, OIr. *siur*, MW *chwaer* ‘sister’; *sesloigē > *seloigē > OIr. *selaig* ‘felled’).

*z, the voiced allophone of *s before voiced stops, remained as such in Proto-Celtic. In Gaulish and British, *zd became *tt (e.g., *k^uezdi^o > Gallo-Lat. *pettia* ‘piece’, OIr. *cuít* ‘portion’, MW *peth* ‘thing’). Otherwise, *z fell together with lenited *δ in Insular Celtic.

1.4. Resonants

Syllabic resonants were resolved into sequences of resonant + vowel or vowel + resonant. For the treatment of syllabic resonants before laryngeals, see section 1.5 on laryngeals below. Otherwise, the syllabic liquids *l̥* and *r̥* became *li* and *ri* before non-continuants (i.e., non-sibilant obstruents and nasals): *l̥ > *Li /C_{T,N}; e.g., *k^ur̥tus > *k^uritus > OIr. *cruth*, MW *pryd* ‘form’; *k^ur̥mis > *k^urimis > OIr. *cruim*, MW *pryf* ‘worm’). It is widely believed that *l̥ became *aL before *n, but *l̥neh₂ti > *tlināti > OIr. *tlenaid* ‘takes away’ leaves hardly another option than the explanation chosen here. Furthermore, alternative explanations are available for other alleged cases of *l̥n > *aLn, e.g., MW *carn* ‘hoof’ < *karnom < *k^ur̥snom (Hill 2012). In any case, in all other contexts, all syllabic resonants (liquids and nasals), developed an *a* in front of them (*C_{RC} > *CaRC; *ŋ- > *an- > negative prefix OIr. *an-*, MW *an-*; *dek̥m̥ > *dekam > Cib. *tekametinam*, G *decametos*, OIr. *deich*^N ‘10’; *t̥rstus > *tarstus > OIr. *tart* ‘thirst’; *m̥r̥uos > *maruos > OIr. *marb*, W *marw* ‘dead’; *h₂erh₃u^{r̥} > *aru^{r̥} > OIr. *arbor* ‘corn’). Note that this change happened before the loss of laryngeals (*u^{l̥}Hos > *ualos > G *-ualus*, Early OIr. *-ual* ‘ruler’, *sm̥h₂elis > *samalis > OIr. *samail*, MW *haul* ‘alike’) and before the loss of *s* in certain clusters (*p̥r̥ksketi > *p̥r̥sketi > *p̥arsketi > *p̥arketi > OIr. *-airc*, MW *arch-* ‘asks’).

After the syllabic resonants had been resolved into bisegmental sequences, the resonants remained remarkably stable in Old Celtic, apart from minor assimilations. Nasals assimilated in their place of articulation to the following consonant (*√temk- ‘coagulate’ → *tonke- [ŋk] > OIr. *tocaid*, MW *tynghaf* ‘destine’; *h₃ng^uen [ŋg^u] > *amben [mb] > OIr. *imb*, MBr. *amanen* ‘butter’; *ŋ₂-perish₂eh₂ > *am-pheristā > OIr. *amares* ‘disbelief’). *ln became *ll (e.g., *h₂olnos > *ollos > G *ollo*, OIr. *oll*, MW (h)oll ‘all, great’; *uelnHmon- → *uellamnos ‘ruler’ > G *Vallaunus*, OIr. *Follomain*). For resonants in combination with *s, see 1.2 above. Taken together, these changes introduced a phonemic opposition between single and geminate or long resonants into the system. Word-internally, *m suffered a minor erosion by being lost before *u (*m̥u > *u; e.g., *komuēros > *kouūros > G *Couirus*, MW *cywir* ‘true, truthful’).

Clusters containing a labial segment + *n developed in rather complex manners. First, Proto-Celtic *b became *β word-internally before *n (*Vbn > *Vβn). In a very rare change, *u became the fricative *β before *n (“McCone’s Law”, McCone 1996: 49), e.g. *h₂euh₂-on- >> *aun- >> OIr. *amnair* ‘maternal uncle’; *ph₂ounos > *p̥oβnos > G *-ob/mnos*, W *ofn*, OIr. *omun* ‘fear’. This change was perhaps restricted to contexts where no front vowel preceded the cluster, in order to account for its possible counterexamples W *clun* ‘hip, haunch’ < *k̥le/ounis (unless a loan from Latin *clūnis*), OIr. *cúan* ‘litter’ < *k̥euneh₂ (unless from *keHupneh₂ ‘heap’) and OIr. *búan* ‘permanent’ < *b^heuHnos).

Seeing that it probably occurred only after **a* and **o*, it is obvious that this change must stand in some causal relationship with **φ* > **u/}{a,o}_n* (see above). Eventually **β*, which had arisen in these ways, and inherited **m* fell together before **n* in a nasalized, fricative sound which will be written **μ* (**C_[+lab, +vcd]* > **C_[+fric, +nas]* /*V_n*; e.g. **d^hubnos* > **duμnos* > G *dub/mno-*, *dumno-*, OIr. *domun*, MW *dwfn* ‘world’; **h₂imnos* ‘similar, alike’ > **iμnos* > G *Essib/mnus*, OIr. *emon* ‘twin’ MBret. *effn*, Co. *ewn* ‘even, just’ (?; pers. comm. C. Scheungraber). Due to the lack of pertinent Celtiberian evidence, this change cannot be demonstrated to be Common Celtic. A different, but superficially similar change is probably exclusively Continental Celtic: **m* becomes > **u* before **n* when it is preceded by **a* or **o* (only Celtiberian examples for the latter; e.g., **h₂ekμnos* > **akamnos* > G *acauno-* ‘rock’; **u_lHμnih₂* > **ulamni* > Cib. *launi* ‘wife < *lady’?, dat.sg. of *men*-stems **-omnej* > Cib. infinitive *ambitinkounei*). The Welsh name *Caswallawn*, which looks like another instance, could rather be a loan from G *Vallaunus* ‘ruler’ < **uelamnos* ← **uelnHmon-*.

1.5. Laryngeals

Because of many contradictory or inconsistent developments, the outcomes of the PIE laryngeals is unquestionably the most controversial area of early Celtic diachronic phonology. Laryngeals are the only phonetic subsystem of early Celtic where vowels and consonants interact on a grand scale. No attested Celtic language preserves laryngeals directly, but it is possible that some segmental reflexes may have still been present in Proto-Celtic. This chapter basically follows the ideas of Zair (2010), the latest and most comprehensive study of laryngeal reflexes in Celtic. Even though only the major points of continuing controversy are explicitly highlighted as such, it should be borne in mind that diverging opinions will be found for almost all individual developments presented below.

It is uncontroversial that the laryngeals coloured adjacent PIE **e* (**h₁e* > **e*, **h₂e* > **a*, **h₃e* > **o*) and lengthened a preceding vowel, i.e. fundamentally PIE **e* and **o* (**eh₁* > **ē*, **eh₂* > **ā*, **eh₃* > **ō*; **oH* > **ō*). At the end of words, laryngeals lengthened all vowels (**CVH#* > **CṼ#* and **CIH#* > **Cī#*). **h₃* appears to have voiced a voiceless consonant already in PIE, as evidenced by Skt. *pībati*, Lat. *bibit* (with analogical voicing of the reduplication), and G *ibetis*, OIr. *ibid* ‘drinks’ (**piph₃eti* > **φibeti*). These changes took place before any other Proto-Celtic changes affected the vowels, and the laryngeals were consequently lost.

Laryngeals behaved differently depending on their position in the word (word-initially, first syllable, non-initial syllables). These differences may be relevant to the question of the position of the Proto-Celtic accent. Word-initially, the most plausible scenario is that **h₂* vocalized as **a* before liquids (**h₂rtko-* > **arχto-* > G *art-*, OIr. *art* ‘bear’), but that otherwise all laryngeals in initial position were lost without leaving observable traces. This loss occurred before syllabic resonants were resolved into sequences of resonants + vowels, thus feeding the input for the latter development (*#h₁(₁₃₂)RC* > *#rC*; there are only a few examples: **h₁r^geti* > **rigeti* → OIr. *rigaid*, *regaid* ‘will go’; uncertain **h₁lmo-* > **limo-* > MIr. *lem* ‘elm-tree’).

In the first syllable between two stops or **s*, laryngeals were vocalized as **a* (*#CHC* > *#CaC*; e.g., **ph₂tēr* > **φatīr* > G *atir*, OIr. *aithir* ‘father’; **sh₂gjeti* > **sagjeti* > OIr.

saigid seeks'; **d^hsnos* 'belonging to the religious sphere' > G *dannos* 'official person [?]', and the same applies when the first consonant is a resonant (#*RHC* > #*RaC*; **mh₂tis* 'ripe' > OIr. *maith*, MBret. *mat* 'good'; perhaps **mh₂ntih₂* > **mantī* > OIr. *méit* 'greatness, magnitude', MW *meint* 'size, amount' shows the same treatment). The evidence is too uncertain to decide whether #*IHC* gave #*ĪC* or #*ĪaC* (or both, depending on the language or the laryngeal).

The reflexes of laryngeals after a consonant + a semisyllabic segment are among the most controversial chapters of Celtic historical phonology. The regular result of **CVHC* (as well as **CVHCC*, the environment of the alleged IE "Wetter-Regel") and **CIHC* (including **CIHC* < **CHIC* via metathesis) was very probably **CVC* and **CĪC* (**seh₂lom* > **sēlom* > **sīlom* > OIr. *sil*, MW *hil* 'seed, race'; **ieh₂tus* 'going, passage' > **iātus* > OIr. *áth* 'ford'; **k^urih₂tos* > **k^urītos* > G *-pritom*, OIr. *·críth* 'was bought'; **d^huh₂nom* > **dūnom* > G *-dunum*, OIr. *dún*, MW *din(as)* 'hill-fort'; **seh₂ul-* 'sun' → **sh₂ulis* > **suh₂lis* > **sūlis* > OIr. *súil* 'eye'). However, the number of laryngeal reflexes that do not conform to these general rules is considerable. For the not infrequent examples of **CIHC* > **CĪC* without lengthening of the preceding syllabic nucleus (e.g. **uiHros* > **uiros* > Cib. *uiros*, OIr. *fer*, MW *gwr* 'man'; **ġ^huHtus* 'calling' > **gutus* > G *gutu-*, OIr. *guth* 'voice'), one has to reckon with some sort of shortening rule, often subsumed under the name of Dybo's Rule (Dybo 1961). In its original formulation, Dybo's Rule referred to a shortening of long vowels in pretonic syllables in several IE branches. Perhaps the apparent shortening is rather to be viewed as a rule of laryngeal loss in such contexts. One concern with the theory is that often the pre-Celtic position of the accent cannot be determined, so the proof of the validity of Dybo's Law is circular. Apparent laryngeal-less allomorphs of roots could also have been abstracted from other (e.g., prevocalic) contexts where the laryngeals would have been lost regularly.

The development of laryngeals in sequences of the type **CRHC* is even more disputed. One theory (Matasović 2009: 5) holds that the erstwhile result was **CRaHC*. In pretonic syllables, the laryngeal was then lost without a trace according to Dybo's Law and a short vowel remained, i.e. **CRāC*. Elsewhere, the laryngeal was lost with compensatory lengthening, resulting in **CRāC*. According to a different theory (Schrijver 1995: 168–191), **CRHT* (*T* = voiceless obstruent) became **CRāT*, but all other cases of **CRHC* developed to **CRāC*; instances of **CRāT* have to be explained as full-grade formations. De Bernardo Stempel (1987: 40–43) suggests that **CRHCV* gave **CRāCV*, while **CRHCC* gave **CRāCC*. Another theory holds that all **CRHC* became **CRāC*, but that apparent results of the type **CRāC* are due to analogical (i.e. morphologically motivated) shortenings (McCone 1996: 52–53) or "morphological zero grades" (Joseph 1982: 54). Isaac (2007: 21–59) proposes a complex theory of laryngeal developments whereby the absence or presence of laryngeal reflexes is governed by suprasegmental phonetic features. However, his theory is disproven by many counterexamples (see Stifter 2011: 13–14). Finally, Zair proposes the following rules: **C₁RHC₂* gave **CRāC* except where *C₁* was a continuant (including **φ* < **p* which implies that the laryngeals were still present at an early stage of Proto-Celtic!) and *C₂* a stop, in which case a short vowel resulted, i.e. **SRHT* > **SRāT*. When a consonant cluster followed the laryngeal, the outcome seems to be a short vowel, i.e. **CRHCC* > **CRāCC*. Ultimately, all theories are aporistic in that they are unable to account for all the forms without internal contradiction. Unsatisfying as it may sound, there may be no grand unified theory of Celtic laryngeal developments. It seems that no theory works without taking arbitrary recourse

to at least some sort of morphologically peculiar or innovatory formations, or without invoking some version of Dybo's Rule. Perhaps the attested state of affairs is also partially the result of complex prehistoric borrowings and interferences between different pre-Celtic or Common Celtic dialects, or reflects different sandhi variants.

The outcomes of laryngeals are different in non-initial syllables, where they seem to have been less stable. When followed by a vowel, a sequence of resonant plus laryngeal behaved like a syllabic resonant (*CRHV > *CrV > *CaRV; e.g., *u_lHos > *-ualos > G -ualus, Early OIr. -ual 'ruling'; *trh₂es > OIr. tar, dar 'across'). When a high vowel preceded, a homorganic glide developed (*IHV > *IIV; e.g., *priHos > *φrijos > G rio- (?), MW rydd 'free'; *suh₁eti > *suu₁eti > *sou₁eti > OIr. soid 'turns', MW amheuaif 'I doubt, disagree'; *h₂iūH₁kos > *iūuankos > G Iouinco-, OIr. oac, MW ieuanc 'young'). *VIHV may have given *VIIV, but the evidence is very thin (*b^hoiHos > *boiios > G Boii; *h₂reiHom > *-reijom > OIr. díre, MW dirwy 'fine, honour-price').

Laryngeals were lost before *i, in initial and non-initial syllables alike, i.e. *CRHiV > *Cr_iV and *CIHiV > *C_iV (e.g., *k_lh₁io- > *k_lio- → *kaliākos > OIr. cailech, MW keilyawc 'cock'; *b^huH₁eti > *bui₁eti > *b(u₁)i₁eti > G biietutu, Cib, bionti (?), OIr. biid, MW byd 'is [wont to be]') and *VRHiV > *VR_iV (e.g., *h₂erh₃eti > *ari₁eti > OIr. airid, W arddaf 'plough'). The development before *u is rather unclear (*CRHuV > *CR_uV and *CIHuV > *C_uV according to Zair, but the evidence is limited and ambiguous). In non-initial syllables, laryngeals may have been lost before *u, but retained before *u₁ (e.g. *selh₁ueh₂ > *seluā > G -selua, OIr. selb, MW helw 'possession'; *tenh₂u₁os > *tenau₁os > *tanau₁os > OIr. tanae, MW teneu 'thin'). This is in accordance with a more general rule: Zair suggests that laryngeals were lost in non-initial syllables before single stops, n, or u (*VCH{T,n,u} > *VC{T,n,u}; e.g., *uerH₂geh₂ > *uergā > OIr. ferg 'anger, wrath'; h₃reiHnos > *reinos > Gallo-Lat. Rhenus 'Rhine', OIr. rían 'ocean'), but were vocalized as *a before the other resonants and before two consonants (*VCHR_(-n) > *VcaR_(-n) and *VCHCC > *VCaCC; e.g. *telh₂mō(n) 'bearer' > *telamū > *talamū > OIr. talam 'earth'; *h₃reiHtros > *reiatros > OIr. riāthor, MW raeadyr 'torrent, waterfall'). Not a lot of the evidence is straightforward, though, and other factors, like De Saussure's Law or morphological restructurings, may also have played a role.

In second position of compounds, as well as perhaps in reduplicated formations, laryngeals seem to have been lost (e.g., *-g_hh₁os 'conceived by' > *-gnos > OIr. diminutive suffix -án, -én; cf. also the divergent treatment of *g_hh₃tos > *gnātos > OIr. gnāth, MW gnawt 'usual', but in second position *-gnātos > MW yn₁gnat 'magistrate, judge').

1.6. Vowels

1.6.1. Short vowels

The short vowels inherited from PIE were augmented by those that arose from the changes affecting syllabic resonants and laryngeals (see 1.4–1.5 above). The instances of *a were slightly increased by vowels arising in anaptyxis between word-initial obstruent clusters (*#TT > *#TaT; e.g., *pt₁- 'wing' > *qatar > MW adar 'birds'). Otherwise, the short vowels remained largely stable in Old Celtic. Only minor developments affected them: *e and *u became *o before *u (*teuHteh₂ > *teutā > *toutā > Lep.

toutas, Cib. *toutas*, OIr. *túath*, MW *tud* ‘people’; **supnos* > **suφnos* > **suuinos* > **soy-nos* > OIr. *súan*, W *hun* ‘sleep’), but **u* remained if an **i* followed the **u* (**druuīdes* > OIr. *druid* ‘druids’). The resulting diphthong **ou* is frequently spelled <eu> in Latinate sources. This is unlikely to be more than a purely graphic convention, and does not reflect retained archaic **eu*. In a probably Common Celtic, not Proto-Celtic development, **e* was assimilated to a short **a* of the following syllable in words of more than two syllables if a resonant intervened (**eRa* > **aRa* = “Joseph’s Rule” after Joseph 1982; e.g., **terh₁trom* > **teratrom* > **taratrom* > G *taratrum*, OIr. *tarathar*, MW *taradyr* ‘auger’).

Except for Celtiberian, which provides no evidence in this matter, **e* developed a mid-high allophone **i* before a tautosyllabic nasal, which appears written as <i> and <e> in the Old Celtic sources (e.g., **kenh₁tus* > **kentus* > G *cintu-*, OIr. *cét-*, MW *kynt-* ‘first’). There is evidence for a similar raising of **a* to **æ* before nasals, especially in Gaulish and Old Irish, but these may be independent, *einzel sprachlich* phenomena. Since **a* before **nn* < **sn* was not affected by this change, it introduced a marginal new phoneme into the system. Where **e* had come to stand before **a* or **o* in hiatus, it was raised to **i*.

Word-finally, sequences of the type **-ŷns#* remained unaltered apart from morphologically motivated analogical changes. Loss of the **n* with compensatory lengthening of the vowel was a strictly *einzel sprachlich* phenomenon (Griffith 2005).

Perhaps a marginal opposition between oral and nasal vowels developed very early; otherwise the different treatment of PC **kangsman* > OIr. *céimm* ‘step’ vs. PC **bragsman* > OIr. *broimm* ‘fart’ is hard to explain (assuming that the reduction of the word-internal clusters was a Proto-Celtic development, see 1.2 above).

1.6.2. Long vowels

The most notable development affecting the long vowels is the reduction of vowel qualities from a pre-Celtic system of five long vowels (*ā, ē, ī, ō, ū*) to a minimal one distinguishing only three qualities on the extreme points of the vowel triangle (*ā, ī, ū*). This new system was created in two steps: **ō* became **ū* in final syllables and in monosyllabic words, but **ā* elsewhere (1. **ō* > **ū*/(C)#; 2. **ō* > **ā*; e.g., **kūō(n)* > **kū* > OIr. *cú*, W *ki* ‘dog’; **ui(H)rōns* > **uirūns* > OIr. acc. pl. *fīru* ‘men’; **moh₁ros* > **māros* > G *-marus*, OIr. *már*, MW *mawr*, MBret. *meur* ‘big’; **doh₃nus* > **dānus* > OIr. *dán*, MW *dawn* ‘gift’). In contrast to this, **ē* became **ī* in all positions (e.g., **h₃rēg₁ni₂* > **rīgani* > G *rigani*, OIr. *rigain* ‘queen’, MW *riein* ‘young girl’; **ph₂tēr* > **φatīr* > G *atir*, OIr. *aithir* ‘father’). However, it has also been suggested that Celtiberian retains **ē* unchanged everywhere (Prósper 2007: 80). Since shortening of the type **b^hērst* > **bīrst* > OIr. *birt*, MW *-byrth* ‘carried’ (see below) presupposes **ē* > **ī*, it would then follow from this that the Ostthoff shortening seen in most Celtic languages must have occurred after the branching off of Celtiberian.

Long vowels suffered a series of shortenings in word-final position and before clusters. Even before the Celtic reduction of long vowel qualities, long vowels were shortened in the last syllable before a nasal (**V*: > **ŷ*/*N*#; e.g., thematic gen. pl. **-ōm* > **-om* > G *-on*; hystero-kinetic *n*-stem nom. sg. **-ēn* > **-en+s* > **-ēs* > OIr. *-e*; **g^uoum* >

*g^uōm > *bom > *bon+en > OIr. acc. sg. *boin* ‘cow’; perhaps *g^uenh₂ > *g^uēn > *ben > OIr. *bé^N*. This change apparently did not affect Celtiberian (Eska 2006). After the reduction of long vowel qualities, long vowels underwent Osthoff shortening before clusters of resonants and consonants (*V:RT > *V̄RT; e.g., *kōr(H)dhe > *kārde > *karde > OIr. *(fo)caird* ‘has put’; *sih₂m-de- ‘this’ → *sīnd-os > *sindos > demonstrative/article G *indas*, OIr. *in*, MW *yr*). Apparent instances of shortening before other clusters (typically *TR or *TC) are sporadic (e.g., *puHtros ‘rotten’ > *putros > OIr. *othar* ‘sick, ill’; *sueh₂du-/ueh₂ ‘sweet’ > *suadu-/uā > personal names G *Suadu-*, OIr. *Sadb*) and may be due to other factors like laryngeal loss. They do not attest to a more extended rule of Osthoff shortening (cp. the retained length in *steh₂tleh₂ > *stātā > OIr. *sál*, MW *sawdl* ‘heel’; *kēsreh₂ > *kīsrā > OIr. *cír* ‘comb’).

1.6.3. Vowels + glides

The short *i*-diphthongs **ai*, **eī*, **oi*, which had arisen after the loss of the laryngeals, remained as such in the early Celtic languages. In interaction with **i*, **e* was subject to various changes which may be Common Celtic and need not go back to Proto-Celtic: 1. in non-initial syllables between consonants, **ie* became **i* (e.g., **gabīeti* > **gabiti* > G *gabi*, OIr. *gaibid*); 2. in non-initial syllables, **eīe* became **ī* (e.g., *-*eīes* > *-*īs* > ending of *i*-stem nom. pl. G <-εις> = /-īs/, OIr. *-i*); and 3. tautosyllabic **eī* became **ē* (**deīuos* > **dēuos* > G *deuo-*, OIr. *día*, MW *dwy* ‘god’). The latter is certainly a relatively late Common Celtic development since attestations of the diphthong can be found in the historical record. As for the *u*-diphthongs, **au* and **ou* remained as such, while **eu* and **uu* had probably fallen together with **ou* already in Proto-Celtic (see 1.6.1 above). Among the long diphthongs, one development is notable: word-final *-*āi* perhaps fell together with *-*ai*. All other changes affecting short and long diphthongs are confined strictly to the individual languages.

2. The most important developments of Irish in approximate order

Stress falls on the first syllable

- **-m#* > **-n#*
- **-oi#*, **-ai#* > **-ī#*
- **t^s* > **ss*, **χs* > **ss*
- **o* > **a* (in final syllables and in composition vowels)
- **e* > **i* / ‘C_sV
- **a* > *æ*/NC
- **ænn* > **enn*, **æn#* > **en#*
- **e* > **æ*/γ (except when following syllable contained **i*)
- **CīV* > **CiiV*
- **VC.uV* > **VC.βV* (C = single voiced consonant)
- **ou*, **au* > **ō*
- **-i#* > ∅

- $*i > \emptyset$
- Nasalization: voiceless stops merge with a preceding homorganic nasal and become the corresponding geminate voiced stops ($*NT > *D$), preceding $*a$, $*e$, $*i > *ē$.
- $*-Ṽ_{ns} > *-Ṽ_S$
- Early syncope of post-post-tonic vowels between homorganic consonants.
- Irish lenition: in intervocalic position (word-internally or across word boundaries) or preceded by a vowel and followed by a resonant, single voiceless stops become the corresponding fricatives; in the same positions, $*u$ is lost and single $*s$ becomes $*h$, which is subsequently also lost except word-initially ($*T > *θ$, $*m > *μ$, $*s > *h$, $*u > \emptyset/V_ [R]V$). All final fricatives ($*θ$, $*s$) $> *h$.
- Haplology: $*\#(C)a/oC_xeliC_xV > *\#(C)a/ojC_xV$ (the disappearing consonant must be lenited).
- $*z > *δ$
- Unstressed long vowels are shortened, except in final syllables before $*h\#$.
- Raising: $*e > *i$, $*o > u$ / $\#C_1C_2$ ($C_2 =$ a single voiced consonant or $*nd$, $*ng$, $*mb$); this change only occurs in initial syllables.
- $*æng > *mg$,
- $*æNC > *aNC/_V_{[+ \text{back}]}$, $*æNC > *iNC/_V_{[+ \text{front}]}$
- u -Infection: a short vowel becomes an u -diphthong before a syllable with $*u$; there are several restrictions to this rule.
- 1st Palatalization of single consonants: $*C > *C' / \{ae, \check{e}, i, j, \check{u}\} _ \{ae, \check{e}, \check{i}\}$, with the exception of labials and velars which are not palatalized after $*u$. $*C > *C' / \{\check{a}, \check{o}\} _ \check{i}$. Probably also $*C > *C' / \# _ \{\check{e}, \check{i}\}$.
- $*u > *ü/_C'$
- Lowering: $*i, j > *e$, $*u > *o/_C\{\check{a}, \check{o}\}$
- $*a > *o$, $*i > *u$, $*ē > \bar{e} / \#K^u _$; then $*K^u > *K$
- $*-h\#$ and $*-n\#$ are shifted to the following word and may consequently disappear altogether.
- $*\#u-$ and $*-hu-$ $> *f$
- $*-e\#$ and $*-i\# > *-i$ (strongly palatalizing vowel).
- 2nd Palatalization: $*C > *C' / _ i\#$
- Apocope: $*-Ṽ\# > \emptyset$
- $*-V:\# > *-Ṽ\#$
- Syncope and 3rd palatalization: starting from the beginning of the word, the vowel of every second non-final syllable is deleted (very few exceptions, usually before a cluster of the shape $*θC$). If the syncopated vowel was $*e$, $*i$, $*ü$, all surrounding consonants become palatalized; if the vowel was $*a$, $*o$, $*u$, all surrounding consonants become non-palatalized.
- $*æ > *a/_C'$
- $*æ > *e$
- $*u' > *i$
- Loss of fricatives before resonants: $*VθRV > *V:RV$ (with several exceptions).
- Epenthetic $*ə$ is inserted before resonants that had become syllabic due to syncope or apocope.
- Geminate stops become single stops.
- Reduction in quality of unstressed non-final vowels: $\check{V} > ə$ and $*Ṽu > u$.
- Reduction in quality of proclitic vowels: $e > a$; $u > o$.

- $\bar{o} > \acute{u}a$ (except before g), $\bar{e} > \acute{í}a$ (before non-palatalized consonants).
- $-l\#, -n\# > -ll\#, -nn\#$ in an unstressed syllable that begins with single r, l, n, β or m , maybe also nn (MacNeill's Law).
- Voiceless stops or fricatives are voiced at word-boundaries unless they stand in an accented syllable (McCone's Law).

3. The most important developments of British in approximate order

The very complex developments among the vowels will be only roughly outlined.

- $*-m\# > *-n\#$
- $*-n\# > \emptyset$ in words of more than one syllable.
- $*-oi\#, *-ai\# > *-i\#$
- $*k^u > *p$
- $*oRa > *aRa$
- $*uo > *ua$
- $*ie > *ia$
- $*-i\# > \emptyset$
- $*ai > *e$
- $*st > *t^s$, but $*sst > *st$ (?)
- $*t^s > *ss$
- $*\chi s > *\chi$
- $*i > *\delta/V_V$
- $*ou, *au > *\bar{o}$
- Reduction of proclitic vowels and of vowels in final syllables.
- $*s > *\Sigma$ (a sibilant different from Lat. s)
- $*\bar{u} > *\bar{i}$
- $*\bar{o}, *oi > *\bar{u}$
- Final a-affection: $*i > *e, *u > *o/_C\check{a}(C)\#$
- $*\bar{u} > *\bar{i}$
- British lenition: in intervocalic position (word-internally or across word boundaries) or preceded by a vowel and followed by a resonant, single voiceless stops become the corresponding voiced stops ($*T > *D/V_ (R)V$), $*m > *\mu$.
- $*\bar{a} > *\bar{o}$ (later $> W\ o, aw$; Br. and Co. $*\bar{o}$)
- Final i -affection: $*a, *e, *o, *u > W\ ei/y$, Br. and Co. e before $*\bar{i}$ and $*i$ (but not $*i\bar{a}$) in final syllables; $*e$ is also affected by $*i$ in that position.
- $*ND > *NN$: sequences of nasal + homorganic voiced stop become geminate or long nasals.
- $*\gamma > *u/_u$
- $*o > *u/_N$
- $*e > *i/_N$
- $*\bar{u} > *\bar{i}$
- Apocope: final syllables are lost except those ending in $*r$ and $*nt$.
- Syncope of compositional vowels (sometimes retained in proper names) and sporadic syncope of pretonic vowels.

- $*\Sigma > *h$, $*\Sigma u > *χu$
- Provection: voiced stops and fricatives become voiceless before $*h$.
- Spirantization: $*TT > *θ$, $*RT > *Rθ$ and $*RD > *Rθ$, i.e. geminated or long voiceless stops, and/or voiceless stops that had survived all previous changes, become the corresponding voiceless fricatives; voiced stops (and $*m$) after resonants become the corresponding voiced fricatives.
- $*ijá > W aia$, Br. and Co. *oia*
- $*Rγ > W *Ri$ and $*Rə\#$
- $*γ > *i/_\{l,r,n\}$, $*χt > *iθ$, Lat. $*χs > *is$
- $*jū > *ou$
- $*ē > W *ui$, Br. and Co. $*ɔi$
- Pretonic $*i$ and $*u > *ə$, which later become W *y*, Br. and Co. *e*.
- The quantity opposition among vowels is given up.
- Internal *i*-affection: $*u, *o, *a, *e > W e/ei$, Br. and Co. *e* before $*ī, *i, *i$ or a vowel affected by final *i*-affection.
- $*γ > \emptyset$
- $*\#u > *\#g^u$
- $*NT > *N^Th > *Nh$ in W
- $*\#sC > *əsC$ in W
- Various diphthongs containing rounded vowels + $*u > *ëü$ in W.
- Accent shift from final to penultimate syllables.

4. References

De Bernardo Stempel, Patrizia

1987 *Die Vertretung der indogermanischen liquiden und nasalen Sonanten im Keltischen*. Innsbruck: Institut für Sprachwissenschaft der Universität.

Dybo, Vladimir A.

1961 Sokraščenie dolgot v kelīto-italijskix jazykax i ego značenie dlja balto-slavjanskoj i indoevropejskoj akcentologii [The shortening of long vowels in Italo-Celtic and its significance for Balto-Slavic and Indo-European Accentology]. *Voprosy Slavjanskogo Jazykoznanija* 5: 9–34.

Eska, Joseph F.

2006 The genitive plural desinence in Celtic and dialect geography. *Die Sprache* 46: 229–235.

Griffith, Aaron

2005 $*-n(C)s$ in Celtic. *Die Sprache* 45: 44–67.

Hill, Eugen

2012 Silbische Liquiden vor Nasalen im Inselkeltischen und das Problem der Nasalpräsentien vom Typ air. *sernaid*, kymr. *-sarnu*. *Keltische Forschungen* 5 [2010–2012]: 157–184.

Isaac, Graham

2007 *Studies in Celtic Sound Changes and their Chronology*. Innsbruck: Institut für Sprachen und Literaturen der Universität.

Jackson, Kenneth H.

1953 *Language and History in Early Britain: a Chronological Survey of the Brittonic Languages First to Twelfth Century A.D.* Edinburgh: Edinburgh University Press.

Joseph, Lionel S.

1982 The treatment of $*CRH-$ and the origin of $*CaRa-$ in Celtic. *Ériu* 32: 31–57.

- Kim, Ronald
2008 The Celtic Feminine Numerals '3' and '4' Revisited. *Keltische Forschungen* 3: 143–167.
- Matasović, Ranko
2009 *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
- McCone, Kim R.
1996 *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Changes*. (Maynooth Studies in Celtic Linguistics 1). Maynooth: The Department of Old Irish, St. Patrick's College.
- Pedersen, Holger
1976 [1909] *Vergleichende Grammatik der keltischen Sprachen*. Volume 1. *Einleitung und Lautlehre*. Strassburg: Trübner. [Göttingen: Vandenhoeck & Ruprecht.]
- Prósper, Blanca María
2007 *Estudio lingüístico del plomo celtibérico de Iniesta*. Salamanca: Universidad de Salamanca.
- Schrijver, Peter
1995 *Studies in British Celtic Historical Phonology*. (Leiden Studies in Indo-European 5). Amsterdam: Rodopi.
- Schrijver, Peter
1997 On the Nature and Origin of Word-Initial *H-* in the Würzburg Glosses. *Ériu* 48: 205–227.
- Schumacher, Stefan
2004 *Die keltischen Primärverben. Ein vergleichendes, etymologisches und morphologisches Lexikon*. Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Sims-Williams, Patrick
2003 *The Celtic Inscriptions of Britain. Phonology and Chronology, c. 400–1200*. (Publications of the Philological Society 37). Oxford: Blackwell.
- Sims-Williams, Patrick
2007 Common Celtic, Gallo-Brittonic and Insular Celtic. In: Pierre-Yves Lambert and Georges-Jean Pinault (eds.), *Gaulois et Celtique Continentale*. Geneva: Droz, 309–354.
- Stifter, David
2011 Review of Isaac. 2007. online at: http://www.univie.ac.at/lexlep/wiki/Stifter_2011. [An abbreviated version appeared as: Stifter, David. 2012. Review of Isaac 2007. *Journal of Celtic Linguistics* 14: 111–150.]
- Ternes, Elmar (ed.)
2011 *Brythonic Languages Past and Present*. Bremen: Hempen.
- Thurneysen, Rudolf
1946 *A Grammar of Old Irish*. Translated by D. A. Binchy and Osborn Bergin. Dublin: Dublin Institute for Advanced Studies.
- Zair, Nicholas
2010 The Reflexes of the Proto-Indo-European Laryngeals in Celtic. Oxford University Ph.D. dissertation.

David Stifter, Maynooth (Ireland)

69. The morphology of Celtic

- | | |
|---------------|----------------------|
| 1. Sources | 5. Pronouns |
| 2. Nouns | 6. Personal pronouns |
| 3. Adjectives | 7. Verbs |
| 4. Numerals | 8. References |

1. Sources

The morphology of Celtic is best known from the Insular Celtic (IC) languages, i.e. those attested in medieval times on the British Isles and in Brittany, especially Old Irish (OI), but also Middle Welsh (MW) and to a lesser extent Middle Breton (MB) and Middle Cornish (MC). The fragmentary transmission of all Continental Celtic (CoC) languages (see Vath and Ziegler, this handbook) severely limits our knowledge about their morphology. While for nouns we can at least establish some paradigms, though often incomplete, this is utterly impossible for the verbal complex. Moreover, our poor understanding of Celtiberian (Cib.), Gaulish (G), and Lepontic (L) texts often does not allow an unambiguous interpretation of a given form. As a consequence, this article is based mainly on material from the Insular Celtic languages, especially from Old Irish and Middle Welsh as the earliest non-fragmentary sources of Goidelic and British respectively. Forms from Continental Celtic languages are included where they contribute to the reconstruction of Proto-Celtic morphology. Where the forms cited are found in standard grammars, references to those are omitted. They are Untermann and Wodtko (1997) for Celtiberian; Lambert (2003) for Gaulish; Thurneysen (1946), McCone (1994, 2005) and (1997) for Old Irish; Ternes (2011) for all British languages; Morris Jones (1913) and Evans (1964) for Middle Welsh; Lewis and Piette (1990) for Middle Breton; and Lewis (1990) for Middle Cornish.

2. Nouns

2.1. The Continental Celtic languages as well as Old Irish preserve all three genders, whereas the neuter is lost in British. As to number, no unambiguous dual forms are attested so far in Continental Celtic. Old Irish has separate dual forms, which are always accompanied by the numeral. In British, the dual only survives in a limited number of set expressions, numeral and noun forming a word unit as in MW *deuglust* ‘two ears’, MB *daoulagat* ‘two eyes’, MC *dywluef* ‘two hands’. Synchronically, the form of the noun is identical with the singular. Singular nouns with a collective meaning and certain fossilized plurals can form a singularive in British by adding MW *-yn/-en*, MB, MC *-en*, cf. MB *guez* ‘forest’ → *guezenn* ‘tree’ (cf. Wodtko, this handbook: 4.2).

2.2. Proto-Celtic continued all PIE cases. However, reflexes of the vocative are found only in Irish, whereas the ablative is preserved only in Celtiberian, and the existence of the locative in Gaulish is doubtful. In Irish, dative, locative, and instrumental fell together

to form a single syncretistic case traditionally called dative, which morphologically can continue any of the three cases, cf. dat. sg. *fiur* ‘to a man’ < * $\bar{u}i$ < PIE dat. sg. * $\bar{o}j$, dat. sg. *toimte* ‘to a thought’ < * $\bar{t}ijon-i$ < PIE loc. sg. * \bar{i} , dat. pl. *máthraib* ‘to mothers’ < * $\bar{b}is$ continuing PIE instr. pl. * \bar{b}^hi . The British languages have lost separate case forms and have only one form for the singular and one for the plural, both usually continuing the nominative, but sometimes also an oblique case.

2.3. Old Irish, Gaulish, and Celtiberian distinguish several nominal stem classes, *o-*, *ā-*, *ī-*, *i-*, *u-* and consonantal stems being common to all of them. The *ā*-stems continue the PIE * eh_2 -stems, whereas both *devī*-type * ih_2 -stems and *vr̥kī*-type * iH -stems have fallen together in the category of Celtic *ī*-stems. In Old Irish, stems in * $\bar{i}jo-$ and * $\bar{i}jā-$ form separate classes. In British, the stem class to which a noun formerly belonged is visible only in the plural if at all. Former *o*-stems are often characterized by *i*-affection caused by the nom. pl. ending * \bar{i} < * $\bar{o}i$, e.g. MW *march* ‘horse’, pl. *meirch*. In some of the other stem classes, the effects of apocope caused a former suffix to become a plural ending, cf. MW *cat* ‘battle’ < nom. sg. * $\bar{k}at-u-s$, pl. *cadeu* < nom. pl. * $\bar{k}at-ou-es$. However, since some of these endings become productive, the original stem class of any given noun is often impossible to determine.

2.4. The Continental Celtic languages, retaining final syllables, preserve many of the PIE case endings with no or only minor changes. In Irish, on the other hand, apocope leads to the loss of most endings. However, the effects of raising, lowering, and *u*-infection and of palatalization together with the initial mutations shown by following words in close syntactic groups allow one not only to distinguish between cases synchronically, but often also to identify the lost ending from a diachronic point of view. The evidence of Continental Celtic and Old Irish taken together thus enables us to a certain extent to reconstruct Proto-Celtic paradigms.

o-stems had a singular with nom. * $\bar{o}s$ (G, Cib. *-os*, OI *fer* ‘man’ < * $\bar{u}iros$ with lowering), voc. * \bar{e} (OI *fir^L* < * $\bar{u}ire$ with palatal /*r*’/), acc. * $\bar{o}m$ (G, Cib. *-om*, OI *fer^N* < * $\bar{u}irom$), dat. * $\bar{u}i$ < PIE * $\bar{o}i$ (G, Cib. *-ui*, OI *fiur^L* with *u*-infection), loc. * $\bar{e}i$ (Cib. *-ei*) and abl. * $\bar{u}d$ < PIE * $\bar{o}d$ (Cib. *-uz*). A variety of forms is found in the gen.: * \bar{i} (Gaul. *-i*, OI *fir^L* < * $\bar{u}ir̄$ with palatal /*r*’/) is cognate with Latin *-ī*, whereas the Lepontic form *-oiso* most likely reflects PIE * $\bar{o}sjō$ (cf. Solinas 1995: 350–352). The isolated Celtiberian ending *-o* is obscure. In the plural we find the pronominal nominative ending * $\bar{o}i$ (G *-oi*, *-i*, OI *fir^L* < * $\bar{u}ir̄$), whereas the old nominal ending is preserved in the Old Irish vocative *-u* < * $\bar{u}s$ < PIE * $\bar{o}s$. In the prehistory of Irish this latter ending fell together with acc. * $\bar{u}s$ (OI *firu* < * $\bar{u}ir̄s$), which led to the peculiarity of Old Irish that the form of the accusative plural is used as vocative in all other stem classes as well. Pre-Irish * $\bar{u}s$ and G *-us* either continue PC * $\bar{u}s$ < * $\bar{o}s$ < PIE * $\bar{o}ms$, or, more likely, a remodeled Proto-Celtic ending * $\bar{u}ns$ < * $\bar{o}ns$ (cf. Griffith 2005: 50 f.). In the genitive, the ending * $\bar{o}m$ continued in G (*-on*) and OI (*fer^N* < * $\bar{u}irom$) was either taken over from the athematic stems or else arose regularly from PIE * $\bar{o}m$ by shortening (see Stifter, Phonology of Celtic, this handbook: 1.6.2). Celtiberian, on the other hand, has *-um* < PIE * $\bar{o}m$. The dat. pl. Lep. *-opos* (cf. Solinas 1995: 343 f.) is from PIE * $\bar{o}b^hos$, while Cib. *-ubos* shows an unexpected *-u-* probably of analogical origin. The PIE instr. pl. * $\bar{o}is$ is possibly continued by G *-us/-ouç* (cf. Lambert 2003: 55), whereas OI *-(a)ib*

< *-obis, used as dative plural, is an analogical creation based on the instrumental ending of athematic stems.

The singular of \bar{a} -stems was nom. * \bar{a} (Cib., G *-a*, OI *túath^L* ‘tribe’ < **toṽtā*), acc. *-am < PIE * $\bar{a}m$ (Cib. **-am**, G *-am*, *-an*, *-em*, *-en*, *-im*, OI *túaiθ^N* < *-en < *-am) and gen. * $\bar{a}s$ (Cib., G *-as*). Gaulish has an alternative ending *-ias* in the genitive, probably borrowed from the $\bar{i}/\bar{i}\bar{a}$ -stems, and the same form seems to be reflected in OI *túaiθe* < *-i $\bar{a}s$. In the dative, Old Irish has *túaiθ^L* < *-ī < * $\bar{a}\bar{i}$ or *-ai_i*, which could go back to PIE dat. * $\bar{a}h_2ai$ or loc. * $\bar{a}h_2i$. Possibly both cases are represented in Celtiberian, both written **-ai**, which could stand for /-ai/ or /-āi/. The Gaulish forms in *-ai*, *-i* are syntactically datives and therefore probably go back to * $\bar{a}\bar{i}$. Remarkable is the creation of a new ablative **-az** in Celtiberian on the model of thematic **-uz** < * $\bar{u}d$ < PIE * $\bar{o}d$. Similar forms are found in other stem classes as well: **-iz** with *i*-stems, **-ez** with consonantal stems. In the plural, PIE nom. * $\bar{a}h_2as$ and acc. * $\bar{a}s$ fell together in PC * $\bar{a}s$ (Cib., G *-as*, OI *túatha^h* < * $\bar{a}s$). Unambiguous forms for the genitive are lacking in Continental Celtic (for Celtiberian cf. Untermann and Wodtko 1997: 402⁸⁴), OI *túaiθ^N* reflects an ending with final nasal preceded by a short vowel, possibly *-om. The dative G *-aβo* reflects PIE * $\bar{a}h_2b^hos$ with analogical loss of the final -s, probably due to contamination with the instr. pl. PIE * $\bar{a}h_2b^hi$ > G *-abi* (only in the pronoun *eiabi*). The reverse seems to have happened in the prehistory of Old Irish, since *túathaib*, used as dative plural, must continue * $\bar{a}bis$.

For the other stem classes a few remarks must suffice. In the genitive singular of the consonantal stems, *-os < PIE *-os seems to have been generalized in Old Irish (*rig* ‘of a king’ with non-palatal final < * $\bar{r}\bar{i}gos$) and probably Gaulish (*-os*), whereas we find both **-es** and **-os** in Celtiberian, though the latter is much more common. A notable archaism is found in the Old Irish neuter *n*-stems, which have *anmae* ‘of a name’ < * $\bar{m}\bar{e}s$ < * $\bar{m}en$ -s with the ending in the zero grade. The dat. sg. PIE * $\bar{e}i$ seems to be continued by Cib. **-e** and **-ei** (also used as loc.) and possibly by G *-i*, if to be read as /-ī/. However, the latter could also continue the PIE loc. * \bar{i} , which also underlies the Old Irish “short” datives of the type *toimte* ‘to a thought’ < * $\bar{t}\bar{i}on$ -i (cf. McCone 1978) with Insular Celtic apocope of final * \bar{i} . In the dative plural, G *-bo* shows the same loss of -s as in the \bar{a} -stems, as opposed to Lep. **-pos**. Examples of the instrumental plural in *-bi*, *-be* < PIE * \bar{b}^hi are also found in Gaulish.

3. Adjectives

3.1. With regard to gender, number, and case, adjectives show the same categories as nouns and are basically inflected identically. Most common are adjectives inflecting as *o*-stems when masculine or neuter, as \bar{a} -stems when feminine. The difference between the genders is lost in Cornish and Breton and in many cases also in Welsh. However, the feminine form, going back to nom. sg. * \bar{a} , shows \bar{a} -infection where the root vowel is liable to it, e.g. MW *gwynn* ‘white’, fem. *gwenn*. In Old Irish, *io/iā*-stem adjectives form a separate and quite large group. Much rarer in Old Irish are *i*- and *u*-stem adjectives, which distinguish between masculine/neuter and feminine only in the singular. The feminine forms seem to have been analogically created on the model of the \bar{a} -stems. In addition to these four classes, Old Irish has a handful of consonantal stem adjectives like *té* ‘hot’, nom. pl. *téit*.

3.2. The Insular Celtic languages have four degrees of comparison: positive, comparative, superlative, and equative, the latter expressing equal degree. All of these forms are uninflected and used only predicatively. The comparative OI *-(i)u* continues the nom. sg. **-iūs* < PIE **-iōs*, e.g. *sinu* ‘older’ < **sen-iūs*, and remnants of the same formation are also found in British, e.g. MW *hyn* ‘older’ < **hen-ih* < **-iūs*. In most cases, however, this formation is replaced in British by the productive suffix MW *-ach*, MB *-och* of unclear origin. For a discussion of the Celtic superlative suffixes cf. Cowgill (1970). The most common suffix in Old Irish is *-elam*, corresponding to MW, MB *-haf* and going back to **-isamos*, cf. OI *sinem*, MW *hynaf* ‘oldest’ < **sen-isamos*. The suffix **-isamo-* is also found in toponyms like G *Uxisama* and Cib. **Segisama** and is most likely a contamination of the zero grade **-is-* of the PIE comparative suffix **-ios-* and **-amo-* < PIE **-ṛmo-*. Simple **-amo-* is still occasionally found, e.g. in OI *nessam*, MW *nessaf*, Gaul. *neððamo-* ‘nearest’ or in the Celtiberian personal name **Usama** ‘highest’. In the formation of the equative, Old Irish differs from British in using *-ithir* < **-iseteros* while the latter has MW *-het* < **-isetos* in addition to a prefix *kyn-*.

4. Numerals

4.1. For a detailed discussion of the Celtic numerals cf. Greene (1992). OI *óen-* ‘1’ is uninflected and used only in composition. In British, the corresponding forms are MW, MB *un*, MC *vn*, all going back to PIE **oīno-*. In Breton and Cornish, the numeral develops into an indefinite article. The numerals 2 to 4 are inflected for case and gender in Old Irish, for gender only in British: OI *da*, *dá* m., *di*, *dí* f., MW *deu* m., *dwy* f. ‘2’, OI *tri* m., *téoir* f., MW *tri* m., *teir* f. ‘3’, OI *cethair* m. *cethéoir* f., MW *pedwar* m., *pedeir* f. ‘4’. The numerals 5 to 10 are uninflected: OI *cóic*, MW *pypm* ‘5’, OI *sé*, MW *chwech*, Cib. **sues** ‘6’, OI *secht*, MW *seith* ‘7’, OI *ocht*, MW *wyth* ‘8’, OI *noí*, MW *naw* ‘9’, OI *deich*, MW *dec* ‘10’. In Old Irish, multiples of ten are substantival dental stems: *fiche* ‘20’, *tricho* ‘30’, *cethorcho* ‘40’ etc. A more recent system is reflected by the custom to use multiples of 20 like *da fichit* ‘40’, which has an exact counterpart in British, cf. MW *deu ugeint* ‘40’, *tri ugeint* ‘60’, etc. The numeral for 100 is OI *cét*, MW *cant*, Cib. **kantom**, the one for 1000 OI *míle*, MW *mil*, a loan from Lat. *mīlia*. A peculiarity of Irish is the use of personal numbers formed – except for 2 – by combining the numeral with *fer* ‘man’: OI *oinar* ‘1 person’, *dias* ‘2 persons’, *triar* ‘3 persons’, etc.

4.2. The Proto-Celtic ordinal ‘first’ seems to have been a *u*-stem **kintu-* (cf. Wodtko, this handbook, 1.6) preserved in Gaulish compound personal names like *Cintu-gnatus* ‘firstborn’ and in OI *cét-* ‘first’ used in composition. Built on this stem are OI *cétnae* and MW *kyntaf*, MB *quentaf*, the latter with the superlative suffix *-haf* < **-haṃ*. For ‘second’, **aljo-* or **alno-* are used, continued by OI *aile*, MW *eil*, and G *allos* respectively. No exact correspondence is found for ‘third’: **trito-* is reflected in Gaulish and Celtiberian compound names, whereas MW *trydydd*, MB *trede* go back to **tritijo-* and OI *tris* to **tristo-*. In most other ordinals a suffix **-to-* or **-eto-* is used, cf. G *pinpetos*, OI *cóiced*, MW *pymhet* ‘fifth’, G *sextametos*, OI *sechtmad*, MW *seithuet* ‘seventh’, G *decametos*, Cib. **tekametam**, OI *dechmad*, MW *decuet* ‘tenth’. An older formation for ‘tenth’ is preserved in G acc. *dekantem/n*, probably meaning ‘tithe’.

5. Pronouns

5.1. Celtic has generalized the stem *so- of the suppletive demonstrative pronoun PIE *so-/to-, *to- only surviving in fossilized forms like OI *tó* ‘yes’ < nom./acc. sg. n. *tod. This pronoun seems to have been fully inflected in Celtiberian, attested so far are nom. sg. m. *so* (or gen. sg.?), nom./acc. sg. n. *soz* < *sod, nom. sg. f. *sa*, dat. sg. m./n. *somui*, loc. sg. m./n. *somei* with -m- < *-sm- and gen. pl. m./n. *soisum*. The inflection of the last three forms is typically pronominal and directly comparable e.g. to Ved. dat. sg. *tásmai* and gen. pl. *tésām*. In Insular Celtic and possibly Gaulish, a combination of *so- with the particle PIE *de is found: the acc. sg. *som de was remodeled to a stem *sondo- underlying W *hwnn* m. ‘this’ < *sundos < *sondos, *honn* f. < *sundā < *sondā as well as OI *sund* ‘here’ < dat. sg. *sondūj and possibly also G *onda* with loss of the initial s- (Schrijver 1997: 29f. and 48). The Insular Celtic languages have developed a definite article reconstructed as *sindo-, which loses its initial *s- in proclitic position to give OI *in(d)*, MW *y(r)*, MB, MC *an*. A possible Continental Celtic reflex of this is G acc. pl. f. *indas*. The origin of *sindo- is much disputed, cf. the discussion by Schrijver (1997: 44–89). In Old Irish, the article combines with prepositions to form a word unit, and in this position the original *s- is retained unless the preposition ended in a vowel, cf. *issin(d)* ‘in the’. Old Irish has no inflected demonstrative pronoun but uses the definite article in combination with various particles (Thurneysen 1946: 299–302), e.g. *int-í* m. ‘this one’, *in lebor-so* ‘this book’, *in ben-sin* ‘the above-mentioned woman’, *in fer tall* or *in fer ucut* ‘yonder man’. The combination of *í* and one of these enclitic particles can be used substantivally as in *int-í-sin* ‘the above’ or adjectivally as in *in fer í-siu* ‘this man’.

5.2. No unambiguous forms of the interrogative pronoun are attested in Continental Celtic to date. In Insular Celtic, it has largely lost its inflection. OI *cía*, MW *pwý* < PIE *k^heǵ ‘who?’ refers to persons, OI *cid* ‘what?’ < PIE *k^hid plus an additional element to things. In addition to these stressed forms there is unstressed OI *ce*, *ci*, MW *pa*, *py*, invariable for case, gender and number, used adjectivally and in Old Irish also in combination with verbs in the position of a preverbal particle. In its indefinite use, PIE *k^ho- is contained in OI *nech*, MW *neb*, G acc. *nepo-* < PC *nek^ho- ‘some(one)’, a combination with the negative particle *ne, and OI *cách*, MW *pawb*, G *papo-* < PC *k^hāk^ho- ‘every(one)’, which probably had its starting point in nom. sg. *k^hā k^hos ‘someone (f.) and someone (m.)’ (cf. McCone 2003: 174).

5.3. Celtiberian is the only Celtic language to have an inflected relative pronoun, which continues PIE *H₁jo-. Attested so far are nom. sg. m. **ios**, acc. sg. m. **iom** (possibly also used as a conjunction), dat. sg. m./n. **iomui**, nom./acc. pl. n. **ia**, gen. sg. or acc. pl. f. **ias** and acc. sg. f. **iam**. The Insular Celtic languages as well as Gaulish form relative clauses using originally conjunctive particles like *jo as in OI 3. pl. *bertae* ‘who carry’ < *beronti-jo, MW *yssyd* ‘who is’ < *esti-jo, G 3. pl. *dugiöntiio* ‘who worship (?)’; *k^he as in OI neg. *nach-* ‘who does not’ before infixed pronouns (cf. 6.2); and *de as in OI neg. *nad^L* ‘who does not’ (for an overview on relative marking in Insular Celtic and Gaulish cf. McCone 2006: 247–276).

6. Personal pronouns

6.1. Evidence for personal pronouns is largely confined to the Insular Celtic languages. The independent forms are uninflected and in Old Irish used only as predicative nominatives; the British languages use them also in object position. 1. sg. OI *mé* (emphatic *mēsse*), MB *my*, *me*, MC *me* continue PIE acc. sg. **me* while MW *mi* either goes back to the variant PIE **mē* or is influenced by 2. sg. *ti*, 1. pl. *ni* etc. Two forms are found in the 2. sg.: OI *tú* (emphatic *tússu*) < PIE **tu* and MW *ti* < **tū* < PIE **tuH*, both continuing nominatives. The plural forms 1. OI *sní*, MW *ni*, 2. OI *sí*, MW *chwi* go back to **snī*(s) and **sūī*(s) respectively and seem to be due to a remodeling of the inherited forms (cf. McCone 1994: 187 = 2005: 259 f.). No dual forms survive. The forms of the third person m. and n. are supplied by the anaphoric pronoun PIE **i-/e-*, though the exact preforms are a matter of debate (cf. Schrijver 1997: 54–61). 3. sg. m. OI *é* could continue a nom. **eǵ* or possibly **es*, whereas MW, MB, MC *ef* go back to acc. **em* plus an additional element. The neuter OI *ed* is from **ed* plus a suffixed element. For 3. pl. OI *é*, MW *wy*, nom. pl. **eǵes* and acc. pl. **ens* have been suggested as preforms. The feminine forms OI *sí*, MW *hi* < **sī* probably reflect PIE **sih*₂. Doubling of the personal pronouns results in emphasis, cf. MW 1. sg. *mivi*, 2. sg. *tidi* etc., OI 3. sg. f. *sisi*, 1. pl. *snisni*. In Irish, *-som*, probably from **somo-* ‘same’, was used for emphasis in the 3. sg. m. and the 3. pl. (both *é-som*), which led to the remodeling of 1. sg. **meme*, 2. sg. **tutu* to actually attested *messe* and *tussu* respectively (cf. McCone 1994: 189 = 2005: 265; for an alternative explanation of *-se* and *-su* as deictic particles, cf. Griffith 2010). The second element of these emphatic personal pronouns can also be used on its own in both British and Old Irish to emphasize verbs or other pronouns as in OI 1. pl. *tiagmi-ni* ‘we go’, MW *hyny elwyf ui* ‘until I go’.

6.2. Object pronouns are enclitic in Insular Celtic and are either infixated between a preverb or preverbal particle and the verb or else suffixed to the verb itself, cf. OI *da mbeir* ‘brings him’, *beirthe* ‘carries him’. The forms of the infixated pronouns are 1. sg. OI *-m^L*, MW *'m* < **mu*, 2. sg. OI *-t^L*, MW *'th* < **tu*, 1. pl. OI *-n*, MW *'n* < **nos*, 2. pl. OI *-b*, MW *'ch*, *'weh* < **sues* (cf. McCone 1994: 193 = 2005: 271 f.). In the third person, forms of PIE **so-* (cf. 5.1) are used in the plural, cf. OI, MW, MC *-s* < acc. pl. **sons*, as well as the feminine singular, cf. OI, MB, MC *-s* < **sām* (cf. Schrijver 1997: 21 f.). In Middle Welsh, *-s* is also used as a masculine after certain preverbal particles (cf. Evans 1964: 55). The masculine and neuter singular forms, on the other hand, are formed from PIE **e-*: OI *-a^N*, MW *'e*, *'y*, MB *'n* < **em* and OI *-a^L* < **ed* (cf. Schrijver 1997: 54–56). In Old Irish, the same forms are used as suffixed pronouns, e.g. 1. sg. *beirthium* ‘carries me’ < **bereti-mu*, 2. Sg. *beirtheut* < **-tu*, 3. Sg. m./n. *beirthe* < **-em/*-ed*, 3. pl. *beirtheus* < **-sūs* (cf. McCone 1994: 194 = 2005: 273 f.). The set of Old Irish infixated pronouns discussed above is called class A and is used after preverbs and particles originally ending in a vowel. After preverbs ending in a consonant, set B is used, which inserts the particle **de* between preverb and pronoun as in *fordom chain* ‘teaches me’ < **uorde-mu-* (cf. McCone 1994: 193 f. = 2005: 273). Class C infixated pronouns are used in relative clauses and after certain conjunctions; they are identical in origin to class B, with the difference that the inserted particle **de* is lenited or nasalized according to the element to which it is attached (cf. McCone 1994: 194 = 2005: 273).

6.3. In Insular Celtic, pronouns dependent on a preposition form a word unit with the latter, which leads to forms known as prepositional pronouns or conjugated prepositions. The forms of the pronouns are in many cases the same as the infix/suffix pronouns discussed in 6.2 (for an overview cf. McCone 1994: 191 = 2005: 268). Accusative and dative are distinguished only in the third persons in Old Irish and not at all in British. Examples are OI 1. sg. *indium* ‘in me’, 2. sg. *indiut*, 3. sg. acc. m./n. *ind*, f. *inte*, dat. m./n. *and*, f. *indi*, 1. pl. *indiunn*, 2. pl. *indib*, 3. pl. acc. *intiu*, dat. *indib*; MW 1. sg. *ragof* ‘before me’, 2. sg. *ragot*, 3. sg. m. *racdaw*, f. *racdei*, 1. pl. *ragom*, 2. pl. *ragoch*, 3. pl. *racdut*. In Welsh, the 3. pl. is later remodeled to *-unt* under the influence of the verbal endings.

6.4. The Insular Celtic languages have no inflected possessive pronouns but use instead the old genitive of the personal pronouns. There are two sets, one of unstressed and one of stressed forms. The latter are rare, since other constructions are preferred to express possession. They are 1. sg. OI *muí*, MW *meu*, 2. sg. OI *taí*, MW *teu*, 3. sg. m. OI *aí*, MW *eidaw*, f. OI *aí*, MW *eidi*, 1. pl. OI *náthar*, MW *einym*, 2. pl. OI *sethar*, MW *einwch*, 3. pl. OI *aí*, MW *eidu(nt)*. The unstressed possessive pronouns can form a word unit with preceding prepositions. The independent forms are OI 1. sg. *mo^L*, 2. sg. *do^L*, 3. sg. m./n. *a^L*, f. *a^h*, 1. pl. *ar^N*, 2. pl. *far^N*, *for^N*, 3. pl. *a^N*; MW 1. sg. *vy(n)^N*, 2. sg. *dy*, 3. sg. m. *y^L*, f. *y* (spirant mutation), 1. pl. *an*, *yn*, 2. pl. *ach*, *awch*, *ych*, 3. pl. *eu*, rarely *y*. The forms of the third persons go back to genitive forms of the PIE anaphoric pronoun *e- (cf. 6.1 and 6.2): in Old Irish, 3. sg. m./n. *es̄jo, f. *es̄jās, 3. pl. *ējsom fall together in their stressed form as *aí*, but are kept distinct by the mutations they cause on the following word as unstressed *a^L*, *a^h* and *a^N* respectively (cf. McCone 1994: 188 = 2005: 262). In Middle Welsh, *eid-* (with /ð/) in the stressed forms *eidaw*, *eidi*, *eidu(nt)* goes back to *es̄jV, otherwise these are modeled after the conjugated prepositions (6.3) (cf. Schrijver 1997: 57 f. and 61 f.).

7. Verbs

7.1. The Celtic verb distinguishes three persons and two numbers, singular and plural. Insular Celtic has lost the dual, and no dual forms are so far attested in Continental Celtic. As to voice, Old Irish has an active and a passive; the latter has morphologically separate forms only in the third persons, the first and second persons being supplied by the third singular and the appropriate infix object pronoun. In British, the third singular of the original passive survives as an impersonal form. In Old Irish, forms of the PIE middle are continued by verbs with deponent inflection. These are syntactically active and can form their own passive. No verb has active and deponent inflection at the same time. In the active, passive, and deponent, special relative forms are found in Old Irish in the third persons and in the first plural, which go back to a combination of the verb form and a relative particle *jo (cf. 5.3). These are used in relative clauses when an absolute verb form (7.2) is required. With all other persons and in other syntactic positions, different strategies of building relative clauses are employed. A remnant of relative verb forms in British is MW *yssyd* ‘who is’.

7.2. One of the most striking characteristics of the verb in Insular Celtic is the existence in most categories of two sets of endings, called absolute and conjunct. Absolute endings are used with simple verbs when they are not preceded by any preverbal (e.g. negative, interrogative) particle, conjunct endings in all other circumstances; thus e.g. OI *beirid* ‘carries’ with absolute, but *ní beir* ‘does not carry’ or compound *do· beir* ‘brings’ with conjunct ending. While this dichotomy is still alive in Old Irish, the British languages lost it early, absolute forms only surviving in Old Welsh, early Middle Welsh, and Old Breton and only in the third persons. The creation of these two sets of endings is doubtless connected with the verb-initial word order the Insular Celtic languages have developed. While early attempts at explaining the distinction between absolute and conjunct took the PIE primary and secondary endings as a starting point (cf. Meid 1963, with an overview of earlier research 10–52), it was Cowgill (1975) who first tried to explain both sets from the PIE primary endings, positing a general Insular Celtic apocope of *-i to explain the conjunct variants, e.g. OI *-beir* < IC *beret < PC *bereti. This view is now widely accepted. Controversial is the question of why in the forerunners of the absolute forms the *-i of the primary endings was retained. Cowgill himself (1975) posited a general sentence particle *es present in every Insular Celtic main clause in Wackernagel position, i.e. following the verb if sentence initial, but following the preverb or preverbal particle otherwise. This *es, as an enclitic, formed an accentual unit with a preceding verb, thus causing the *-i of the ending to be retained and leading to absolute endings like OI *beirid* < IC *bereti-es. The alleged Insular Celtic sentence particle *es was later identified as continuing PIE *eti ‘and’ by Schrijver (1997: 146–156), a view accepted e.g. by Schumacher (2004: 97–114). On the other hand, the existence of such a general sentence particle *es or *eti in Insular Celtic is disputed, mostly on phonological grounds, by McCone (most recently 2006 with references to earlier work), who claims that the retention of *-i in the forerunners of the absolute endings was not caused by any specific particle, but by any element in Wackernagel position.

A further important distinction is made in Old Irish between deuterotonic and prototonic forms of compound verbs. In deuterotonic forms, the first preverb is proclitic and the stress lies on the root or a second preverb as in *do· beir* ‘brings’. In the prototonic forms, the stress shifts to the first preverb as in *ní tabair* ‘does not bring’. This shift of stress results in different syncope patterns (cf. Stifter, *Phonology of Celtic*, this handbook) and thus often leads to prototonic forms that resemble their deuterotonic counterpart very little, e.g. prototonic *-díltai* ‘denies’ beside deuterotonic *do· shuindi* (cf. Thurneysen 1946: 534–536).

The distribution of deuterotonic and prototonic forms is the same as that of absolute and conjunct forms: deuterotonic forms stand in sentence-initial position, whereas a verb assumes prototonic forms after a preverbal particle. Prototonic forms are also used in the imperative, unless an infix pronoun is present.

In certain positions, Old Irish employs a semantically empty preverbal particle *no-*. It is used only with simple verbs when no other preverbal particle is present. It precedes all forms of the imperfect, the conditional, and the past subjunctive, which have no absolute forms. It is also used to form relative clauses where no relative verb form is available (7.1) and to support infix pronouns (6.2).

In Old Irish, almost any verb form can be used with what since McCone (1997: 91 f.) is called the augment. It has potential or resultative force, so that augmented preterites are sometimes called perfects (for details cf. McCone 1997: 89–161). Most verbs use

ro- < PIE **pro* as their augment, but **ad-*, **kom-* and **uss-* also occur. The British languages have a similar system, but with only one augment, i.e. MW *ry*, MB *ra-*, MC *ro-* corresponding to OI *ro-*.

7.3. In Old Irish, the verb is organized into five stems: the present stem, from which present indicative, imperfect indicative, and imperative are formed; the subjunctive stem, which supplies present subjunctive and past subjunctive; the future stem with future and conditional (secondary future); the active preterite stem; and finally the stem of the preterite passive. The British languages differ in two respects: they have lost the future stem, but on the other hand have developed a pluperfect based on the active preterite stem.

7.3.1. The various primary present stem classes posited for Proto-Celtic are described by Schumacher (2004: 36–47). They have merged in Old Irish into three classes of strong verbs and three of hiatus verbs. The latter are synchronically characterized by a stem ending in a vowel and thus by a hiatus between stem and ending. Historically, they arose through the loss of various intervocalic consonants belonging either to the root or the suffix. In addition, Old Irish has two classes of weak verbs of mostly secondary origin, resulting in eight present stem classes altogether. In contrast, there is basically only one way of inflecting a regular verb in the British languages, the personal endings employed having their origin in various stem classes. For some of the endings there are several allomorphs.

The first class of Old Irish strong verbs, called S1 (the terminology used here is McCone's, cf. 1997: 25), consists of old thematic stems. It is characterized by an alternation between palatal and non-palatal quality of the root-final consonant, which reflects the quality of the thematic vowel, cf. 3. sg. *-beir* 'carries' with palatal /-r'/ < *beret < PIE *b^her-e-ti, 3. pl. *-berat* with non-palatal /-r-/ < *beront < PIE *b^her-o-nti. Beside stems built with the simple thematic vowel, this class also contains formations with PC *-ske/o- < PIE *-skē/o- like OI *-naisc* 'binds' < *nad-ske-ti and thematized nasal presents like *-ding* 'presses' < *ding-e-ti ← PIE *d^hi-ne-ġ^h-ti. S2 presents are characterized by a palatal root-final consonant throughout and go back to formations with the thematic suffix *-je/o-, cf. OI *-gaib* 'takes' < *gab-je-ti. In contrast, S3 presents have a non-palatal root-final consonant throughout. They continue nasal presents from roots ending in a laryngeal like *-ben* 'hits' < *binati ← PIE *b^hi-n-h₂-. The first class of weak verbs, W1, was originally characterized by a stem in *-ā-, which resulted in a paradigm with non-palatal root-final consonant throughout and an ending *-a* in the 3. sg. conjunct. This class includes factitives to adjectives like *-mára* 'praises' < *mārāti from *már* 'big, great' and denominatives to nouns like *-rann* 'divides' < *rannāti from *rann* 'part'. The W2 presents, built with a suffix *-ī- and characterized by palatal root-final consonant throughout and a third singular conjunct in *-i*, have three sources: denominative formations like *-rími* 'counts' < *rīmīti from *rím* 'number', causatives like *ad·suidi* 'holds back' < *sodīti < PIE *sod-eje-ti, and PIE essives like OI *-ruidi* 'blushes' < *rudīti ← PIE *rud^h-h₁jé-.

As for the endings, a short overview must suffice here, concentrating on the conjunct variants. The PIE thematic endings are well represented, cf. OI 1. sg. *-biur* 'I carry' < PC *berū < PIE *-ō, 2. sg. *-bir* < *beris < PC *beresi, 3. sg. *-beir* < *bereθ < PC *bereti, 1. pl. *-beram* < *beromah < PC *beromosi, 2. pl. *-beirid* < *bereθih < PC *beretesi, 3.

pl. *-berat* < *berod < PC beronti. The 2. sg. **-esi* > IC **-isi* is also continued by MW *-y*, older *-yd* (with /ð/) < **-iji*. In the first singular we find *-u* in G *delgu* ‘I contain’, which might, however, have to be interpreted as subjunctive in view of the alternative ending *-umi* also attested in Gaulish. This is formed by adding the athematic ending **-mi* onto thematic **-ū*. An identical form is continued by the rare variant MW *-if*, e.g. *kenif* ‘I sing’ < *kanīmi < PC *-ūmi*. Athematic endings are found in the Old Irish S3 presents, cf. 1. sg. abs./conj. *benaim* ‘I hit’ < *binami+ptcl. (with unlenited *-m* introduced from the copula *am* ‘I am’ < *emmi < PIE **h₁es-mi*), 2. sg. abs./conj. *benai* < *binasi+ptcl., 3. sg. *-ben* < *binaθ < PC *binati*, 1. pl. *-benam* < *binaṃah < PC *binamosi, 2. pl. *-benaid* < *binaθih < PC *binatesi*, 3. pl. *-benat* < *binad < PC *binanti. In the Middle Welsh standard paradigm, athematic endings on stems in **-a-* are continued in the first singular and third plural, cf. MW 1. sg. *caraf* ‘I love’ < **-aui*, 3. pl. *carant* < **-ant(i)*. The inflection of the Old Irish deponents is characterized by endings in *-r*, e.g. 3. sg. *-midethar* ‘judges’ < *meditor < *medjetor, 3. pl. *-midetar* < *medijodor < *medijontor. For the passive, the same endings are used except for the third singular in classes S1 and S3, which lacks the *-th-* < **-t-*, cf. *-berar* ‘is carried’ < *beror (for a detailed discussion of passive and deponent endings cf. Cowgill 1985). The third singular of the passive is continued in British as an impersonal, cf. MW *kenitor* < *kan-ītor ‘one sings’. In the Old Irish imperfect, there is only one set of endings for active and deponent verbs. The origin of most of these endings is obscure, with the notable exception of 3. sg. *-d* from the PIE secondary middle ending **-to*, cf. *no·bered* ‘used to carry’ < **-bereto*. Traces of this ending are also found in British in the same function, cf. MW *gwyd(y)at* ‘used to know’ < **-a-to*. In the imperative, the second singular has a zero ending, cf. OI *beir* ‘carry!’ < *bere, *ben* ‘hit!’ < *bina, the second plural has **-te*, cf. OI *beirid* < *berete, *benaid* < *bina-te. In the third singular, a form *-tuz* is attested several times in Celtiberian, which continues PC **-tūd* < PIE **-tōd*. In Gaulish, on the other hand, 3. sg. *biitetutu* and 3. pl. *biontutu* ‘let him/them hit (?)’ seem to contain *-e-tu* and *-o-ntu* < PIE **-tu/-ntu* respectively, either with reduplicated **-tu* or an added particle. Neither **-tūd* nor **-tu* can account for the Old Irish form *beired*, since both would have caused *u*-infection. It is therefore best explained as deriving from **beretō* < **-toū* and compared to Germanic forms like Gothic *bairadau* ‘let him carry!’. The same must then hold true for OI 3. pl. *berat* < *berontō and MW impersonal *caret* ‘let him love’ < **-e-tō*.

7.3.2. Apart from two or three verbs that continue the subjunctive of a PIE root aorist (cf. Schumacher 2004: 48 f.), there are two main formations of the subjunctive stem in Old Irish: strong verbs with roots ending in a dental or velar take an *s*-subjunctive (e.g. 3. sg. *geiss* from *guidid* ‘prays’), all other strong verbs as well as all weak verbs take an *ā*-subjunctive (e.g. 3. sg. *beraid* from *beirid* ‘carries’). The British languages form their subjunctive with a suffix *-h-*, which is, however, often lost for phonological reasons (e.g. 3. sg. W *carho/caro* from *caraf* ‘I love’). There are also traces of a formation corresponding to the Old Irish *s*-subjunctive, cf. MW 3. sg. *gwaress* ‘may help’ comparable to OI *fo·ré* ‘id.’. It was Rix (1977: 151–154) who first realized that both subjunctive formations of Old Irish can be derived from the same source. While it had long been established that the *s*-subjunctive goes back to a formation consisting of the full grade root and a suffix **-se/o-*, the Old Irish *ā*-subjunctive had formerly been compared to the *ā*-subjunctive of Latin. However, Rix showed that it can be derived from earlier **-āse/o-*, which has the additional advantage of allowing the Old Irish *ā*-subjunctive to

be equated with the British *h*-subjunctive (cf. McCone 1991: 98–104; Schumacher 2004: 51 f.). Thus Insular Celtic formed its subjunctive by adding *-se/o- to the full-grade root if the latter ended in a dental or velar, but *-ase/o- if it ended in a resonant. The additional *-a- can be explained as originating in roots with a final laryngeal: a form of the structure *CeRa-se/o- < PIE *CeRH-se/o- was metanalyzed as *CeR-ase/o- and the suffix variant *-ase/o- then spread to roots of the structure CeR as well (cf. Rix 1977: 153; Schumacher 2004: 49 f.). *-āse/o- was subsequently probably remodelled to *-āse/o- under the influence of the weak verbs in *-ā- and possibly of the suffix of the future, on which cf. 7.3.3 (cf. McCone 1991: 111 f.; Schumacher 2004: 50 f.). A thematic *s*-subjunctive is probably also reflected in two Celtiberian forms, **robiseti** < *-bid-se-ti and **ambitseti** < *-dig-se-ti (exact meaning uncertain; cf. Schumacher 2004: 224 and 276). The source of the Celtic subjunctive in *-se/o- is the subjunctive of the PIE *s*-aorist according to some scholars (cf. e.g. McCone 1991: 63 f.), an unreduplicated desiderative according to others (cf. e.g. Rix 1977: 152).

The endings of the present subjunctive are identical in origin to those of the thematic present indicative, cf. e.g. OI *s*-subj. 2. sg. *-geiss* ‘you may pray’ < *g^hessis < PC *g^hedsesi, 3. pl. *-gessat* < *g^hessont < PC *g^hedsonti, *ā*-subj. 3. sg. *-bera* < *bepāθ < *berāhet < IC *berāset, 3. pl. *-berat* < *berād < *berāhont < IC *berāson̄t. However, in the third singular of the *s*-subjunctive an athematic ending was introduced, probably under the influence of the preterite (cf. Rix 1977: 152; McCone 1991: 55–80): 3. sg. OI *-gé* < *geh < *g^hess instead of expected *g^hesset < PC *g^hedseti. The endings of the past subjunctive are the same as those of the imperfect.

7.3.3 The origins of the Old Irish future were discussed in detail by McCone (1991: 137–182). The future of the strong verbs goes back to the PIE reduplicated desiderative and is characterized by reduplication with *i*, zero-grade root and a thematic suffix *-se/o- (e.g. 3. pl. *fo·lilsat* from *fo·loing* ‘endures’). Similar to the situation in the subjunctive (7.3.2), roots with final resonant take the suffix variant *-āse/o-, which arose regularly in roots with a final laryngeal, cf. *at·béla* < *beβlāθ < *biβlāhet < PC *biβlāseti < PIE *g^ui-g^uh₁-se-ti from *at·baill* ‘dies’. As in the subjunctive, a form like *biβlā-se-ti was metanalyzed as *biβl-āse-ti and gave rise to the suffix variant *-āse/o-, which then spread to roots ending in resonants. As exemplified by *at·béla*, various phonological and analogical processes often obscure the original reduplication. The Old Irish weak verbs form their future with a suffix containing *-f-* and probably going back to *-īsuā-, cf. 3. sg. *-máirfea* from *máraid* ‘praises’. The origin of this formation is obscure; for a suggestion cf. McCone (1991: 176–182) with references to earlier attempts. A connection with the Latin *b*-future must be excluded for phonological reasons. The endings of the future correspond exactly to those of the subjunctive, including the analogical athematic third singular, cf. *fo·lil* < *liluh < *liluss instead of expected *lilusset < *li-lug-se-ti from *fo·loing* ‘endures’. The conditional is formed by adding the endings of the imperfect to the future stem.

7.3.4. In Insular Celtic (or possibly Proto-Celtic) the PIE aorist and perfect merged into a single category expressing past tense and traditionally called preterite. In Old Irish, three main types are to be distinguished: the suffixless preterite continuing the PIE perfect, the *t*-preterite and the *s*-preterite, which both go back to the PIE *s*-aorist. The regular preterite of the British languages corresponds to the Old Irish *s*-preterite. In Old Irish,

this type is formed by all weak and most hiatus verbs. It is characterized by a suffix *-s-* < **-ss-*, which arose in the third singular of the *s*-aorist, first in verbs with roots ending in **-a* < PIE **-H*. On the basis of the 3. sg. **anass* (OI *-an* ‘waited’) < **ana-s-t* (from PIE **h₂enh₁* ‘breathe’) a new paradigm was created, in which **anass-* served as a new stem; only the 3. sg. **anass* itself remained unchanged, thus becoming synchronically endingless (cf. Watkins 1962: 177 f.; Schumacher 2004: 66 f.). This type of preterite then spread to weak verbs in Old Irish and practically ousted all other types in British. Watkins (1962: 156–174) was the first to clearly show that the Old Irish *t*-preterite and the traces of a corresponding formation in British also derive from the PIE *s*-aorist (cf. also Schumacher 2004: 61–66). The *t*-preterite is only found with roots ending in *r*, *l*, *m*, or a velar. As in the *s*-preterite, the basic form is the third singular, where the original *-s-* was regularly lost as in OI *-gart* < **gar-s-t* from *gairid* ‘calls’. The resulting *-t* was reinterpreted as a suffix and the stem **gart-* generalized throughout the paradigm (for details cf. Schumacher 2004: 61–66). The vocalism of forms like *birt* ‘carried’ < **bīrt* < **bīrst* < PIE **b^hēr-s-t* shows that in the singular, Proto-Celtic inherited a stem with lengthened grade (cf. Schumacher 2004: 63 f.). Old Irish strong verbs not forming a *t*-preterite have a suffixless preterite, which goes back to the PIE perfect with reduplication and *o*-grade of the root in the singular. The vowel used for the reduplication is **e* as a rule, but **i* or **u* respectively if the zero grade of the root contains *i* or *u*. The reduplication is basically preserved as in 3. sg. *gégain* < **geγove* < PC **g^ue-g^uon-e* from *gonaid* ‘kills’. However, it can be obscured by phonological processes as in *fích* < PC **ūi-uojk-e* from *fichid* ‘fights’ or *-cúalae* ‘heard’ < **koχloue* < PC **ku-klou-e*. A special formation is found in verbs with a root structure CeT: they form their suffixless preterite from a stem CāT-, cf. 3. sg. *gáid* ‘prayed’ < **gāde* < PC **g^uād-e*. These forms are probably related to perfects with a weak stem CēT- in other languages, cf. Latin 3. sg. *frēgit* ‘broke’, Gothic 3. pl. *brekun* ‘broke’; for details cf. McCone (1986: 236 f.) and Schumacher (2004: 74–76).

The endings added to the newly created stems of the *s*- and *t*-preterite were thematic, cf. 1. sg. OI *-márus* ‘praised’ < **-āss-ū*, MW *kereis* ‘loved’ < **karassī* < **-ass-ū* or 3. pl. OI *-ansat* ‘remained’ < **anass-ont*, *-gartat* ‘called’ < **gart-ont*. An exception is the third singular, which stayed synchronically endingless in the conjunct form of the Old Irish *t*- and *s*-preterites, cf. OI *-acht* ‘drove’ < PC **ag-s-t*, *-an* ‘remained’ < **anass*. In British, however, a thematic ending was introduced in the third singular of all *s*- and *t*-preterites, cf. MW *ysgaras* ‘separated’ < **-ass-et*, *aeth* ‘went’ < **axt-et*. New absolute endings were created in the *s*- and *t*-preterite both in Irish and British, cf. 3. sg. OI *anais* < **-ass-i+ptcl.*, MW *keressyt* < **karass-eti+ptcl.* These must be analogical to the endings of the present, since there was no place for primary endings in the PIE aorist. There is no distinction between absolute and conjunct endings in the suffixless preterite, one and the same form being used in all positions. The endings go back to those of the PIE perfect, cf. 1. sg. *gegan* < **geγova* < PC **g^ue-g^uon-a* < PIE **-h₂a*, 3. sg. *gégain* with palatal final < **geγove* < PC **g^ue-g^uon-e*. An exclusively British innovation is the pluperfect, formed by adding the endings of the imperfect to the preterite stem as in MW 3. sg. *carassei* ‘had loved’.

7.3.5. The passive of the preterite is formed from a separate stem, which is based on the PIE verbal adjective in **-to-* (for details cf. Cowgill 1983: 104–106; McCone 1994: 172 = 2005: 231 f.). This was probably originally used with the copula as a periphrastic

passive similar to Lat. *factus est* ‘was made’. Since the copula could easily be omitted, the adjective alone was reinterpreted as a finite form. The root shows zero grade where possible, cf. OI 3. sg. *-breth* ‘was carried’ < *britos < PIE *b^hṛ-to-s, 3. sg. OI *-fess*, MW *gwys* ‘is known’ < *uissos < PIE *uid-tos.

Old Irish also has a 3. pl. *-bretha*, *-fessa* < *britās, *wissās, which must continue the plural of the feminine, whereas in British there are traces of a third plural with the verbal ending *-nt*: OB *guedint* ‘were seen’, MW *llesseint* ‘were killed’.

7.4. It is a peculiarity of the Insular Celtic languages that in the present indicative the verb ‘be’ has two different stems, an unmarked one, PC *es- (< PIE *h₁es-) and a marked one, PC *b_{uijē}/o- (< PIE *b^huh₂-jē/o-) with habitual force (for details cf. Schumacher 2007: 186–190). Thus there is an opposition between OI *is*, MW *ys* < PC *esti ‘is (now)’ as opposed to OI *-bi*, MW *byd* < PC *b_{uijēti} ‘is (usually)’. Outside the present, all forms are derived from the root of the marked stem. It is possible that this dichotomy goes back as far as Proto-Celtic, since both stems are also attested in Gaulish, cf. 1. sg. *imi*, *imu* < *esmi on the one hand, 2. pl. *biiete* (indicative or imperative) on the other hand. In Old Irish there is a further distinction between the copula and the substantive verb. The habitual present as well as all non-present forms of both are derived from the root underlying *b_{uijē}/o-, the formal differences between the two arising from the fact that the substantive verb is stressed whereas the copula is not. In the non-habitual present, the forms described above are those of the copula, while those of the substantive verb are derived from a suppletive stem *tā-jē/o- (from PIE *stah₂ ‘stand’), cf. 3 sg. *at-tá* ‘is (now)’.

7.5. The PIE present participle was lost in Insular Celtic (or Proto-Celtic) as a living category; traces of its former existence are lexicalized items like OI *carae* ‘friend’ < *karant- (cf. Wodtko, this handbook: 3.2). On the development of the *to*-adjective into a preterite passive cf. 6.3. Old Irish, Breton, and Cornish continue a formation in *-tjō- used as a passive past participle (cf. Pedersen 1913: 408–410), cf. OI *márthae* ‘praised’, MB *caret*, MC *kerys* ‘loved’. The strong verbs of Old Irish use the same stem formation as for the preterite passive, cf. pret. pass. *-cleth* < *klitos < *k_l-tos, past part. *cleithe* ← *klitijos from *ceilid* ‘hides’. In Old Irish, the participle is inflected like an *io/iā*-stem adjective (cf. 3.1). The formation in *-tjō- is not found in Welsh, which instead has a form in *-etic* < *ā-tīko- called verbal adjective in the same function, cf. MW *caedic* ‘loved’. All Insular Celtic languages have a verbal of necessity or gerundive (cf. Pedersen 1976 [1913]: 410). It is uninflected and used only predicatively; the stem formation is the same as that of the past participle, cf. OI *márthai* ‘to be praised’, MW *caradwy*, MC *caradow* ‘to be loved’.

Insular Celtic is unique among Indo-European languages in not having developed a morphologically defined infinitive. Instead, an abstract noun called verbal noun is attached to every verb. The formation of these verbal nouns varies greatly, a whole range of suffixes being used (cf. Thurneysen 1946: 444–455; Gordon 2012; and Stüber 2015: 103–132 for Old Irish; Schumacher 2000 for Welsh). In contrast to the Insular Celtic languages, Celtiberian seems to have developed an infinitive in *-unei*, added to the root as in *taunei* or to a thematic stem as in *ambitinkounei*. These forms seem to continue fossilized datives of an n-stem formation, either of neuter heteroclitics in *-uer/n-, or, more likely, of neuters in *-men- (cf. Stüber 1998: 78).

8. References

- Cowgill, Warren
 1970 Italic and Celtic superlatives and the dialects of Indo-European. In: George Cardona, Henry M. Hoenigswald, and Alfred Senn (eds.), *Indo-European and Indo-Europeans*. Philadelphia: University of Pennsylvania Press, 113–153.
- Cowgill, Warren
 1975 The origins of the Insular Celtic conjunct and absolute verbal endings. In: Helmut Rix (ed.), *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.–14. September 1973*. Wiesbaden: Reichert, 40–70.
- Cowgill, Warren
 1983 On the prehistory of Celtic passive and deponent inflection. *Ériu* 34: 73–111.
- Cowgill, Warren
 1985 On the origin of the absolute and conjunct verbal inflexion of Old Irish. In: Bernfried Schlerath and Veronica Rittner (eds.), *Grammatische Kategorien: Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft, Berlin, 20.–25. Februar 1983*. Wiesbaden: Reichert, 109–118.
- Evans, D. Simon
 1964 *A Grammar of Middle Welsh*. Dublin: The Dublin Institute for Advanced Studies.
- Gordon, Randall Clark
 2012 *Derivational Morphology of the Early Irish Verbal Noun*. PhD dissertation UCLA: Los Angeles.
- Greene, David
 1992 Celtic. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: De Gruyter, 497–554.
- Griffith, Aaron
 2005 **-n(C)s in Celtic*. *Die Sprache* 45: 44–67.
- Griffith, Aaron
 2010 Die Etymologie der 1. und 2. Person Singular der altirischen notae augentes. In: Karin Stüber, Thomas Zehnder, and Dieter Bachmann (eds.), *Akten des 5. Deutschsprachigen Keltologensymposiums*. Vienna: Praesens, 109–121.
- Lambert, Pierre-Yves
 2003 [1994] *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*. 2nd edn. Paris: Errance.
- Lewis, Henry
 1990 *Handbuch des Mittelkornischen*. Deutsche Bearbeitung von Stefan Zimmer. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Lewis, Henry and J. R. F. Piette
 1990 *Handbuch des Mittelbretonischen*. Deutsche Bearbeitung von Wolfgang Meid. Innsbruck: Institut für Sprachwissenschaft der Universität.
- McCone, Kim
 1978 The dative singular of Old Irish consonant stems. *Ériu* 32: 26–38.
- McCone, Kim
 1986 From Indo-European to Old Irish: Conservation and innovation in the verbal system. In: D. Ellis Evans, John G. Griffith, and Edward Martyn Jope (eds.), *Proceedings of the 7th International Congress of Celtic Studies*. Oxford: D. E. Evans (Distributed by Oxbow Books), 222–266.
- McCone, Kim
 1991 *The Indo-European Origins of the Old Irish Nasal Presents, Subjunctives and Futures*. (Innsbrucker Beiträge zur Sprachwissenschaft 66). Innsbruck: Institut für Sprachwissenschaft der Universität.

McCone, Kim

- 1994 An tSean-Ghaeilge agus a Réamhstair [On Old Irish and its prehistory]. In: Kim McCone, Damian McManus, Cathal Ó Háinle, Nicholas Williams, and Liam Breatnach (eds.), *Stair na Gaeilge in ómós do Pádraig Ó Fiannachta* [The history of the Irish language in tribute to Patrick Fenton]. Maynooth: The Department of Old Irish, St. Patrick's College, 61–219.

McCone, Kim

- 1997 *The Early Irish Verb*. 2nd edn. (Maynooth Studies in Celtic Linguistics 1). Maynooth: Department of Old Irish, National University of Ireland.

McCone, Kim

- 2003 Old Irish na nní: A case of quid pro quo? *Celtica* 24: 168–181.

McCone, Kim

- 2005 *L'irlandese antico e la sua preistoria*. Translated by Elisa Roma. Alessandria: Edizione dell'Orso.

McCone, Kim

- 2006 *The origins and development of the Insular Celtic verbal complex*. (Maynooth Studies in Celtic Linguistics 6). Maynooth: Department of Old Irish, National University of Ireland.

Meid, Wolfgang

- 1963 *Die indogermanischen Grundlagen der altirischen absoluten und konjunkten Verbalflexion*. Wiesbaden: Harrassowitz.

Morris Jones, J.

- 1913 *A Welsh Grammar Historical and Comparative*. Oxford: The Clarendon Press.

Pedersen, Holger

- 1976 [1913] *Vergleichende Grammatik der keltischen Sprachen*. Vol. II. Göttingen: Vandenhoeck & Ruprecht. [Strassburg: Trübner.]

Rix, Helmut

- 1977 Das keltische Verbalsystem auf dem Hintergrund des indo-iranisch-griechischen Rekonstruktionsmodells. In: Karl Horst Schmidt and Rolf Ködderitzsch (eds.), *Indogermanisch und Keltisch. Kolloquium der Indogermanischen Gesellschaft am 16. und 17. Februar 1976 in Bonn*. Wiesbaden: Reichert, 132–158.

Schrijver, Peter

- 1997 *Studies in the history of Celtic pronouns and particles*. (Maynooth Studies in Celtic Linguistics 2). Maynooth: The Department of Old Irish, National University of Ireland.

Schumacher, Stefan

- 2000 *The Historical Morphology of the Welsh Verbal Noun*. (Maynooth Studies in Celtic Linguistics 4). Maynooth: The Department of Old Irish, National University of Ireland.

Schumacher, Stefan

- 2004 *Die keltischen Primärverben. Ein vergleichendes, etymologisches und morphologisches Lexikon*. Unter Mitwirkung von Britta Schulze-Thulin and Caroline aan de Wiel. Innsbruck: Institut für Sprachen und Literaturen der Universität.

Schumacher, Stefan

- 2007 Die Deutschen und die Nachbarstämme: Lexikalische und strukturelle Sprachkontaktphänomene entlang der keltisch-germanischen Übergangszone. In: Hans Hablitzel and David Stifter (eds.), *Johann Kaspar Zeuß im kultur- und sprachwissenschaftlichen Kontext (19. bis 21. Jahrhundert)*. Wien: Praesens, 176–207.

Solinas, Patrizia

- 1995 Il Celtico in Italia. *Studi Etruschi* 60: 311–408.

Stüber, Karin

- 1998 *The Historical Morphology of n-Stems in Celtic*. (Maynooth Studies in Celtic Linguistics 3). Maynooth: The Department of Old Irish, National University of Ireland.

Stüber, Karin

2015 *Die Verbalabstrakta des Altirischen*. (Munich Studies in Historical Linguistics 15). Bremen: Hempen.

Ternes, Elmar (ed.)

2011 *Brythonic Celtic – Britannisches Keltisch. From Medieval British to Modern Breton*. (Münchener Forschungen zur historischen Sprachwissenschaft 11). Bremen: Hempen Verlag.

Thurneysen, Rudolf

1946 *A Grammar of Old Irish*. Translated by D. A. Binchy and Osborn Bergin. Dublin: The Dublin Institute for Advanced Studies.

Untermann, Jürgen and Dagmar S. Wodtko

1997 *Monumenta Linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert.

Watkins, Calvert

1962 *Indo-European origins of the Celtic verb*. I. *The sigmatic aorist*. Dublin: The Dublin Institute for Advanced Studies.

Karin Stüber, Würzburg (Germany)

70. The syntax of Celtic

- | | |
|--|--------------------------|
| 1. Introduction | 4. Verbal morphosyntax |
| 2. Word classes | 5. Clausal configuration |
| 3. Nominal and pronominal morphosyntax | 6. Abbreviations |
| | 7. References |

1. Introduction

Syntax is the component of Celtic grammar that is least well understood. This chapter focuses upon the surface syntax of the Celtic languages, which differs considerably across the ancient and medieval languages. The medieval Goidelic and Brittonic languages are very well attested and will generally receive the greatest amount of attention in this chapter, but data from Celtiberian, Cisalpine Celtic (here used to include both traditionally termed “Leponitic” and “Cisalpine Gaulish”, cf. Eska 1998 and 2010: 24), and Transalpine Celtic (excluding Tartessian, which, *pace* Koch 2009, 2011, 2013, is in my view not a Celtic language [cf. Eska 2013a, 2013b, 2014]) will be cited as relevant to reveal the substantial amount of change that occurred in the prehistory of the Insular Celtic languages.

2. Word classes

The inflected word classes of the Celtic languages are those inherited from proto-Indo-European: nouns, pronouns, adjectives, and verbs. Non-inflected word classes likewise

<https://doi.org/10.1515/9783110523874-025>

Stüber, Karin

2015 *Die Verbalabstrakta des Altirischen*. (Munich Studies in Historical Linguistics 15). Bremen: Hempen.

Ternes, Elmar (ed.)

2011 *Brythonic Celtic – Britannisches Keltisch. From Medieval British to Modern Breton*. (Münchener Forschungen zur historischen Sprachwissenschaft 11). Bremen: Hempen Verlag.

Thurneysen, Rudolf

1946 *A Grammar of Old Irish*. Translated by D. A. Binchy and Osborn Bergin. Dublin: The Dublin Institute for Advanced Studies.

Untermann, Jürgen and Dagmar S. Wodtko

1997 *Monumenta Linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert.

Watkins, Calvert

1962 *Indo-European origins of the Celtic verb*. I. *The sigmatic aorist*. Dublin: The Dublin Institute for Advanced Studies.

Karin Stüber, Würzburg (Germany)

70. The syntax of Celtic

- | | |
|--|--------------------------|
| 1. Introduction | 4. Verbal morphosyntax |
| 2. Word classes | 5. Clausal configuration |
| 3. Nominal and pronominal morphosyntax | 6. Abbreviations |
| | 7. References |

1. Introduction

Syntax is the component of Celtic grammar that is least well understood. This chapter focuses upon the surface syntax of the Celtic languages, which differs considerably across the ancient and medieval languages. The medieval Goidelic and Brittonic languages are very well attested and will generally receive the greatest amount of attention in this chapter, but data from Celtiberian, Cisalpine Celtic (here used to include both traditionally termed “Leponitic” and “Cisalpine Gaulish”, cf. Eska 1998 and 2010: 24), and Transalpine Celtic (excluding Tartessian, which, *pace* Koch 2009, 2011, 2013, is in my view not a Celtic language [cf. Eska 2013a, 2013b, 2014]) will be cited as relevant to reveal the substantial amount of change that occurred in the prehistory of the Insular Celtic languages.

2. Word classes

The inflected word classes of the Celtic languages are those inherited from proto-Indo-European: nouns, pronouns, adjectives, and verbs. Non-inflected word classes likewise

<https://doi.org/10.1515/9783110523874-025>

are those inherited from proto-Indo-European: adverbs, adpositions, preverbs, particles, connectives, complementizers, and interjections.

3. Nominal and pronominal morphosyntax

3.1. Cases

The Continental Celtic languages demonstrate that the earliest Celtic continued all of the proto-Indo-European cases, at least in the singular number of the thematic flexion: nominative, accusative, dative, genitive, ablative, instrumental, locative, and vocative. In Goidelic, the dative, ablative, instrumental, and locative have merged and are conventionally labeled as “dative”, while in the Brittonic languages case flexion has disappeared (for some isolated nominal forms in these languages that continue inflected case forms, see Pedersen 1913: 72 and Koch 1991: 113–114; cf. also Padel 2013: 121–124).

3.1.1 Nominative case

The principal functions of the nominative case are to indicate the agent or experiencer in active sentences, the patient or recipient in passive sentences, and the predicative complement in copular sentences:

- (1) a. Cisalpine Celtic (RIG *E-5 = CIS 142 = CIM 277)
ṗTeKnaTi TruTiKni KarniTu arTuaś Koisis T|rṗTīKṗṗ
 A. D.GEN.SG set up.3.SG.PRET stone.ACC.PL K. D.NOM.SG
 ‘K. D. set up the burial stones of A. D.’
- b. Transalpine Celtic (RIG *G-154)
ονατιοοννου σο νεμετος κομμου εσκεγγιλου
 V.DAT.SG PROXIMATE.NOM.SG sacred.NOM.SG K. E.INSTR.SG
 ‘This sacred (object was offered) to V. by K. E.’
- c. Old Irish (Wb. 1^a1)
air is dia do cāch
 for COP.3.SG.PRES god.NOM.SG to everyone.DAT.SG
 ‘for he is God unto everyone.’

For other functions of the nominative case in Old Irish, see Thurneysen (1946: 156).

3.1.2 Accusative case

The principal functions of the accusative case are to indicate the patient in active sentences and as the complement of an adposition:

- (2) a. Transalpine Celtic (RIG G–27)
ουηβουμαρος δεδε παρανοου βρατου δεκαντεμ
 V.NOM.SG offer.3.SG.PRET T.DAT.SG gratitude.INSTR.SG tithe.ACC.SG
 ‘V. offered a tithe to T. with gratitude.’
- b. Old Irish (Wb. 3^{a5})
tri óen pheccad la adam
 through one sin.ACC.SG by A.ACC.SG
 ‘through one sin by A.’

For other functions of the accusative case in Old Irish, see Thurneysen (1946: 157–158).

3.1.3 Dative case

The principal function of the dative case as attested in Continental Celtic is to indicate the beneficiary/recipient:

- (3) a. Cisalpine Celtic (CIS 3 = CIM 34)
slaniai uerKalai Pala
 S.DAT.SG V.DAT.SG memorial-stone.NOM.SG
 ‘The memorial stone for S. V.’
- b. Transalpine Celtic (RIG L–10)
LICNOS CONTEXTOS IEVRV ANVALONNACV CANECOSEDLON
 L. C.NOM.SG offer.3.SG.PRET A.DAT.SG *canecosedlon*.ACC.SG
 ‘L. C. offered (this) *canecosedlon* to A.’

In Old Irish, it is principally used as the complement of adpositions, but other common functions are in comparative constructions and in apposition to personal pronouns:

- (4) a. Wb. 28^{b21}
re n-airite gráid iar mbathius
 before receive.VN.DAT.SG orders.GEN.SG after baptism.DAT.SG
 ‘before receiving orders (and) after baptism’
- b. Wb. 23^{a14}
ni diliu nech limm alailiu
 NEG.COP.3.SG.PRES dear.COMP anyone.NOM.SG with.1.SG another.DAT.SG
 ‘No-one is dearer to me than another.’
- c. Wb. 7^{a8}
hé-som triuss
 3.MASC.NOM.SG-TOP third.DAT.SG
 ‘he as third’

For other functions of the dative case in Old Irish, see Thurneysen (1946: 160–162). On Old Irish forms such as *-som* (“*notae augentes*”) as topics, see Griffith (2011a).

3.1.4 Genitive case

The principal function of the genitive case as attested in Continental Celtic is to indicate patronymy and, in Celtiberian, affiliation to a family group. There is at least one token in Transalpine Celtic in which it indicates possession. Other tokens of genitive nominals likely modify another nominal, but the uncertainty of translation does not allow a definitive analysis:

- (5) a. Celtiberian (MLH K.16.1)
TiřTanoř aPuloCum leTonTunoř Çę(niř) PęliCioř
 T.NOM.SG A.GEN.PL L.GEN.SG son.NOM.SG Beligos.NOM.SG
 ‘T. of the A.-family-group, son of L., a Belgian’
- b. Transalpine Celtic (RIG L–13)
MARTIALIS DANNŌTALI IEVRV VCVETE SOSIN CELICNON
 M.NOM.SG D.GEN.SG offer.3.SG.PRET U.DAT.SG PROXIMATE edifice.ACC.SG
 ‘M. son of D. offered this edifice to U.’
- c. Transalpine Celtic (RIG G–13)
εσκεγγολατι ανιατειοç ιμμι
 E.GEN.SG inalienable-property.NOM.SG be.1.SG.PRES
 ‘I am the inalienable property of E.’

In Old Irish, the genitive is principally used to modify another nominal, but is also the case of the complement of verbal nouns and certain adjectives:

- (6) a. Wb. 5^o14
icc in domuin
 salvation.NOM.SG DEF.GEN.SG world.GEN.SG
 ‘the salvation of the world’
- b. Wb. 5^o16
oc tuiste dũile hi tossuch
 at create.VN.DAT.SG element.GEN.PL in beginning.DAT.SG
 ‘in creating (the) elements in (the) beginning’
- c. Wb. 13^o8
am irlam techte martre cach
 COP.1.SG.PRES ready.NOM.SG go.VN.GEN.SG martyrdom.GEN.SG every
dia
 day.GEN.SG
 ‘I am ready to go to martyrdom every day.’

For other functions of the genitive case in Old Irish, see Thurneysen (1946: 158–159).

3.1.5 Ablative case

The ablative case is attested only in Celtiberian. Its only attested function is to indicate the origin of an individual in the Celtiberian onomastic formula:

- (7) MLH K.1.1 B4–5
aPulu lousoCum ušeisunoś PinTiś aCainas
 A.NOM.SG L.GEN.PL U.GEN.SG *bindis*.NOM.SG Akaina.ABL.SG
 ‘A. of L.-family group, (son of) U., a *bindis* from Akaina’

3.1.6. Instrumental case

The instrumental case is clearly attested in Transalpine Celtic to express the means by which an action is accomplished and the agent in a passive sentence:

- (8) a. RIG L–100
brixlja anderon
 magic.INSTR.SG infernal.GEN.PL
 ‘by the magic of infernal (deities)’
- b. RIG *G–154
ουατιουσνουι σο νεμετος κομμου εσκεγγιλου
 V.DAT.SG PROXIMATE.NOM.SG sacred.NOM.SG K. E.INSTR.SG
 ‘This sacred (object was offered) to V. by K. E.’

Transalpine Celtic possesses several forms in *-bi/-be* < IE instr. pl. **-b^his* whose status is indeterminate. It is clear from the token in (9) that this desinence was able to be used as a dative (Eska 2003: 112–115):

- (9) RIG L–13
MARTIALIS DANNŌTALI IEVRV VCVETE SOSĪN CELICNON ETIC
 M.NOM.SG D.GEN.SG offer.3.SG.PRET U.DAT.SG DEM edifice.ACC.SG CONN
GOBEDBI DVGIONTĪO VCVETIN IN ALISIĀ
 smith.DAT.PL serve.3.PL.PRES.REL V.ACC.SG in Alisia.²LOC².SG
 ‘M. son of D. offered this edifice to U. and to the smiths who serve U. in Alisia.’

The contexts in which forms such as *SVIOREBE* (RIG L–6), *mesamobi* (RIG L–66 A4), and *ejabi* (RIG L–98 1^{b9}) occur are not sufficiently clear to determine whether they function as datives or instrumentals.

Villar (1993–1995) proposes that several forms in *-u* in Celtiberian coin legends, listed in (10), are instrumentals that indicate toponyms as passive agents (thus, ‘(coin minted) by X’). Prósper (2011) discusses other Continental Celtic forms which she would identify as instrumentals:

- (10) a. **ColouñioCu** (MLH A.67.1)
 b. **eCualaCu** (MLH A.63.4)
 c. **oilauñu** (MLH A.56.2)
 d. **Tamaniu** (MLH A.79)
 e. **TaPaniu** (MLH A.90)

3.1.7. Locative case

The locative case is employed to indicate location. There are only a few fairly certain tokens attested in Continental Celtic languages:

- (11) a. Celtiberian (MLH K.1.1 A9)
śárniCiei ‘at *Sarnikiom’ (cf. thematic gen. sg. **śárniCio** [MLH K.1.1 A1, 11])
- b. Transalpine Celtic (RIG L–79)
IN ALIXIE
 in Alisia.LOC.SG
 ‘in Alisia’

3.1.8. Vocative case

The function of the vocative case is to identify the person(s) being addressed:

- (12) a. Transalpine Celtic (RIG L–66 A7)
gnate ‘O son!’
- b. Old Irish (Wb. 1^c8)
á huili duini .i. a cach duini
 VOC-PTLC all.VOC.SG man.VOC.SG i.e., VOC-PTLC every man.VOC.SG
 ‘O all man, i.e., O every man!’

Lambert (2003a: 57) suggests that several Transalpine Celtic *ā*-stem forms in -<a> in inscriptions engraved upon spindle whorls could be interpreted as in the vocative case.

3.2. Gender and number

3.2.1. Gender

All of the Continental Celtic languages and Old Irish retain the masculine, feminine, and neuter genders inherited from Indo-European. The neuter was lost in Middle Irish and in the Brittonic languages.

3.2.2. Number

While there is substantial evidence in Old Irish nominal flexion for the dual number beside the singular and plural, it has been lost in Brittonic. There is no certain evidence for the dual number in the Continental Celtic languages, though this is almost certainly owing to accident of attestation.

In Old Irish and the Brittonic languages, many plural forms continue old collectives. In Brittonic, the plurals that continue collectives possess singulative forms.

3.3. Head-complement configuration

3.3.1. Adjectives

The Continental Celtic languages are too fragmentarily attested and often in what are likely formulaic constructions for there to be certainty in the configuration of adjectives and nouns. In Celtiberian, one often finds adjectival forms preceding the head noun **Cař** as in (13a), but the adjectival token in (13b) follows its head noun. In Transalpine Celtic, adjectives seem most often to follow their head nouns, as in (13c–d):

- (13) a. MLH K.25.1
uirouiaCa Cař
- b. MLH K.25.1
TiriCanřam PeřCuneTaCam
- c. RIG G–153
τοοουτιο{v}ς ναμιασσις
 tribesman.NOM.SG Nimes.ADJ.NOM.SG
 ‘citizen of Nimes’
- d. RIG L–112
NATA VIMPI
 girl.NOM.SG pretty.NOM.SG
 ‘pretty girl’

In this last passage it is possible that the NP is inflected in the vocative case.

In Old Irish prose, with a small number of exceptions having variable position (notably *uile* ‘all’ and *sain* ‘separate’), attributive adjectives follow their head noun. Predicative adjectives, on the other hand, precede the noun. Both agree in case, number, and gender:

- (14) a. Wb. 17^{b2}
it mōra na bretha
 COP.3.PL.PRES great.FEM.NOM.PL DEF.FEM.NOM.PL judgment.FEM.NOM.PL
 ‘The judgments are great.’
- b. Wb. 13^{b5}
ar is miad mōr ind
 for COP.3.SG.PRES honor.NEUT.NOM.SG great.NEUT.NOM.SG DEF.NOM.SG
apstalacht
 apostleship.NOM.SG
 ‘For the apostleship is a great honor.’

In the Brittonic languages, most attributive adjectives follow their head noun, but it is normal for some to precede, e.g., MW *prif* ‘chief’, MBr. *ber* ‘short’. In Middle Welsh, gender agreement is the rule, but number is not; in Middle Cornish and Middle Breton, such agreement has all but disappeared. Predicative adjectives precede the noun and abide by the same agreement rules as for attributive adjectives:

- (15) a. Middle Breton (TPMB 17.3)
an scriptur glan
 DEF writing pure
 ‘Saint Writing’
- b. Middle Welsh (EWGP i 1.1)
llym awel, llum brin
 sharp wind bare hill
 ‘The wind is sharp, the hill is bare.’

3.3.2. Determiners

3.3.2.1 Articles

Articles are not certainly attested in the Continental Celtic languages. In Old Irish and the Brittonic languages, they always precede their head noun. In Old Irish, they agree for case, number, and gender, but such agreement has been lost in the Brittonic languages. In all of the Insular Celtic languages, they may be suffixed to adpositions and, in the Brittonic languages, to connectives, as well:

- (16) a. Old Irish (Wb. 12^{b5})
in corp uile
 DEF.MASC.ACC.SG body.MASC.ACC.SG whole.MASC.ACC.SG
 ‘the whole body’
- b. Middle Cornish (BK 1908)
a 'n goys ryal
 of DEF blood royal
 ‘of the royal blood’

3.3.2.2. Demonstratives

Demonstrative pronouns are attested in the Continental Celtic languages, in which they precede their head nouns and agree with them in case, number, and gender:

- (17) a. Celtiberian (MLH K.6.1)
śTań **CofTiCań**
 PROXIMATE.FEM.ACC.SG *kortika*.FEM.ACC.SG
 ‘this *kortika*’
- b. Transalpine Celtic (RIG *G–154)
σο νεμετος
 PROXIMATE.MASC.NOM.SG sacred.MASC.NOM.SG
 ‘this sacred (object)’

In the Insular Celtic languages, demonstrative pronouns are clitics that are attached to their head noun. It is important to note that the definite article always co-occurs with these clitics:

- (18) a. Old Irish (Wb. 5^c14)
ind rín so
 DEF mystery.NOM.SG PROXIMATE
 ‘this mystery’
- b. Middle Welsh (PKM 63.14)
y pryd hwnnw
 DEF creature DISTAL
 ‘that creature’

3.3.2.3. Possessive pronouns

Possessive pronouns are not certainly attested in the Continental Celtic languages. One possible token follows its head noun in Transalpine Celtic (19a), though, as Watkins (1999: 541) notes, its placement likely may be to create a ring composition with the imperative verb *MONI* ‘come!’ at the head of the inscription. Another token in a partially Latinized inscription (the Celticity of which, however, has been disputed, cf. Blom 2009: 19) suggests that the unmarked placement was before its head noun (19b):

- (19) a. RIG L–119
BVÐÐVTTON IMON
 ?penis?.ACC.SG 1.SG.POSS
 ‘my ?penis?’
- b. Marcellus Empiricus viii 171
in mon derco
 in 1.SG.POSS eye.ABL.SG
 ‘in my eye’

In the Insular Celtic languages, possessive pronouns are clitics that precede their head nouns. In all of the Insular Celtic languages, they may be suffixed to adpositions and, in the Brittonic languages, to connectives, as well:

- (20) a. Old Irish (Wb. 9^b5)
ní im aughtortás féin
 NEG in.1.SG.POSS authority.DAT.SG. RFLX
 ‘not in my own authority’
- b. Middle Welsh (Owein 3–4)
Gwenhwyuar a’e llawuorynyon
 G. CONN.3.SG.POSS.F handmaiden.PL
 ‘G. and her handmaidens’

c. Middle Breton (AMSG 16)

da roe
 2.SG.POSS King
 ‘your king’

3.3.3. Genitive complements

It appears that genitive complements, outside of the onomastic formula, precede their head nouns in Celtiberian, while unmarked genitive complements follow their head noun in Transalpine Celtic and in Old Irish. In the Brittonic languages, which do not inflect for case, nouns that function as genitive complements follow their head noun:

(21) a. Celtiberian (K.3.3)

TIASO TOGIAS
 N.GEN.SG N.ACC.PL

b. Transalpine Celtic (RIG L–100)

brixtja anderon
 magic.INSTR.SG infernal.GEN.PL
 ‘by the magic of infernal (deities)’

c. Old Irish (Wb. 15°12)

etarscarad coirp et anme
 severance.NOM.SG body.GEN.SG CONN soul.GEN.SG
 ‘the severance of body and soul’

d. Middle Cornish (ACD ii 32.389)

the 'n beth men
 to DEF grave stone
 ‘to the tomb of stone’

A special feature of the Insular Celtic languages is that, with a few exceptions in Old Irish and Middle Welsh (cf. Ó Gealbháin 1991), when a definite genitive complement is present, the article does not appear before a definite head noun:

(22) a. Old Irish (Wb. 5°14)

icc in domuin
 salvation.NOM.SG DEF.GEN.SG world.GEN.SG
 ‘the salvation of the world’

b. Middle Welsh (PKM 10.16)

penn y orsed
 head DEF mound
 ‘the top of the mound’

3.3.4. Topic/auxiliary pronouns

Clitic pronominal elements based upon the demonstrative stem **so-* known as “*notae augentes*” occur in Old Irish to express topicality. They are inflected for person, number, and gender (in the third person singular) and may be attached to nouns, verbs, and conjugated prepositions. Among the Brittonic languages, clitic personal pronouns probably served to express emphasis at an early period in Welsh, after which, as in Cornish and Breton, they served as markers of agreement. When employed with nouns, clitic personal pronouns must function as possessives:

- (23) a. Old Irish (Wb. 27^o22)
am cimbid-se
 COP.1.SG.PRES captive.NOM.SG-TOP.1.SG
 ‘I am a captive.’
- b. Middle Welsh (BD 44.8)
y uab ynteu
 3.SG.MASC.POSS son 3.SG.MASC
 ‘his son’
- c. Middle Breton (VSN 168)
dauedouch huy
 towards.2.PL 2.PL
 ‘towards you’

3.3.5. Numerals

3.3.5.1. Cardinal and ordinal numerals in Continental Celtic

Numerals that are not part of onomastic forms are not frequently attested in the Continental Celtic languages. There are some cardinal numerals attested in Celtiberian that precede their head nouns, and Transalpine Celtic evinces a full range of ordinal numerals from one to ten that follow their head nouns:

- (24) a. Celtiberian (MLH K.1.1 A4)
CanTom śanCiliśTařa
 100 *sanklistra*.ACC.PL
 ‘100 *sanklistra*.’
- b. Transalpine Celtic (GLG 12.1 = RIG L–29)
tuθos suexos
 allotment.NOM.SG sixth.NOM.SG
 ‘the sixth allotment’

3.3.5.2. Cardinal numerals in Insular Celtic

In Old Irish and the Brittonic languages, cardinal numerals precede their head noun. ‘Two’, ‘three’, and ‘four’ are inflected for case in Old Irish and gender in Old Irish and Middle Welsh. Nouns are inflected for the dual after ‘two’ and the plural after ‘three’ and higher in Old Irish. In the Brittonic languages, singular nouns are generally found after numerals higher than one, though there are many tokens of nouns inflected for the plural. In Old Irish and Middle Welsh, cardinal numerals other than the decades from eleven are constructed with the adposition ‘upon’. In Middle Cornish and Middle Breton, an older system compounding the single digit and ‘ten’ was preserved, e.g., MCorn. *pymthek* ‘15’. A trace of this system is attested in Old Irish, and there are more vestiges in Middle Welsh. Middle Cornish and Middle Breton adopt the adpositional system from ‘21’ upwards. In all of the Insular Celtic languages, an adpositional phrase with ‘of’ or ‘from’ may follow the cardinal numeral. This construction usually has a partitive sense in Old Irish. In this construction, the Brittonic languages employ plural nouns:

- (25) a. Middle Welsh (CA 425)
rwg dwy vydin
 between 2.FEM army
 ‘between two armies’
- b. Old Irish (MIB 104.16)
dā staid ar fícht
 2.MASC.NOM stadium.NOM.PL on 20
 ‘22 stadia’
- c. Old Irish (Wb. 7^b11)
óin di airchinchib assia
 1 from leader.DAT.PL Asia.GEN.SG
 ‘one of the leaders of Asia’
- d. Middle Breton (VSN 324)
try á diziou
 3.MASC of day.MASC.PL
 ‘three days’

3.3.5.3. Ordinal numerals in Insular Celtic

In Old Irish, ordinal numerals are inflected as adjectives and, aside from *tánise* ‘second’, precede their head nouns (but *aile* ‘other’ can also mean ‘second’ and may precede its head noun). The Brittonic languages inflect ‘third’ and ‘fourth’ for gender. Ordinal numerals follow their head nouns:

- (26) a. Old Irish (Wb. 14^d3)
isin técht tánisi
 in.DEF.DAT.SG come.VN.DAT.SG second.DAT
 ‘on the second coming’

- b. Old Irish (MI. 118^d18)
in chethramad accuis
 DEF.NOM.SG fourth cause.NOM.SG
 ‘the fourth cause’
- c. Middle Welsh (BD 144.15)
tryde ran yr enys
 third.FEM part DEF island
 ‘a third part of the island’

3.3.6. Adpositions

Among the Continental Celtic languages, adpositions are attested construed with a noun in Celtiberian and Transalpine Celtic. Celtiberian possessed both postpositions and prepositions, while in Transalpine Celtic, as in the Insular Celtic languages, adpositions are always prepositions:

- (27) a. Celtiberian (MLH K.1.1 A4)
ToCoĩTei eni
 T.LOC.SG at
 ‘at T.’
- b. Celtiberian (MLH K.1.1 A6)
enTařa Tiriř maTuř
 within 3 *matus*.ACC.PL
 ‘within three *matus*’
- c. Transalpine Celtic (RIG L–79)
IN ALIXIE
 in Alisia.LOC.SG
 ‘in Alisia’
- d. Old Irish (Wb. 29^a28)
fri laa bráth
 to day.ACC.SG judgment.GEN.SG
 ‘to the day of judgment’
- e. Middle Cornish (BK1230)
gans pobyl Christ
 by people C.
 ‘by the people of C.’

3.3.7. Complement clauses

It appears that complement clauses follow their head nouns in all Celtic languages, though the Celtiberian evidence is tentative:

- (28) a. Celtiberian (MLH K.1.3)
 [_{NP} **riśaTioCa leśTeŕa** [_{RC} **ia TaŕaCuai nouisa ausanTo eśCeninum Tanio-CaCue ōiśum alPana**]]
- b. Transalpine Celtic (RIG L–13)
 [_{NP} *GOBEDBI* [_{RC} *DVGIIONTILO* *ŲCVETIN IN ALISIĪA*]]
 smith.DAT.PL serve.3.PL.PRES.REL V.ACC.SG in Alisia.^{?LOC?}.SG
 ‘to the smiths who serve U. in Alisia’
- c. Old Irish (Wb. 6^d11)
is [_{NP} *h *]
 COP.3.SG.PRES 3.MASC.NOM.SG
 [_{RC} *no-don-nerta-ni*]]
 PV-1.PL.OBJ-strengthen.3.SG.PRES.REL-TOP.1.PL
 ‘It is he who strengthens us.’
- d. Middle Welsh (PKM 2.15)
 [_{NP} *yr erchwys* [_{RC} *a l dyssei y carw*]]
 DEF pack REL.PTCL kill.3.SG.PL.PERF DEF stag
 ‘the pack that had killed the stag’

4. Verbal morphosyntax

4.1. Number and person

4.1.1. Number

The Insular Celtic languages continue only the singular and plural numbers. Subject NPs that are inflected for the dual in Old Irish are construed with a plural verb. There is no evidence for dual verbal flexion in Continental Celtic as presently attested.

4.1.2. Person

All of the Celtic languages continue the three persons of verbal inflexion inherited from Indo-European.

4.2. Voice

There is a distinction between active and passive voice in the verb. Morphology suggests that passive forms are attested in the Continental Celtic languages, e.g., Celtib. **PinToŕ** (MLH K.1.1 A10) and Transalp. Celt. *tixsintor* (RIG L–98), but uncertainties of translation do not permit us to exclude the possibility that they are deponents. Transalp. Celt. *MARCOSIOR* [RIG L–117], however, appears to be a certain token of a deponent in Conti-

mental Celtic. In the Insular Celtic languages, there are forms only for the third person, though in Brittonic the plural is attested early. First and second person constructions employ the third singular form of the verb impersonally with an attached object agreement element:

- (29) a. Old Irish (Wb. 15^a3)
in bértar epistli uaín
 INTERROG bear.3.PL.FUT.PASS letter.NOM.PL from.1.PL
 ‘Shall letters be brought from us?’
- b. Middle Cornish (ACD 2.1)
en tas a nef y ’m gylwyr
 DEF father of heaven PTCL 1.SG call.3.SG.PASS
 ‘The Father of Heaven I am called.’

Monovalent verbs also have passive forms, which are used impersonally:

- (30) a. Old Irish (Wb. 9^a23)
righthir cuccuib
 go.3.SG.FUT.PASS to.2.PL
 ‘One will go to you.’
- b. Middle Breton (VSN 148)
pan vezer aman ganet
 when be.3.SG.PRES.PASS here be-born.PSTPTCPL
 ‘when one is born here’

4.3. Mood

The Celtic languages distinguish three moods: indicative, subjunctive, and imperative. There is good evidence for all three in the Continental, as well as Insular, Celtic languages.

4.3.1. Indicative mood

The indicative mood is employed to make declarative statements:

- (31) a. Transalpine Celtic (RIG L–50)
neđđamon delgu linda
 neighbor.GEN.PL hold.1.SG.PRES.IND drink.ACC.PL
 ‘I hold the drinks of neighbors.’
- b. Old Irish (Wb. 14^d38)
ni cuingem lóg ar precepte
 NEG seek.1.PL.PRES.IND reward.ACC.SG 1.PL.POSS teaching.GEN.SG
 ‘We do not seek a reward for our teaching.’

- c. Middle Welsh (CLIH 7.43^c)
gwercheidw *Llam yr Bwch Lloryen*
 guard.3.SG.PRES.IND Ll. y B. Ll.
 ‘Ll. guards Ll. y B.’

4.3.2. Subjunctive mood

The principal functions of the subjunctive mood in the Celtic languages are, on the basis of the Insular Celtic languages, to express wishes, commands, potentiality, purpose, and irrealis. (For other functions of the subjunctive mood in Old Irish and the Brittonic languages see Pedersen 1913: 312–320 and 320–331, respectively.) There are a number of subjunctive verbs attested in Celtiberian and Transalpine Celtic, but it is usually not possible to be certain of their precise function.

4.3.2.1. Wishes

Tokens of the subjunctive mood to indicate a wish are:

- (32) a. Old Irish (Wb. 31^{a2})
d-a-ro-lgea *dia* *doib*
 PV-3.SG.OBJ·PFCTV-forgive.3.SG.PRES.SUBJ God.NOM.SG to.3PL
 ‘God forgive it to them.’
- b. Middle Welsh (PKM 30.11)
duw a rodo *da ywch*
 God PTCL give.3.SG.PRES.SUBJ good to.2.PL
 ‘May God give good to you.’

4.3.2.2. Commands

The subjunctive mood may be employed to express commands. In root clauses in Old Irish, it usually has the sense of a future imperative:

- (33) a. Old Irish (Wb. 5^{d39})
do:gné-su *maith* *fris-som*
 do.2.SG.PRES.SUBJ-TOP.2.SG good.ACC.SG to.3.SG.MASC-TOP.3.SG.MASC
 ‘You will do good to him.’
- b. Middle Breton (VSN 388)
doe, guir roe tron, ra ’m pardono
 God true king heaven PTCL 1.SG.OBJ pardon.3.SG.PRES.SUBJ
 ‘Let God, the true king of heaven, pardon me!’

4.3.2.3. Potentiality

The subjunctive mood can be employed to express the potential for events to occur:

- (34) a. Old Irish (Wb. 10
- ^d
- 24)

mani pridag at·bél ar ocht et
if.NEG preach.1.SG.PRES.SUBJ die.1.SG.FUT from cold.ACC.SG CONN
gorti
hunger.ACC.SG
‘If I do not preach, I will die from cold and hunger.’

- b. Middle Welsh (PKM 3.5–6)

a manac ditheu y mi pa furyf y gallwyf
CONN tell.2.SG.IMPV 2.SG to 1.SG what form PTCL be-able.1.PRES.SUBJ
hynny
DISTAL
‘And tell me how I may be able to do that!’

4.3.2.4. Purpose

The subjunctive mood may be employed to express purpose:

- (35) a. Old Irish (Wb. 11
- ^a
- 7)

ar mbad irlamu de don buáith
so-that COP.3.SG.PRET.SUBJ ready.COMP from for.DEF.DAT.SG victory.DAT.SG
‘so that he might be more ready for the victory’

- b. Middle Breton (VSN 1156)

euit maz guili
for that see.2.SG.SUBJ
‘so that you may see’

4.3.2.5. Irreality

The subjunctive mood can be employed to express that which is not real or which is impossible:

- (36) a. Old Irish (Wb. 14
- ^b
- 15)

ní tabir dia forn-ni didiu
NEG put.3.SG.PRES God.NOM.SG upon.1.PL-1.PL.TOP then
fochith nád fochomolsam
suffering.VN.ACC.SG NEG endure.1.PL.PFTV.PRES.SUBJ.REL
‘God does not put upon us, then, suffering which we cannot endure.’

b. Middle Welsh (PKM 4.3)

nit oes yndi nep ni 'th adnappo
 NEG be.3.SG.PRES.INDEF in.3.SG.FEM anyone NEG 2.SG know.3.SG.PRES.SUBJ
 'There is not anyone in it who does not know you.'

4.3.3. Imperative mood

The imperative mood is employed in commands and prohibitions. It is well attested in the Continental Celtic, as well as the Insular Celtic, languages. In Celtiberian, the attested evidence continues the so-called future imperative. It is noteworthy that imperative verbs are not fronted to the left periphery of the clause, as is common cross-linguistically, in Celtiberian, whereas they are in the other Celtic languages:

(37) a. Celtiberian (MLH K.1.1 A10)

TeCameTam TaTus
 tithe.ACC.SG give.3.SG.FUT.IMPV
 'Let him offer a tithe.'

b. Transalpine Celtic (RIG L-119)

MONI GNATHA GABI BVDDVITTON IMON
 come.2.SG.IMPV girl.VOC.SG take.2.SG.IMPV 'penis'.ACC.SG 1.SG.POSS
 'Come, girl, take my 'penis'!'

c. Old Irish (Wb. 23^e11)

dénid a n-as-berar frib
 do.2.PL.IMPV DEM.REL say.3.SG.PRES.PASS to.2.PL
 'Do that which is said to you!'

d. Middle Welsh (PKM 25.12-13)

nac ewch bellach hynny
 NEG go.2.PL.IMPV far.COMP DISTAL
 'Do not go beyond that!'

4.4. Tense

The tense system of the Insular Celtic languages includes the present, imperfect, and preterite. Old Irish also has future and secondary future (= conditional) tenses, while the Brittonic languages have a pluperfect tense. Old Irish and Middle Welsh, furthermore, possess a consuetudinal present tense in the verb 'be'. The Continental Celtic languages provide evidence for the present, future, and preterite tenses.

4.4.1. Present tense

The principal function of the present tense is to express action occurring in present and indefinite time, including action proceeding up to the present:

- (38) a. Transalpine Celtic (RIG L–120)
GENEŦĀ IMI DAGA VIMPI
 girl.NOM.SG COP.3.SG.PRES good.NOM.SG pretty.NOM.SG
 ‘I am a good and pretty girl.’
- b. Old Irish (Wb. 18^c10)
berit in soscéle
 bring.3.PL.PRES DEF.ACC.SG gospel.ACC.SG
 ‘They bring the gospel.’
- c. Middle Welsh (RM 126.28)
ys gwers yd wyf yn keissaw
 COP.3.SG.PRES while PTCL be.1.SG.PRES PTCL seek.VN
 ‘I have been looking for a while.’

For other functions of the present tense in Middle Welsh, see Evans (1964:108–109); for Middle Breton, see Hemon (1975: 252–253). In Middle Cornish, as well as Middle Welsh, the present tense can also express futurity.

4.4.2. Consuetudinal present tense

The principal function of the consuetudinal present tense is to express action that is habitual. In the Brittonic languages, it is also employed to express futurity:

- (39) a. Old Irish (16^d8)
biuu-sa oc irbáig dar far
 be.1.SG.CONS-PRES-1.SG.TOP at boast.VN.DAT.SG over 2.PL.POSS
cenn-si fri maccidóndu
 head.ACC.SG-2.PL.TOP towards Macedonian.ACC.PL
 ‘I am boasting on your behalf to the Macedonians.’
- b. Middle Welsh (Owein 305)
a minneu a vydaf ar yr esgynuaen racko
 CONN 1.SG PTCL be.1.SG.CONS-PRES on DEF mounting-block yonder
 ‘And I will be on the mounting block yonder.’

4.4.3. Imperfect tense

In Old Irish, the principal function of the imperfect tense is to express repeated or usual action in the past. In the Brittonic languages, it also expresses continuation in the past, and, in Middle Welsh and Middle Cornish, can express conditionality:

- (40) a. Old Irish (Wb. 15^a18)
do·gnithe a n-asbered Moysi
 do.3.SG.IMPF.PASS DEM.REL say.3.SG.IMPF M.NOM.SG
 ‘That which M. used to say used to be done.’

- b. Middle Welsh (WM 433.8)
ual y kyrchei ef y bont
 as PTCL approach.3.SG.IMPF 3.SG.M DEF bridge
 ‘as he approached the bridge’
- c. Middle Welsh (RM 268.29)
ny wydyat hi beth a wnaei
 NEG know.3.SG.IMPF 3.SG.F what PTCL do.3.SG.IMPF
 ‘She did not know what she would do.’

For other functions of the imperfect tense in Middle Welsh, see Evans (1964: 109–111).

4.4.4. Preterite tense

The function of the preterite tense is to indicate a completed action or state in the past:

- (41) a. Cisalpine Celtic (CIS 65 = CIM 180)
uvamoKozis Plialeθu uvlTiauiuPos ariuonePos siTeś
 ‘U. B.NOM.SG U. A.DAT.PL *sidets*.ACC.PL
TeTu
 give.3.SG.PRET
 ‘U. B. gave *sidets* to the U. A.’
- b. Old Irish (Wb. 21^c22)
ní fitir cid muntar nime
 NEG.3.SG.OBJ know.3.SG.PRET even family.NOM.SG heaven.GEN.SG
 ‘Not even the family of heaven knew it.’
- c. Middle Cornish (BK 311)
der thowgys e tathorhas
 by-means-of deity PTCL rise-again.3.SG.PRET
 ‘Through deity he arose again.’

4.4.5. Pluperfect tense

The principal function of the Brittonic pluperfect tense is to indicate the completion of an action prior to present time:

- (42) Middle Welsh (PKM 90.16)
ny welsei neb ar wr dremynt druanach
 NEG see.3.SG.PLUPF anyone on man sight wretched.COMP
 ‘No-one had seen a more wretched sight on man.’

4.4.6. Future tense

The principal function of the Old Irish future tense is to indicate action at a future time:

(43) Old Irish (Wb. 1^a3)

húare ro·creitset ardlathi in betho
 because PV·believe.3.SG.PRET chief-prince.NOM.PL DEF.GEN.SG world.GEN.SG
cretfed cách iarum
 believe.3.SG.FUT everyone.NOM.SG then
 ‘Because the chief princes of the world believed, everyone will believe then.’

4.4.7. Secondary future

The principal functions of the Old Irish secondary future are to indicate an action which is completed in the future from a past point of view and to indicate an action which could, should, or would happen under appropriate circumstances:

(44) a. Old Irish (Wb. 7^a2)

is díim-sa tairrchet
 COP.3.SG.PRES of.1.SG-TOP.1.SG prophesy.3.SG.PASS.PRET.REL
ad·cichitis genti per me
 see.3.PL.SEC-FUT Gentile.NOM.PL per me
 ‘It is of me that it has been prophesied that Gentiles would see *per me*.’

b. Old Irish (Wb. 13^b3)

mad áill dúib cid accaldam
 if.COP.3.SG.PRET.SUBJ desire.NOM.SG to.2.PL even converse.VN.NOM.SG
neich díib da·rigénte
 anyone.GEN.SG of.3.PL PV-3.SG.OBJ.do.2.PL.SEC.FUT
 ‘If you even desire to converse with any of them, you would be able to do it.’

4.5. Periphrasis

The Insular Celtic languages possess a periphrastic construction in which a finite form of the verb ‘be’ is construed with an aspectual particle plus verbal noun. The Brittonic languages also make periphrastic constructions with a finite form of the verbs ‘be’, ‘have’, or ‘do’ with a verbal noun or past participle:

(45) a. Old Irish (Wb. 26^b14)

ní-r bommar utmuill oc foigdi
 NEG-PV be.1.PL.PRET restless.NOM.PL at beg.VN.DAT.SG
 ‘We are not restless in begging.’

b. Middle Welsh (WM 408.7)

y mae gvedy mynet gyd a Gwenhwyuar y
 PTCL be.3.SG.PRES after go.VN together with G. to.3.SG.F
hystauell
 chamber
 ‘She has gone with G. to her chamber.’

c. Middle Breton (MM 1497)

heulyet hon eux
 follow.PSTPRTCP 1.PL be.3.SG.PRES
 ‘We have followed.’

4.6. Agreement

4.6.1. Subject agreement

It is probable that verbs agree with their subject for person and number in the Continental Celtic languages, though there are no certain tokens of full subject NPs agreeing with the verb outside of the third singular. In Old Irish and Old Brittonic, subject NPs agree with their verb for person and number. In the Middle phase of the Brittonic languages, the verb normally occurs in its third singular form when it precedes its subject, save for when the subject is a personal pronoun, though this is not the rule in early Middle Welsh:

(46) a. Cisalpine Celtic (CIS 65 = CIM 180)

[uvamoKozis Plialeθu]_i uvITiauiopos ariuonePos siTeś
 U. B.NOM.SG U. A.DAT.PL *sidets*.ACC.PL
TeTu_i
 give.3.SG.PRET
 ‘U. B. gave *sidets* to the U. A.’

b. Old Irish (Wb. 5^c6)

ní-m charat_i-sa [ind fir]_i hore
 NEG-1.SG.OBJ love.3.PL.PRES-1.SG.TOP DEF.NOM.PL man.NOM.PL because
pridchim soscele do gentib
 preach.1.SG.PRES gospel.ACC.SG to Gentile.DAT.PL
 ‘The men do not love me because I preach the gospel to the Gentiles.’

c. Middle Cornish (BK 864)

me_i a s nask_i thys war un dro
 1.SG PART 3.PL.OBJ yoke.3.SG.PRES for.2.SG over one moment
 ‘I shall yoke them for you in a moment.’

4.6.2. Object agreement

Eska (2009–2010) and Griffith (2011b) argue that the morphemes traditionally known as infixes and suffixed pronouns are, in fact, object agreement affixes. Part of the basis

register contexts to create phonological ornamentation via alliteration. Thus, in the following passage the placement of the verb in final position in the clause creates an extended /k/-alliteration:

- (48) AM 6.43–44
ath- márcathu fri crícha comnámát
 PV great-battalion.ACC.PL to border.ACC.PL hostile-neighbor.GEN.PL
-cuirethar
 dispatch.3.SG.PRES
 ‘He dispatches great battalions to the borders of hostile neighbors.’

This construction likely does not directly continue the proto-Indo-European situation, but is an innovation based upon the Bergin’s Rule construction, which places the normally clause-initial verb in medial or final position (Eska 2007).

4.9. Clitics

There are numerous clitics in the Celtic languages, including pronouns, preverbs, particles, negators, adpositions, connectives, and complementizers. They may have a narrow scope and be attached to the words with which they are directly construed or a broader, sentential scope, in which case they most frequently occur after the first stressed form or constituent in the clause.

In the fragmentarily attested Continental Celtic languages, there is evidence for only a small range of clitics. Celtiberian attests only connective =**Cue** and disjunctive =**ue**, which evidently can be attached to either all members of a coordination or the final one, and the negator **ne**=, which is attached directly to its verb. Cisalpine Celtic attests the connective =**Pe** and two object pronouns, which occur in clause-second position. Transalpine Celtic likewise attests negators attached to their verbs as proclitics, various emphasizing pronouns attached to their verbs as enclitics, and object pronouns, which occur in clause-second position. While it is clear that adpositions in Celtiberian were stressed, it appears that they often were clitics in Transalpine Celtic.

In the Insular Celtic languages, all of the word categories listed above occur as clitics. Preverbs frequently are proclitic in Old Irish, but this is the situation in only the oldest stages of the Brittonic languages.

5. Clausal configuration

5.1. Unmarked configurations

5.1.1. With full NP object

The unmarked clausal configuration of the various Celtic languages differs among them. Celtiberian and early Cisalpine Celtic continue the Subject-Object-Verb order reconstructed for proto-Indo-European:

- (49) a. Celtiberian (MLH K.1.1 A7–8)
CuśTaiCoś afsnaś CuaTi
 NOM.SG ACC.PL 3.SG.PRES
- b. Cisalpine Celtic (CIS 65 = CIM 180)
uvamoKozis Plialeθu uvlTiauoPos ariuonePos siTeś
 ‘U. B.NOM.SG U. A.DAT.PL *sidets*.ACC.PL
TeTu
 give.3.SG.PRET
 ‘U. B. gave *sidets* to the U. A.’

In later Cisalpine Celtic and Transalpine Celtic, the unmarked configuration is Subject-Verb-Object:

- (50) Transalpine Celtic (RIG L–10)
LICNOS CONTEXTOS IEVRV ANVALONNACV CANECOSEDLON
 L. C.NOM.SG offer.3.SG.PRET A.DAT.SG *canecosedlon*.ACC.SG
 ‘L. C. offered (this) *canecosedlon* to A.’

In Old Irish and Old Brittonic, the unmarked configuration is Verb-Subject-Object:

- (51) a. Old Irish (Wb. 12^c22)
ro-cluinethar cách in fogur
 hear.3.SG.PRES everyone.NOM.SG DEF.ACC.SG sound.ACC.SG
 ‘Everyone hears the sound.’
- b. Old Welsh (OWWM 77)
prinnit hinnoid .iiii. aues
 buy.3.SG.PRES DISTAL four *aues*
 ‘That buys four birds.’

In the Middle phase of the Brittonic languages, however, the unmarked configuration is verb-second, with the initial constituent generally indicating topicality:

- (52) a. Middle Welsh initial subject (PKM 2.3)
ef a welei uarchauc
 3.SG.M PTCL see.3.SG.IMPF rider
 ‘He could see a rider.’
- b. Middle Welsh initial object (PKM 15.11)
digawn rydodet ymman
 sufficiency PFCTV.put.IMPRSL herein
 ‘A sufficiency has been put herein.’
- c. Middle Welsh initial PP (PKM 1.24)
at y cwn y doeth ef
 to DEF hound.PL PTCL come.3.SG.PRET 3.SG.M
 ‘To the hounds he came.’

In Middle Brittonic negative clauses, both verb-initial and verb-second configurations occur:

- (53) a. Middle Cornish negated initial verb (BK 10)
ny wyllyn da na drog
 NEG see.1.SG.IMPV good nor bad
 ‘I saw neither good nor bad.’
- b. Middle Cornish negated verb second (BK 46)
ow negys ny wothvethys
 1.SG.POSS business NEG know.2.SG.CONV-PRES
 ‘My business you will not know.’

5.1.2. With pronominal NP object

In later Cisalpine Celtic and Transalpine Celtic, it is noteworthy that when the object of a clause is a clitic pronoun, the verb occurs in initial position in the clause and hosts the clitic:

- (54) a. Cisalpine Celtic (CIM 233)
to-me-declai obalda natina
 PV-1.SG.OBJ-set-up.3.SG.PRET O.NOM.SG daughter.NOM.SG(.DIM)
 ‘O., (their) (dear) daughter, set me up.’
- b. Transalpine Celtic (GLG 14.20–21 = RIG L–31)
siox-t-i albanos panna extra tud(du)
 add.3.SG.PRET-3.PL.OBJ A.NOM.SG vessel.ACC.PL beyond allotment.DAT.SG
ccc
 300
 ‘A. added vessels beyond the allotment (in the amount of) 300.’

Although some doubt the Celticity of (54a), Eska and Wallace (2011) defend its status as a Celtic epigraphic document. Their reading differs from that of CIM.

5.2. Examples of marked movement

5.2.1. Continental Celtic

In these languages, with their robust nominal and verbal morphology, movement may occur for various pragmatic information-structuring purposes, as well as for purposes of literary ornamentation:

- (55) a. Movement of verb to left edge of clause: Celtiberian (MLH K.1.1 A6)
ašeCaTi_i [a]mPiTinCounci sTena t_i
 3.SG.PRES.SUBJ VN.DAT.SG ACC.PL

- b. Movement of noun to left edge of clause: Transalpine Celtic (RIG L-3)
 [RATĪN BRIVĀTIOM]_i FRONŪV TARBETIS[o]NĪOS IE{I}VRV t_i
 fort.ACC.SG B.GEN.PL F. T.NOM.SG offer.3.SG.PRET
 ‘F. T. offered the fort of the B.’
- c. Postposition of subject to right edge of clause; leftward extraction of genitive NP: Cisalpine Celtic (RIG *E-5 = CIS 142 = CIM 277)
 [[aT]eKnaTi TruTi[K]ni]_j t_i [Kar]niTu loKan t_j [Ko[i]s]is
 A.D.GEN.SG bury.3.SG.PRET urn.ACC.SG K. D.NOM.SG
 [Tr]uTiKnos]_i
 ‘K. D. buried the urn of A. D.’
- d. Left dislocation; clitic doubling:
 Cisalpine Celtic (RIG *E-2 = CIS 141 = CIM 100)
 [aKisios arKaToKo{K}maTereKos]_i To-šo_j-KoTe_i
 A.NOM.SG PV-3.SG.OBJ-give.3.SG.PRET
 [aTom TeuoχToniϕn]_j
 boundary.ACC.SG god-man.GEN.PL
 ‘A. A. gave the boundary of gods and men.’

5.2.2. Insular Celtic

5.2.2.1. Topicalization

In Old Irish and Old Brittonic, left dislocation indicates topicalization. The topicalized NP is often resumed with a pronoun in the clause, particularly if it is a notional object:

- (56) a. Old Irish (Wb. 14^b17)
 [a ndu:gniat ar magistir]_i is ferr
 what do.3.PL.PRES.REL 1.PL.POSS master.NOM.PL COP.3.SG.PRES good.COMP
 dún a_i dénum
 to.1.PL 3.SG.POSS.NT do.VN.NOM.SG
 ‘That which our masters do, it is better for us to do it.’
- b. Old Welsh (OWWM 77)
 [ir pimphet eterin diguormechis lucas]_i hegit hunnoid_i
 DEF five.ORD bird add.3.SG.PRET.REL L. go.3.SG.PRES DISTAL
 ‘The fifth bird that L. added, that one goes.’

5.2.2.2. Focalization

In Old Irish and the Brittonic languages, clefting indicates focalization:

(57) a. Old Irish (Wb. 4^c15)

ar ba miscuis at-roilliset dib
 for COP.3.SG.PRET hatred.NOM.SG deserve.3.PL.PRET.REL two.DAT.PL
línaib
 number.DAT.PL
 ‘For it was hatred that they had both deserved.’

b. Old Welsh (Comp 12–13)

is aries isid in arcimeir
 COP.3.SG.PRES A. COP.3.SG.PRES.REL in opposite
 ‘It is A. that is opposite.’

5.2.2.3. Alternative subject and object positions in Early Irish

Mac Giolla Espaing (1980) demonstrates that subject and object NPs may be postposed to the end of the clause for a variety of reasons, including:

- (58) a. They are heavy, either modified by a relative clause or otherwise lengthy.
 b. They function as verbal noun phrases.
 c. They indicate “emphasis”.
 d. They indicate that two consecutive actions are contrasted.
 e. They reintroduce old information into the discourse after a lengthy absence.
 f. They indicate a change in focus.

Additionally, Lash (2014) demonstrates that in addition to the usual position of the subject immediately after the verb which may be followed by a demarcating adverb, there is also a subject position that follows a demarcating adverb:

(59) a. MT 142.14–15

mad do-airli sale dano ind laim oc praind
 if fall.3.SG.PRES spittle.NOM.SG ADV in hand.ACC.SG at meal.DAT.SG
 ‘Now, if spittle falls into the hand at a meal ...’

b. LGTH 316.17–18

Do-ber dano rí Locha Léin a gāill
 give.3.SG.PRES ADV king.NOM.SG L. L.GEN.SG 3.SG.POSS.M hostage.ACC.SG
do rí[g] Ciarraige fria folta
 to king.DAT.SG C.GEN.SG towards.3.SG.POSS.M obligation.ACC.PL
tēchti
 proper.ACC.PL
 ‘The king of L. L. gives his hostage to the king of C. in accord with his proper obligations.’

Lash proposes that the position before the adverb encodes old information such as a familiar topic or an aboutness topic, while the position after the adverb encodes new information such as a switch topic (speaker/hearer old, but context new) or presentational focus.

6. Abbreviations

ACD	Norris (1859)	MLH K	Untermann (1997: 349–722)
AM	Kelly (1976)	MM	Ernault (1914)
AMSG	Ernault (1935)	MT	Gwynn and Purton (1911–1912)
BD	Lewis (1942)	Owein	Thomson (1968)
BK	Thomas and Williams (2007)	OWWM	Lambert (2003b)
CA	Williams (1961)	PKM	Williams (1951)
CIM	Morandi (2004)	RIG E	Lejeune (1988: 1–54)
CIS	Solinas (1995)	RIG G	Lejeune (1985)
Comp	Falileyev (2008: 97–104)	RIG L–1–16	Lejeune (1988: 55–194)
EWGP	Jackson (1935)	RIG L–18–139	Lambert (2002)
GLG	Marichal (1988)	RM	Evans and Rhys (1887)
LGTH	Meyer (1912)	TPMB	Hemon (1981)
MIB	Bergin (1932)	VSN	Ernault (1887)
MI.	Stokes and Strachan (1903: 7–483)	Wb.	Stokes and Strachan (1903: 499–712)
MLH A	Untermann (1975)	WM	Evans (1907)

7. References

- Bergin, Osborn
1932 A mystical interpretation of the *Beati. Ériu* 11: 103–106.
- Blom, Alderik H.
2009 Gaulish in the formulae of Marcellus of Bordeaux? Methodological considerations. In: Stefan Zimmer (ed.), *Kelten am Rhein. Akten des dreizehnten Internationalen Keltologiekongresses ii, Sprachen und Literaturen*. Mainz am Rhein: von Zabern, 13–24.
- Ernault, Émile
1887 La vie de Saint Nonne. *Revue celtique* 8: 230–301, 405–491.
- Ernault, Émile
1914 *Le miroir de la mort, poème breton du xv^{ième} siècle*. Paris: Champion.
- Ernault, Émile
1935 *L'ancien mystère de Saint-Gwénolé*. Rennes: Plihon.
- Eska, Joseph F.
1998 The linguistic position of Lepontic. *Proceedings of the Berkeley Linguistics Society. Special session on Indo-European subgrouping and internal relations* 24S: 2–11.
- Eska, Joseph F.
2003 On syntax and semantics in Alise-Sainte-Reine (Côte-d'Or), again. *Celtica* 24: 101–120.

- Eska, Joseph F.
2007 Bergin's Rule. Syntactic diachrony and discourse strategy. *Diachronica* 24: 253–278.
- Eska, Joseph F.
2009 The emergence of the Celtic languages. In: Martin J. Ball and Nicole Müller (eds.), *The Celtic languages*. 2nd edition. London: Routledge, 22–27.
- Eska, Joseph F.
2009–2010 Where have all the object pronouns gone? *Zeitschrift für celtische Philologie* 57: 25–47.
- Eska, Joseph F.
2013a Review of Koch 2009. *Kratylos* 58: 58–67.
- Eska, Joseph F.
2013b Review of Koch 2011. *Kratylos* 58: 68–73.
- Eska, Joseph F.
2014 Comments on John T. Koch's Tartessian-as-Celtic enterprise. *The Journal of Indo-European studies* 42: 428–438.
- Eska, Joseph F. and Rex Wallace
2011 Script and language at ancient Voltino. *Alessandria* 5: 93–113.
- Evans, D. Simon
1964 *A grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies.
- Evans, J. Gwenogvryn
1907 *The White book Mabinogion; Welsh tales and romances reproduced from the Peniarth manuscripts*. Pwllheli: J. Gwenogvryn Evans.
- Evans, J. Gwenogvryn and John Rhys
1887 *The text of the Mabinogion; and other Welsh tales from the Red Book of Hergest*. Oxford: J. Gwenogvryn Evans.
- Falileyev, Alexander I.
2008 *Le vieux-gallois*. Potsdam: Université de Potsdam.
- Griffith, Aaron
2011a The genesis of the animacy hierarchy in the Old Irish *notae augentes*. In: Thomas Krisch and Thomas Lindner (eds.) unter redaktioneller Mitarbeit von Michael Crombach und Stefan Niederreiter. *Indogermanistik und Linguistik im Dialog*. Wiesbaden: Reichert, 182–191.
- Griffith, Aaron
2011b Old Irish pronouns. Agreement affixes vs. clitic arguments. In: Andrew Carnie (ed.), *Formal approaches to Celtic linguistics*. Cambridge: Cambridge Scholars Publishing, 65–93.
- Griffith, Aaron
2015 Degrees of agreement in Old Irish. In: Jürg Fleischer, Elisabeth Rieken, and Paul Widmer (eds.), *Agreement from a diachronic perspective*. Berlin: De Gruyter, 165–188.
- Gwynn, Edward J. and Walter J. Purton (eds.)
1911–1912 The Monastery of Tallaght. *Proceedings of the Royal Irish Academy* 29C: 115–179.
- Hemon, Roparz
1975 *A historical morphology and syntax of Breton*. Dublin: Dublin Institute for Advanced Studies.
- Hemon, Roparz
1981 *Trois poèmes en Moyen-Breton*. Dublin: Dublin Institute for Advanced Studies.
- Jackson, Kenneth H.
1935 *Early Welsh gnomic poems*. Cardiff: University of Wales Press.
- Kelly, Fergus (ed.)
1976 *Audacht Morainn*. Dublin: Dublin Institute for Advanced Studies.

- Koch, John T.
 1991 Gleanings from the *Gododdin* and other Early Welsh texts. *Bulletin of the Board of Celtic Studies* 38: 111–118.
- Koch, John T.
 1996 Some thoughts on the Gaulish inscription from Larzac. In: Wolfgang Meid and Peter Anreiter (eds.), *Die grösseren altkeltischen Sprachdenkmäler*. Innsbruck: Institut für Sprachwissenschaft der Universität, 37–40.
- Koch, John T.
 2009 *Tartessian. Celtic in the south-west at the dawn of history*. Aberystwyth: Celtic Studies Publications.
- Koch, John T.
 2011 *Tartessian 2. The inscription of Mesas do Castelinho. ro and the verbal complex. Preliminaries to historical phonology*. Aberystwyth: University of Wales Centre for Advanced Welsh and Celtic Studies.
- Koch, John T.
 2013 *Tartessian. Celtic in the south-west at the dawn of history*. 2nd edn. Aberystwyth: Celtic Studies Publications.
- Lambert, Pierre-Yves
 2002 *Recueil des inscriptions gauloises ii/2, Textes gallo-latins sur instrumentum*. Paris: Centre National de la Recherche Scientifique.
- Lambert, Pierre-Yves
 2003a *La langue gauloise. Description linguistique, commentaires d'inscriptions choisies*. 2nd edn. Paris: Errance.
- Lambert, Pierre-Yves
 2003b The Old Welsh glosses on Weights and Measures. In: Paul Russell (ed.), *Yr hen iaith. Studies in early Welsh*. Aberystwyth: Celtic Studies Publications, 103–134.
- Lash, Elliott
 2014 Subject positions in Old and Middle Irish. *Lingua* 148: 278–308.
- Lejeune, Michel
 1985 *Recueil des inscriptions gauloises i, Textes gallo-grecs*. Paris: Centre National de la Recherche Scientifique.
- Lejeune, Michel
 1988 *Recueil des inscriptions gauloises ii/1, Textes gallo-étrusques. Textes gallo-latins sur pierre*. Paris: Centre National de la Recherche Scientifique.
- Lewis, Henry
 1942 *Brut Dingestow* [Dingestow Chronicle]. Cardiff: University of Wales Press.
- Mac Giolla Espaig, Dónall
 1980 Aspects of variant word order in Early Irish. *Ériu* 31: 28–38.
- Marichal, Robert
 1988 *Les graffites de La Graufesenque*. Paris: Centre National de la Recherche Scientifique.
- Meyer, Kuno (ed.)
 1912 The Laud genealogies and tribal histories. *Zeitschrift für celtische Philologie* 8: 291–338.
- Morandi, Alessandro
 2004 *Celti d'Italia II. Epigrafia e lingua*. Rome: Spazio Tre.
- Norris, Edwin
 1859 *Ancient Cornish drama*. Oxford: Oxford University Press.
- Ó Gealbháin, Séamus
 1991 The double article and related features of genitive syntax in Old Irish and Middle Welsh. *Celtica* 22: 119–144.

Padel, Oliver

- 2013 Aneirin and Taliesin. Sceptical speculations. In: Alex Woolf (ed.), *Beyond the Gododdin. Dark Age Scotland in medieval Wales*. St. Andrews: The Committee for Dark Age Studies, University of St. Andrews, 115–152.

Pedersen, Holger

- 1913 *Vergleichende Grammatik der keltischen Sprachen*. Vol. II. *Bedeutungslehre (Wortlehre)*. Göttingen: Vandenhoeck & Ruprecht.

Prósper, Blanca María

- 2011 The instrumental case in the thematic noun inflection of Continental Celtic. *Historische Sprachforschung* 124: 250–267.

Solinas, Patrizia

- 1995 Il celtico in Italia. *Studi etruschi* 60: 311–408.

Stokes, Whitley and John Strachan

- 1901 *Thesaurus Palaeohibernicus. A collection of Old-Irish glosses, scholia, prose, and verse*. Vol. 1. *Biblical glosses and scholia*. Cambridge: Cambridge University Press.

Stüber, Karin

- 2009 *Der altirische do-Infinitiv – eine verkannte Kategorie*. Bremen: Hempen.

Thomas, Graham and Nicholas Williams (eds.)

- 2007 *Bewnans Ke. The life of St. Kea. A critical edition with translation*. Exeter: University of Exeter Press.

Thomson, R. L. (ed.)

- 1968 *Owein or Chwedyl Iarles y ffynnawn* [Owein or the tale of the countess of the fountain]. Dublin: Dublin Institute for Advanced Studies.

Thurneysen, Rudolf

- 1946 *A grammar of Old Irish*. Translated by D. A. Binchy and Osborn Bergin. Dublin: Dublin Institute for Advanced Studies.

Untermann, Jürgen

- 1975 *Monumenta linguarum Hispanicarum I. Die Münzlegenden*. Wiesbaden: Reichert.

Untermann, Jürgen and Dagmar S. Wodtko

- 1997 *Monumenta linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert.

Villar, Francisco

- 1993–1995 El instrumental en celtibérico. *Kalathos* 13–14: 325–338.

Watkins, Calvert

- 1999 Two Celtic Notes. In: Peter Anreiter and Erzsébet Jerem (eds.), *Studia Celtica et Indogermanica. Festschrift für Wolfgang Meid zum 70. Geburtstag*. Budapest: Archaeolingua Alapítvány, 539–543.

Williams, Sir Ifor (ed.)

- 1951 *Pedeir Keinc y Mabinogi* [The four branches of the mabinogi]. 2nd edn. Cardiff: University of Wales Press.

Williams, Sir Ifor (ed.)

- 1961 *Canu Aneirin* [The poems of Aneirin]. 2nd edn. Cardiff: University of Wales Press.

Joseph F. Eska, Blacksburg, VA (USA)

71. The lexicon of Celtic

- | | |
|--------------------------|-------------------|
| 1. The inherited lexicon | 4. Word formation |
| 2. Loan words | 5. References |
| 3. Specific vocabulary | |

1. The inherited lexicon

1.1. The lexicon of the Celtic languages is well known from the medieval Insular Celtic traditions of Old and Middle Irish, Middle Welsh and Breton, and to a lesser extent Middle Cornish, and, moreover, from the living Goidelic and British Celtic (Brit.) languages of our own day. It is far less known for the Celtic languages of antiquity, including all the Continental Celtic languages. For these, indigenous inscriptions are the most important source (excepting Galatian, which has no indigenous texts), but unfortunately the meaning of words transmitted in such texts is often unknown and can only be established with a greater or lesser degree of certainty with the help of etymological arguments. Additional information comes from Celtic words transmitted in the context of non-Celtic languages, as substratum words or in glossaries, and particularly from proper names, which are an important source for Continental Celtic (CoC) as well as for the early stages of the Insular Celtic (IC) languages. The “meaning” of a proper name, or of the common noun that was behind it, can, again, usually only be approximately determined by means of comparative linguistics and will often remain ambiguous. Qualifications like “probably”, “perhaps”, “possibly”, drawing attention to such uncertainties, should always be on the reader’s mind, even if, for reasons of space, they will not always be used in the description of onomastic and Continental Celtic materials.

Due to the fragmentary attestation of languages like Galatian, Lepontic, and Celtiberian, not to mention Cumbrian, it is much easier to attribute a given word to the Proto-Celtic (PC) than to the Common Celtic (CC) lexicon. A word preserved from PIE in one Celtic language will necessarily have existed also in PC, but as far as CC is concerned, it is almost impossible to find word equations attested from all over the Celtic speaking world, notwithstanding the fact that such equations may safely be supposed to have been frequent in the ancient Celtic lexicon. One typical example is O(ld)I(rish) *már, mór*, M(iddle)W(elsh) *mawr*, M(iddle)B(reton) *neur*, M(iddle)C(ornish) *muer* < PC **māros* ‘great, big’, which is also found as an element of personal names (PN’s) in e.g. G(aulish) Σεγο-μαρος, Lep(ontic) DSg. **latu-marui** (or G ?), Gal(atian) Ζμερτο-μαρος*, and probably as first member of the compound **maro-midom** in the C(elt)ib(erian) lead inscription from Cuenca (cf. Lorrio and Velaza 2005: 1040). As such examples are rare, the focus in what follows will be on Insular Celtic and on Gaulish, the best attested Continental Celtic language, under which heading unspecified Continental Celtic material is also included.

Etymological dictionaries for Celtic languages are Matasović (2009) (Proto-Celtic), Vendryes, Bachellery, and Lambert (1959–) (OI and M[iddle] I[rish]), Delamarre (2003 [2001]) (G), Wodtko (2000) (Cib.); for a discussion of many Lepontic words cf. Lejeune (1971) and Morandi (2004), for Galatian cf. Freeman (2001); the British languages are covered in part by Campanile (1974) (O[ld] C[ornish]), Falileyev (2000) (O[ld] W[elsh]),

Fleuriot (1964) (O[Id] B[reton]) and Deshayes 2003 (B); for personal names cf. also Evans (1967), Meid (2005), for toponyms see Falileyev et al. (2010). An account of the Celtic lexicon according to the Swadesh-list has been provided by Elsie (1979) for British and by Lucht (2007) for Old Irish.

1.2. The lexicon of Celtic is organized according to the inherited word classes of nouns, adjectives, verbs, pronouns, numerals, particles and place words, the latter having evolved into prepositions, preverbs, and local adverbs at least in the Insular Celtic languages. The following are examples of Celtic lexical items according to word classes and semantic fields.

PIE kinship terms in words denoting the core family:

OI *athir* ‘father’, G DPl. *atrebo* < **ph₂tér-*, OI *máthir* ‘mother’, G *matir* < **máh₂ter-*. In British these words are replaced by WB *tad*, C *tas* ‘father’ and WBC *mam(m)* ‘mother’, but the older word for ‘mother’ underlies OB *motr-ep*, OW Pl. *modr-eped*, MC *moder-eb* ‘aunt, mother’s sister’. Inherited words for ‘brother’ and ‘sister’ are OI *bráthir*, MW *brawd*, MC *braud*, MB *breuzr*, cf. the G derivative *Bratronos*, and OI *siur*, MW *chwaer*, MC *huir*, MB *hoar*, and perhaps G IPl. *suiorebe*. The PIE word for ‘daughter’ is attested in G *duxtir*, Cib. NPl. **tuateres**, and in OI *Der* as an element of personal names. It has been replaced by OI *ingen*, MWB *merch*, MC *mergh* ‘girl, daughter’. ‘Son’ is OI *macc*, Ogam GSg. MAQQI < **mak^uk^uo-*, MWBC *mab* < **mak^uo-*, cf. G *maponos*; Cib. /*gentis*/ may be a word for ‘child, (immediate) descendant (of either sex)’, certain reflexes of **suH-nú-* are as yet lacking.

Other inherited Celtic terms for persons are: OI *ben*, GSg. *mná* ‘woman’ < **g^uénh₂-*, cf. G N or APl. *mnas* and name element *-bena*. OI *fer*, MW *gwr*, MB *gour*, MC *gur*, Cib. UIROS ‘man’, G *Viro-*, Lep. **uiro**[in personal names < **uiHró-*. The alternative word for ‘man’, **h₂nér-*, left traces only in MW *ner* ‘lord’, OI *ner* ‘boar’ and the Gaulish divine name *Nerios**. ‘Person, human’ is OI *duine*, MW *dyn*, MCB *den*, cf. the second element of G **teuo-γtoni[o]n** ‘gods and men’ < PC **gdonjo-*, a derivative of the PIE word for ‘earth’. To the legal sphere belong OI *rí*, GSg. *rig*, MW *rhi* ‘king’ < PIE **(-)h₃rég-*, also amply attested as a name element in G *-rix*, Gal. *-pυξ*. The fem. counterpart is OI *rígain* f. *ī* ‘queen’, MW *rhiein* ‘young woman, girl’, cf. G *rigani* (< **h₃rég-η-ih₂-*, cf. McCone 1998: 8). OI *túath* f. *ā* ‘people, tribe’, MW *tud* ‘people, country’, MB *tud*, MC *tus* ‘people’, G **toutas** and Cib. **toutam**[correspond to words for ‘people’ in e.g. Goth. *þiuda*, Lith. *tautà*; *tout-* also occurs as a name element, cf. in particular the personal names G *Toutona*, Cib. *Toutonus* and Goth. *þiudans* ‘king’ < **teu(H)tono-*.

Among designations for body parts are e.g. OI *lám* f. *ā*, MW *llaw*, OB *lom*, OC *lof* ‘hand’, cf. Gr. *παλάμη*, Lat. *palma*, OE *folm*; OI *dét* n. *nt*, MW, OB *dant*, OC *dans* ‘tooth’ cf. Ved. *dánt-*, Lat. *dens*, etc.

Names for animals are e.g. OI *ech* m. *o* ‘horse’ < **h₁ék^uo-*, cf. the Gaulish and Galatian name element *epo-* and an apparent derivative of it in the Ogam personal name EQOD[, MW *ebawl*, OC *ebol*, B *ebeul* ‘foal’; OI *art*, MW *arth* ‘bear’ < **h₂ít^{ko}-*, cf. a derivative in the Gaulish divine name *Artio*.

Well attested are also e.g. OI *mid* n., *u* OW *med*, MC *meth*, B *mez* ‘mead’, cf. the personal names Cib. **među-kenos**, G *μεδου-πειξ* < PIE **méd^hu-*; OI *treb* f. *ā* ‘house, farm, holding’, OB *treb*, MW *tref*, MC *tre* ‘dwelling place’ < **treb-* ‘homestead, dwelling’, cf. the Celtiberian toponym **konterbia**, *Contrebia*, OBrit. divine name *Contrebi* and OBrit. and G tribal name *Atrebates*.

The word for ‘god’, OI *día*, OW *duiu*, OC *duy*, B *doue* continues **deǵuó-*, cf. G *Devo-gnata*, Gal. $\Delta\eta\iota\sigma\text{-}\tau\alpha\rho\sigma$ (PNs), Cib. **teiuo-reikis**, and the Lep. PN **teu** < **deǵu-ōn-*.

1.3. Inherited adjectives can be illustrated by the following examples: OI *núae*, MW *newydd*, OB *nouuid*, MC *newyth* ‘new’, G toponym *Novio-dunum* < **néu-jo-*; OI *lethan*, MW *llydan* MBC *ledan* ‘broad, wide’, G *Litano-* in proper names, Cib. family name **litanokum** < **p_lth₂-no-*; OI *riiad*, OWBC *rud* ‘red’, G *roud-* in personal names < **(h₁)rouǵ^h-o-*; OI *sen*, OWBC *hen* ‘old’, cf. Gal. PN $\Sigma\epsilon\nu\tau\alpha\mu\omicron\varsigma^*$ (?), G *seno-* in personal names < **sén-o-*; OI *oac*, OC *iouenc*, MB *iouanc* ‘young’, MW *ieuanc*, and G PN *Iovinc-atus* < **h₂j-u-h₃ŋ-kó-*; OI *béo*, MWC *byw*, MB *beu* ‘alive, living’, cf. Lep. PN (DSg.) **piuonei** < PC **b₁yo-* < **g^uih₃-uó-*; OI *słán* ‘whole, sound’, cf. the abstract *sláine* ‘soundness’ (also personal name and river name), Cib. toponym Abl.Sg. **slaniađ**, Lep. PN DSg. **slaniai**, G *slano-* (?) < **s_lH-no-*.

Irish, to the present day, preserves the prefixes *so-* ‘good, well’ < **h₁su-* and *do-* ‘bad’ < **du-* ← **dus-* for the formation of compounds.

1.4. The inherited verbal lexicon of Celtic, consisting of primary verbs, has been described in Schumacher (2004). Continental Celtic has yielded relatively few verbs so far, none being attested for Galatian (nor for Ogam Irish). Due to the fragmentary attestation of ancient Celtic and the tendency to inflect verbs according to a unified stem-based pattern in the “middle” stages of Insular Celtic (MWBC, also MI), recognizably primary verbs that can be ascribed to Proto-Celtic number only about 200. Moreover, attested Continental Celtic verbs are not always etymologically transparent, cf. e.g. Schumacher (2004: 741 ff.) on G *auot* (and variants). A well attested verbal paradigm in Celtiberian consists of 3Sg. **dideti**, 3Pl. **didonti**, impv. **tatud**, infinitive **taunei** which may continue forms of the root **doh₃-* ‘give’ or **d^heh₁-* ‘put’, both of which have left rather few remnants in Insular Celtic (cf. Schumacher [2004: 265, 278 f.]). Celtiberian and Gaulish show a verbal stem *bi-(i)e/o-* in Cib. **bionti**, **atibion**, G *biuetutu*, *biontutu*, but the meaning and etymology, and consequently the identity of these forms remains ambiguous.

The majority of primary verbs in Insular Celtic have clear cognates in other IE languages (cf. Wodtko 2007: 107 ff. with references). Among them are verbs of broad attestation in many IE branches, such as OI WB *is* ‘is’ < **h₁és-ti*; OI *beirid* ‘carries’, MW *beryt*, OB *ber* ‘flows’ < **b^hér-e/o-*; *aigid* ‘drives’, OW *hegit*, OB *egit*, MC *a* ‘goes’ < **h₂ág-e/o-*; OI *airid*, MW *ardd* ‘ploughs’ < **h₂árh₃-je/o-*; OI *saidid* ‘sits’, MB *hez*, MC *heth* ‘leaves off’ < **séd-e/o-*; OI *-raig* ‘arises’ < **h₃rég-e/o-*, cf. G 1Sg. *regu*; OI *-sissedar* ‘stands’, cf. Cib. S1STAT < **si-stáh₂-*; OI *gainithir* ‘is born’, cf. MW *geni*, MB *guenell* ‘to be born’, MC *genys* ‘born’ < **ǵ_hh₁-jéló-*; OI *sechithir* ‘follows’ < **sék^u-(i)e/o-*.

Primary verbs frequently show stem formations paralleled in other IE languages, as e.g. OI *beirid*, *airid*, *-sissedar* above; in some cases a particular stem formation is confined to Celtic, as in the nasal presents OI *roindid* ‘dyes’ (cf. **reud^h-* ‘redde’, LIV 508). Sometimes a primary verb is found in Celtic only, the root being attested in nominal formations elsewhere, cf. e.g. OI *reithid*, MWB *ret*, MC *res* ‘runs’ (< **ret-e/o-*, LIV 507), with widespread nominal forms meaning ‘wheel’, like Lat. *rota*, elsewhere. Primary verbs lacking cognates in other IE languages completely are rare (cf. e.g. Schumacher [2004: 198 f.] on OI *aingid* ‘protects’, vn. *anacul*, cf. the G name element $\text{Ave}\chi\tau\lambda\omicron\text{-}$).

1.5. Among inherited pronominal stems are OI *aile*, MWB *eil*, ‘other’, MC *yl* ‘one of two’ < **al̥io-*, beside which OI *all-*, MW *all-* (as F[irst]M[ember] of compounds), G *allos*, *Allo-* (in proper names) point to earlier **al-no-*. OI *cách*, OW *paup*, OB MC *pop* ‘each, everyone’ < earlier **k^uāk^uo-* is built on **k^uo-*, cf. G *papon*. Celtiberian, Gaulish, and Insular Celtic show pronominal stems based on **so-*; Cib. **ios**, **ia**, **iomui**, etc. continue the inflected relative pronoun; uninflected **-jo* in Gaulish and Insular Celtic relative forms is based on the same stem. On personal pronouns see Stüber (this handbook); no certain reflex of a first person singular pronoun corresponding to Gr. ἐγώ, Lat. *egō* has yet been found in Celtic.

1.6. Celtic numerals continue the PIE forms (cf. Greene 1992). OI *mīle*, MWB *mil*, MC *myl* ‘1,000’ are loans from Lat. *mīlia*. ‘One’ is OI *oín*, MWBC *un* < **oīno-*. The ordinals ‘first’, OI *cétnae*, prefixed *cét-*, OWB *cisemic*, MW *cyntaf*, MB *quentaf*, MC *kynsa*, cf. G PNs *Cintu-gena*, *Cintusmus* ‘first’, are derived from **ken-* ‘spring from’; ‘second’ is OI *aile* (and *tánaise*), MWB *eil*, G *a(l)los* (lit. ‘other’, cf. 1.5).

1.7. Among inherited particles cf. Cib. **kue**, QUE, Lep. **-pe**, OI infix *-ch-* < PIE **k^ue* ‘and’, the negative Cib. **ne-**, G *ne* (?) < **ne* ‘not’, also the first element of OI *ní*, *ni*, OWB *ni* ‘not’. The combination of both particles appears in Cib. **nekue** ‘and not, nor’, cf. OI *na*, *nach-*, MWCB *na(c)*. Disjunctive Cib. **-ue** ‘or’ < **ue*, cf. perhaps G *-ue*, OI *fá* ‘or’; OI *nó*, MW *neu* ‘or’ may also contain **-ue* (cf. the discussion in Schrijver 1997: 160). G *eti* ‘likewise, in addition’ is probably from PIE **eti*, as is the first element of G *etic* ‘and’. The Old Irish preverbal particle *no-* and infix *-d-* (with pronouns) seem to be from PIE **nu*, and **de* respectively. At least in Insular Celtic, particles like **de*, and allegedly **eti*, play a major role in the verbal/pronominal system, but reflexes have become rather opaque and belong more properly to the grammar than to the lexicon (cf. Schrijver 1994; McCone 2006: 271–276, 225–245).

1.8. Celtic languages have adpositions and preverbs developed from PIE local adverbs. Prepositions are found in Gaulish and Insular Celtic; Celtiberian may have a postposition **eni** (if this is not adverbial). The languages just mentioned also have preverbs. In Galatian and Lepontic corresponding elements are attested only as first members in composition with nominal forms, all proper names. For an overview of Celtic preverbs cf. Schumacher (2004: 83 f.). The following are a selection:

OI pvb. *ad-*, e.g. *ad-gair* ‘accuses’, OW prp. *ad* ‘to’, G *ad-*, e.g. *ad-garion*, Gal. *Ad-iatorix* < **ad*.

OI pvb. *aith-*, pretonic *ad-* ‘again, re-’, e.g. *aithesc*, MW *atep* ‘answer’, OI *ad-gainethar* ‘is reborn’, G *ate-*, e.g. PN **ate-knatos***, Cib. verb **ati-bion**, < **ati-*.

OI pvb. *air-*, *ar-*, prp. *ar^L* ‘for, before’, e.g. *airchinn* ‘narrow side of a rectangle, head, end’, cf. MW *ar-bennig* ‘chief, excellent’, G *are-*, e.g. (in Lat.) *arepennem* ‘semiiumgerum’, Cib. **arei** (?) < **pr_h₂i*.

OI prp. *co^N* ‘with’, pvb. *con-*, cf. OW *com-*, *cim-*, OC *con-*, *chef-*, OB *cem-*, G *com-con-*, e.g. toponym *Con-date*, Cib. *kom-*, *kon-*, e.g. toponym **konterbia**, *Contrebia*, cf. OBrit. divine name *Con-trebi*, MI *con-treba* ‘inhabits’ < **kom-*.

OI prp. *i^N*, pvb. *in-*, OW prp. *in*, MWC *yn*, MB *en* ‘in’, cf. G prp. *in* (?) < **en*, and cf. Cib. **eni**, OI *in-*, e.g. *in-chinn* ‘brain’, *in-gen* ‘daughter’, Ogam INI-GENA < **eni*

OI *ess-amain*, MW *eh-ofn* ‘fearless’, cf. OI pvb. *as-*, prp. *a^G* ‘out of’, G *Ex-obnos*, Lep. **es-opnio** (PNs), Cib. **es-ankios** < *eǵ^{hs}.

OI prp. and pvb. *imb^L* ‘around’, OWB *am-*, G PN Αμβι-τῶν[τοῦ], Gal. tribal name *Ambi-touti*, Cib. verb **ambi-tiseti** < *h₂(a)mb^{hi}.

OI pvb. *ro-*, e.g. *ro-fítir* ‘knows’, MW *ry*, MBC *re*, G divine name *Ro-smerta*, Cib. verb **ro-biseti** < *pro.

OI prp. *fo^L* ‘under’, pvb. *fo-*, e.g. *fo-reith*, OW *guo-rit*, MC *gueres* ‘helps’, cf. OB *-uuoret*, G PN *Vo-reto-virius* < *upo.

OI prp. and pvb. *for* ‘on’, OWB *guor(-)*, OC *gur-*, G *Ver-*, e.g. PN *Ver-cingeto-rix*, cf. Cib. derivative VERAMOS, VORAMOS ~ ‘highest’ < *uper (influenced by *uo < *upo?).

Other derivatives from local adverbs are e.g. Celtiberian personal name and toponym **usama**, *Uxama* < *up-s-ṛmmo- ‘highest’, G and OBrit. toponym *Uxel(l)o-dunum*, cf. OI *úasal*, MW *uchel*, MB *uhel*, MC *huhel* ‘high, elevated’ < *eup-s-el-o- (?), G *uxedia* ~ ‘higher (?)’; *up- may also be the first element of Lep. **uvamokozis**.

2. Loan words

2.1. Celtic, like any IE language, has a few well attested words which lack a convincing IE etymology. For some of these borrowing from an unknown source into an early stage of Celtic has been suggested, e.g. for G *curmi*, OI *cuirm*, MW *cwrwf*, OC *coruf* ‘beer’, cf. the Cib. family name **kurmilokum**, but also for OI *brocc* ‘badger’, *mucc* ‘pig’ and their cognates (see 3.1 and cf. McCone 2005: 404, 409). Obviously, possible loans from an unknown language are very difficult to identify. This problem also stands in the way of tracing possible substratum words from Pre-Celtic languages in the British Isles (cf. e.g. Schrijver 2005).

2.2. In historical times Latin was the most important contact language for all Celtic languages with the exception of Galatian, where Greek influence dominated. Other contact languages were Etruscan and Ligurian in Northern Italy, Iberian and the forerunner of Basque in the Iberian Peninsula, Aquitanian in south-western Gaul, Germanic in the Rhineland, and languages like Pannonian in the east. Yet for all the Continental Celtic areas just listed, contact phenomena are not easy to describe in terms of loan words. As only a tiny part of the lexicon of all these languages has come down to us, it is difficult to specify the degree to which loan words had acquired a definable position in the lexicon. Language contact can rather be observed by the appearance of foreign names in indigenous Celtic inscriptions or in connection with Celtic patronymics or family names in Latin inscriptions, cf. e.g. Cib. **biurtilaur alaskum**, G *Martialis Dannotali* (cf. McCone 2005: 398 f.); the graffiti from La Graufesenque and some other Gaulish texts point to a bilingual environment, where both Latin and Gaulish were used (cf. Adams 2003: 184 ff., 687 ff.; Blom 2010–2012); the Lepontic use of a patronymic suffix *-al-* may have been encouraged by Etruscan models (McCone 2005: 396). Interference of a Latin inflectional ending is to be suspected, e.g. in Cib. CARORUM.

2.3. Latin loan words entered the Insular Celtic languages (at least) from the time of the Roman conquest of Britain and continued to do so through the Middle Ages and

beyond, because of the importance of Latin as the language of Christianity and as a prestigious lingua franca in Western Europe. The number of Latin loan words in British is estimated at about 800 (Jackson 1953: 76 with n. 3). The majority of Latin loans entered Irish with the introduction of Christianity, partly via British Latin intermediary forms. The Insular Celtic languages have also adopted Latin loan suffixes; for Latin borrowings into Welsh cf. Zimmer (1990, 2002), for those into Irish, McManus (1983, 1984).

Germanic loans came into the languages of the British Isles at a later stage. Cf. Marstrander (1915) for a collection of Scandinavian words in Irish, often related to seafaring; cf. Piette for French loanwords in Middle Breton.

3. Specific vocabulary

3.1. Here belong words which can be seen as characteristically Celtic by virtue of being well attested in a number of Celtic languages. Some of these have no established etymology (e.g. PC **dago-* ‘good’), but the majority are words that can be related to a PIE root yet may show a particular word formation, a specialized meaning, or simply great frequency as core members of the Celtic lexicon.

The following nouns illustrate “typically” Celtic words:

-brig- ‘height’, *-dūnom* ‘fortress’, and *magō-* ‘plain’ are frequent elements of toponyms, all preserved as common nouns in Insular Celtic languages. *-brig-*, e.g. in the Cib. toponym **ner tobis** /nerto-briχs/, is a root noun from **b^hreg^h-* ‘rise’, cf. OI *bri*, GSg. *breg* f. ‘hill’. It is latinized as *-briga*, but cf. WBC *bre* ‘hill’ from an indigenous Celtic form **brigā*. Toponyms in *-brig-* are most frequent in the Iberian Peninsula, *-dūnom*, which lives on in OI *dún*, W OB *din*, is more frequent elsewhere, as is *magō-*, cf. OI *mag* n. *s* ‘plain’, OB *ma*, WC *-ma* (cf. Rix 2001: 1 ff.; Sims-Williams 2006: 307 f., 328–330; Wodtko 2000: 278 f. with references). OI *mruig*, GSg. *mrogo*, and WBC *bro* ‘land’ reflect PC **mrogi*, cf. Gal. PNs like Βρογι-μααρος*, Βρογι-ταρος, toponyms *Eco-brogis*, Εριγο-βρογίς, G tribal names *Allo-brogēs*, *Nitio-brogēs*. Schol. Juven. 8,234 explains the name of the Allobroges: *brogae Galli agrum dicunt*. Words for ‘world’ are OI *domun* m. *o* < **d^hub-no-*, MW *elfyd* (< earlier **albiō-*), and OI *bith* m. *u*, MW *byt*, OCB *bit* < PC **g^uitu-* (< **g^uih₃-tu-*), cf. G PN *Dumno-rix*, divine name *Albio-rix* (epithet of Mars), tribal name *Bitu-rigēs*, Gal. PNs Βίτο-γνατος*, Αλβίο-ριξ, and perhaps Δομνε-κλειος.

OI *nemed* ‘privileged person, privilege, sanctuary’, OW *-nivet*, OB *-nimet* (in toponyms), G νεμητον, and *nemedo* in a Latin context in Celtiberia are from PC **nemeto-*, attested for Galatian by Strabo’s reference to a meeting place Δρυνεμετον (see Freeman 2001: 83 f.); cf. names like the toponym *Nemetobriga* in the north-west of the Iberian Peninsula. See Morandi (2004: 540) for a possible Lep. PN DSG. **nemetalui** (if this is the correct reading).

Typically Celtic designations for body parts are OI *cenn* n. *o*, MW OBC *pen(n)* ‘head’, cf. the G names Πεννο-ουινδος, *Penne-locos*, and *are-pennis** (1.8) < earlier **k^uenno-*; OI *dorn* m. *o*, MW *dwrn*, B *dorn*, *dourn* ‘fist, hand’, and G *durno-* in the toponym *Durno-magus* and PN *Dago-durnus* < PC **durno-* (< **d^hur-n-h₁-o-?*, cf. Wodtko 2007: 106 n. 54); OI *taul* n. *u* ‘protuberance, boss’, OWBC *tal* ‘front, forehead’ and the PNs G Σαμο-ταλος, **tano-taliknoi**, Cib. **talukokum** may be from PC **talū-*.

Among the specifically Celtic words denoting animals are OI *brocc* m. *o*, WBC *broch* ‘badger’, cf. the G toponym *Broco-magus*; OI *luch*, GSg. *lochad* f. ‘mouse’, MW *llygot*, B *logod* ‘mice’, OC singulative *logoden* ‘mouse’, cf. perhaps the G PN *Lucotios*; OI *molt* m. *o* ‘ram, wether’, MW *mollt* m., OC *mols* ‘wether’, B *maout* ‘sheep’ < PC **molto-*, cf. the G divine name *Moltinus*. OI *mucc* f. ‘pig’, MW OB *moch*, MC *mogh* ‘pigs’ is generally equated with the G epithet (DSg. in Lat.) *Mocco* of Mercury, pointing to a PC **mokku-*. The word for ‘bull’, OI *tarb* m. *o*, MW *tarw*, MC *tarow*, MB *taru*, and G *tarvos* (cf. perhaps the Gal. PN $\Delta\eta\iota\text{-}\tau\alpha\rho\varsigma$), goes back to **taruo-* as opposed to **tauro-* in Lat. *taurus*, Gr. $\tau\alpha\upsilon\rho\varsigma$, etc. *tauro-* appears in G PNs like *Donno-taurus*, and Celtic inscriptions from Spain contain the toponym *tarvo-duro-* and the PN (GSg.) **tauro**. Forms in *tauro-* have been claimed to be non-Celtic but could be archaisms according to Evans (1967: 261).

Celtic shares with Germanic (e.g. OE *widu*) a word for ‘tree, wood, forest’, OI *fid* m. *u*, OW *guit*, OB *-guid*, OC singulative *guiden* < **uid^hu-*, cf. perhaps the G PN *Viducus*, tribal name *Vidu-casses*, Cib. family name **uiduskikum**. A more characteristic word for ‘tree, wood’ is OI *crann* n. *o*, MWBC *pren(n)*, and G *prene* ‘arbores grandem’ (cf. McManus 1992: 205 f.).

Among Celtic names for specific trees are OI *ibar* m. *o* ‘yew’ < PC **eburo-*, well attested in early Celtic onomastics, cf. OBrit. toponym *Eburacum*, G PN $\text{E}\beta\upsilon\rho\upsilon\rho\varsigma$, toponym *Eburodunum*, tribal names *Eburones*, *Ebuovices*, Gal. PN $\text{E}\beta\upsilon\rho\upsilon\rho\eta\upsilon\varsigma$ *, Cib. PNs (GSg.) *Eburi*, *Eburianus*, and most likely **ebursunos**. The apparent cognates MW *efwr*, B *evor*, however, mean ‘cow-parasit, hogweed’ and ‘buckthorn’, respectively. Another word for the ‘yew’ is OI *éo* m., W *yw*, cf. the personal names Ogam GSg. IVA-GENI, G *Ivo-rix*, and OC *hiu-in* gl. *taxus*, B *iv-in* ‘yew’. OI *beithe*, MW *bedw*, MB singulative *bezuenn* mean ‘birch’, cf. OC *bedewen* gl. *populus*. These words are based on PC **betu-* < **g^uet-u-* ‘resin’ (cf. Pliny’s remark on the birch: *bitumen ex ea Galli excoquant*, n.h. 16,75). Corresponding forms may underlie the personal names *Betua*, *Betuca*, *Betunia* in Spain and perhaps G **petua** in Italy. OI *fern* f. *ā* ‘alder’, MW MoB *gwern* ‘alders’, OC singulative *guernen* < **u^hernah₂-* is probably also the first element of the G river name *Vernodubrum*. The second element corresponds to MW *dwfr*, OC MoB *dour* < **d^hub-ro-*, the usual word for ‘water’ (OI *dobur* ‘water’ is confined to compounds, toponyms, and glossaries; its normal meaning is ‘dark’). *Vernodubrum* is commonly understood as ‘alder-water’.

Celtic has some characteristic words denoting people. Here belong the names of the Celtic professions ‘druid’ and ‘bard’, cf. OI *druí* m., D MW *dryw* (rare, usually in the meaning ‘wren’) < **dru-uid-*; the Latin and Greek adaptations NPl. *druīdēs*, *druīdae*, $\delta\rho\upsilon\iota\delta\alpha\iota$ attest the word for Continental Celtic. OI *bard* m. *o*, MW *bard(d)*, OC *barth*, MB *barz*, and the Latin and Greek forms *bardus*, NPl. $\beta\acute{\alpha}\rho\delta\alpha\iota$ are from PC **bardos*, which has been explained from a compound **g^urH-d^hh₁-ó-* (with a root noun, like Ved. *gír-*, OAv. *gar-* ‘song, hymn’ as FM, cf. Schrijver 1995: 143 f.; NIL 100, 109 n. 26). OI *tigern* m. *o* ‘lord’, Ogam TIGIRN, MW *teyrn*, OW *-tigirn*, OB *-tiern* (in personal names), and the G toponym *Castrum T(h)igernum* < PC **tigerno-* ‘lord’; OI *gobae* m. *n*, MWB OC *gof* (f) ‘smith’ point to a base **gobann-* ‘smith’, cf. the names of mythological persons, OI *Goibniu*, MW *Gofannon*, G divine name DSg. *Gobanno*, personal names like *Gobannitio*; a variant stem **gobet-* ‘smith’ is commonly seen in G *gobedbi* (cf. Stüber 2005: 35 f.; Blažek 2008).

To the semantic field of war and peace belong OI *cath* m. *u*, cf. Ogam names like CATTU-VIR, OWB *cat*, MC *cas* ‘battle, fight’. *katu-* is frequent as an element of personal names, cf. G *Catu-marus*, Gal. Κατο-μαρος, and perhaps Hispano-Celtic names like *Catuenus*; cf. Morandi (2004: 653 f.) on possible **katua** in Northern Italy. OI *nāmae* m. *nt* ‘enemy’ has cognates in Gaulish personal names like *Namanto-bogius* and possibly in Cib. **namato** [Díaz/Jordán 2006: 261 f.]. OI *búaid* n. *i* ‘victory’, MW *budd* ‘profit’, OB *bud* gl. bradium < PC **bou̯di-* appears as a name element in G PNs *Boudi-latis*, *Boudillus*, Gal. PN Βουδο-ρις, tribal name *Touto-bodiaci*. It may be attested in G *boudi* on the Lezoux plate. MW *tanc*, OW *tagc* ‘peace’ seems to be related to the name element *tanc(o)-* in G PNs Ταυκο-λατις, *Tanconus*, and very frequent *Tancinus* in Lusitania.

3.2. Among characteristically Celtic adjectives G *-māros* and its cognates have already been mentioned (1.1). Next in terms of broad attestation among various Celtic languages comes OI *find olā*, OWB *guin(n)*, OC *guyñ* ‘white, bright, fair’ < PC **u̯indo-*, cf. G PNs *Vindus*, fem. *Vinda*, Lep. **alko-uinos**, with derivatives in the Galatian toponym *Vindia*, divine name (DSg.) Ουνιδεινω, and possibly the name of the mountain *Vindius* in Northern Spain.

Restricted to Insular Celtic and Geltic, but well attested there, is the adjective for ‘good’, OI *dag-* (as first member of a compound), MWBC *da*, cf. fem. *daga* on a Gaulish spindle whorl, and personal names like Δαγο-λιτους, *Bitu-daga*.

The verb OI *caraid* ‘loves’, MW *caru*, MC *care*, MB *caret* ‘to love’ and all related words point to a Celtic root **kār-* as opposed to apparently related **kah₂-ro-* in Lat. *cārus* ‘dear’, Goth. *hors* ‘adulterer’ etc. **kar-* seems to underlie abbreviated **kar** on Celtiberian *tesserae hospitales*; it appears in many proper names in Spain, Gaul and beyond, where, however, it can be difficult to distinguish Celtic from corresponding Latin names. A characteristic formation is the *nt*-stem in G PNs *Carantus*, *Carantodia*, OI *carae* m. *nt* ‘friend’, MW *car*, Pl. *carant*, *kereint* ‘relative, friend’, MB *car*, Pl. *querent*; cf. also the PN G *Veni-carus*, OI *Fin-char*, OW *Gun-car*, MB *Guen-gar*.

3.3. Celtic shares some exclusive lexical isoglosses with other IE languages. Thus OI *marc* n. *o*, MW OC *march*, OB *marh* ‘horse’, attested for Galatian as ASg. μάρκων in the explanation of τριμαρκισία, a three-horse battle group, by Pausanias, finds cognates only in Germanic, cf. OE *mearh*, OHG *marah*, ON *marr*. OI *gíall* m. *o*, OC *guistel*, OB *guuistl* ‘hostage’, MW *gwystl* ‘hostage, surety’, and perhaps the G PN (GSg. in Lat.) *Con-geistli*, may be from **g^heid^h-tlo-*, sharing the meaning and word formation of OHG *gīsal*, OE *gīsl*, ON *gísl* (cf. Schrijver 1995: 405 f. with n. 1). Common semantic developments of Celtic and Germanic are seen in the words for ‘oath’, OI *oíth* m. *o* ‘oath’, cf. OW *an-utonou* gl. *periuria*, Goth. *aiþs*, OE *āð*, OHG *eid* ‘oath’ < **h₂ói-to-* ‘a going’ and in MW *rhydd*, OBC *rid* ‘free’, cf. Goth. *freis*, OE *frēo*, OHG *frī* ‘free’ < **priH-ó-* vs. Ved. *priyá-*, OAv. *frīia-* ‘dear’ (cf. Schumacher 2007: 176 ff.).

Isoglosses of Celtic and Italic are e.g. the verbs OI *gaibid* ‘takes’, Lat. *habēre* ‘to have’ < root(-shape) **g^hHb-* (LIV 195); MW *hoedl*, OB *hoetl* (in PNs) ‘lifetime, age’, Lat. *saeculum* ‘age’ < earlier **saj^h-tlo-*. OI *nathir*, GSg. *nathrach* f. ‘snake’ and MW *traw* ‘beyond’ are generally taken as cognates of Lat. *natrix* f. ‘snake’ and *trāns* ‘past, over’ rather than loan words (Stüber 2012: 410 ff.; Schumacher 2012).

4. Word formation

4.1. Celtic word formation patterns frequently continue inherited PIE elements and rules. Nouns and verbs can be formed by composition and suffixation, in some cases composition can be described as prefixing. Studies of nominal word formation are e.g. de Bernardo Stempel (1999), Irlinger (2002), Uhlich (1993), Wodtke (1995) for Irish; Schumacher (2000), Zimmer (2000) for Welsh; Russell (1990), Stüber (1998) for Celtic.

Nominal composition is well attested all over the Celtic speaking world. Continental Celtic evidence comes mostly from proper names. Compound personal names are amply attested in Gaulish, Galatian, and the Insular Celtic languages, but are not so frequent in Lepontic and are really rare in Celtiberian. Celtiberian, however, provides sufficient examples of nominal compounds in toponyms and in the common lexicon, so that a restriction on composition seems to affect personal names only and was not a feature of the language as such. In Insular Celtic some compositional types are replaced by syntactic collocations, cf. e.g. MW *Din Orben* vs. an old type of compound like OBrit. *Branodunum*. In general, however, nominal composition is well preserved into the medieval Insular Celtic period.

For a detailed discussion of nominal composition types cf. Uhlich (1993: 78–120) and Zimmer (2000: 1–270); only a broad classification can be given here.

For endocentric compounds with noun or adj. as a first member cf. the toponyms in *-brig(a)*, *-dūnom* and *-magos* (3.1), e.g. Hispano-Celtic *Nemeto-briga*, G *Litano-briga*, *Lug(u)-dunum*, *Novio-dunum*, *Blato-magus* (cf. OI *bláth* ‘flower’), *Novio-magus*; cf. also e.g. OI *ár-mag* ‘battle field’, MW *aer-fá* ‘battle’, OB Pl. *air-maou*.

Dvandva compounds are seen in G **teuoxtoni[o]n** ‘gods and men’, and OI *gaisced* n. o ‘weapons’ (*gai* ‘spear’ and *sciath* ‘shield’). Dvandvas are more frequent in adjectival compounds in Old Irish and Middle Welsh, cf. e.g. OI *find-chass* ‘fair and curly’, MW *hir-lwys* ‘long and comely’ (*glwys*); the G PN *Dago-marus* may belong here, if it is to be interpreted as ‘good and great’.

Exocentric compounds with a prefix like *so-* ‘good’ (< **h₁su-*) or *dí-* ‘without’ (< **dē* ‘away from’) as first member can inflect as *i*-stems in Old Irish, when the S(econd) M(ember) was originally an *o-* or *ā-*stem, cf. *so-chenéol i* ‘of good family’ (*cenél* n. o), *dí-lmain i* ‘free’ (*loman* f. *ā* ‘leash’); however, more often a bahuvrīhi will either preserve the stem class of the second member or reinforce the adjectival meaning of the compound by means of the possessive suffix *-a/ech*, cf. e.g. *coin-chenn* ‘dog-headed’ (also PN, cf. *cenn* n. o ‘head’) and *coin-chennach*.

“Armstrong compounds” or “reversed bahuvrīhis” are frequent in Celtic (see Uhlich 1997), cf. G, Lep., Gal. names in *-maros*, like Βρογι-μαραρος*, also e.g. OI *cenann*, MW *pen-wyn* ‘white-headed’, G PN Πεννο-ουινδος, Lep. PN **alko-uinos**. Evidence of this type from Celtiberian is lacking.

Verbal governing compounds (Uhlich 2002) could be formed with root nouns as second member, cf. the tribal names OBrit. *Ordo-vices* ‘hammer fighters’ (cf. OI W *ord* ‘hammer’ + **meik-* ‘overcome’), Gal. Τεκτο-σαγες ‘striving for property’ (root **seh₂g-* ‘track, trace’); compounds in *-rix*, like G PN *Dumno-rix*, may originally belong here, as ‘world ruling’, being only later reanalysed as *tatpuruṣas* ‘world’s ruler’ (McCone 1998: 7). The type Gr. *δρυ-τόμος*, with *o*-grade *o*-stem as a second member, does not seem to be frequent in Celtic.

Prepositional governing compounds can be exemplified by OI *essamain i* ‘fearless’ (*ess-* ‘out of’ + *omun m. o* ‘fear’) and its cognates (see 1.8); this type is not productive in Insular Celtic. However, in Insular Celtic some prepositions have developed into prefixes with a privative or gradient function, cf. OI *ro-*, MW *ry-* ‘(too) much, exceedingly’ with adjectives like OI *romár* ‘very big’, MW *ry uawr* ‘too big’. This function is assumed also for Continental Celtic, where the semantics of prefixes are difficult to determine. On preverbs see 1.8, 4.5.

4.2. Celtic nominal stem-formation shows a number of inherited, primary suffixes, but primary formations are no longer productive in the Insular Celtic languages, and productivity for Continental Celtic cannot be proven. Thus OI *recht m. u*, ‘law’ is from **h₃reg₂-tu-*, a deradical abstract in **-tu-* from **h₃reg₂-* ‘stretch, direct’, cf. MW *reith*, MB *reiz* ‘law’, personal names G *Rextu-genos*, Cib. **retukenos**, *Rectugenus*; while **-tu-* remains the productive suffix for verbal nouns of denominative verbs in Old Irish, primary formations are no longer possible. This is also the case with primary **-ti-* (cf. Irslinger 2002: 233), **-tiHon-*, **-men-*, and **-es-* (cf. Stüber 1998: 141 f., 80 f., 2002: 62), all found in older formations, and with all the Insular Celtic agent noun suffixes. An innovated suffix **-īmā* has become productive in Welsh verbal nouns (cf. Schumacher 2000: 129 ff.).

Adjectival abstracts in OI *-e*, MW *-edd*, MB *-ez* < **(i)jah₂-* are frequent, cf. OI *sláine* f. ‘soundness’ and its cognates (1.3). Celtic shares with Latin and Gothic a suffix continuing earlier **-tūt-*, which forms masculine deadjectival abstracts, cf. OI *oíntu*, GSg. *oíntad m.* ‘unity’, and OW *duiu-tit* ‘divinity’; on abstracts in OI *-as*, *-us* see Hill (2006). Agent nouns in **-te/or-* have not been identified. The productive Insular Celtic formation deriving agent nouns from verbal nouns is in OI *-(a)id*, MW *-iad* < **-jati-*, probably a variant of earlier *-ati-* as in G *ναμιασταις* ‘from Nîmes, belonging to Nîmes’ and the name of the Galates (McCone 1995: 7). Agent nouns in *-mon*, *-amon*, *-jamon* (m. *n*-stems) exist in Old Irish, cf. e.g. *brithem* ‘judge’, GSg. *brithemon* (*breth* ‘judgment’), but the formation is obsolete in historical times (cf. Remmer 2002/2003, 2004). The suffix OI *-aige* < **-sag-ijo-s*, as in *scélaige* ‘story-teller’ (*scél* ‘story’), is related to denominative verbs in OI *-(a)igidir* (Joseph 1987: 140 ff.); a more original meaning ‘to strive for’ is preserved in some British formations, e.g. MW *cynutai* ‘gatherer of firewood’ (*cynnud*) and possibly in the G PN *Curmi-sagios* ‘striving for beer’. MW *-ydd*, OC B *-it* < **-ijo-* appears in desubstantival agent nouns like MW *prydydd* ‘poet’ (cf. Russell 1989: 37); corresponding Old Irish stems in *-e* are mostly confined to second members of compounds, e.g. *sed-guine* ‘deer-slayer’ (Breatnach 1983).

Individualizing *n*-stems are well attested in Continental Celtic personal names, cf. Cib. **aiu** (fem. **aia**), **amu** (fem. **ama**), and **stenu** (fem. **stena**), which may be a short form of a longer name like **steniontes**, as may be **tirtu** to **tirtanos** or **tirtuiois**, and **statu** to **statulu**. For Lep. cf. **teu** (1.2), DSg. **piuonei** (1.3).

A suffix *-ono/ā* appears in some Continental Celtic derivatives denoting persons, cf. an equivalent in W *-on* for G *Matrona* : W *Modron*, *maponos* : W *Mabon*, Cib. *Vironus*

:

W *gwron* ‘hero’ (cf. Stüber 2004).

British Celtic has a suffix MW *-yn*, *-en*, B *-enn*, C *-en* to form singulatives (cf. Irslinger 2010), cf. MW *ser-en* ‘star’ (*ser* ‘stars’), OC *ster-en*, B *gwez-enn* ‘tree’ (*gwez* ‘trees’).

4.3. Adjectives are productively formed with various *-Vko-* suffixes in OI *-a/ech*, MW *-og*, *-ig* (and others), also frequent in Continental Celtic (cf. Russell 1990). In Celtiberian *ko-* suffixes form family names, among other denominative derivatives. Gaulish has a patronymic suffix *-ikno-*, e.g. in **tanotaliknoi**, in addition to *-io-*. Superlatives from adjectives are formed by Celtic **-isamo-*, cf. e.g. the Cib. toponym *Segisama*, (Stüber, this handbook). *-amo-* appears with local adverbs, cf. MW *uchaf*, Cib. **usama** (1.8). **-tero-* is most clearly preserved in British deadjectival abstracts in *-der* (Pedersen 1976 [1913]: 43 f.).

4.4. For verbal stem formations see Stüber (this handbook); among primary present stems thematic and nasal presents are well represented in Insular Celtic, information on Continental Celtic is insufficient. Denominative verbs could be formed with the inherited suffixes **-ah₂-* and **-eǵéló-* (> **-ī-*) in Insular Celtic, cf. OI *nert* ‘strength’ → *nertaid*, *-nerta* ‘stengthens’ and *rím* ‘number’ → *rímid*, *-rími* ‘numbers’, respectively. G **karnitus**, Lep. **karite** may belong here. Denominative *ī-* verbs have coalesced with former causatives and iteratives in **-éǵe/o-*, which do not seem to have been very productive to judge from the Insular Celtic remains; for causatives, iteratives, and nasal presents in British see Schulze-Thulin (2001). A Celtic innovation are denominatives in OI *-(a)igid-ir*, MWBC *-(h)a-* < **-sag-ǵe/o-*, reanalysed from **-ǵe/o-* derivatives of nouns in **-sag-* (Joseph 1987). The formation has not yet been identified in Continental Celtic, but its starting point can be seen in compounds like G, Gal. tribal name *Tectosages*, G PN *Curmisagios*.

4.5. A most important device of forming deverbative verbs at least in Insular Celtic languages was by adjoining preverbs to the verbal stems. Preverbatation is also attested for Gaulish and Celtiberian, cf. e.g. G *ni-tixsintor*, Cib. **ambi-tiseti** (cf. 1.8); however, Continental Celtic evidence comes mostly from deverbal nouns used as personal names (cf. 1.7, see Wodtko 2013). In Old Irish the verb can take up to five preverbs, and two or three are not rare (McCone 2006: 177 ff.).

5. References

- Adams, James N.
 2003 *Bilingualism and the Latin Language*. Cambridge: Cambridge University Press.
- de Bernardo Stempel, Patrizia
 1999 *Nominale Wortbildung des älteren Irischen*. Tübingen: Niemeyer.
- Blažek, Václav
 2008 Celtic ‘smith’ and his colleagues. In: Alexander Lubotsky, Jos Schaecken, and Jeroen Wiedenhof (eds.), *Evidence and counter-evidence. Essays in honour of Frederik Kortlandt*. Vol. 1. Amsterdam: Rodopi. 67–86.
- Blom, Alderik H.
 2010–12 Language at La Graufesenque: The evidence of the potters’ graffiti. *Keltische Forschungen* 5: 13–53.
- Breatnach, Liam
 1983 On the agent suffix *-e* in Irish. *Ériu* 34: 194.
- Campanile, Enrico
 1974 *Profilo etimologico del Cornico Antico*. Pisa: Pacini.

- Delamarre, Xavier
2003 [2001] *Dictionnaire de la langue gauloise. Une approche linguistique du vieux-celtique continental*. 2nd edn. revised and expanded. Paris: Errance.
- Deshayes, Albert
2003 *Dictionnaire étymologique du breton*. Douarnenez: Le Chasse-Marée.
- Díaz, María Antonia and Carlos Jordán
2006 Dos téseras de hospitalidad procedentes de Fitero (Navarra). *Palaeohispanica* 6: 257–266.
- Elsie, Robert
1979 *The position of Brittonic*. Ph.D. thesis, Bonn.
- Evans, D. Ellis
1967 *Gaulish Personal Names*. Oxford: The Clarendon Press.
- Falileyev, Alexander
2000 *Etymological Glossary of Old Welsh*. Tübingen: Niemeyer.
- Falileyev, Alexander with Ashwin E. Gohil and Naomi Ward
2010 *Dictionary of Continental Celtic Place-Names*. Aberystwyth: CMCS.
- Fleuriot, Léon
1964 *Dictionnaire des gloses en vieux breton*. Paris: Klincksieck.
- Freeman, Philip
2001 *The Galatian Language*. Lewiston, NY: Edwin Mellen.
- Greene, David
1992 Celtic. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. Berlin: De Gruyter, 497–554.
- Hill, Eugen
2006 Ein germanisch-keltisches Suffix für Nominalabstrakta. *Münchener Studien zur Sprachwissenschaft* 62(2002): 39–70.
- Irslinger, Britta S.
2002 *Abstrakta mit Dentialsuffixen im Altirischen*. Heidelberg: Winter.
- Irslinger, Britta S.
2010 Les dérivés gallois, cornique en *-yn /-en*, breton en *-enn* et irlandais en *-ne*: fonction et sémantique. *La Bretagne Linguistique* 15: 43–81.
- Jackson, Kenneth H.
1953 *Language and History in Early Britain: A Chronological Survey of the Brittonic Languages First to Twelfth Century A.D.* Edinburgh: University Press.
- Joseph, Lionel
1987 The origin of the Celtic denominative formation in **-sag-*. In: Calvert Watkins (ed.), *Studies in Memory of Warren Cowgill (1929–1985). Papers from the Fourth East Coast Indo-European Conference, Cornell University, June 6–9, 1985*. Berlin: De Gruyter, 113–159.
- Lejeune, Michel
1971 *Lepontica*. Paris: « Les Belles Lettres ».
- LIV = Helmut Rix, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer
2001 *Lexikon der indogermanischen Verben*. 2nd, corrected edn. Wiesbaden: Reichert.
- Lorrio, Alberto José and Javier Velaza
2005 La primera inscripción celtibérica sobre plomo. *Palaeohispanica* 5: 1031–1048.
- Lucht, Martina
2007 *Der Grundwortschatz des Altirischen*. Ph.D. thesis, Bonn. http://hss.ulb.uni-bonn.de/diss_online/phil_fak/2007/lucht_martina/teil1.pdf [Last accessed 8 December 2007].
- Marstrand, Carl
1915 *Bidrag til det norske Sprogs Historie i Irland* [Contributions to the history of the Norse language in Ireland]. Kristiana: Dybwad.

- Matasović, Ranko
2009 *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
- McCone, Kim
1995 OIr. *senchae*, *senchaid* and preliminaries on agent noun formation in Celtic. *Ériu* 46: 1–10.
- McCone, Kim
1998 ‘King’ and ‘Queen’ in Celtic and Indo-European. *Ériu* 49: 1–12.
- McCone, Kim
2005 Mögliche nicht-indogermanische Elemente in den keltischen Sprachen und einige frühe Entlehnungen aus indogermanischen Nachbarsprachen. In: Gerhard Meiser and Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, Halle a. S.* Wiesbaden: Reichert, 395–435.
- McCone, Kim
2006 *The origins and development of the Insular Celtic verbal complex*. (Maynooth Studies in Celtic Linguistics 6). Maynooth: Department of Old Irish, National University of Ireland.
- McManus, Damian
1983 A chronology of the Latin loan-words in Early Irish. *Ériu* 34: 21–71.
- McManus, Damian
1984 On final syllables in the Latin loan-words in Early Irish. *Ériu* 35: 137–162.
- McManus, Damian
1992 Varia III: On the vocalism of Old Irish *crann*. *Ériu* 43: 205 f.
- Meid, Wolfgang
2005 *Keltische Personennamen in Pannonien*. Budapest: Archaeolingua.
- Morandi, Alessandro
2004 *Celti d’Italia II. Epigrafia e lingua*. Rome: Spazio Tre.
- NIL: Dagmar S. Wodtke, Britta Irslinger, and Carolin Schneider
2008 *Nomina im Indogermanischen Lexikon*. Heidelberg: Winter.
- Pedersen, Holger
1976 [1913] *Vergleichende Grammatik der keltischen Sprachen*. Vol. II. Göttingen: Vandenhoeck & Ruprecht. [Strassburg: Trübner.]
- Piette, J. R. F.
1973 *French loanwords in Middle Breton*. Cardiff: University of Wales Press.
- Remmer, Ulla
2002/2003, 2004 Das indogermanische Suffix *-mon-* im Altirischen. (1. Teil) *Die Sprache* 43: 171–211. (2. Teil) *Die Sprache* 44(2004): 26–69.
- Rix, Helmut
2001 Zur Verbreitung und Chronologie einiger keltischer Ortsnamentypen. In: Gerhard Meiser (ed.), *Helmut Rix. Kleine Schriften*. Bremen: Hempen (Originally published in 1954), 1–9.
- Russell, Paul
1989 Agent suffixes in Welsh: native and non-native. *Bulletin of the Board of Celtic Studies* 36: 30–42.
- Russell, Paul
1990 *Celtic Word Formation. The Velar Suffixes*. Dublin: The School of Celtic Studies, Dublin Institute for Advanced Studies.
- Sadovski, Velizar and David Stifter (eds.)
2012 *Iranistische und indogermanistische Beiträge in Memoriam Jochem Schindler (1944–1994)*. Vienna: Österreichische Akademie der Wissenschaften.
- Schrijver, Peter
1994 The Celtic adverbs for ‘against’ and ‘with’ and the early apocope of **-i*. *Ériu* 45: 151–189.

- Schrijver, Peter
1995 *Studies in British Celtic Historical Phonology*. (Leiden Studies in Indo-European 5). Amsterdam: Rodopi.
- Schrijver, Peter
1997 *Studies in the history of Celtic pronouns and particles*. (Maynooth Studies in Celtic Linguistics 2). Maynooth: Department of Old Irish, National University of Ireland.
- Schrijver, Peter
2005 Varia I. More on non-Indo-European surviving in Ireland in the first millennium AD. *Ériu* 55: 137–144.
- Schulze-Thulin, Britta
2001 *Studien zu den urindogermanischen o-stufigen Kausativa/Iterativa und Nasalpräsentien im Kymrischen*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Schumacher, Stefan
2000 *The Historical Morphology of the Welsh Verbal Noun*. (Maynooth Studies in Celtic Linguistics 4). Maynooth: Department of Old Irish, National University of Ireland.
- Schumacher, Stefan
2004 *Die keltischen Primärverben. Ein vergleichendes, etymologisches und morphologisches Lexikon*. Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Schumacher, Stefan
2007 Die Deutschen und die Nachbarstämme: Lexikalische und strukturelle Sprachkontaktphänomene entlang der keltisch-germanischen Übergangszone. In: Hans Hablitzel and David Stifter (eds.), *Johann Kaspar Zeuß im kultur- und sprachwissenschaftlichen Kontext (19. bis 21. Jahrhundert)*. Vienna: Praesens, 176–207.
- Schumacher, Stefan
2012 Mittelkymrisch *tra, tros, traws, traw*, altirisch *trá* und Verwandtes. In: Sadovski and Stifter (eds.), 361–375.
- Sims-Williams, Patrick
2006 *Ancient Celtic Place-Names in Europe and Asia Minor*. Oxford: Wiley-Blackwell.
- Stüber, Karin
1998 *The Historical Morphology of n-stems in Celtic*. (Maynooth Studies in Celtic Linguistics 3). Maynooth: Department of Old Irish, National University of Ireland.
- Stüber, Karin
2002 *Die primären s-Stämme des Indogermanischen*. Wiesbaden: Reichert.
- Stüber, Karin
2004 Individualisierendes *-on- in Namen von Göttern und Göttinnen. *International Journal of Diachronic Linguistics* 1: 1–17.
- Stüber, Karin
2005 *Schmied und Frau. Studien zur gallischen Epigraphik und Onomastik*. Budapest: Archaeolingua.
- Stüber, Karin
2012 Zur Frage des Italo-Keltischen: Erkenntnisse aus der Erforschung der Partikeln. In: Sadovski and Stifter (eds.), 403–417.
- Uhlich, Jürgen
1993 *Die Morphologie der komponierten Personennamen des Altirischen*. Witterschlick: Wehle.
- Uhlich, Jürgen
1997 Der Kompositionstyp ‘Armstrong’ in den indogermanischen Sprachen. *Historische Sprachforschung* 110: 21–46.
- Uhlich, Jürgen
2002 Verbal governing compounds (synthetics) in Early Irish and other Celtic languages. *Transactions of the Philological Society* 100: 403–433.

- Vendryes, Joseph, Edouard Bachellery, and Pierre-Yves Lambert
1959 ff. *Lexique étymologique de l'irlandais ancien*. Dublin: The School of Celtic Studies.
- Wodtko, Dagmar S.
1995 *Sekundäradjektive in den altirischen Glossen. Untersuchungen zur präfixalen und suffixalen Wortbildung*. (Innsbrucker Beiträge zur Sprachwissenschaft 81). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Wodtko, Dagmar S.
2000 *Monumenta Linguarum Hispanicarum V.1: Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Reichert.
- Wodtko, Dagmar S.
2007 Das Verb im Lexikon: Indogermanisch und Irisch. *International Journal of Diachronic Linguistics* 4: 91–133.
- Wodtko, Dagmar S.
2013 Preverbs and Personal Names. In: Juan Luis García Alonso (ed.), *Continental Celtic Word Formation: The Onomastic Data*. Salamanca: Ediciones Universidad de Salamanca.
- Zimmer, Stefan
1990 Dating the loanwords: Latin suffixes in Welsh (and their Celtic congeners). In: Alfred Bammesberger and Alfred Wollmann (eds.), *Britain 400–600: language and history*. Heidelberg: Winter, 263–281.
- Zimmer, Stefan
2000 *Studies in Welsh Word-formation*. Dublin: The School of Celtic Studies.
- Zimmer, Stefan
2002 Latin and Welsh. In: Lea Sawicki and Donna Shalev (eds.), *Donum grammaticum. Studies in Latin and Celtic Linguistics in Honour of Hannah Rosén*. Leuven: Peeters, 395–406.

Dagmar S. Wodtko, Freiburg (Germany)

72. The dialectology of Celtic

- | | |
|-----------------------|---|
| 1. Preface | 8. Insular Celtic |
| 2. Proto-Celtic | 9. Goidelic vs. Brittonic |
| 3. Hispano-Celtic | 10. Other varieties of Continental Celtic |
| 4. Nuclear Celtic | 11. Conclusions |
| 5. Cisalpine Celtic | 12. References |
| 7. Transalpine Celtic | |

1. Preface

1.1. The interrelationship of the Celtic languages has become a controversial topic in recent years. Previously, the phylogenetic structure of the Celtic language family was based upon but two criteria, viz., the treatment of proto-IE */k^w/, either continued unchanged or labialised to /p/, and the treatment of the proto-IE syllabic nasals, either as

- Vendryes, Joseph, Edouard Bachellery, and Pierre-Yves Lambert
 1959 ff. *Lexique étymologique de l'irlandais ancien*. Dublin: The School of Celtic Studies.
- Wodtko, Dagmar S.
 1995 *Sekundäradjektive in den altirischen Glossen. Untersuchungen zur präfixalen und suffixalen Wortbildung*. (Innsbrucker Beiträge zur Sprachwissenschaft 81). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Wodtko, Dagmar S.
 2000 *Monumenta Linguarum Hispanicarum V.1: Wörterbuch der keltiberischen Inschriften*. Wiesbaden: Reichert.
- Wodtko, Dagmar S.
 2007 Das Verb im Lexikon: Indogermanisch und Irisch. *International Journal of Diachronic Linguistics* 4: 91–133.
- Wodtko, Dagmar S.
 2013 Preverbs and Personal Names. In: Juan Luis García Alonso (ed.), *Continental Celtic Word Formation: The Onomastic Data*. Salamanca: Ediciones Universidad de Salamanca.
- Zimmer, Stefan
 1990 Dating the loanwords: Latin suffixes in Welsh (and their Celtic congeners). In: Alfred Bammesberger and Alfred Wollmann (eds.), *Britain 400–600: language and history*. Heidelberg: Winter, 263–281.
- Zimmer, Stefan
 2000 *Studies in Welsh Word-formation*. Dublin: The School of Celtic Studies.
- Zimmer, Stefan
 2002 Latin and Welsh. In: Lea Sawicki and Donna Shalev (eds.), *Donum grammaticum. Studies in Latin and Celtic Linguistics in Honour of Hannah Rosén*. Leuven: Peeters, 395–406.

Dagmar S. Wodtko, Freiburg (Germany)

72. The dialectology of Celtic

- | | |
|-----------------------|---|
| 1. Preface | 8. Insular Celtic |
| 2. Proto-Celtic | 9. Goidelic vs. Brittonic |
| 3. Hispano-Celtic | 10. Other varieties of Continental Celtic |
| 4. Nuclear Celtic | 11. Conclusions |
| 5. Cisalpine Celtic | 12. References |
| 7. Transalpine Celtic | |

1. Preface

1.1. The interrelationship of the Celtic languages has become a controversial topic in recent years. Previously, the phylogenetic structure of the Celtic language family was based upon but two criteria, viz., the treatment of proto-IE */k^w/, either continued unchanged or labialised to /p/, and the treatment of the proto-IE syllabic nasals, either as

/aN/ or /eN/. It is now known that the former is a phonologically trivial and easily repeatable sound change and, hence, unimportant for purposes of phylogenetic classification, and that the proto-Celtic realization of the latter was uniformly */aN/, instances of /eN/ in not only Old Irish, but sporadically also in Continental Celtic, being secondary developments. What were the two traditional linchpins upon which the phylogeny of the Celtic language family was based, in fact, are not diagnostic of anything at all.

1.2. In the wake of this determination, no consensus about the diachronic evolution of the Celtic language family has emerged, especially at its earliest stages, owing to the fragmentary state of the documentation. In fact, given the variation that exists in the Continental Celtic linguistic record, numerous scholars, including Koch (1992), Garrett (1999), Isaac (2005), and Sims-Williams (2007), have come to abandon the idea that proto-Celtic descended as a unitary language from proto-Indo-European, but, instead, arose by the convergence of a number of closely related languages. Others, such as De Bernardo Stempel (2013) and McCone (2001), maintain a phylogenetic approach, as do I in this chapter, which adopts a computational methodology in the spirit of Ringe, Warnow, and Taylor (2002), Nakhleh, Ringe, and Warnow (2005), and Warnow et al. (2006). As such, individual innovatory linguistic features are here designated as “characters”.

1.3. In anticipation of the following exposition, I note that the following terminology is adopted. “Hispano-Celtic” designates the Celtic of the Iberian Peninsula. “Cisalpine Celtic” designates all of the Celtic attested on the Italian side of the Alps, both “Lepontic” and “Cisalpine Gaulish”; see further 5. “Transalpine Celtic” designates all of the Celtic on the French side of the Alps. “Goidelic” and “Brittonic” are employed in the standard way. Of course, one must assume some considerable degree of dialectal variation within these large groupings, but the fragmentary state of the data and the relatively early stage of the research which deals with such questions render finer categorizations premature. As will be seen, it is commonly the case that the branches and leaves of the Celtic family tree rest upon only a small number of characters.

2. Proto-Celtic

2.1. At the outset, an important distinction must be made between linguistic changes that propagated throughout the entirety of the proto-Celtic speech community and those which did not, i.e., issues of dialect geography must be taken into consideration. For example, the shortening of long vowels before final nasals, which did not affect the sector of the speech community that later became Hispano-Celtic, must be ordered before the raising of proto-IE */o:/ to proto-Celt. */u:/ in final syllables, which affected all of Celtic. This is the only way to account for *Hisp.-Celt. gen. pl. -um* = *-u:m* beside *Cisalp. Transalp. Celt. -on* and the Old Irish genitive plural, which continues **-on*, all = **-/on/* (Schrijver 2006: 53; Eska 2006).

3. Hispano-Celtic

3.1. Hispano-Celtic is characterized by a number of innovations which it does not share with the remainder of the Celtic language family, e.g.:

1. a) The extension of **-d* (continued by the character <ſ> = <s> in the Celtic adaptation of the Iberian script) in the ablative singular desinence throughout all stem classes after the thematic flexion, e.g., thematic-stem **uſamus** (Untermann 1975: 72), *ā*-stem **aŕeCoŕaTas** (Untermann 1975: 52.1), *i*-stem **PilPilis** (Untermann 1975: 73), consonant-stem **ſeCoPiriCes** (Untermann 1975: 89), etc. (all toponyms).
- b) The development of *-o* as the thematic genitive singular desinence, e.g., **aulo** (Untermann and Wodtko: 0.2) ‘an idionym’; cf. nom. sg. **auloſ** (Untermann and Wodtko: 1.3 i 55).
- c) The generalization of /u:/-vocalism throughout the entire masculine *n*-stem paradigm after the nominative singular, e.g., nom. sg. **uſeisu** (Untermann and Wodtko: 1.1 B7) ‘an idionym’, gen. sg. **uſeisunoſ** (e.g., Untermann and Wodtko: 1.3 ii 15), dat. sg. **uſeisunei** (Lorrio and Velaza 2005).

Villar (1997: 939–940) provides a fuller list, some items of which are controversial. These innovations, of course, tell us nothing about the phylogeny of the Celtic language family.

3.2. What does provide evidence that Hispano-Celtic was the first language to depart from the proto-Celtic speech community are a number of characters that it continues unaltered from proto-Celtic, whereas all of the other Celtic languages have innovated. They are described in 4.1.

4. Nuclear Celtic

4.1. Following the departure of what was to become Hispano-Celtic from the proto-Celtic speech community, the next major node in the family tree will be termed “Nuclear Celtic”, for it excludes Hispano-Celtic on the basis of features long associated with the Celtic language family which it does not share. The characters which validate Nuclear Celtic are as follows:

2. a) Proto-IE **/st/* > the tau Gallicum phoneme, which frequently has been interpreted as an affricate /tʃ/ (e.g., Lambert 2003: 45–46), though Eska (1998a) thinks it more likely to be a retracted coronal fricative /θθ/. Cf. Cisalp. Celt. nom. sg. pron. **iſos** (Solinas 1995: 119 = Morandi 2004: 106) < **isto-* and Latinized Transalp. Celt. dat. sg. **DIRONAE** (e.g., CIL xiii 4498) ‘a theonym’ < **h₂ster-* ‘star’ beside Hisp.-Celt. acc. sg. **PouſTom** (Untermann and Wodtko: 1.1 A6) ‘cow stable’ < **g^wou-sto-*. At a later date, the tau Gallicum phoneme was simplified to /s(s)/, e.g., OIr. MW *glas* ‘blue, green’ < **glasto-*.
- b) Proto-IE **/ej/* > /e:/ in non-final position. Cf. Cisalp. Celt. **Teu** (Solinas 1995: 119 = Morandi 2004: 106) ‘an idionym’ < **dejuō* and Transalp. Celt. *devo-* ‘an

onomastic element' < **deīuo-* 'god' beside *Hispano-Celt. Teiuořeiciś* (Untermann and Wodtko: 6.1) 'an idionym' < **deīuo-*, as well as *OIr. sciath*, *MW ysgwyd* 'shield' < **sk^wē-to-* < **sk^wej-to-*.

- c) The shortening of long vowels before final nasals; cf. 2.1.
- d) The grammaticalization of the connective **to* (cf. *OHitt. ta*). While in *Hispano-Celtic* this form continues to function as a connective in the large inscription from Peñalba de Villastar (Untermann and Wodtko: 3.3; Ködderitzsch 1985: 212–213, 1996: 149), it functions only as a host for pronominal clitics in *Cisalpine Celtic*, e.g., **To=šo=KoTe** (RIG *E-2 = Solinas 1995: 141 = Morandi 2004: 100) 'he gave it', after which it evolved to the preverb attested by *OIr. do-* and *MW dy-* (Eska 2007).

Other characters which validate *Nuclear Celtic* are in the process of developing in *Cisalpine Celtic*:

- e) The rise of the thematic genitive singular in *-ī*. Cf. *Cisalpine Celtic. aśKoneTi* (Solinas 1995: 21 = Morandi 2004: 38) 'an idionym', *Transalpine Celtic. addedillj* (RIG L-100) 'a patronym', *OIr. fir* 'man' < **uirī*, *Early MW kyrd* 'army' < **korī* (Koch 1991: 114), beside *Hispano-Celtic*, in which only *-o* is known. In the earliest *Cisalpine Celtic*, the thematic genitive singular desinence is *-oiso* ← **-osjo*, e.g., **Ploiso** (Solinas 1995: 80 = Morandi 2004: 153), after gen. pl. pron. **-oīsōm* as attested by *Hispano-Celt. śoiśum* (Untermann and Wodtko: 1.3 Ü).
- f) The syncretism of the dative singular desinence with that of the locative singular in consonant-stems. Cf. *Cisalpine Celtic. Kuaśoni* (Solinas 1995: 20 = Morandi 2004: 26) 'an idionym', *Transalpine Celtic. EPAĐATEXTORICI* (RIG L-6) 'an idionym', *OIr. rīg* 'king' < **rīgi* beside *Hispano-Celt. uśeisunei* (Lorrio and Velaza 2005) 'an idionym'. *Cisalpine Celtic* also attests inherited dat. sg. **-ej*, e.g., **Piuonei** (Solinas 1995: 26 = Morandi 2004: 36) 'an idionym'.

5. Cisalpine Celtic

5.1. The second language to branch off from the Celtic family tree is *Cisalpine Celtic*. It is here taken to be composed of both of what have traditionally been termed "Leponitic" and "Cisalpine Gaulish". Though conventionally assumed to be discrete languages, the differences between them are largely due to the disparity in their periods of attestation, "Leponitic" being attested from at least the early fifth century BCE on, while "Cisalpine Gaulish" is attested from no earlier than ca. 150 BCE, and the fact that "Leponitic", being attested in a small, mountainous area in the northern Italian lake district, likely represents a geographically peripheral and conservative dialect. Thus differences such as *Lep. -/m/* vs. *Cisalpine Gaul. -/n/* in final position and the presence of the patronymic formant *-alo-* in *Leponitic* (perhaps borrowed from neighboring *Raetic*), but not in *Cisalpine Gaulish*, are due merely to temporal and geographical factors (Eska 1998b; cf. Uhlich 1999, 2007).

5.2. On the other hand, "Leponitic" and "Cisalpine Gaulish" share such characters as the regular effacement of nasals before voiceless plosives and heteromorphemic voiced plosives, e.g., *Lep. PiuoTialui* (Solinas 1995: 3 = Morandi 2004: 34) 'a patronym' <

*g^w*iu-ont-* and Cisalp. Gaul. **KuiTos** (RIG E-1 = Solinas 1995: 140 = Morandi 2004: 97) ‘an idionym’ = Lat. *Quintus* and the assimilation of /nd/ > /nn/ intramorphemically, e.g., Lep. **alKouinos** (Solinas 1995: 21 = Morandi 2004: 38) ‘an idionym’ < **yindo-* and Cisalp. Gaul. **anoKoPoKios** (RIG E-1 = Solinas 1995: 140 = Morandi 2004: 97) ‘a patronym’ < **ando-*, and the evolution of the prefix **eks-* > *ess-* as seen in Lep. **esoPnio** (Solinas 1995: 127 = Morandi 2004: 71) ‘an idionym’ and Cisalp. Gaul. **esaneKoTi** ‘a patronym’ (RIG E-1 = Solinas 1995: 140 = Morandi 2004: 97). Following the departure of Cisalpine Celtic, the remaining node will be termed Core Celtic.

6. Core Celtic

6.1. Following the departure of what was to become Cisalpine Celtic from the Celtic speech community, the next major node in the family tree will be termed “Core Celtic”, for it includes characters long identified as diagnostic of Celtic on the basis of Goidelic and Brittonic. The characters which validate Core Celtic are as follows:

3. a) Proto-IE */eĵ/ > /e:/ in final position. Cf. Transalp. Celt. *i*-stem dat. sg. VCVETE (RIG L-13) ‘a theonym’ < **eĵ* < **eĵeĵ* via haplology, OIr. *cía*, OW *pui* ‘who?’ < **k^we* < **k^weĵ* beside Hisp.-Celt. *n*-stem dat. sg. **uŕeisunei** (Lorrio and Velaza 2005) ‘an idionym’ and early Cisalp. Celt. *n*-stem dat. sg. **aTilonei** (Solinas 1995: 12 = Morandi 2004: 13) ‘an idionym’.

b) The merger of the *ā*-stem and *ī*-stem paradigms. Cf. Transalp. Celt. *ā*-stem nom. sg. *paulla* (RIG L-98 1^a10), gen. sg. *paullias* (RIG L-98 1^a12) ‘an idionym’, OIr. *ā*-stem nom. sg. *túath* < **tōtā* < **toutā*, gen. sg. *túaithe* < **tōtīās* < **toutīās* beside Hisp.-Celt. *ā*-stem nom. sg. AIA (e.g., CIL ii 5798), gen. sg. *aias* (Untermann and Wodtko: 1.3 ii 29) ‘an idionym’ and Cisalp. Celt. *ā*-stem nom. sg. **aŕmina** (Solinas 1995: 122 = Morandi 2004: 94) ‘an idionym’, gen. sg. **TouTas** (RIG E-1 = Solinas 1995: 140 = Morandi 2004: 97) ‘tribe’.

A potential third character is:

c) The evolution of an uninflected clitic subordinator =*ĵo*, though the lack of evidence from Cisalpine Celtic which bears upon it leaves room for the possibility that it may be relevant, instead, to the validation of the Nuclear Celtic node. Cf. Transalp. Celt. DVGIJONTI=JO (RIG L-13) ‘who serve’, OIr. *bertae* ‘who bear’ < **beronti=ĵo*, OW *issid* ‘who is’ < **hesti=ĵo* beside Hispano-Celtic nom. sg. masc. stressed **ioŕ** (Untermann and Wodtko: 1.1 A10), dat. sg. masc. **iomui** (Untermann and Wodtko: 1.1 A7), nom. sg. fem. or nom. pl. **ia** (Untermann and Wodtko: 1.3 Ü), and acc. pl. fem. **ias** (Untermann and Wodtko: 1.1 A8).

Other characters which validate Core Celtic are in the process of developing in Transalpine Celtic:

d) The weakening and loss of intervocalic /s/. While intervocalic /s/ is normally preserved in Transalpine Celtic, e.g., ESVS (RIG *L-14) ‘a theonym’, there are some tokens which evince its loss, e.g., dat. pl. SVIOREBE (RIG L-6) ‘sister’ < **syesor-*; cf. OIr. *siur*, MW *chwaer*.

e) The syncretism of the dative plural desinenence with that of the instrumental plural. Cf. Transalp. Celt. GOBEDBI (RIG L-13) ‘to the smiths’, OIr. *feruib* ‘man’ (dat. pl.) < **b^hi(s)* beside Hisp.-Celt. **afeCoŕaTaCuPoŕ** (Untermann and Wodtko: 6.1)

‘an ethnic name’ and Cisalp. Celt. **ariuonePos** (Solinas 1995: 65 = Morandi 2004: 180) ‘a ?patronym?’. Transalpine Celtic also attests inherited dat. pl. *-bo* < **-b^hos*, e.g., *ματρεβο* (RIG G-64) ‘to the mothers’.

7. Transalpine Celtic

7.1. The third language to branch off from the Celtic speech community is Transalpine Celtic. It is validated by its participation in the Core Celtic characters described in 6 and its failure to do so in the characters which distinguish the Insular Celtic node on the family tree.

8. Insular Celtic

8.1. Following the separation of what was to become Transalpine Celtic, the next, and last, major node in the family tree will be termed “Insular Celtic” on the basis of geography. The characters which validate Insular Celtic are as follows:

4. a) The evolution of the dual system of verbal flection. This is best known from Old Irish, as, e.g., in the present indicative paradigm of simplex *beirid* ‘bears’, in which absolute flexion is employed in absolute clause-initial position and conjunct flection when the verb is preceded by a “conjunct” particle, among which are included negators, complementizers, connectives, and preverbs:

(i)	Absolute	Conjunct
1. sg.	<i>biru</i>	<i>biur</i>
2.	<i>biri</i>	<i>bir</i>
3.	<i>beirid</i>	<i>beir</i>
1. pl.	<i>bermai</i>	<i>beram</i>
2.	<i>beirthe</i>	<i>beirid</i>
3.	<i>berait</i>	<i>berat</i>

and in the present indicative paradigm of compound *do·beir* ‘give’, in which deuterotonic stress is employed in absolute clause-initial position and prototonic stress when the verb is preceded by a conjunct particle:

(ii)	Deuterotonic	Prototonic
1. sg.	<i>do·'biur</i>	<i>'tabur</i>
2.	<i>do·'bir</i>	<i>'tabair</i>
3.	<i>do·'beir</i>	<i>'tabair</i>
1. pl.	<i>do·'beram</i>	<i>'taibrem</i>
2.	<i>do·'beirid</i>	<i>'taibrid</i>
3.	<i>do·'berat</i>	<i>'taibret</i>

There are vestiges of this system in Brittonic, too, as exemplified by the Middle Welsh gnomic maxim *trenghit golut, ny threingk molut* ‘wealth perishes, fame does not perish’, with absolute *trenghit* vs. conjunct *treingk*. However the evolution of this dual system is to be accounted for – there are several competing

theories –, the uniqueness of this character is robust enough by itself to validate the legitimacy of the Insular Celtic node in the Celtic family tree.

- b) The grammaticalization of the verbal adjective in **-to/ā-* as the passive preterite member of the verbal paradigm. Cf. OIr. *·slass*, MW *llas* ‘was killed’ < **slad-to-* beside Hisp.-Celt. **ConsCiliTom** (Untermann and Wodtko: 1.1 A3) ‘?cut up?’ and Transalpine Celtic idionyms such as Latinized CINTVGNATVS (e.g., AE 1990, 695) ‘first born’.

8.2. Many scholars prefer to group Transalpine Celtic and Brittonic together to the exclusion of Goidelic, thus establishing what has been labeled as a “Gallo-Brittonic” node in the Celtic family tree. Many of the linguistic features shared by Transalpine Celtic and Brittonic, most in the realm of phonology, however, are natural and easily repeatable, and are probably to be ascribed to areal developments. The evolution of the dual system of verbal flexion shared by Goidelic and Brittonic, evidence for which is completely lacking in Transalpine Celtic, on the other hand, is so unusual and distinctive as to guarantee the diagnosis of an Insular Celtic node in the Celtic family tree.

9. Goidelic vs. Brittonic

9.1. There is little controversy concerning the separating out of the Goidelic and Brittonic languages. Goidelic divided into a western branch which became Irish and an eastern branch which later became Scottish Gaelic and Manx after its speakers expanded into Scotland and the Isle of Man in the fifth century CE. Brittonic is now thought to have remained a dialect continuum longer, Old Welsh, Old Cornish, and Old Breton not being truly distinct languages, but dialects of what may be labeled “Old Brittonic” (Schrijver 2011). They eventually differentiated into a northern branch represented by Welsh and a southwestern branch composed of Cornish and Breton. Most of the linguistic developments that distinguish Goidelic from Brittonic and their respective daughter languages from each other are phonological. These may be easily found in Pedersen (1909). There are some morphological differences, many of these triggered by phonological changes, e.g., the preservation of a nominal case system in Goidelic as opposed to all but the barest vestiges in Brittonic (Koch 1991: 113–115). Within the syntactic component of the grammar, it is noteworthy that the Brittonic languages shifted to verb-second clausal configuration as they transitioned from their Old to their Middle stages (the best explanation in print is Manning 2001), but Welsh then later reverted back to verb-initial order (Willis 1998).

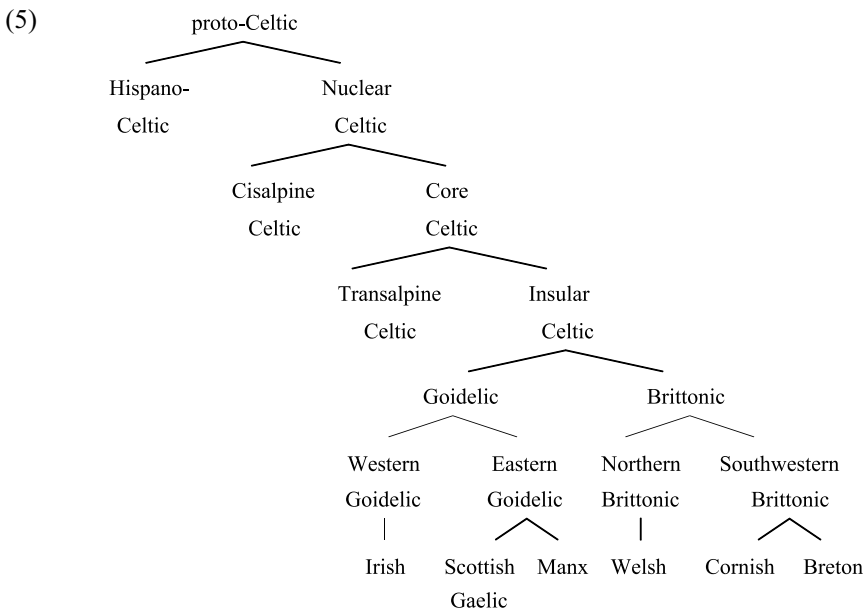
10. Other varieties of Continental Celtic

10.1. There are very modest remains of Continental Celtic attested in eastern Europe, variously known as Noric or Eastern Celtic, and Asia Minor, where Galatian was located, almost all of it onomastic (see Eska 2013a). The linguistic features that can be garnered from these records suggest that, by and large, these languages were similar to Transalpine Celtic.

10.2. Two fragmentarily attested languages from the Iberian Peninsula have been claimed to be Celtic. Untermann (1985) argues that Lusitanian, attested in the west of the Peninsula, is an archaic form of Celtic, but the *communis opinio* is decidedly against this (e.g., Prósper 2002: 429–431). Correa (1992) and Untermann (1995) have also tentatively proposed that Tartessian, attested in the extreme southwest of the Peninsula, may be Celtic, an idea now forcefully taken up by Koch (2009, 2011), but which is strongly disputed by de Hoz (2011: 588) and Eska (2013b, 2013c, 2014).

11. Conclusions

11.1. The phylogenetic structure of the Celtic family tree thus takes the following shape:



This array of relationships, of course, is provisional and subject to revision upon the future discovery of further linguistic data.

12. References

AE = *L'année épigraphique*.

De Bernardo Stempel, Patrizia

- 2013 From Indo-European to the individual Celtic languages. In: Dónall Ó Baoill, Donncha Ó hAodha, and Nollaig Ó Muraíle (eds.), *Saltair saíochta, sanasaíochta agus seachais. A Festschrift for Geróid Mac Eoin*. Dublin: Four Courts Press, 25–42.

CIL = *Corpus inscriptionum Latinarum*. Berlin: Reimer.

- Correa, José Antonio
 1992 La epigrafía tartesia. In: Dieter Hertel and Jürgen Untermann (eds.), *Andalusien zwischen Vorgeschichte und Mittelalter*. Cologne: Böhlau, 75–114.
- Eska, Joseph F.
 1998a Tau Gallicum. *Studia Celtica* 32: 115–127.
- Eska, Joseph F.
 1998b The linguistic position of Lepontic. *Proceedings of the Berkeley Linguistics Society* 245: 2–11.
- Eska, Joseph F.
 2006 The genitive plural desinence in Celtic and dialect geography. *Die Sprache* 46: 229–235.
- Eska, Joseph F.
 2007 On the prehistory of OIr. do-, MW dy-, etc. *International journal of diachronic linguistics and linguistic reconstruction* 4: 185–206.
- Eska, Joseph F.
 2013a A salvage grammar of Galatian. *Zeitschrift für celtische Philologie* 60: 51–63.
- Eska, Joseph F.
 2013b Review of Koch (2009). *Kratylos* 58: 58–67.
- Eska, Joseph F.
 2013c Review of Koch (2011). *Kratylos* 58: 68–73.
- Eska, Joseph F.
 2014 Comments on John T. Koch's Tartessian-as-Celtic enterprise. *Journal of Indo-European studies* 42: 428–438.
- Garrett, Andrew
 1999 A new model of Indo-European sub-grouping and dispersal. In: Steve S. Chang, Lily Liaw, and Josef Ruppenhofer (eds.), *Proceedings of the twenty-fifth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 146–156.
- de Hoz, Javier
 2011 *Historia lingüística de la Península Ibérica en la antigüedad II. El mundo ibérico prerromano y la indoeuropeización*. Madrid: Consejo Superior de Investigaciones Científicas.
- Isaac, Graham R.
 2005 Insular Celtic vs. Gallo-Brittonic. An empirical and methodological question. In: William Gillies and Dennis W. Harding (eds.), *Celtic connections II. Archaeology, numismatics, historical linguistics*. Edinburgh: University of Edinburgh Press, 190–202.
- Koch, John T.
 1991 Gleanings from the Gododdin and other Early Welsh texts. *Bulletin of the Board of Celtic Studies* 38: 111–118.
- Koch, John T.
 1992 'Gallo-Brittonic' vs. 'Insular Celtic'. The inter-relationships of the Celtic languages reconsidered. In: Gwennolé Le Menn (ed.), *Bretagne et pays celtiques. Langues, histoire, civilisation: mélanges offerts à la mémoire de Léon Fleuriot*. Saint-Brieuc: Presses Universitaires Rennes, 471–495.
- Koch, John T.
 2009 *Tartessian. Celtic in the south-west at the dawn of history*. Aberystwyth: Celtic Studies Publications.
- Koch, John T.
 2011 *Tartessian 2. The inscription of Mesas do Castelinho. ro and the verbal complex. Preliminaries to historical phonology*. Aberystwyth: University of Wales Centre for Advanced Welsh and Celtic Studies.

Ködderitzsch, Rolf

- 1985 Die große Felsinschrift von Peñalba de Villastar. In: Hermann M. Ölberg and Gernot Schmidt (eds.), *Sprachwissenschaftliche Forschungen. Festschrift für Johann Knobloch zum 65. Geburtstag am 5. Januar 1984 dargebracht von Freunden und Kollegen*. Innsbruck: Institut für Sprachwissenschaft der Universität, 211–222.

Ködderitzsch, Rolf

- 1996 Noch einmal zur großen Inschrift von Peñalba de Villastar. In: Wolfgang Meid and Peter Anreiter (eds.), *Die grösseren altkeltischen Sprachdenkmäler*. Innsbruck: Institut für Sprachwissenschaft der Universität, 149–155.

Lambert, Pierre-Yves

- 2002 *Recueil des inscriptions gauloises II/2. Textes gallo-latins sur instrumentum*. Paris: Centre National de la Recherche Scientifique.

Lambert, Pierre-Yves

- 2003 *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*. 2nd edn. Paris: Errance.

Lambert, Pierre-Yves and Georges-Jean Pinault (eds.)

- 2007 *Gaulois et celtique continental*. Geneva: Droz.

Lejeune, Michel

- 1985 *Recueil des inscriptions gauloises I. Textes gallo-grecques*. Paris: Centre National de la Recherche Scientifique.

Lejeune, Michel

- 1988 *Recueil des inscriptions gauloises II/1. Textes gallo-étrusques. Textes gallo-latins sur pierre*. Paris: Centre National de la Recherche Scientifique.

Lorrio, Alberto J. and Javier Velaza

- 2005 La primera inscripción celtibérica sobre plomo. *Palaeohispanica* 5: 1031–1048.

Manning, H. Paul

- 2001 *Typology, history and 'teratology'. The rise and fall of the 'abnormal' main clause with relative structure in p-Celtic*. Unpublished University of Chicago doctoral dissertation.

McCone, Kim

- 2001 Celtibérico, celta continental y celta común. In: Francisco Villar and Maria Pilar Fernández Álvarez (eds.), *Religión, lengua y cultura prerromanas de Hispania*. Salamanca: Ediciones Universidad de Salamanca, 483–494.

Morandi, Alessandro

- 2004 *Celti d'Italia II, Epigrafia e lingua*. Rome: Spazio Tre.

Nakhleh, Luay, Don Ringe, and Tandy Warnow

- 2005 Perfect phylogenetic networks. A new methodology for reconstructing the evolutionary history of natural languages. *Language* 81: 382–420.

Pedersen, Holger

- 1909 *Vergleichende Grammatik der keltischen Sprachen*. Vol. I. *Einleitung und Lautlehre*. Göttingen: Vandenhoeck & Ruprecht.

Prósper, Blanca Maria

- 2002 *Lenguas y religiones prerromanas del occidente de la Península Ibérica*. Salamanca: Ediciones Universidad de Salamanca.

RIG G see Lejeune 1985.

RIG E see Lejeune 1988: 1–54.

RIG L-1-L-16 see Lejeune 1988: 55–194.

RIG L-18-L-139 see Lambert 2002.

Ringe, Don, Tandy Warnow, and Ann Taylor

- 2002 Indo-European and computational cladistics. *Transactions of the Philological Society* 100: 59–129.

Solinas, Patrizia

- 1995 Il celtico in Italia. *Studi etruschi* 60: 311–408.

- Schrijver, Peter
2006 Review of Gerhard Meiser, *Veni, vidi, vici. Die Vorgeschichte des lateinischen Perfectsystems. Kratylos* 51: 46–64.
- Schrijver, Peter
2011 Old British. In: Elmar Ternes (ed.), *Brythonic Celtic – Britannisches Keltisch. From Medieval British to Modern Breton*. (Münchener Forschungen zur historischen Sprachwissenschaft 11). Bremen: Hempen, 1–84.
- Sims-Williams, Patrick
2007 Common Celtic, Gallo-Brittonic and Insular Celtic. In: Lambert and Pinault (eds.), 309–354.
- Uhlich, Jürgen
1999 Zur sprachlichen Einordnung des Lepontischen. In: Stefan Zimmer, Rudolf Ködderitzsch, and Arndt Wigger (eds.), *Akten des zweiten deutschen Keltologen-Symposiums*. Tübingen: Niemeyer, 277–304.
- Uhlich, Jürgen
2007 More on the linguistic classification of Lepontic. In: Lambert and Pinault (eds.), 373–411.
- Untermann, Jürgen
1975 *Monumenta linguarum Hispanicarum I. Die Münzlegenden*. Wiesbaden: Reichert.
- Untermann, Jürgen
1985 Lusitanisch, Keltiberisch, Keltisch. *Veleia* 2–3: 77–91.
- Untermann, Jürgen
1995 Zum Stand der Deutung der ‘tartessischen’ Inschriften. In: Joseph F. Eska, R. Geraint Gruffydd, and Nicolas Jacobs (eds.), *Hispano-Gallo-Brittonica. Essays in honour of Professor D. Ellis Evans on the occasion of his sixty-fifth birthday*. Cardiff: University of Wales Press, 244–259.
- Untermann, Jürgen and Dagmar S. Wodtke
1997 *Monumenta linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert, 347–722.
- Villar, Francisco
1997 The Celtiberian language. *Zeitschrift für celtische Philologie* 49–50: 898–949.
- Warnow, Tandy, Stephen N. Evans, Donald Ringe, and Luay Nakhleh
2006 A stochastic model of language evolution that incorporates homoplasy and borrowing. In: Peter Forster and Colin Renfrew (eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaeological Research, 75–87.
- Willis, David W. E.
1998 *Syntactic change in Welsh. A study of the loss of verb-second*. Oxford: The Clarendon Press.

Joseph F. Eska, Blacksburg, VA (USA)

73. The evolution of Celtic

1. Introduction
2. The stress accent, its position and effect
3. The evolution of regularization
4. Language contact and the evolution of the Celtic languages
5. References

- Schrijver, Peter
2006 Review of Gerhard Meiser, *Veni, vidi, vici. Die Vorgeschichte des lateinischen Perfectsystems. Kratylos* 51: 46–64.
- Schrijver, Peter
2011 Old British. In: Elmar Ternes (ed.), *Brythonic Celtic – Britannisches Keltisch. From Medieval British to Modern Breton*. (Münchener Forschungen zur historischen Sprachwissenschaft 11). Bremen: Hempen, 1–84.
- Sims-Williams, Patrick
2007 Common Celtic, Gallo-Brittonic and Insular Celtic. In: Lambert and Pinault (eds.), 309–354.
- Uhlich, Jürgen
1999 Zur sprachlichen Einordnung des Lepontischen. In: Stefan Zimmer, Rudolf Ködderitzsch, and Arndt Wigger (eds.), *Akten des zweiten deutschen Keltologen-Symposiums*. Tübingen: Niemeyer, 277–304.
- Uhlich, Jürgen
2007 More on the linguistic classification of Lepontic. In: Lambert and Pinault (eds.), 373–411.
- Untermann, Jürgen
1975 *Monumenta linguarum Hispanicarum I. Die Münzlegenden*. Wiesbaden: Reichert.
- Untermann, Jürgen
1985 Lusitanisch, Keltiberisch, Keltisch. *Veleia* 2–3: 77–91.
- Untermann, Jürgen
1995 Zum Stand der Deutung der ‘tartessischen’ Inschriften. In: Joseph F. Eska, R. Geraint Gruffydd, and Nicolas Jacobs (eds.), *Hispano-Gallo-Brittonica. Essays in honour of Professor D. Ellis Evans on the occasion of his sixty-fifth birthday*. Cardiff: University of Wales Press, 244–259.
- Untermann, Jürgen and Dagmar S. Wodtke
1997 *Monumenta linguarum Hispanicarum IV. Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Reichert, 347–722.
- Villar, Francisco
1997 The Celtiberian language. *Zeitschrift für celtische Philologie* 49–50: 898–949.
- Warnow, Tandy, Stephen N. Evans, Donald Ringe, and Luay Nakhleh
2006 A stochastic model of language evolution that incorporates homoplasy and borrowing. In: Peter Forster and Colin Renfrew (eds.), *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaeological Research, 75–87.
- Willis, David W. E.
1998 *Syntactic change in Welsh. A study of the loss of verb-second*. Oxford: The Clarendon Press.

Joseph F. Eska, Blacksburg, VA (USA)

73. The evolution of Celtic

1. Introduction
2. The stress accent, its position and effect
3. The evolution of regularization
4. Language contact and the evolution of the Celtic languages
5. References

1. Introduction

Other chapters have dealt with the specifics of the development of the Celtic languages and how they are related. The present chapter adopts a different perspective and starts with a series of questions: How have the Celtic language evolved from our earliest evidence onwards? What were the main drivers of that evolution? The aim is take several large themes and consider their impact on the internal evolution of the Celtic languages. Complete coverage is not possible but different sections consider phonology, morphology, and the effects of language contact. Given that we are discussing evolution, most of the presentation centers on the insular Celtic languages where we have a longer documented history to work with.

2. The stress accent, its position and effect

The position of the stress in the Continental Celtic languages remains uncertain (see de Bernardo-Stempel 1995; Schrijver 1995: 20–21). This is in part due to the relatively short time-span of the evidence, especially for Celt-Iberian, such that in effect we only gain a snapshot of the languages over two to three hundred years of development, when it takes a longer period for us to be able to see the effects of a stress-accent or the effects of a change of position of the accent. Unlike Celt-Iberian Gaulish does at least provide evidence over a longer duration as we see the effects of the accent in place-names in later Gaul and France. The evidence, however, is contradictory displaying outcomes of both a penultimate and an ante-penultimate accent, and sometimes in the same form: e.g. *Redon* < *Redónes* : *Rennes* < *Rédones*, *Berry* < *Bituríges* : *Bourges* < *Bitúríges*, *Nîmes* < *Némausus* : *Nemours* < *Nemáusus*. It has been argued that the ante-penultimate pattern is innovatory, having been influenced by Greek and Latin speakers for whom an ante-penultimate stress was acceptable (Falc'hun 1981). A similar pattern may also be found in personal names like *Cintusmus* < **kintusamos* (which shows apocope of the penultimate syllable which therefore could not have been stressed). The French reflexes of common nouns of Celtic origin are less easy to use in this connection; there are numerous examples which indicate a non-penultimate (often initial) stress of the original but, in all cases, this could be the ante-penultimate stress of Gallo-Latin, e.g. *berle* < *bérula* (< Gaulish *berura*), *breuil* < *brógilo*, *vautre* < *vétragus*, etc. (on Gaulish loanwords in French, see Lambert 2003: 187–203). If so, then, by a process of elimination it would appear that the native Gaulish pattern was penultimate (Schrijver 1995: 21).

Brittonic languages likewise show a regular penultimate stress pattern, and this tendency was sufficiently strong that after apocope of final syllables the stress shifted back onto the new penultimate in most dialects (except for Vannetais Breton [below 2.2.4]). However, there is some evidence to show that Common Brittonic inherited an initial stress accent. Two examples (where shorter and longer forms can be contrasted) will make the point (for further discussion, see Schrijver 1995: 18–19). First, beside MW *nei* '(first) cousin' < Common Celtic **ne(φ)u:ts* < IE **nepo:ts* (where the first syllable might be described as penultimate or initial), MW *keifn* 'third cousin' can only go back to **kom-ne(φ)u:ts* with initial stress, as a penultimate-stressed ***kom-ne(φ)u:ts* would have given ***kyfnei* (with reduction of the pretonic **kom-*). Similarly, MW *eil*, MC *yl*,

Breton *eil* ‘second’ < *alijos (cf. OIr. *aile*) and *MW arall* ‘other’ < *alalijos show different treatments of /li/ which is probably related to the position of the accent (i.e. /li/ > /A/ following an unstressed vowel): *alijos vs. *alalijos (Schrijver 1995: 19). To judge from Latin loanwords like *papiliō* > *MW pebyll*, *Aprīlius* > *Ebrill* which show the same reflex as *arall*, the initial stress was still in place at the time when the earliest Latin loanwords were being borrowed into Brittonic, *Aprīlius* being crucial as the Latin stress could never have been on the initial syllable in this word. If Brittonic originally had an initial stress accent, then this would fit with Goidelic where the stress has remained firmly on the initial syllable (see 2.1.1 for some modern exceptions), though it remains unclear why and how the stress accent in Brittonic shifted (see 2.2.4 below).

We now consider the effects of the stress accent in Goidelic and Brittonic in greater detail. It should however be remembered that the different phonological outcomes of initial and penultimate stress are only visible on words of more than three syllables.

2.1. Goidelic

The effect of a persistent initial stress accent was to bring about significant changes to the vocalism and syllabic structure of Goidelic. The survival of inscriptions in Ogam script from the 5th century CE onwards allow us to gain a sense of the changes as they preserve a stage of the language before these wholesale changes had taken place (or at least before they had affected the written form); thus Ogam *RITUVECCAS* : *Rethech*, *TOVISACI* (gen. sg.) : OIr. *toisech* ‘leader’, etc. (McManus 1991: 103, 108; Ziegler 1994: 226, 237 respectively). What follows considers how the stress accent affected different aspects of the phonology.

2.1.1. Vocalism

Goidelic languages inherited a contrastive pattern of long and short vowels (see Stifter, this handbook) which were still phonemically contrastive in all syllables. One effect of the stress accent was to reduce unstressed long vowels to the equivalent short vowel (McCone 1996: 110), e.g. L *molīna* > OIr. *muilenn* where the *-e-* is the outcome of an unaccented reduced short /i/ lowered to /e/ before the final /a/ in the next syllable (/i:/ could not have been lowered). As we shall see (2.1.2 below), however, long vowels in final closed syllables ending in /h/ seem to have survived longer. In Archaic Old Irish, therefore, long and short vowels were only contrastive in initial stressed syllables of polysyllabic words and in monosyllabic words. The reduction of unaccented long vowels must have taken place before the loss of final syllables since, as in the example above, a short /i/ arising from the shortening was open to vowel affection from the vowel in the final syllable. At a later stage, after the loss of final syllables, new long vowels arose in unaccented syllables from clusters of spirant and resonant, e.g. *anatlo- > */anaθl/- > /'ana:l/ OIr. *ánal* ‘breath’ (cf. *MW anadl*), etc. It is worth noting also that the connection between vowel length and the stress accent was long-lasting. One of the features of the Munster dialect of Modern Irish is that in certain environments the stress shifts to the secondary long syllable, notably in derivatives in *-án* and *-ín* (especially

when they are disyllabic [Doyle 1992: 115–132]). Conversely, in more northern dialects of Irish and in Scots Gaelic there is a tendency in certain environments to reduce *-/a:n/* to *-/an/* (which was not reduced to *-/ən/*).

The persistence of the effects of the stress accent on unaccented syllables is traceable too in Middle Irish: while in Old Irish the quality of the short vowels in unaccented syllables remained intact, by late Old Irish and early Middle Irish the quality of unaccented short vowels was reduced to */ə/*. Thus the final syllables in Old Irish *céle* (nom. sg.), *céli* (gen. sg.), *céliu* (dat. sg.) would have been distinct: */'ke:l̪e/*, */'ke:l̪i/*, and */'ke:l̪u/* respectively. But the evidence of rhyme shows that in Middle Irish they would all have been pronounced */'ke:l̪ə/* and spelt indiscriminately *céle*, *céli*, etc. The consequences for the case system (and especially for those declensions where some of the distinctions were marked by final vowels) were potentially chaotic and was one of the triggers for the extension of the consonant-stem forms in the nom. and acc. plurals, thus MIr. *céileda* replacing OIr. *céli* (Breatnach 1994: 246).

2.1.2. Loss of most final syllables (apocope)

The gradual loss of final syllables likewise seems to have been, at least in part, a consequence of the initial stress accent. The process seems to have been long and slow (McCone 1982: 24–25), and the rate may have been determined by the number of syllables in the word. The changes were clearly happening over the period that the Ogam inscriptions were being produced as it is possible to see some of the changes reflected in the inscriptions; for example, the fate of the original genitive singular feminine ending *-/ja:s/* can be charted through the inscriptions: *MAQI ERCIAS*, *MAQI RITEAS*, *MAQI RITE* (cf. OIr. *túaithe*), phonologically corresponding to *-/ja:s/* > *-/i̯a:s/* > *-/i̯jas/* > *-/e̯jah/* > *-/e(i̯)a #h-/* > *-/e/*. It is clear that a significant element in this (in addition to the resegmentation of *-/h/* [**-/s/*] and *-/n/* onto the beginning of the following word) was the reduction of long vowels in closed syllables and the complete loss of short vowels in absolute final position, e.g. nom. sg. *fer* ‘man’ < **/'uera #h-/* < **/'uerah/* < **/'ueras/* < IE **/'uīros/* : acc. pl. *fīru* < **/'uīru #h-/* < **/'uīru:h/* < **/'uīru:s/* < IE **/'uīro:s/*.

2.1.3. Syncope

After the loss of final syllables, the immediately post-tonic syllable, which by now could only contain a short vowel, was lost through syncope. Again the changes can be best illustrated using contrasting pairs, e.g. (syncopated vowel in **bold**) OIr. *samail* ‘similar’ / *'saɲal̪i/* : *cosmail* ‘similar’ / *'koʃmal̪i/* (< **co-samail*); OIr. *torad* (early OIr. *toreth*) ‘fruit’ / *'torəð/* (earlier *'toreθ/*) : *toirthech* ‘fruitful’ / *'tor̪iθ̪eχ/* < **toreta:ko-* (in addition the syncope of a front vowel caused palatalization of the whole cluster; on palatalization, see Stifter, *The phonology of Celtic*, this handbook). Apocope and syncope together had the effect of reducing four-syllable Common Celtic words to disyllabic words in Old Irish. In some cases the patterns of syncope could have consequences for the morphological patterns; for example, Common Celtic had an adjectival suffix **-od̪i-* (> W *-aid̪*), but the form of the suffix in Old Irish was determined by the syllable count of the base:

e.g. *rigdae* ‘royal’ /’ri:γðe/ (< *ri:g-odjō-) : *rí* ‘king’ (< *ri:g-), but *bliadnaide* ‘annual’ /’bliadnaðjē/ (< *bleidan-odjō-) : *bliadain* ‘year’ (on this suffix, see Wodtko 1995: 224–311).

A further effect of syncope was to create new consonantal clusters especially in words which originally had had four or more syllables: thus schematically ‘CV₁CV₂CV₃C V₄(C) > ‘CV₁CV₂CV₃C (by apocope) > ‘CV₁CCV₃C (by syncope). The consonantal clusters which arose were sometimes subject to adjustment in order to produce an articulatorily acceptable cluster, and in some cases this adjustment might involve epenthesis; for example, OIr. *comlann* ‘complete’ (a compound of *com-* and *lán* ‘full’) shows the post-tonic shortening of /a:/ to /a/ (and also the strengthening of the nasal at the end of an unaccented syllable [MacNeill’s Law]). The denominative verb based on *comlann* was *comalnaitir* ‘fulfills’ /’koμalnaθjirj/ < /’koμlanaθjirj/; here the original second syllable was syncopated, thus creating a cluster -/μln/- which required epenthesis, thus /’koμaln/-. The result is that the formal link between the stem of the derivative and the base adjective was weakened, if not broken. A similar but more wide-ranging case involving deponent verbs has been discussed by Ó Cruaíoch (1997) where divergent syncope patterns resulted in the development of distinct morphological categories (cf. also Stifter 2011 for developments in the nominal system where the expected syncope patterns are absent).

Finally, it is worth noting that consonants were not immune to the effects of the stress accent. It is clear that at the end of unstressed syllables changes took place to consonants which must be related to the absence of stress; for example, unvoiced and voiced fricatives tend to fall together in this position favouring the less strongly articulated voiced variant, e.g. (to take an example used earlier) CC *toreto- > early OIr. *toreth* /’toreθ/ ‘fruit’ > OIr. *torad* /’torað/ (> MIr. /’torəð/) beside *toirthech* ‘fruitful’, where /θ/ was preserved; similar mergers have been proposed for the gutturals and labials, e.g. *léicfea* ‘he will leave’ : *léiciub* ‘I shall leave’ (McCone 1996: 133–134).

2.1.4. Pretonic reductions

So far the discussion has concentrated on the effect of the initial stress patterns on the following syllables. However, it is clear that elements which immediately preceded the stressed syllable (and had some close connection with it) were also vulnerable to reduction and in some cases quite substantial reduction, though they have received less attention than their post-tonic counterparts.

The most common instances are found in the pretonic article and possessive pronouns. The forms of the Irish article, OIr. *in(d)*, *inna*, etc. (later *an*, *na*, etc.) are best explained as pretonic reductions of *sindos, *sinda, etc. (cf. Gaulish *indas*), the main feature of the reduction being the loss of /s/- (probably via /h/-). These forms may be contrasted with the accented pronoun *sin* ‘this/that’ which has preserved the /s/-; though the details of the relationship between these forms are unclear (see Schrijver 1997: 44–50), nevertheless they do seem to reflect different forms of the same particle. Likewise, the possessive pronouns, OIr. *mo* ‘my’, *do* ‘your (sg.)’, and *a* ‘his, her, their’ show considerable reduction in contrast to their stressed counterparts, *muí*, *taí*, and *aí*, both probably deriving from a proximate *moϥe, *teϥe, *eșjō (m.) / eșjā:s (f.), respectively. In other words, associated with their unaccented position is a reduction in articulation.

Another environment where we see forms moving from an unstressed to a stressed position is in the alternation between deuterotonic and prototonic verbs (for details, see Stifter, this handbook, 7.2). Essentially in compound verbs like OIr. *do·beir* ‘gives’, *do·gaib* ‘seizes’, the stress is on the verbal stem (deuterotonic), thus /doˈberʲ/, /doˈgavʲ/, but when the verb is preceded by a negative or interrogative particle, or it is in the imperative or verbal noun form, the stress shifts on to the preverb, thus *ní tabair* /niːˈtavarʲ/, *ní digaib* /niːˈdigavʲ/. What this pair of examples nicely illustrates is that the two preverbs (*tu [Schrijver 1995: 17; Eska 2007]) and *di are only distinguished in the prototonic forms where they are stressed; in unaccented position they fall together as /do-/. Preverbs and prepositions tend to be closely related in form (Russell 1988), and both raise interesting questions of stress. As these examples show, there is a clear relationship between the forms: *tu (+ stress) : do (- stress), *di (+ stress) : do (- stress). However, it is not clear that the distinction is binary. One function of preverbs in deuterotonic verbs is to support infixed pronouns, thus *dos·beir* ‘hands them over’, where the enclitic 3pl. pronoun -s occupies second position. It would follow that the *do* here cannot be completely unstressed or it would not be able to support the pronoun. Moreover, such collocations are frequent in initial position in the sentence, where they are not supported by some preceding stressed element. Similarly the prepositions *do* ‘to’ and *di* ‘from’, the forms corresponding to the preverbs noted above, also support enclitic articles or possessive pronouns, e.g. *dond fhiur* ‘to the man’ (dat. sg.), *dim thig* ‘from my house’, etc. Such forms suggest that preverbs and prepositions may in certain contexts carry a secondary stress, sufficient for hosting enclitics, but still weak enough for them to undergo the reduction of articulation associated with unstressed elements, e.g. /t/ > /d/, lowering of vowels, loss of /s/, etc. This is particularly evident in the later language when compared with Old Irish, e.g. OIr. *toichim* (verbal noun of *do·cing* ‘advances’ < *to-king(s)men) > *dochum* ‘towards’ > MÍr. *chum*, *immallei fri* ‘along with’ > *malle re* (Manx *marish*), *agus* ‘and’ > *as* /əs/, etc.

Such reduction can be substantial. One of the more complex areas where phonological reduction occurs has to do with personal names. Many Irish personal names involve assertions of patronymy and filiation (*mac* ‘son’, *ingen* ‘daughter’) or more complex relationships (OIr. *auē* ‘grandson’, *nad* ‘nephew [?]’). The evidence for such forms extends from the Ogam inscriptions, e.g. *MAQI CAIRATINI AVI INEQAGLAS*, *LUGUD-ECCAS MAQI MUCOI NETA SEGAMONAS*, OIr. *Nad Buidb*, *Nad Féic*, *Nad Ferb*, etc., to the modern languages, and it is clear from the later forms that these elements, which almost always preceded a personal name in the genitive, thus ‘son of X’, etc., had at least a secondary reduced stress (if it was not completely unstressed); good evidence for the pretonic (if not atonic) nature of these elements can be found in Manx surnames which often show the *mac* element reduced to a single /k/- or /kw/-, thus *Kermode* < *mac* (or perhaps *meic*) *Dhíarmaid*, *Quiggin* < *mac hUiginn*, *Clague* < *mac Liaig* (Kneen 1937). The modern *ní* ‘daughter’ (the parallel to *mac* for female names) has been derived from *ingen hui* (Ogam *INIGENA AVI*), lit. ‘the daughter of the grandson of ...’ (O’Brien 1973). Likewise, at an early stage Ogam *MAQI MUCOI* has been suggested as the source of OIr. *moccu*, *maccu* ‘descendant’ (used only as part of names) (Byrne 1994–1995; de Bernardo Stempel 1991). A striking example is the element *nad* (Ogam *NIOTTA*, *NETTA*), e.g. *Nad Fraích*, probably derived from an oblique form of OIr. *níae* (gen. sg. *níad*) ‘nephew’, which seems to have been fixed in names in this reduced form, even though it is used in the nominative; on reductions in personal names, see Russell (2015: 82–83).

2.2. Brittonic

As was noted in 2 above, there is some evidence that Brittonic originally had an initial stress accent, which might have been preserved into the period when Latin loanwords were being borrowed. It cannot have lasted very long into the common era before it shifted to a penultimate accent, as one of the early changes which distinguish Brittonic from Goidelic, *-/i/- > -/ð/-*, where *-/i/-* follows a stressed vowel (thus, *-/i̯io/- > W -ydd*), is dependent on the stress accent being penultimate; the change was dated to the 4th–5th century by Jackson (1953: 353–354). It is probable anyway that the shift from an initial to a penultimate stress was gradual, and of course disyllabic words would not have been affected as the initial and penultimate syllable would have been one and the same. How and why the accent shifted is unclear and will not detain us here. It is conceivable that contact with Latin (with its penultimate/antepenultimate pattern) was a factor, but the work on the cause of this shift remains to be done. Most of what follows, then, considers the effect of the penultimate stress accent and the consequence of the shift of that accent onto the new penultimate syllable after the loss of final syllables. One point worth making here is that generally the effects of the Brittonic stress accent were less destructive to the vocalism and syllable structure of the languages than the initial stress accent in Goidelic. As such, that might suggest that it was perhaps less strong or perhaps involved an element of pitch in addition to stress.

2.2.1. Vocalism

In contrast to Goidelic where the initial stress brought about wholesale reductions of long vowels and syncope in the following post-tonic syllables, unaccented vowels were relatively unaffected, except in Welsh, where pretonic */i/* and */u/* were reduced to */ə/*, e.g. *W ychen* ‘oxen’ */ʔəχen/* (earlier pre-shift */əʔχen/*) but *C ohan*, *B ouchen*, *oc’hen* < IE *uksén(es), cf. Skt. *ukṣāṇas*; *W ynyd* ‘Lent’ */ʔənɪd/* (earlier */ə'nɪd/*) but *C enes*, *B ened* ‘the beginning of Lent’ < L *initium* (Schrijver 1995: 161–162) The reduction must have occurred after the split between Welsh and the south-west Brittonic languages while the stress was still on the final syllable and before the shift of the accent, the effect of which was to move the stress back on to a syllable that had been originally unstressed and which might contain reduced vowels as in the examples quoted. The later changes of */i/* and */o/* are due to *i*-affection from a high front vowel in the following syllable and not from any kind of reduction. For cases of reduction of the vowels in final syllables after the accent shift, see 2.2.4 below.

2.2.2. Loss of final syllables

The nature of the Goidelic evidence (and especially the survival of inscriptions in Ogam) allows us to chart the gradual decay of final syllables alongside the preservation of a case-system and distinct declensional inflections. But in Brittonic it is easy to suppose that the loss of final syllables and the loss of case and declension are related; for discussion, see Jackson (1953: 618–633); Koch (1982–1983); Russell (2011a: 144–147).

While it is certainly difficult to plot the slow decay of case endings in Brittonic, this may in part be a problem of evidence, but partly too because in Brittonic the final syllables would have been in immediately post-tonic position, and it is conceivable that their reduction and loss might have happened quite rapidly after the stress shifted onto the penultimate syllable. Recent discussion (following Koch 1982–1983) has sought to distinguish loss of case and declension from the loss of final syllables, arguing that the Brittonic case system was already reduced in the number of cases (and perhaps also through the merger of declensional types) before final syllables were lost (for some of the possible processes involved, see Russell 2011a: 145–146, where it is suggested that, parallel to developments in later Irish, Brittonic might have been reduced to subject/object case vs. a genitive with other functions taken over by prepositions).

2.2.3. Syncope

Just as the effect of the stress accent was relatively mild in terms of vowel reductions and changes, so it was in relation to syncope. It is clear that composition vowels (the vowel separating the two elements of a compound) were lost very early irrespective of the position of the accent (Jackson 1953: 643–644; Sims-Williams 2003: 115–132, 2013). For example, while the Welsh personal name *Maelgwn* < *mayl-gun- < *maglo'kunos, shows the loss of a pretonic composition vowel, in *Cyndeyrn* < *kuno-ti'gernos, the composition vowel is lost even though it is not immediately pretonic (for an intermediate stage, cf. *Cundigeorn* [Durham *Liber Vitae*]). Only in certain literary contexts were composition vowels preserved (perhaps in some cases under the influence of the Latin version of the name [Russell 2004]), e.g. OW *Dinocat* < *du:no'katos, *Cunedda(g)* < *kuno'dagos). At a later stage, syncope seems to have happened sporadically, and clear-cut, systematic examples are hard to come by; beside the Middle Welsh personal name *Ceretic* we find a syncopated *Certic*, both from *Coro'ticos or Cara'ticos. Clearer examples come from Latin loanwords, though we have to be clear that the syncope is Brittonic and not Late Latin (examples of the latter usually involve syncope of the penultimate, e.g. 'musculus > W *mwsgl* 'muscle', 'calidus > W *call* 'clever', 'populus > W *pobl* 'people', 'solidus > W *swllt*, OC *sols*, B *saout* 'shilling'); clear examples of Brittonic syncope are ones where the outcome is different from that which would have been produced by Latin syncope, e.g. MW *pylgeint* 'dawn' < L *pullicantio* (where L ***pulcanti-* would have given ***pylcheint*), *bendith* 'blessing' < *benedict-* (not ***ben-nith* < L **benedict-*), *mellith* 'curse' < *maledict-* (not ***mellith* < L *maldict-*), etc. Generally, syncope, seems to have been resisted if the outcome were to produce an over-heavy consonant cluster; thus, *Gwrtheyrn*, OW *Guorthigirn*, OB *Gwrđiern* < *uorti'gernos (not ***Gwrthyrn* if, as might be expected, the -/i/- were to be syncopated), and we might compare the population-group name *Gwerthrynion* < *uortige'rn(i)ones where some form of syncope (and metathesis) probably did take place (cf. also the OE Bede's *Wyr-tzeorn* with OE initial stress beside Bede's *Uurtigermus*).

2.2.4. Accent shift

With the loss of final syllables, the originally penultimate accent was now final. Although in the above sections it has been suggested that the effects of the accent were

relatively mild, it is clear that the phonotactics of the Brittonic languages seem to have tended towards a penultimate stress. The precise processes of the shift are unclear (as they often are with accent shifts), but it is clear that in all the Brittonic languages the accent shifted back from the old penultimate to the new. Apart from the general evidence of the penultimate stress in the modern languages, specific evidence for the shift itself is to be found in two features. First, originally unstressed syllables (in which the vocalism was reduced in some way [for examples, see above 2.2.4]) ended up carrying the accent. Secondly, originally stressed syllables lost the accent and underwent a secondary vowel reduction, e.g. MW *marchauc* ‘horseman’ was originally stressed /ma'rxauɡ/ < *ma'rka:kos, hence the full diphthong on the final syllable, but the shift of the stress to the new penultimate led to the reduction of /au/ to /o/, thus later W *marchog* /'marxog/.

The date of the accent shift has been much debated (Jackson 1953: 682–689; Watkins 1972; Sims-Williams 1991: 79; and most recently Schrijver 1998–2000) with proposals for Welsh ranging from the eighth century to the twelfth. What we do know is that the shift occurred independently in Brittonic languages and may have worked through the languages in slightly different ways. However, the debate has received some recent focus by Schrijver’s demonstration (1998–2000) that the Old Welsh glosses on the Martianus Capella manuscript (Cambridge, Corpus Christi College MS 153) show consonantal spellings consistent with a post-shift situation; if so, then the change was at least underway by the late ninth century in Welsh, though it was not necessarily operating in the same timescale in Breton and Cornish (Jackson 1967: 79–81). The difficulty of course with dating such shifts is that they are only gradual; the Martianus Capella evidence consists mainly of single-word glosses and brief phrases, what would be less clear in the evidence is how the stress patterns might have changed in longer phrases and clauses. It is not unreasonable to suppose a long period of variation such as in the pronunciation of English *controversy*, *kilometer*, etc. The one exception to the accent shift was the Vannetais dialect of Breton which seems to have resisted the full-scale shift, e.g. *kloom* ‘dove’ (beside *kolom* /'kolom/ in other dialects) < *columba* (Jackson 1967: 79–84); the causes have been debated but a significant factor certainly must have been the geographical adjacency of French with its final stress (Schrijver 1995: 16–22, 2011b: 386–7).

3. The evolution of regularization

The chapters in this volume provide numerous examples from many different periods of the regularization of morphological categories and the eradication of irregularity. For example, in the earliest stages of the IE languages we can detect the analogical spread of thematization in nominal and verbal paradigms, at the expense of athematic formations, as a device for keeping stems and endings distinct. Such developments were particularly common in the most paradigmatic parts of the language which would have been subject to general analogical and levelling pressures, and less so, for example, in adverbs which typically preserve archaisms because they have become isolated in the system. The Celtic languages were no exception to such tendencies especially within nominal and verbal paradigms.

3.1. The nominal paradigm

As was noted above (2.1.2, 2.2.2), the probable effect of the stress accent was the reduction and eventual loss of final syllables in the Insular Celtic languages. Despite this, the Goidelic languages maintained a case system, though Brittonic languages did not. In both language groups, however, we can observe a progressive regularization of the nominal system which gradually prioritizes number-marking over case-marking.

3.1.1. Goidelic

Celtic languages inherited a full-blown case and declensional system as is reconstructible from the other IE languages. In the insular Celtic languages the number of cases was reduced to five: nominative, vocative, accusative, genitive, and dative; in terms of number, a singular, dual, and plural was preserved, though the dual was lost in the earliest phases of Common Brittonic and within early Old Irish. Likewise preserved were patterns of declension reflecting the different stem formations. Thus it is conventional (e.g. Thurneysen 1946: 176–217) to distinguish vowel-stem declensions (*-a-*, *-o-*, *-iō-*, *-ia-*) from consonant stems (*-i-*, *-u-*, *-s-*, *-t-*, *-d-*, *-n-*, etc.), even if the description of *a-*, *o-*, *iō-*, *ia-*, and *s-*stems is anachronistic, as none of these formants is visible (except in the Continental Celtic languages). For consonant-stem declensions, the stem-marker was usually lost in the nominative singular but generally preserved in the oblique cases (much as in other languages), e.g. OIr. *carae* ‘friend’ < *kare/ant-s, gen. sg. *carat* < *kare/ant-as (< IE *-os), and nom. pl. *carait* < *kare/ant-es; *cing* ‘hero’ < *kinget-s, gen. sg. *cingid* < *kinget-as (cf. Gaul. *Cingeto-*), etc. Case distinctions were maintained in Old Irish and onwards through a combination of distinctions between palatal and non-palatal final consonants (determined originally by the vowel in the apocoped syllable: *-Ci and *-Ce > -Cⁱ beside *-Ca, *-Co and *-Cu > -C), and by the nature of the mutation caused both to the next word (if in close combination with it, e.g. an adjective or a genitive), and to its own initial consonant if the preceding word, e.g. an article, was in close combination; thus, OIr. *fer* ‘man’ (< *uīras < *uīros) : *fer n-álaind* ‘a handsome man’ (acc. sg.) (< *uīran < *-on) : *fir bháin* ‘of a fair man’ (< *uīri ba:ni); *in brat* ‘the cloak’ < *sindos brattas (< *-os) : *in mrat* ‘the cloak’ (acc. sg.) < *sindon brattan (< *-on) : *ind bhrait* ‘of the cloak’ < *sindi bratti, etc.

Already, however, in Old Irish it is possible to observe blurring of some of these distinctions and the analogical spread of the predominant declensional patterns (cf. Thurneysen 1946: 196–197). This is most noticeable in the early period in those declensions where the distinctive consonantal marker was lost, such as in *i-*, *u-*, and *s-*stems; for example, the genitive singular of both *i-* and *u-*stems is *-o* (later also *-a*), e.g. *i*-stem *súlo* ‘eye’s’ and *u*-stem *crotho* ‘shape’s’. This is more easily explained by assuming that the *u*-stem *-oys (Ogam *-OS*) spread to the *i*-stem declension than by trying to reconstruct an *i*-stem gen. sg. in *-ojs (see McCone 1994: 116–17, 2005: 123–5). Elsewhere the distinction between *i-* and *u-*stems simply results in a distinction between palatal final consonants and front vowels in *i*-stems and non-palatal consonants and back vowels in *u*-stems, e.g. nom. sg. *u*-stem *cruth* ‘shape’ : *i*-stem *súil* ‘eye’, nom. pl. *crothae* : *súili*, acc. pl. *cruthu* : *súili*, etc. Further blurring of declensional patterns occurred when *u-*

stems gradually lost their distinctive *u*-coloured vowel; thus beside *cruth* < *k^writu- (W *pryd*) : gen. sg. *crotho* (earlier *cretho*), forms like *tomus* ‘measuring’ < *to-med-tus and *fiuss* ‘knowledge’ < *uid-tus, began to be treated as *o*-stems with genitive singular in *tomuis* and *fiss* beside the regular *toimseo* and *fesso*.

More substantive changes happened to these declensional patterns in late Old Irish and Middle Irish when (as noted above 2.1.1) the quality of vowels in final unaccented syllables began to blur, first by the merging respectively of *-o/* and *-u/* and of *-e/* and *-i/*, and then eventually of all of them to *-ə/*. This was particularly problematic for the original *-io-* and *-ia-* stems where many of the distinctions between cases were carried by the final vowels (Greene 1974; Hickey 1985). The response to this potential confusion is illuminating as it appears that the prime concern seems to have been to preserve the distinction between singular and plural even at the expense of less clear distinctions between cases; hence the spread of clear plural markers from the consonant stems, and especially in Irish the use of a dental stem ending in the nominative and accusative plural *-eda* *-eðə/*, e.g. OIr. nom. sg. *slige* ‘path’ : nom. pl. *sligi* : MIr. *sligeda*, OIr. *gille* ‘lad’ : *gilli* : MIr. *gilleda*, etc. However, this restructuring (which effectively speeds up the breakdown of declensional patterns) was only briefly successful since by the twelfth century intervocalic *-ð/-* had fallen together with *-ɣ/-* (lenited *-g/-*) to give non-palatal *-ɣ/-* but palatal *-j/-*. As testimony to an ongoing pre-occupation with number it is also worth noting that the reflexes of this ending survived into the modern language as *-aí*, but several dialects, including Connacht, developed more complex nominative/accusative plural markers such as *-acha(i)* and *-anna(i)* (Ó Siadhail 1989: 164).

3.1.2. Brittonic

The loss of case distinctions in all the Brittonic languages left them in a rather different position from the earlier Goidelic languages in that the crucial distinction to maintain was that of number. At the point where the Brittonic languages were separating out after the loss of final syllables, two types of plural distinction were possible: first, original *o*-stems had a nominative plural in **-i* (< **-oi*) which caused raising of the preceding vowel, e.g. W *gŵr* ‘man’ < *uīros : pl. *gwyr* < *uīri, *bard* ‘poet’: pl. *beirdd*, *twrch* ‘boar’ : pl. *tyrch*, etc. Secondly, consonantal stem declensional forms were usually non-parasyllabic and so the plural forms were marked by an extra syllable, in origin the stem-marker; Brittonic plural forms could therefore be marked by a range of syllabic markers, e.g. MW *-eu*, MnW *-au*, CB *-ou* < **-oues* (*u*-stem), WCB *-(i)on* < **-ones* (*n*-stem [cf. Stüber 1998: 29–30]), WCB *-ant* (*-nt*-stems), etc. (Schrijver 2011a: 41–44). Some of these endings, especially when preceded by *-j/-*, could also cause vowel affection, e.g. W *mab* ‘son’ : pl. *meibion*. Another source of plural markers was from phonological developments to inherited *io-*, *ia-*, and *s*-stems within Brittonic; the reflexes of these plural endings produced a series of endings containing *-ð/*, e.g. W *-oedd* < **-eja* (and also **-esa*, which fell together with it), *-edd* < **-ejes*, etc.

However, even in the earliest phases of the attested languages it is very difficult to match the plural forms of various nouns with their historically original stem formation; MW *cad* : pl. *cadeu* (CB *cadou*), *car* : pl. *carant* are rare examples of continuity. What is clear is that there was a major regularization and analogical spread of particular mark-

ers far beyond their historical extent. Thus, plurals in W *-eu/-au*, CB *-ou* (original *u*-stems) and in *-(i)on* (original *n*-stems) became very productive at the expense of other well-marked suffixes. It is clear from the patterns of Middle Welsh that there was no single one-to-one replacement of one marker with another; many nouns could form their plural with a range of markers and suffixes, e.g. *angel* ‘angel’ : pl. *engyl*, *englylon*; *assen* ‘donkey’ : pl. *essyn*, *assenoed*; *boly* ‘sack’ : pl. *byly*, *bolyeu*; *cerd* ‘song’ : pl. *kyrd*, *kerdeu*; *mor* ‘sea’ : pl. *myr*, *moroed*; *nant* ‘valley, stream’ : *neint*, *nanheu* (MnW *nentydd*), etc. In one particular case of a plural suffix with a very narrow semantic range we can observe the spread of the marker in action: W *-od* (probably in origin extracted from *llygod* ‘mice’) is productive as a plural marker of the names of animals; but it was not nearly so productive (if at all) in the medieval period, e.g. *cath* ‘cat’ : pl. MW *catheu*, MnW *cathod*; *baed* ‘boar’ : pl. MW *baid*, MnW *baeddod*; *draenog* ‘hedgehog’ : pl. MW *draenogyon*, MnW *draenogod*; *eliffant* ‘elephant’ : pl. MW *eliffantieit*, MnW *eliffantod* (new loanwords also take the *-od* marker, e.g. *teigrod* ‘tiger’); and the suffix has even spread to a few names of nationalities: *Gwyddelod* ‘Irish people’, *Ffrancod* ‘French people’.

With the generalization of some plural markers, there were inevitably some casualties. One particularly interesting one is W *-awr* (which probably spread from the Latin ending *-ārium* or *-āria*). In early Welsh it is already restricted to verse with only *gwaywawr* ‘spears’ (: sg. *gwayw*) being frequent in prose, and it does not survive into the modern period (Nurmio 2014). It is a useful reminder of how a wider range of endings might have been preserved for purposes of rhyme and syllable-counting in verse long after disappearing from prose.

While the spread of agglutinative, suffixed plural markers is unsurprising, it is noteworthy that plurals marked by vowel affection also spread beyond their original domain of *o*-stem nouns. L *castellum* (pl. *castella*) was borrowed into Welsh as *castell*, but the plural *cestyll* is probably to be understood as an analogical form based on *o*-stem nouns with pre-suffixal vowels in *-a-* and *-e-*, though it is not impossible that it reflects a Late Latin plural *castelli*. A case which cannot be other than analogical is *cerryg* ‘stones’ (: sg. *carreg*); it is a feminine noun (< **karrika*) and so the *i*-affected plural has to be secondary; the one Old Welsh attestation of a plural *carrecou* is striking in this respect as it is arguably a regular plural formed by a non-native speaker of Welsh (Russell 2012: 206–214).

Although broadly the same plural markers are found in all the Brittonic languages, Breton plural formation is particularly complex and innovative in its productivity (Trépos 1957; Acquaviva 2008: 234–265). One facet of that complexity is that it allows for plurals to be formed on singulatives in a way that cannot happen in Welsh (a handful of cases in the Old Welsh glosses may be due to Breton influence in the glossing process). In Welsh collective nouns, e.g. *adar* ‘birds’, form a suffixed singulative to refer to a single item of that collection, thus *ederyn* ‘bird’. In Breton, however, it is possible to create plurals from singulatives, e.g. *ster* ‘stars’ (cf. W *sêr*) : *steren* ‘star’ (cf. W *seren*) : *sterennou* ‘stars’, etc.

One final point is worth noting. The fact that the same plural markers, OW *-ou*, MW *-eu*, MnW *-au*, CB *-ou*, and WCB *-(i)on* have become productive in the later stages of the Brittonic languages (and this is clear not only from Welsh and Breton but also in what survives in Cornish) suggests that this process of regularization began in Common Brittonic before the split up of the languages, since apocope is usually regarded as the

watershed. The process, then, would have been rather different from a later generalization of certain endings (which also occurred, cf. *-od* above); the earlier process would have involved the morphological resegmentation of stems and endings. How that might have worked is unclear and needs more work (for some preliminary thoughts, cf. Koch 1982–1983).

3.2. The verbal paradigm

The complexities of the Old Irish and probably (on the basis of the Brittonic evidence) the Common Celtic verbal system has generated a huge amount of scholarly literature (for recent discussions [which include full surveys of earlier scholarship], see McCone 2006; Rodway 2013). In particular the so-called absolute-conjunct, deuterotonic-prototonic distinctions (for which see above 2.1.4) introduced a high level of variation (and arguably redundancy) into the system which was to a certain extent gradually leveled out in the later languages, completely so in the Brittonic languages but only partially in the Goidelic languages. A further process of regularization involved the creation of new verbal paradigms based on nouns and adjectives. The productive categories of verbal formation in most languages tend to be denominative and deadjectival, and the Celtic languages are no exception.

3.2.1. Goidelic

The accentual alternations in the Old Irish verb have been discussed in 2.1.4 above. From late Old Irish onwards, and especially during the Middle Irish period, a wholesale restructuring and adjustment of this system took place (for discussion, McCone 1997: 163–241). Some of the reasons for this have already been alluded to: the reduction of all unaccented vowels to /ə/ had the effect that distinctions between various verbal endings were lost (e.g. 2 sg. *marbai* ‘you kill’, 3 sg. *marba* ‘kills’, 3 sg subj. *marbae*, etc. were all pronounced /marvə/ [McCone 1997: 205]) as well as distinctions between the different infixed pronouns (e.g., the preverb in *ro·gab*, *ra·gab*, *ra·ngab* was pronounced /rə/ [McCone 1997: 169]). Already in Old Irish the use of stressed independent object pronouns was beginning to obviate the need for infixed pronouns and the loss of the neuter gender gave rise to the petrification of preverbs containing neuter infixed pronouns, hence the rise of forms like leniting *at·* (McCone 1997: 172–173) and more generally the development of lenition in compound verbs in main clauses (where in Old Irish they had been unlenited). Given that one of the functions of deuterotonic verbs was to carry infixed pronouns (in addition to the important function of marking relative clauses, etc.), the reduction in their use led to a drop in frequency of deuterotonic verbs. Cumulatively we see the rise of a generalized verbal stem in original compound verbs often based on the statistically more frequent prototonic stem: thus, OIr. *do·sluindi* ‘denies’ (: prototonic *·diltai*) was gradually replaced by forms of a new simple verb *diltaid* (McCone 1997: 207–209); *do·léici* ‘casts’ (: *·teilci*) by *teilcid*, etc. While for some verbs the prototonic stem was the basis for generalization, in others the verbal noun provided the base, e.g. OIr. *do·fich* ‘avenges’ (: *·dich*) : verbal noun *dígal* yielding MIr. *díglaid*

(McCone 1997: 193), though in many cases it is difficult to tell the difference between the source of the generalized stem. Another regularizing process in Middle Irish was the creation of a standard, all-purpose 3 sg. ending *-(e)ann* for all verbs which was derived from forms such as *·éireann* ‘sells’ (prototonic form of *as·ren*); this was generalized in Irish but not in Manx or Scottish Gaelic (McCone 1997: 205–208).

Given the massively increased frequency of weak verbs with their regular and predictable pattern of tense forms (*s*-preterite, *f*-future, *a*-subjunctive), the various categories of strong verbs (many inherited from IE) with their irregular (in Irish terms) morphological patterns (reduplicated preterites, reduplicated *s*-futures, long vowel preterites, etc.) became increasingly restricted to unproductive categories of irregular verbs. This kind of shifting of categories is already detectable in Old Irish, where we see *s*-preterite markers creeping into strong verbs, e.g. *·arroímsat* ‘(who) have accepted’ for *·arroítatar*, and the spread of *f*-futures, e.g. *benaid* ‘strikes’ : future *·bi* replaced by *benfa* (McCone 1997: 209–210). One class of verbs that was particularly vulnerable to this kind of shift was the so-called “hiatus” verbs (McCone 1997: 24–25). In many respects their development is analogous to that of *i*- and *u*-stem nouns (3.1.1 above) in that the stem-final consonants of these verb (*-u/-*, *-i/-*, or *-s/-*, e.g. *soíd* ‘turns’ [*< *sou-*], *gniíd* ‘does’ [*< *gniġ-*], *cíid* ‘weeps’ [*< *k^weis-*]) were lost very early in the history of Irish, and so they gradually shifted from strong-verb inflection (with a stem-final consonant) to weak-verb (with a stem-final vowel). The process was very slow and made up of numerous small changes tending in the same direction, that of weak verb inflection.

In many languages another way of regularizing the verbal paradigm is to replace (or at least supplement) primary verbs with verbs based on nouns and adjectives in the sense of ‘be X’ (where X = noun or adjective), ‘become X’, ‘make something X’, etc. Many such verbs are found in Old Irish, e.g. *marb* ‘dead’ : *marbaid* ‘kills’ (lit. ‘makes something dead’), but one particularly striking type, which must have arisen in Common Celtic, as it is also found in Brittonic and in Gaulish, involves creating verbs with a suffix which is in OIr. *-aigidir* and in Brittonic languages *-ha-* *< *sag-* ‘seek’ (cf. OIr. *saigid*, W *haeddu*, etc.) (Joseph 1987). The semantics suggest that in origin the first element of such compounds was nominal, ‘seeking X’; that sense is still preserved in a group of Middle Welsh verbs, e.g. *cardotau* ‘seek alms’ (*< cardod* ‘alms’ + *-ha-*), *gwreica* ‘seek a wife’ (*< gwreig* ‘wife’) (the *-h-* [*< *-s-*] devoicing the preceding voiced stop). However, at a relatively early period, particularly in Irish and then later in Brittonic, the sense ‘seek’ was bleached out to the extent that the ending could be added to both nouns and adjectives, thus OIr. *béo* ‘living’ : *·béoigedar* ‘makes alive, vivifies’, *ainm* ‘name’ : *·ainmigedar* ‘names’, W *ysgafn* ‘light’ (of weight) : OW *scamnehgint* ‘they lighten’, etc. Corresponding nominal forms (originally *< *sagio-*) are also attested, e.g. MW *cynutai* ‘one who seeks kindling’, the Gaulish personal name *Curmisagios* lit. ‘beer-seeker’, and perhaps OIr. *cennaige* ‘merchant’ (*< *k^wenno-sagio-* ‘bargain-hunter’) (Uhlich 2002: 417–418).

3.2.2. Brittonic

The Brittonic languages preserve sufficient traces of patterns parallel to those found in Old Irish for it to be clear that Common Brittonic inherited an absolute-conjunct pattern

of inflection similar to that in Irish (Rodway 2013: 85–116). In addition to the infixing of pronouns, both direct object, e.g. MW *pan y'th weleis gyntaf* 'when I first saw you', *minheu a'e kymeraf* 'I shall accept it', and indirect object, e.g. *y gwr a'm rodes y gwin* 'the man gave me the wine', there are some traces of the *berid* : *beir* type of alternation in some quasi-proverbial expressions, e.g. *trenghit golut, ni threingk molut* 'wealth perishes, praise does not perish' (Schumacher 2011: 170–175). Similar third person forms in *-it* or *-yt* are also attested in Old Welsh and Old Breton, e.g. OW *prinit* 'buys', OB *doguolouit* (Fleuriot 1964: 300), and exploited in medieval Welsh verse as a way of providing an extra syllable. What we do not know for certain is how much of the complexity of the Old Irish patterns were Common Celtic in origin, and so inherited into Brittonic, and how much was innovated in Goidelic. However, other features, such as preservation of the suffixed relative particle in MW *yssyd* < **esti-jo* which is cognate with Old Irish *asa* (cf. also OIr. *bertae* 'whom they carry, who [3 pl.] carry' < **beront-jo*), suggest that the basics of the system were inherited into Brittonic.

If so, it clearly underwent considerable remodeling which was in many cases not so different from what occurred in Middle Irish. For example, apart from a very few surviving instances of reduplicated preterites, e.g. *cigleu* 'I/he heard', *adwaen* 'he knows' (< **ate-gegna*), *s*-preterites dominate the scene in Brittonic (Schumacher 2011: 163–165). Even though within Welsh a productive non-sigmatic preterite 3 sg. ending, MW *-awd*, MnW *-odd* arose, *s*-markers remained productive elsewhere in the Welsh paradigm and everywhere in the Breton and Cornish preterite (Rodway 2013: 154–165). Stem classes seem to have effectively collapsed with hints of the original distinctions in certain endings, e.g. 3 sg. preterite *-as*, *-es*, *-is*, *-wys*, 1 sg. present *-af*, but also *-if*, etc. Broadly, the complexity of the strong verb paradigms attested in Old Irish is not to be found in Brittonic. Furthermore, strikingly absent from Brittonic languages is a clearly and simply marked future tense. Old Irish preserved a system of *s*-markers, reduplication (and long stem-vowels deriving from an earlier reduplication), and an *f*-marker in weak verbs. While Middle Welsh has preserved forms in *-hawt* (which may reflect an original **-sa:/-*) in verbs which can be interpreted as futures (Schumacher 1995; Isaac 2004), these are lost in the later language where essentially the old synthetic presents, e.g. MW *caraf* 'I love', are used to refer to the future; the present is marked by a periphrastic *yr ydwyf yn caru* 'I love' (lit. 'I am loving') (Schumacher 2011: 211–213). That there was no clearly marked future inherited from Common Celtic is indicated by the fact that Breton and Cornish mark futures in different ways. While in Welsh the original synthetic present was taken over for future usage, in Breton the present subjunctive performed a similar function (Schrijver 2011b: 402); such usage can occasionally be found in Welsh but generally in the medieval language subjunctives are used in contexts of indefiniteness. In Middle Cornish present forms can also refer to the future but, probably on the model of English *will*, Cornish forms a future using *myn* 'wants' with a verbal noun, e.g. *ny a vyn formye an bys* 'we shall create the world' (lit. 'we want to form the world') (Williams 2011: 330). The innovatory forms of the future tense attested in Brittonic languages raise an awkward question about the propriety of reconstructing future tenses for Celtic. While the Old Irish *f*-future has long been recognized as an Irish innovation, it has been argued that the reduplicated *s*-future, e.g. *gigsea* 'I shall pray', is an IE inheritance with cognates adduced from Vedic, etc. (McCone 1991: 137–182). But if such formations are only attested in Goidelic, and neither in Brittonic nor in Continental Celtic, a doubt should arise as to whether they are inherited. Another way of thinking

about these forms is that they were not inherited as futures but perhaps as a desiderative, but that does not explain why they are not detectable in Brittonic. It might be argued that the reduplication was not preserved in order to maintain a clear distinction from reduplicated preterites, and the *s*-suffix tended to disappear or merge into the stem-final consonants. Whatever may be the case, it is important to take the Brittonic evidence, or the absence of it, into account when reconstructing verbal categories.

A final example of the regularization which is unique to Brittonic may be considered. All the Insular Celtic languages show compounds of the verb 'to be' using the same preverbs as are used with all other verbs, e.g. W *dyfod* 'come', C *devos*, B *devout* (< *tu-but-), W *gorfod* 'overcome', C *gorfos*, B *gorvout* (< *uor-but-; cf. OIr. *for-bí* 'remains' [consuetudinal present to *for-tá*]). In addition, all Brittonic languages have a different type of 'be'-compound where in some tenses of the pair of 'knowing' verbs the stem is suffixed by parts of the verb 'to be', e.g. W *gwybod*, MC *gothvos*, B *gouzvout* < *uid-but-; W *adnabod*, C *aswonvos*, B *aznavout* < *ate-gna-but- (Schumacher 2011: 196–198; Schrijver 2011b: 407–410). Both verbs have inherited irregular paradigms in the present and imperfect indicatives, e.g. W *gwyr* 'knows', *gwyd(y)at* 's/he knew', but beyond these tenses the tense and mood markers are carried by the appended verb 'to be'. Given the notorious irregularities associated with paradigms of verbs of knowing in many languages, the use of 'to be' to carry the paradigm is best interpreted as a shift towards regularization of the paradigm. In terms of the present discussion, however, what is particularly interesting is how, to a certain extent in Welsh but more so in Cornish (e.g. *wharfos* 'happen' [cf. MW *chweir*], *tal-* 'pay'; note also the number of verbs where a 3rd sg. future was made in this way, e.g. *care* 'love': *caruyth*, *clewas* 'hear': *clewuyth* [cf. W *clybot* 'hear' beside the usual *clywed*], *gweles* 'see': *guylvuyth*), and even more so in Breton (e.g. *hoaruout*, *c'hoarvezout* 'happen'; *talouout*, *talvezout* 'be worth'; *deurvout*, *teurvezout* 'wish'; *eme* 'say'; *falvezout* 'lack, desire'; *dleout* 'owe, have to') the use of a conjugating suffixal 'to be' spread to other verbs, especially those expressing a state and, above all, to verbs which otherwise would have had irregular conjugations.

4. Language contact and the evolution of the Celtic languages

Throughout their documented history the Celtic languages have been spoken in ongoing contact with speakers of other languages, whether of Etruscan and Latin in northern Italy, of Greek and Latin (and perhaps other Italic languages) in Gaul, of Greek, Latin, Iberian and Punic (and probably others) in Spain, of Latin and Germanic languages in Britain and Ireland, etc. The most visible effect of this contact has been on the lexicon (Wodtko, *The lexicon of Celtic*, this handbook: 2). Thus collections and discussions of Latin loanwords in Irish, Welsh, and Breton (Vendryes 1902; Lewis 1943; Haarmann 1970, 1973; McManus 1983, 1984a); English loanwords in Welsh (Parry-Williams 1923); French loanwords in Breton (Piette 1973); etc. are staples of the philological discussion of the Celtic languages. Conversely, too, Celtic loanwords in other languages, e.g. French, Latin, English and other Germanic languages, and Irish in Hiberno-English have their place in the historical discussion of those languages. The traditional focus on the lexicon, however, has served to obscure other kinds of contact phenomena. In recent years a greater awareness of these other aspects of language contact have begun to

impinge on the work of Celtic linguists, and the results could potentially be very important for our understanding of the evolution of the Celtic languages. Essentially, features which may have been previously assumed to be genetic (and therefore diagnostic of a genetic relationship) may turn out to be the result of contact; for example, we may take the case of the possible “Gallo-Brittonic” node in the family tree of Celtic which has been much debated in recent years (Koch 1992; Schrijver 2007b; Sims-Williams 2007a). On the other hand, it has been argued (Schrijver 1995: 463–465) that the evidence may be better explained as contact between speakers of Gaulish and Brittonic as the Roman Empire expanded through Gaul into Britain (not least if some of the legions taken to Britain were recruited in Gaul and Spain). Recent work by Schrijver (2005, 2014: 122–157) has further suggested that there may have been substrate influence from British and Gaulish on the phonology of Northern French and Dutch. Furthermore, it has been suggested that some of the differences between the Brittonic languages may reflect varying degrees of influence from Latin and possible changes in the relationship between British and Latin during the Roman and post-Roman periods (Schrijver 2002, 2007a, 2009; cf. also Russell 2011a). What follows therefore in this section is not so much a survey or discussion of work completed and ideas fixed as an exploration of the potential for innovative work.

One particular area of interest has been the application of scholarship on language contact (both multi- and bilingualism) to ancient and medieval languages. One consequence is that there has been a shift from a crude sub-/superstrate view of language contact towards an appreciation of the complexities and messiness of real-language contexts which must have obtained in the past as well (for a good introduction, see Mullen 2012). In essence, there are instances (and Roman and post-Roman Britain may be one) where it may be more fruitful to think in terms of long periods of balanced and functional bilingualism with influences moving in both directions between the languages in question (Russell 2011a: 142–143). Similarly, in the early medieval period we can get glimpses of the ways in which speakers of Celtic languages influenced their spoken Latin beyond just the borrowing of lexical items. We may instance the use of *L contra* after verbs of speaking (i.e. *legere contra* ‘speak to’) which is modeled on the use of OIr. *fri*, MW *wrth*, etc. ‘against’ after verbs of speaking in Celtic languages; and the use of ordinal numbers to mean ‘one of *n*’ rather than simply ‘*nth*’. Such features are sometimes detectable both in Hiberno-Latin and in Latin from early medieval Britain in authors such as Asser, a Welshman who wrote a life of King Alfred, though other writers, such as Gildas, were too polished in their Latin to give anything away. The early medieval Christian inscriptions from western Britain seem to provide evidence for spoken Latin and the changes it may have undergone. Although their significance has been debated, their testimony has been held to be useful (Adams 2007: 616–620; Charles-Edwards 1995: 715–720, 2013: 95–115). For the earlier, Roman and immediately post-Roman period, there is relatively little evidence to work with; the Vindolanda tablets are potentially useful but, given the mix of nationalities in the area of Hadrian’s Wall at that period, its utility may be limited beyond recording the occasional borrowing (Adams 1992, 1995; Russell 2006, 2011b). Likewise, the Bath curse tablets offer tantalizing glimpses, though it is difficult to assess the significance of what we find in them (Mullen 2007a, 2007b).

4.1. Lexicon

As noted, loanwords are perhaps the most obvious effect of contact and collections of such words abound. Interesting though they are, such collections tend to be under-exploited. Generally, most work has tended to concentrate on the phonology and how the loanwords might help us to understand the phonological developments in the Celtic languages themselves. The advantage of loanwords is that we think we know their starting points rather better than those of the reconstructions posited for native words. For example, the borrowing into Irish at different periods of words containing Latin /p/ has formed a significant plank in our understanding of the development of /k^w/ and the rise of a new /p/ in Insular Celtic; cf. early examples, such as L *puteus* ‘well’ > OIr. *cuithe*, *pallium* ‘veil’ > *caille*, beside later examples, e.g. *pācem* ‘peace’ > *póc* ‘kiss’, etc. (McManus 1983: 21–27); likewise, the Latin suffix *-ārius* was borrowed early as *-aire* (before the shortening of unaccented vowels [see 2.1.1]), e.g. *notaire* ‘professional scribe’ < *notārius* and later as *-óir*, e.g. *caindleóir* < *candelārius* (McManus 1983: 32–33). However, we cannot always be certain whether a given word was borrowed from Classical Latin or from a later, more colloquial variety (for the debate about the status of Latin loanwords in Brittonic, see McManus 1984b for a useful discussion).

4.2.1. Nominal morphology

There is more to the borrowed lexicon than just loanwords. For example, one neglected area of study has to do with the degree of penetration of loanwords into a language: how far were they embedded in the language? One way of testing this is by considering borrowed derivatives: does the suffix spread beyond the core group of loanwords and, if so, how far? For example, in Middle Welsh it is clear that the adjectival suffix *-awl* (later *-ol*), which almost certainly derives from L *-ālis*, spread far beyond its original Latinate domain to become the default productive suffix in modern Welsh (Russell 1990: 125–131). On the other hand, *-gl* (< L *-Vculu-*), which forms nominal derivatives in Latin, barely seems to have spread beyond its core group, e.g. *ffagl* ‘torch’ < *facula*, *magl* ‘stain’ < *macula*, *caregl* ‘cup, chalice’ < *calicium*, *ffenigl* ‘fennel’ < *fenicula*, *dysgl* ‘dish’ < *disculum*; and it is questionable whether we should think of it as a suffix at all in Welsh. Reflexes of L *-illus/-illa*, however, e.g., *porchell* ‘piglet’, *padell* ‘dish’, *gradell* ‘griddle’ < *cratella* (BL *gratella*), etc., also as diminutives and terms for tools, e.g. *gwaell* ‘awl’, *bachell* ‘corner, nook’, *rhathell* ‘rasp’, sit between the two because of their productive extension to small semantic niches. Much more work is necessary before we understand the penetrability of loanword morphology into the Brittonic languages.

4.2.2. Verbal morphology

One striking example of Latin influence on verbal morphology has been the development of a pluperfect tense in Brittonic. Unlike Goidelic, all Brittonic languages have such a tense, e.g. MW *carassei*, MB *carse* ‘he had loved’. It has been argued, probably rightly, that this tense arose on the model of the Latin pluperfect (Mac Cana 1976; but cf. Russell

2011a: 150–153) and, since its structure (*s*-preterite stem and secondary [imperfect] endings) is the same in all Brittonic languages, it almost certainly arose in Common Brittonic. Although Mac Cana (1976) argued that it was modelled on pluperfect indicative forms such as *amaverat*, Russell (2011a: 150–153) suggested that a more straightforward model would be a subjunctive such as *ama(vi)sset*. The forms seem to have arisen in the spoken language, and in Late Latin indicative forms like *amaverat* had already been replaced by periphrastic formations. On the other hand, the modal *ama(vi)sset* survived to give imperfect subjunctives in the Romance languages. A telling point in favor of the latter proposal is that in Middle Cornish a pluperfect form, e.g. *carse*, is treated as modal, i.e. ‘he would have loved’, unless prefixed by the perfective particle *re* (cf. Wry, OIr. *ro*), e.g. *re garse* ‘he had loved’. In other words, the unmarked function of these forms seems to be modal, which would fit with a construction modeled on the Latin pluperfect subjunctive.

5. Abbreviations

B	Breton	MnB	Modern Breton
BL	British Latin	MW	Middle Welsh
C	Cornish	OB	Old Breton
CC	Common Celtic	OC	Old Cornish
IE	Indo-European	OIr.	Old Irish
L	Latin	OW	Old Welsh
MB	Middle Breton	Skt.	Sanskrit
MC	Middle Cornish	W	Welsh
MIr.	Middle Irish		

6. References

- Acquaviva, Paolo
 2008 *Lexical plurals. A Morphosemantic Approach*. (Oxford Studies in Theoretical Linguistics 19). Oxford: Oxford University Press.
- Adams, James N.
 1992 British Latin: Notes on the language, text and interpretation of the Bath curse tablets. *Britannia* 23: 1–26.
- Adams, James N.
 1995 The language of the Vindolanda writing-tablets: an interim report. *Journal of Roman Studies* 85: 86–134.
- Adams, James N.
 2007 *The Regional Diversification of Latin, 200 BC–AD 600*. Cambridge: Cambridge University Press.
- de Bernardo-Stempel, Patrizia
 1991 Archaisch Irisch *maccu* als morphologisches Relikt. *Historische Sprachforschung* 104: 205–223.
- de Bernardo-Stempel, Patrizia
 1995 Gaulish accentuation. Results and outlook. In: Joseph F. Eska, R. Geraint Gruffydd, and Nicolas Jacobs (eds.), *Hispano-Gallo-Brittonica. Essays in Honour of Professor D. Ellis*

- Evans on the Occasion of his Sixty-fifth Birthday*. Cardiff: University of Wales Press, 16–32.
- Breatnach, Liam
1994 An Mheán-Ghaeilge [Middle Irish]. In: McCone et al. (eds.), 221–333.
- Byrne, Francis John
1994–5 *Derco*: The female form of *Moccu*. *Éigse* 28 : 42–70.
- Charles-Edwards, Thomas
1995 Language and society among the Insular Celts, AD 400–1000. In: Miranda Green (ed.), *The Celtic World*. London: Routledge, 703–736.
- Charles-Edwards, Thomas
2013 *Wales and the Britons*, 350–1064. (The Oxford History of Wales, vol. 1). Oxford: Oxford University Press.
- Doyle, Aidan
1992 *Noun Derivation in Modern Irish. Selected Categories, Rules and Suffixes*. Lublin: Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego.
- Eska, Joseph F.
2007 On the prehistory of OIr. *do-*, MW *dy-*, etc. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 4: 185–206.
- Falc'hun, François
1981 *Perspectives nouvelles sur l'histoire de la langue bretonne*. Paris: Union Générale d'Éditions.
- Fleuriot, Léon
1964 *Le vieux breton: éléments d'une grammaire*. Paris: Klincksieck.
- Greene, David
1974 Distinctive plural forms in Old and Middle Irish. *Ériu* 25: 190–199.
- Haarmann, Harald
1970 *Der lateinische Lehnwortschatz im Kymrischen*. Romanistische Versuche und Vorarbeiten. Bonn: Romanisches Seminar der Universität.
- Haarmann, Harald
1973 *Der lateinische Lehnwortschatz im Bretonischen*. Hamburger philologische Studien. Hamburg: Buske.
- Hickey, Raymond
1985 Reduction of allomorphy and the plural in Irish. *Ériu* 36: 143–162.
- Isaac, Graham R.
2004 The Old- and early Middle-Welsh 'future' tense: form and function of a moribund category. *Journal of Celtic Studies* 4: 153–170.
- Jackson, Kenneth H.
1953 *Language and History in Early Britain: a Chronological Survey of the Brittonic Languages First to Twelfth Century A.D.* Edinburgh: Edinburgh University Press.
- Jackson, Kenneth H.
1967 *A Historical Phonology of Breton*. Dublin: Dublin Institute for Advanced Studies.
- Joseph, Lionel
1987 The origin of the Celtic denominatives in **sag-*. In: Calvert Watkins (ed.), *Studies in Memory of Warren Cowgill (1929–1985). Papers from the Fourth East Coast Indo-European Conference, Cornell University, June 6–9, 1985*. Berlin: De Gruyter, 113–159.
- Kneen, John J.
1937 *The Personal Names of the Isle of Man*, London, Oxford University Press.
- Koch, John T.
1982–1983 The loss of final syllables and loss of declension in Brittonic. *Bulletin of the Board of Celtic Studies* 30: 201–233.

- Koch, John T.
 1992 'Gallo-Brittonic' vs. 'Insular Celtic': The inter-relationships of the Celtic languages reconsidered. In: Gwennolé Le Menn (ed.), *Bretagne et pays celtique. Langues, histoire, civilisation. Mélanges offerts à la mémoire de Léon Fleuriot*. Saint-Brieuc: Presses Universitaires Rennes, 471–495.
- Lambert, Pierre-Yves
 2003 *La langue gauloise. Description linguistique, commentaire d'inscriptions choisies*. 2nd edn. Paris: Errance.
- Lambert, Pierre-Yves and Georges-Jean Pinault (eds.)
 2007 *Gaulois et celtique continental*. Geneva: Droz.
- Lewis, Henry
 1943 *Yr Elfen Ladin yn yr Iaith Gymraeg* [The Latin element in the Welsh language]. Cardiff: University of Wales Press.
- Mac Cana, Proinsias
 1976 Latin influence on British: the pluperfect. In: John J. O'Meara and Bernd Naumann (eds.), *Latin Script and Letters AD 400–900: Festschrift Presented to Ludwig Bieler on the Occasion of his 70th Birthday*. Leiden: Brill, 194–206.
- McCone, Kim R.
 1982 Further to absolute and conjunct. *Ériu* 33: 1–29.
- McCone, Kim R.
 1991 *The Indo-European Origins of the Old Irish Nasal Presents, Subjunctives and Futures*. (Innsbrucker Beiträge zur Sprachwissenschaft 66). Innsbruck: Institut für Sprachwissenschaft der Universität.
- McCone, Kim R.
 1994 An tSean-Ghaeilge agus a Réamhstair [On Old Irish and its prehistory]. In: McCone et al. (eds.), 61–219.
- McCone, Kim R.
 1996 *Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change*. (Maynooth Studies in Celtic Linguistics 1). Maynooth: Department of Old Irish, St. Patrick's College.
- McCone, Kim R.
 1997 *The Early Irish Verb*. 2nd edn. (Maynooth Monographs 1). Maynooth: An Sagart.
- McCone, Kim R.
 2005 *L'irlandese antico e la sua preistoria*. Alessandria: Edizione dell'Orso.
- McCone, Kim R.
 2006 *The Origins and Development of the Insular Celtic Verbal Complex*. (Maynooth Studies in Celtic Linguistics 6). Maynooth: Department of Old Irish, National University of Ireland.
- McCone, Kim R., Damian McManus, Cathal Ó Háinle, Nicholas Williams, and Liam Breatnach (eds.)
 1994 *Stair na Gaeilge in Ómós do Pádraig Ó Fiannachta* [The history of the Irish language in tribute to Patrick Fenton]. Maynooth: Department of Old Irish, St Patrick's College.
- McManus, Damian
 1983 A chronology of the Latin loan-words in early Irish. *Ériu* 34: 21–72.
- McManus, Damian
 1984a On final syllables in the Latin loan-words in early Irish. *Ériu* 35: 137–162.
- McManus, Damian
 1984b *Linguarum diversitas: Latin and the vernaculars in early medieval Britain*. *Peritia* 3: 151–188.
- McManus, Damian
 1991 *A Guide to Ogam*. (Maynooth Monographs 4). Maynooth: An Sagart.

- Mullen, Alex
 2007a Evidence for written Celtic from Roman Britain: a linguistic analysis of *Tabellae Sulis* 14 and 18. *Studia Celtica* 41: 29–43.
- Mullen, Alex
 2007b Linguistic evidence for ‘Romanization’: continuity and change in Romano-British onomastics. A study of the epigraphic record with particular reference to Bath. *Britannia* 38: 35–61.
- Mullen, Alex
 2012 Introduction: multiple languages, multiple identities. In: Mullen and James (eds.), 1–35.
- Mullen, Alex and Patrick James (eds.)
 2012 *Multilingualism in the Greco-Roman Worlds*. Cambridge: Cambridge University Press.
- Nurmio, Silva
 2014 Middle Welsh *-awr*: the case of the lost plural suffix. *Studia Celtica* 48: 139–170.
- O’Brien, Michael A. (ed. Rolf Baumgarten)
 1973 Old Irish Personal Names. *Celtica* 10: 211–236.
- Ó Cruaíoch, Conchubhar
 1997 Syncope Patterns in Denominative Verbs. *Ériu* 48: 239–264.
- Ó Siadhail, Micheál
 1989 *Modern Irish: Grammatical Structure and Dialectal Variation*. Cambridge: Cambridge University Press.
- Parry-Williams, Thomas H.
 1923 *The English Element in Welsh*. London: Honourable Society of Cymmrodorion.
- Piette, J. R. F.
 1973 *French Loanwords in Middle Breton*. Cardiff: University of Wales Press.
- Rodway, Simon
 2013 *Dating Medieval Welsh Literature. Evidence from the Verbal System*. Aberystwyth: CMCS Publications.
- Russell, Paul
 1988 Preverbs, prepositions and adverbs: sigmatic and asigmatic. *Transactions of the Philological Society* 86: 144–172.
- Russell, Paul
 1990 *Celtic Word Formation: the Velar Suffixes*. Dublin: School of Celtic Studies, Dublin Institute for Advanced Studies.
- Russell, Paul
 2004 Old Welsh *Dinacat, Cunedag, Tutagual*: fossilised phonology in Brittonic personal names. In: John H. W. Penney (ed.), *Indo-European Perspectives in Honour of Anna Morpurgo Davies*. Oxford: Oxford University Press, 447–460.
- Russell, Paul
 2006 *VILBIAM (RIB 154): kidnap or robbery?* *Britannia* 37: 363–367.
- Russell, Paul
 2011a Latin and British in Roman and Post-Roman Britain: methodology and morphology. *Transactions of the Philological Society* 109: 138–157.
- Russell, Paul
 2011b *Uocridem*: a new British word from Vindolanda. *Studia Celtica* 45: 192–197.
- Russell, Paul
 2012 *An habes linguam Latinam? Non tam bene sapio*: views of multilingualism from the early-medieval West. In: Mullen and James (eds.), 193–224.
- Russell, Paul
 2015 *Gwas, Guos-, Gos-*: the reflexes of Brittonic *wo. In: Guillaume Oudaer, Gaël Hily, and Hervé Le Bihan (eds.), *Mélanges en l’honneur de Pierre-Yves Lambert*. Rennes: TIR, 77–89.

- Schrijver, Peter
1995 *Studies in British Celtic Historical Phonology*. (Leiden Studies in Indo-European 5). Amsterdam: Rodopi.
- Schrijver, Peter
1997 *Studies in the History of Celtic Pronouns and Particles*. (Maynooth Studies in Celtic Linguistics 2). Maynooth: Department of Old Irish, National University of Ireland.
- Schrijver, Peter
1998–2000 Geminate spellings in the Old Welsh glosses to *Martianus Capella*. *Études Celtiques* 34: 147–160.
- Schrijver, Peter
2002 The rise and fall of British Latin: evidence from English and Brittonic. In: Markku Filppula, Juhani Klemola, and Heli Pitkänen (eds.), *The Celtic Roots of English*. Joensuu: Faculty of Humanities, University of Joensuu, 87–110.
- Schrijver, Peter
2005 Early Celtic diphthongization and the Celtic-Latin interface. In: Javier De Hoz, Ernesto R. Luján, and Patrick Sims-Williams (eds.), *New Approaches to Celtic Place-names in Ptolemy's Geography*. Madrid: Clásicas, 55–67.
- Schrijver, Peter
2007a What Britons spoke around 400 AD. In: Nick Higham (ed.), *Britons in Anglo-Saxon England*. Woodbridge: Boydell and Brewer, 164–171.
- Schrijver, Peter
2007b Some common developments of Continental and Insular Celtic. In: Lambert and Pinault (eds.), 355–372.
- Schrijver, Peter
2009 Celtic influence on Old English: phonological and phonetic evidence. In: Markku Filppula and Juhani Klemola (eds.), Special Issue on Re-evaluating the Celtic Hypothesis. *English Language and Linguistics* 13: 193–211.
- Schrijver, Peter
2011a Old British. In: Ternes (ed.), 1–84.
- Schrijver, Peter
2011b Middle and Early Modern Breton. In: Ternes (ed.), 359–430.
- Schrijver, Peter
2014 *Language Contact and the Origins of the Germanic Languages*. (Routledge Studies in Linguistics). London: Routledge.
- Schumacher, Stefan
1995 The Middle Welsh absolute endings *-(h)yt*, *-(h)it* and *-(h)awt* and the question of the ‘future’ in Old and early Middle Welsh. *Die Sprache* 37: 54–72.
- Schumacher, Stefan
2011 Mittel- und Frühneukymrisch. In: Ternes (ed.), 85–236.
- Sims-Williams, Patrick
1991 The emergence of Old Welsh, Cornish and Breton orthography, 600–800: the evidence of Archaic Old Welsh. *Bulletin of the Board of Celtic Studies* 38: 20–86.
- Sims-Williams, Patrick
2003 *The Celtic Inscriptions of Britain: Phonology and Chronology, c. 400–1200*. (Publications of the Philological Society 37). Oxford: Blackwell.
- Sims-Williams, Patrick
2007a Common Celtic, Gallo-Brittonic and Insular Celtic. In: Lambert and Pinault (eds.), 309–354. [Reprinted in Sims-Williams 2007b: 1–42.]
- Sims-Williams, Patrick
2007b *Studies on Celtic Languages before the Year 1000*. Aberystwyth: CMCS Publications.

Sims-Williams, Patrick

- 2013 The Celtic composition vowels *-o-* and *-io-*. In: Juan Luis García Alonso (ed.), *Continental Celtic Word Formation. The Onomastic Data*. Salamanca: Ediciones Universidad, 37–50.

Stifter, David

- 2011 Lack of syncope and other nichtlautgesetzlich vowel developments in OIr. consonant-stem nouns. Animacy rearing its head in morphology? In: Thomas Krisch, Thomas Lindner, Michael Crombach, and Stefan Niederreiter (eds.) *Indogermanistik und Linguistik im Dialog. Akten der XIII. Fachtagung der Indogermanischen Gesellschaft vom 21. bis 27. September 2008 in Salzburg*. Wiesbaden: Reichert, 556–565.

Stüber, Karin

- 1998 *The Historical Morphology of n-stems in Celtic*. (Maynooth Studies in Celtic Linguistics 3). Maynooth: Department of Old Irish, National University of Ireland.

Ternes, Elmar (ed.)

- 2011 *Brythonic Celtic – Britannisches Keltisch. From Medieval British to Modern Breton*. (Münchener Forschungen zur historischen Sprachwissenschaft 11). Bremen: Hempen.

Thurneysen, Rudolf

- 1946 *A Grammar of Old Irish*. Translated by Daniel A. Binchy and Osborn J. Bergin. Dublin: Dublin Institute for Advanced Studies.

Trépos, Pierre

- 1957 *Le pluriel breton*. Brest: Emgleo Breiz.

Uhlich, Jürgen

- 2002 Verbal governing compounds (synthetics) in early Irish and other Celtic languages. *Transactions of the Philological Society* 100: 403–433.

Vendryes, Joseph

- 1902 *De hibernicis vocabulis quae a latina lingua originem duxerunt* [On the Irish words which originated from the Latin language]. Paris: Klincksieck.

Watkins, T. Arwyn

- 1972 The accent in Old Welsh – its quality and its development. *Bulletin of the Board of Celtic Studies* 25: 1–11.

Williams, Nicholas

- 2011 Middle and Late Cornish. In: Ternes (ed.), 237–358.

Wodtko, Dagmar S.

- 1995 *Sekundäradjektive in den altirischen Glossen. Untersuchungen zur präfixalen und suffixalen Wortbildung*. (Innsbrucker Beiträge zur Sprachwissenschaft 81). Innsbruck: Institut für Sprachwissenschaft der Universität.

Ziegler, Sabine

- 1994 *Die Sprache der altirischen Ogam-Ischriften*. (Historische Sprachforschung Ergänzungsheft 36). Göttingen: Vandenhoeck und Ruprecht.

Paul Russell, Cambridge (UK)

XII. Tocharian

74. The documentation of Tocharian

1. Documentation
2. Dating
3. References

1. Documentation

There are two Tocharian languages, conventionally dubbed Tocharian A and Tocharian B, which would have been mutually unintelligible in their attested forms but can be derived from a single original Common Tocharian, a distinct branch of Indo-European. The entrenched name “Tocharian” is almost certainly a misnomer, insofar as it suggests any connection (as was once supposed) with the *Tokharoi* of classical sources, who lived in the region of Bactria and spoke a form of Iranian (but see Adams [2000] for an argument that the link can still be maintained, if one allows for language shift in the eastern region).

The evidence for Tocharian consists of written documents from the early mediaeval period found at various sites along the northern Silk Road within the Tarim basin in the Autonomous Region of Xinjiang in China. (A single text from Endere in the south is generally supposed to have been brought there from elsewhere). Some manuscripts are of uncertain provenance, having been acquired locally by collectors from the late nineteenth century onwards, but the majority were recovered by archaeological expeditions (Prussian, French, British, Russian, and Japanese) in the early years of the twentieth century. A more recent chance find (1974) brought to light forty-four fragmentary leaves of a Tocharian A manuscript and offers hope that there is still more material yet to be discovered. The find-places extend from sites around Kucha in the west, through the area around Karashahr (especially Shorchuk just to the south), to the Turfan depression in the east, and the geographical distribution of the remains of the two languages is as follows: Tocharian A manuscripts come from the eastern sector (Shorchuk and Turfan), Tocharian B texts are found throughout the whole area.

Two very fragmentary texts from the Turfan region, perhaps dating to the mid-tenth century CE, are written in Manichean script: one is a hymn to Mani in Tocharian B with an Old Turkic translation, the other a few lines of a hymn to Jesus (von Gabain and Winter 1958; Pinault 2008). All other Tocharian texts are written in a specially adapted version of the Indian *brāhmī* script found in countless Sanskrit manuscripts from the region, having been introduced there along with Buddhism, which became the prevailing culture. Buddhist literature, in the form of translations or adaptations of Indian originals, constitutes a substantial part of the surviving Tocharian texts, and much of the non-literary material (in Tocharian B) has Buddhist connections, for instance monastery accounts or the graffiti of religious pilgrims.

The literary texts are normally written on paper with a reed pen in fine and regular calligraphy. The oblong pages, containing usually four to six lines of writing on each side (sometimes more), are of a traditional Indian type, reflecting palm-leaf manuscripts, while the use of paper clearly points to Chinese influence in the region. The individual pages were collected into books with wooden covers, the binding being no more than a string passed through a hole in the left-hand side of each page. No complete text survives, and even whole pages are relatively rare.

A convenient classification of the literary texts following the categories applied to Sanskrit Buddhist literature has been provided by Georges-Jean Pinault (1989: 14 f., 1991: 240 ff.). Among them are a) *Vinaya* texts concerned with monastic discipline (including rules for the conduct of monks and nuns and also an initiation ritual in Tocharian B with Sanskrit passages, unusually written on birch bark), b) collections of the sayings of the Buddha such as the stanzas that comprise the *Udānavarga*, a very popular work in the region, being attested not only in the original Sanskrit version but also in Tibetan and Chinese translations, c) commentaries on these, including the *Udānālaṅkāra* (a verse text that sets verses from the *Udānavarga* in short narratives to explain the context of the utterances), which is known only in its Tocharian version and of which several fragmentary manuscripts survive, mainly in Tocharian B but also in Tocharian A, d) narratives of the earlier existences of the Buddha (*jātaka* stories), such as the *Punyavanta-Jātaka*, of which there are substantial and well-preserved portions in Tocharian A, e) dramatized versions of similar stories, such as the *Araṇemi-Jātaka*, known from fragments in both languages, and f) poetic texts such as hymns praising the Buddha. There are also scholarly productions such as writing exercises, texts of grammatical instruction, works on astronomy and on medicine.

In Tocharian A the texts are all concerned with religious literature (though there are few *Vinaya* texts), whereas in Tocharian B there is also a mass of other less elevated material pertaining to everyday life. This difference has encouraged a view that Tocharian A survived essentially as a religious language, the sacred texts being recopied for reasons of piety and to earn merit, while Tocharian B was clearly a living language used for instruction in Buddhist doctrines and also for more mundane purposes.

Secular documents in Tocharian B are often written in a more cursive version of the *brāhmī* script, using a paint brush. The writing material is again usually paper, but in larger sheets, sometimes assembled into rolls. There are also documents on wood, most appealingly a number of caravan passes, found at a control post. These consist of two small flat oblong pieces of wood, with notches in each side: the surfaces with the written text (containing details of the constitution of a given caravan and instructions to let it pass) were placed face to face and the whole bound by string, held in place by the notches; the name of the addressee was then written on the outside. The documents on paper include a lone example of Tocharian B original literature in the form of a love poem and a few fragments of historical documents, but mainly relate to administrative and legal affairs – monastery accounts, business letters, reports of legal proceedings (Pinault 1989: 15 f.). To these may be added graffiti of various kinds, names on pottery vessels, and captions on frescoes (Pinault 1993–1994: 172 f.).

The circumstances of the recovery of the Tocharian documents by archaeologists from different countries naturally led to their wide dispersal, and the various national collections contain written material of all kinds. Publication of the texts has been a slow process and is still a long way from completion; furthermore some of the existing edi-

tions have proved less than wholly satisfactory, so that much work remains to be done before a full and trustworthy Tocharian corpus is available.

The several thousand fragmentary documents found by the Prussian expeditions were sent to Berlin, where they are still kept (Malzahn 2007a: 80 ff.). The first extensive set of fragments to be published, in transliteration and with a selection of good photographs, was of texts in Tocharian A (Sieg and Siegling 1921). There followed two volumes of B texts, in transliteration only (Sieg and Siegling 1949, 1953); the first 116 fragments were re-published later (Thomas 1983), but it appears that many readings need to be revised (see Schmidt 1997). A selection of monastery business documents has also been published (Sieg 1950). There remain in the Berlin collection a large number of unpublished texts, many small fragments but some that are more substantial, including the initiation ritual on birch bark, edited only in an unpublished Habilitationsschrift. Nearly all the Berlin texts are now available on the *TITUS* website established at Frankfurt University (see References under TITUS), but for most of the unpublished texts there are only photographs without any accompanying transcription.

The French archaeologists brought their extensive finds from the Kucha region to Paris, but the publication of these texts was for a long time slow and fitful. A few were edited by Sylvain Lévi in various articles and in a book that also contained some fragments from other collections (Lévi 1933; but see Sieg 1938 and Thomas 1971 for corrections). Jean Filliozat published some medical and magical texts, together with a few manuscripts from the London collection (Filliozat 1948; but see Sieg 1955 for corrections). Walter Couvreur published a few texts, often only in translation. It is only with the series of articles from the 1980s onwards by Georges-Jean Pinault (for instance, Pinault 1984) that a proper start has at last been made on making more of the rich material generally available; he has also published the caravan passes and a number of graffiti (Pinault 1987). For a helpful concordance to the Paris collection, including an explanation of the baffling catalogue numbers in the Bibliothèque Nationale and a brief identification of the content of each fragment, see Pinault (2007).

There are some 1500 Tocharian fragments in the British Library (Malzahn 2007a: 84 ff.), some documents coming originally from the explorations of Sir Aurel Stein, others from various collections created by purchase in British India. These have been published only in a doctoral thesis (Broomhead 1964), but photographs of much of the material can be found on the website of the International Dunhuang Project (see References under IDP), together with transliterations of about half the documents, the other half being expected to appear shortly.

The Russian finds, considerably fewer in number, were taken back to St. Petersburg. Most are still unpublished, having only recently been made generally available to western scholars, but for a selection of economic texts, see Pinault (1998). For a complete list of the texts held, see Malzahn (2007a: 91 ff.).

The handful of documents found by the Japanese expedition under Count Otani were removed to Japan, and nearly all are now in Tokyo; most are not yet published (Malzahn 2007a: 93 f.).

Still in China is the manuscript of the Tocharian A dramatic text (the *Maitreyasamiti-Nāṭaka*) found in 1974, which has been fully published by Ji, Winter, and Pinault (1998). There are also a few unpublished documents in museum collections (Malzahn 2007a: 94 f., 107), and there remain *in situ* a number of captions and inscriptions in the painted caves of Buddhist monasteries (cf. Pinault 1993–94).

2. Dating

The caravan passes and certain other Tocharian B documents are dated by regnal years, and already Lévi (1913) was able to show that some of the kings mentioned in these texts could be identified with kings of Kucha mentioned in Chinese records, giving dates in the first half of the seventh century CE. Even greater precision is possible in the case of a graffito (G-Qa 1) dated both by the regnal year of Suvarnadeva and by the year within the Chinese cycle of twelve animals (in fact the year of the Tiger), which establishes the date as 642 (Pinault 1987: 84 f.). It seemed clear to the early editors on general palaeographical grounds that some texts were older, and from the presence of Uighur elements that others were more recent: the range of dates for the Tocharian material as a whole has thus been standardly given as from the sixth century to the eighth.

Another reason for assuming chronological diversity within the corpus is based on observable linguistic differences. Many Tocharian B texts from the western part of the region (especially those from Ming-oi Qizil, the MQ-texts) show forms that appear to be older than those of the standard central dialect, and texts from the eastern part of the region often show more evolved forms (with simplifications of consonant clusters, etc.). These variations had been attributed to synchronic dialectal variation (Winter [1955] 2005), but it was convincingly argued by Stumpf (1990) that there were different chronological layers to be recognized here, and this has now been demonstrated with great precision by Peyrot (2008).

This conclusion is supported by detailed palaeographical analysis, building on the relative chronology established for Sanskrit manuscripts written in the region and going well beyond the rather impressionistic comments of the early editors (Malzahn 2007b). It emerges that there is a strong correlation between archaic ductus and MQ-features, even in texts from outside the western region, so that these must now be accepted as historically earlier. The Tocharian script may have been developed at the beginning of the fifth or at the end of the fourth century, and certainly attained its standard form by the early seventh century, when the Tocharian B language too had progressed to its classical stage.

A new dimension has now been introduced by the ¹⁴C dates offered for some of the Berlin texts (Tamai 2005, see also Adams 2006): for the nine B manuscripts tested, the ¹⁴C dates range from ca. 400 to ca. 900 CE (and in one startling case ca. 1200), largely in accord with the palaeographical dates assignable to the documents; the four A manuscripts range from ca. 700 to ca. 1000 CE. Once testing has been extended to a wider range of manuscripts, there seems good reason to hope that a better understanding of the chronology of the texts, with clear periodization, will at last be achieved, with interesting results for the study of the historical development of the two languages.

3. References

Adams, Douglas Q.

- 2000 Some Observations of Peoples, Places, and Languages in the Tarim Basin in the First Millennium AD. *Tocharian and Indo-European Studies* 9: 1–28.

- Adams, Douglas Q.
2006 Some Implications of the Carbon-14 Dating of Tocharian Manuscripts. *Journal of Indo-European Studies* 34: 381–389.
- Broomhead, J. W.
1964 *A Textual Edition of the British Hoernle, Stein and Weber Kuchean Manuscripts. With Transliteration, Translation, Grammatical Commentary and Vocabulary*. 2 vols. Cambridge University Dissertation.
- Filliozat, Jean
1948 *Fragments de textes koutchéens de médecine et de magie. Texte, parallèles sanskrits et tibétains, traduction et glossaire*. Paris: Adrien-Maisonneuve.
- von Gabain, Annemarie and Werner Winter
1958 *Ein Hymnus an den Vater Mani auf 'Tocharisch B' mit alttürkischer Übersetzung*. (Türkische Turfantexte IX). Berlin: Akademie Verlag. IDP, <http://idp.bl.uk>
- Ji, Xianlin, Werner Winter, and Georges-Jean Pinault
1998 *Fragments of the Tocharian A Maitreyasamiti-Nāṭaka of the Xinjiang Museum, China*. Berlin: Mouton de Gruyter.
- Lévi, Sylvain
1913 Le “Tokharien B”, langue de Koutcha. *Journal Asiatique* 2 (11^e série): 381–421.
- Lévi, Sylvain
1933 *Fragments de textes koutchéens*, Paris: Imprimerie Nationale.
- Malzahn, Melanie
2007a Tocharian Texts and Where to Find Them. In: Malzahn (ed.), 79–112.
- Malzahn, Melanie
2007b The Most Archaic Manuscripts of Tocharian B and the Varieties of the Tocharian B Language. In: Malzahn (ed.), 255–297.
- Malzahn, Melanie (ed.)
2007 *Instrumenta Tocharica*. Heidelberg: Winter.
- Pinault, Georges-Jean
1984 Une lettre de monastère du Fonds Pelliot Koutchéen. *Revue de la Bibliothèque Nationale* 11: 21–33.
- Pinault, Georges-Jean
1987 Épigraphie koutchéenne. In: Chao Huashan, Simone Gaulier, Monique Maillard, and Georges Pinault (eds.), *Sites divers de la région de Koutcha* (Mission Paul Pelliot, Documents archéologiques VIII). Paris: Collège de France, 59–196 with plates XL–XCVI.
- Pinault, Georges-Jean
1989 Introduction au tokharien. *Lalies* 7: 5–224.
- Pinault, Georges-Jean
1991 Les manuscrits tokhariens et la littérature bouddhique en Asie centrale. *Comptes rendus des séances de l'Académie des inscriptions et belles-lettres*: 227–251.
- Pinault, Georges-Jean
1993–1994 Une nouvelle inscription koutchéenne de Qumtara: légende de scènes bouddhiques de Prañidhi. *Bulletin d'Études Indiennes* 11–12: 171–220.
- Pinault, Georges-Jean
1994 Aspects du bouddhisme pratiqué au nord du désert du Taklamakan, d'après les documents tokhariens. In: Fukui Fumimasa and Gérard Fussman (eds.), *Bouddhisme et cultures locales. Quelques cas de réciproques adaptations* (Actes du colloque franco-japonais de septembre 1991). Paris: École Française d'Extrême-Orient, 85–113.
- Pinault, Georges-Jean
1998 Economic and administrative documents in Tocharian B from the Berezovsky and Petrovsky collections. *Manuscripta Orientalia* 4/4: 3–20.
- Pinault, Georges-Jean
2007 Concordance des manuscrits tokhariens du fonds Pelliot. In: Malzahn (ed.), 163–219.

- Pinault, Georges-Jean
2008 Bilingual hymn to Mani. Analysis of the Tocharian B parts. *Studies on the Inner Asian Languages* 23: 93–120.
- Schmidt, Klaus T.
1997 Wie zuverlässig sind unsere tocharischen Textausgaben? Kritische Bemerkungen zu den Editionen der *Tocharische Sprachreste, Sprache B*, von E. Sieg, W. Siegling und W. Thomas und einigen weiteren westtocharischen Textstellen. *Die Sprache* 39 [2000]: 224–238.
- Sieg, Emil
1938 Die Kutschischen Karmavibhaṅga-texte der Bibliothèque Nationale. *Zeitschrift für vergleichende Sprachforschung* 65: 1–54.
- Sieg, Emil
1950 Geschäftliche Aufzeichnungen in Tocharisch B aus der Berliner Sammlung. *Miscellanea Academica Berolinensia* II/2: 208–223.
- Sieg, Emil
1955 Die medizinischen und tantrischen Texte der Pariser Sammlung in Tocharisch B. *Zeitschrift für vergleichende Sprachforschung* 72: 63–83.
- Sieg, Emil and Wilhelm Siegling
1921 *Tocharische Sprachreste*. I. Band: Die Texte, 2 vols. (A: Transcription, B: Tafeln). Berlin: De Gruyter.
- Sieg, Emil and Wilhelm Siegling
1949 *Tocharische Sprachreste, Sprache B. Heft 1: Die Udānālakāra-Fragmente, Text, Übersetzung und Glossar*. Göttingen: Vandenhoeck & Ruprecht.
- Sieg, Emil and Wilhelm Siegling
1953 *Tocharische Sprachreste, Sprache B. Heft 2: Fragmente Nr. 71–633*, aus dem Nachlaß herausgegeben von Werner Thomas. Göttingen: Vandenhoeck & Ruprecht.
- Stumpf, Peter
1990 *Die Erscheinungsformen des Westtocharischen. Ihre Beziehungen zueinander und ihre Funktionen*. (TIES Supplementary Series 2). Reykjavík: Málvísindastofnun Háskóla Íslands.
- Tamai, Tatsushi
2005 Paläographische Untersuchungen und ¹⁴C-Prüfung. Digitalisierung der chinesischen, tibetischen, syrischen und Sanskrit-texte der Berliner Turfansammlung, 2. 6. 2005 – Berlin. <http://www.bbaw.de/Forschung/Forschungsprojekte/turfanforschung/bilder/Tamai.pdf>.
- Thomas, Werner
1971 *Bilinguale Udānavarga-Texte der Sammlung Hoernle* (Abhandlungen der Akademie der Wissenschaften und der Literatur Mainz, Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse 1971, Nr. 7). Mainz: Akademie der Wissenschaften und der Literatur.
- Thomas, Werner
1983 *Tocharische Sprachreste, Sprache B. Teil I: Die Texte, Band 1: Fragmente Nr. 1–116 der Berliner Sammlung, herausgegeben von Emil Sieg und Wilhelm Siegling, neubearbeitet und mit einem Kommentar nebst Register versehen von Werner Thomas*. Göttingen: Vandenhoeck & Ruprecht. TITUS, <http://titus.fkidg1.uni-frankfurt.de/texte/tocharic/tht1.htm>
- Winter, Werner
2005 [1955] A Linguistic Classification of “Tocharian B” Texts. In: *Kleine Schriften / Selected Writings*. Bremen: Hempen, I. 1–10. [*Journal of the American Oriental Society* 75: 216–225.]

75. The phonology of Tocharian

- | | |
|---|-----------------------|
| 1. Introduction | 4. Non-syllabics |
| 2. Suprasegmental features and mechanisms | 5. Consonant clusters |
| 3. Syllabics | 6. References |

1. Introduction

The following is a brief outline of the phonology of Tocharian, which, due to the usual space limitations of handbook articles, offers neither an exhaustive treatment of all Tocharian sound changes nor complete bibliographical coverage but sets as its goal to describe some of the most important phonological traits and developments of the two Tocharian languages. Following this introduction, the treatment comprises three core sections: 2 Suprasegmental features and mechanisms, 3 Syllabics, and 4 Non-syllabics. Each section pursues a double focus, beginning with productive synchronic phonology, moving to less productive patterns, and then closing with historical sound changes. In what follows, superscript “A” and “B” are used to denote Tocharian A and Tocharian B. Superscript “j” indicates palatalization. Abbreviations: PIE = Proto-Indo-European, CToch. = Common Tocharian.

The basic synchronic reference works on Tocharian are Sieg, Siegling, and Schulze (1931), Krause (1952), Krause and Thomas (1960), and Thomas (1964). For the orthography of Tocharian, see Schmidt (1997: 19 f.) and Malzahn (2007: 223–254). A thorough description of the historical phonology of Tocharian is included in Ringe (1996), Pinault (2008: 413–460), and Malzahn (2010: 1–22). For diachronic variants of Tocharian B see Peyrot (2008). A concise 15-page sketch of the grammatical structure of both Tocharian languages is contained in Winter (1998). The basic etymological dictionary is Adams (2013), superseding van Windekens (1976) and Adams (1999) (for various corrigenda and addenda, cf. Winter 2001, 2003, and Hackstein 2003). A dependable but unfortunately incomplete etymological dictionary is Hilmarsson (1996).

2. Suprasegmental features and mechanisms

2.1. Accent and accentual allophony

2.1.1. Tocharian B (West Tocharian)

2.1.1.1. Accentual allophones of /a/ and /ä/ in Tocharian B

Tocharian B distinguishes three central vowels, written < \bar{a} >, <a>, and the so-called shwa or *Fremdvokal* < \ddot{a} >. Their phonetic values can be approximated as [a], [ʌ], and [i̯]. In the central and eastern dialects of Tocharian B, the three vowels may alternate with each other as accentual allophones of the two phonemes /a/ and /ä/ (Ringe 1987: 254–269).

phoneme	stressed allophone	unstressed allophone
/a/	ǎ	a
/ä/	á	ä (in closed syllables) Ø (in open non-final syllables)

Note that native Tocharian orthography lacks accent marks. Accent marks, when noted in the transliteration of West Tocharian forms, are recoverable solely from the accentual allophony of vowels.

Phoneme /a/

Stressed /a/ = <ǎ>	Unstressed /a/ = <a>
^B <i>ǎk-em</i> ‘they lead’ ^B <i>pǎcer</i> ‘father’ ^B <i>pǎke</i> ‘part’	prs. mid. ptcp. <i>ak-émane</i> ‘being led’ gen. pl. <i>pacérams</i> ‘of the fathers’ nom. pl. <i>pakénta</i> ‘parts’

Phoneme /ä/ (shwa, so-called *Fremdvokal*)

Stressed shwa /ǎ/ = <a> (note 1)	Unstressed shwa in closed syllables /ä/ = <ä> (note 2)	Unstressed shwa in open syllables /ä/ = Ø
^B <i>kánte</i> ‘hundred’ ^B <i>yáltse</i> ‘thousand’ ^B <i>cámel</i> ‘birth’ ^B <i>śána</i> ‘wife’ ^B <i>tsáksäm</i> ‘burns’ (trans.)	<i>kánténma</i> ‘hundreds’ <i>yáltсэнma</i> ‘thousands’	<i>cméla</i> ‘births’ <i>śnóna</i> ‘wives’ <i>tskétär</i> ‘burns’ (intr.)

Both the accent and the accentual allophones of the *ǎ*- and *a*-phoneme are crucial to the differentiation of morphemes and lexemes. Note the following near minimal pairs:

- prs. IXa *anáššäm* ‘breathes in’ versus prs. IXb *ánäššäm* ‘lets breathe in’.
- prt. I *takám* ‘we were’ versus sbj. V *tákam* ‘we shall be’.
- prs. VIII *nákštär* ‘perishes’ versus prs. VIII *nǎkštär* ‘scolds’.
- sbj. V 3sg. *tsánka(m)ne* ‘he will rise for him’ versus 3pl. ^B*tsánkam* ‘they will arise’, with weak ablaut grade as in ^B*tsánkoy* ‘may he rise’, ^B*tsánkatsi* ‘to arise’.

Note 1: A Tocharian B sound change causes the shwa vowel *ǎ* to be dropped, “even when it was accented, whenever it occurred between a semivowel and a consonant” (Ringe 1989: 37; cf. differently Malzahn 2010: 6f.), and a preceding Pre-Toch. B **e* to be converted into *a*:

- Pre-Toch. B **ye-yǎku* > ^B*yaiku* ‘commanded’.
- Pre-Toch. B **we-wǎsu* > ^B*aúsu* ‘having put on’.

The sound change also applies to compounds and univerbations, cf.

- Pre-Toch. B **péle yákne* > ^B*pelaikne* ‘way of conduct, dharma’ (cf. Pinault 1995: 19); **té yákne* > ^B*taiknesa* ‘in that way’.
- Pre-Toch. B **kéw wárše* ‘cow male’ > ^B*kaúrše* ‘bull’.

Note 2: Internal syllables of the shape -CVC.C(V)- count as closed syllables. The same holds for word-final CVC#. Thus, the *Fremdvokal* of the ending *-äm* (prs. 3sg. act.) is left unscopated. Syncope occurs before the 2sg. ending, e.g. *áš(o)* ‘you lead’ < **ásätä*, because the underlying form is trisyllabic, ending in a final vowel, as still attested in metrical texts by the spelling *ášto*.

2.1.1.2. West Tocharian peninitial accent

The default accent rule for words with more than two syllables in West Tocharian is that the peninitial syllable hosts the accent. Cf. Krause (1952:10): “Der Akzent der westtochar. Wörter scheint normalerweise auf der ersten Silbe zu ruhen, sofern das betr. Wort ein- oder zweisilbig ist, auf der zweiten Silbe dagegen, wenn es dreisilbig, meist auch, wenn es vier- oder fünfsilbig ist.” [The accent of West Tocharian words normally seems to rest on the first syllable, provided that the word concerned is mono- or disyllabic, but on the second syllable if it is trisyllabic, (and) usually also if it has four or five syllables] Monosyllabics normally bear an accent, unless clitic. Thus,

monosyllabics	ǰ	^B <i>má</i> ‘(do) not’
trisyllabics	xǰx, xǰxx ...	^B <i>pacéra</i> ‘fathers’

For the treatment of disyllabic words, see 2.1.1.3.

Apparent exceptions may be accounted for in various ways. Occasionally disyllabic forms appear to bear the accent on their final syllable but comply with the peninitial accent rule once they are recognized to be underlyingly trisyllabic (a). Further, initially accented trisyllabic forms may recover univerbated syntagms (b, c) or the loss of reduplicating syllables (d).

- a) The verbal endings 1sg. ^B*-m(ä)* from PIE **-mi*, 2sg. ^B*-t(ä/o)*, 1pl. ^B*-m(ä/o)* from PIE *-mes* are underlyingly syllabic: 1sg. prs. act. opt. ^B*pälsköym* = *pälsköymä* ‘may I think’, 1sg. prs. act. opt. ^B*cämpim* = *cämpimä* ‘may I be capable’; 2sg. prs. act. opt. ^B*akóyt* = *akóytä* ‘may you guide’; 1pl. prs. act. ^B*cämpem* = *cämpemo* ‘we can’, 1pl. prt. act. *takäm* = *takämo* ‘we became, were’.
- b) Secondary case forms (perlative, comitative, allative, ablative, locative, causal) are usually treated as syntagms consisting of a word followed by a postposition: *lákne=ne* ‘distress=in’, *śámna=ne* ‘men=in’. The secondary case morpheme (originally and still accentually a postposition) is appended (cf. Winter 1967; Pinault 2011).
- c) The prose form ^B*pánäkte* ‘buddha’, preserves an underlying syntagm *pá(t)=ñäkte* ‘buddha=god’. In the poetic genitive *pud=ñäkténtse*, appositional *ñäkténtse* is still treated as an uncompounded form.
- d) In Tocharian verb stems, the loss of reduplicating syllables resulted in initially accented forms. The initial accent was then redefined as a token of the corresponding causative category (morphologization): prs. Xb ^B*tánmästär* < **tetánmästär*, cf. already Krause and Thomas (1960: 44), and note the pret. abstract noun ^B*lelákäšsor*.

2.1.1.3. Synchronic accent retraction in disyllabics

Synchronically disyllabic words retract the accent to the initial syllable:

disyllabics $x\acute{x} \rightarrow \acute{x}x$ ${}^B{}^*pac\acute{e}r \rightarrow p\acute{a}c\acute{e}r$

“The underlying accent of a (morphophonemically) final syllable will be retracted to the preceding one in the surface realization of a Tocharian B form.” (Winter 1990: 372), e.g., prt. 3sg. act. $*tak\acute{a} > t\acute{a}ka$ ‘(s)he, it became, was’. Note that the underlying final accent surfaces before a following clitic, e.g. prt. 3sg. $tak\acute{a} = ne > tak\acute{a}ne$ ‘(s)he, it became, was to him/her’.

Example: preterite I $tak-\acute{a}$ -:

- | | | |
|-------------|--|------------------------|
| sg. act. 1. | $tak\acute{a}wa$ | ‘I became, was’ |
| 2. | $tak\acute{a}sta$ | ‘you became, were’ |
| 3. | $t\acute{a}ka$, but $tak\acute{a}-ne$ | ‘it was/became to him’ |
| | | |
| pl. act. 1. | $tak\acute{a}m(o)$ | |
| 2. | $tak\acute{a}s(o)$ | |
| 3. | $tak\acute{a}re$ | |

2.1.1.4. Historical accent protraction

In words with three or more syllables, PIE initial accent was protracted by one syllable (Ringe 1987: 258 ff.):

- PIE $*pr\acute{o}ti\acute{h}_3k^uom > {}^Bprats\acute{a}ko$ ‘breast’; cf. Ved. $pr\acute{a}tikam$, Gk. $\pi\rho\acute{o}\sigma\omega\pi\omicron\nu$.
- PIE $*h_2\acute{e}g\acute{o}nti > {}^Bak\acute{e}n-ne$ ‘they lead him/her’, PIE $*b^h\acute{e}romh_1no- > *p\acute{a}r\acute{e}man\acute{e}$, whence by deletion of unstressed shwa in first syllable ${}^Bpr\acute{e}man\acute{e}$.
- PIE $*s(u)\acute{e}k\acute{s}tos > *s^i\acute{a}k\acute{s}t\acute{e}- > *s^i\acute{a}k\acute{s}t\acute{e}-$ (with anaptyxis) $> *s^i\acute{a}k\acute{s}t\acute{e}- > {}^B\acute{s}k\acute{a}ste$ ‘sixth’; cf. Gk. $\acute{\epsilon}k\tau\omicron\varsigma$, Lat. $sestus$, $sextus$.

2.1.1.4.1. Subrules for clausal clitics

Subrule 1: Enclitic words draw the accent. A word followed by an enclitic is stressed on the syllable immediately preceding the enclitic.

- PIE $*h_1\acute{a}l\acute{i}os > {}^B\acute{a}lye$ [$*\acute{a}lye$] $\rightarrow aly\acute{e}=k, aly\acute{e}=k\acute{a}, aly\acute{e}=ko$ [$aly\acute{e}=k\acute{a}$]. This phenomenon has typological parallels, cf. e.g. Latin $virum \rightarrow vir\acute{u}m=que$ (Weiss 2009: 111).

Subrule 2: Enclisis of disyllabic words triggers stress shift to the ultimate syllable, cf. e.g. Toch. B $st\acute{e}, st\acute{a}r, skent\acute{a}r, sp\acute{e}$. These words occur in unstressed position, as is indicated by shwa-deletion in their first syllables, and by the fact that these words are notoriously assigned the unstressed, pre-caesural position in verse (Hackstein 1995: 275 f.).

Subrule 3: Proclisis of disyllabics shows the same rightward accent shift, e.g. adverbs with underlying penultimate accent show oxytone accent when used as conjunctions, cf. e.g. Toch. B interrogative adverb $k\acute{a}tu$ ‘why so?’ \rightarrow conjunction $kat\acute{u}$ ‘for’ (Hackstein 2011: 199 f.).

2.1.1.5. CToch. weakening of *-a-* > *-ä-* in medial post-tonic syllables

In the initially accented causative *sk*-presents of classes IX–XI, post-tonic stem-final *-a-* is weakened to *-ä-*. This rule applies to both Tocharian languages, contrast prs. IXa ^B*anāššäm* ‘breaths’, but Xlb *ánāššäm*; prs. IXa ^B*lkāskau* ‘I see’, but prs. IXb *lákāskemane* ‘showing’; prt. ptc. ^B*kärnau* ‘struck’, but prs. IXb ^B*kárnäššeñca*, ^A*kárnämām* ‘striking’ (Hackstein 1995: 33; Malzahn 2010: 438 ff.).

2.1.2. Tocharian A (East Tocharian)

2.1.2.1. Tocharian A vowel weakening

In East Tocharian, second-syllable *ā* undergoes weakening to *a* in disyllables and both *ā* and *a* are weakened to *ä/Ø* in words of three or more syllables, if the first syllable contains *ā*, *a*, *e*, or *o*; cf. (2) for disyllables, and (4, 6, 8) for trisyllables. In trisyllables, weakened *ä* remains intact in closed syllables (4, 8), but is syncopated in open syllables (6, 8). See Krause and Thomas (1960: 45) §11,2 on disyllables and §11,3 on trisyllables. After *ä/i/u* in the first syllable, weakening fails to occur (1, 3, 7). Likewise, weakening is blocked after *ā* and *e* in the first syllable and before *ä* (and occasionally before *i* and *u*) in the third syllable (5) (cf. Winter 1994). Vowel weakening is a productive rule in East Tocharian, affecting also personal names borrowed from Sanskrit, e.g., nominative *Ānand* ‘Ānanda’ (A20a1), but allative *Ānāndanac* (A20a5). Contrast inhibited vowel weakening in column A and vowel weakening in column B:

Disyllables	
A) No weakening of 2 nd syll. <i>-ā-</i> after <i>-ä/i/u-</i> in first syllable	B) Weakening of 2 nd syll. <i>-ā-</i> to <i>-a-</i> after <i>-ā/a/e/o-</i> in first syllable in disyllabic words
(1) <i>kārsnāš</i> ‘(s)he knows’ <i>kālkāšt</i> ‘you went’ (prt. 2sg.) <i>yuknāš</i> ‘(s)he overcomes’	(2) <i>*kalkāš</i> > <i>kalkaš</i> ‘(s)he will step’ <i>*kārpnāš</i> > <i>kārnaš</i> ‘(s)he descends’ <i>*tākāšt</i> > <i>tākašt</i> ‘you became, were’ <i>*skenāš</i> > <i>skenaš</i> ‘(s)he endeavors’ <i>*pekāt</i> > <i>pekat</i> ‘he portrayed himself’ <i>*kotnāš</i> > <i>kotnaš</i> ‘(s)he splits’ <i>*kropāt</i> > <i>kropat</i> ‘was gathered’
Trisyllables +	
A) No weakening of 2 nd syll. <i>-a/ā-</i> after <i>-ä/i/u-</i> in the first syllable or between <i>-ā/e-</i> in the first syllable and <i>-ä-</i> in the third syllable	B) Weakening of 2 nd syll. <i>-a/ā-</i> to <i>-ä/-Ø-</i> after <i>-ā/a/e/o-</i> in initial and before <i>-ā-</i> in ultimate syllable
(3) <i>māsk-a-mār</i> ‘I am staying’	(4) <i>*pāša-ntāñ</i> > <i>pāšāntāñ</i> ‘guarding’ (nom. pl.)

(5) <i>pāsantār</i> ‘they observe’ <i>pekamāt</i> ‘we wrote (for ourselves)’	(6) * <i>arta-mār</i> > <i>artmār</i> ‘I approve’ * <i>pāsa-mām</i> > <i>pāsmām</i> ‘guarding’ * <i>āka-mām</i> > <i>ākmām</i> ‘leading’
(7) <i>nutāṣlu-ne</i> ‘destruction’ <i>kālpāmār</i> ‘I shall obtain’	(8) * <i>pāplā-ntu</i> > <i>pāplāntu</i> ‘delighted’ * <i>knānā-mām</i> > <i>knānmām</i> ‘knowing’

3. Syllabics

For the vowel systems of Tocharian A and B, see Jasanoff (1978: 29–34), Ringe (1996: xxi–xxiv), Winter (1998: 157 f.), Pinault (2008: 420–422), Malzahn (2010: 1–3).

Tocharian A and B vowels

[i] <i, ī>		[i̥] <ä>		[u] <u, ū>
	[ɛ] <e>	[ʌ] <a>	[o] <o>	
		[a] <ā>		

Tautosyllabic diphthongs are confined to Tocharian B [ai], [oy], [au]; in archaic texts and western dialects of Tocharian B [ey] and [ew] occur instead of [ai] and [au] (cf. Pinault 2008: 416 f.).

3.1. Restitution of vowels in ablaut systems

Two of the most pervasive developments affecting the vowel system are the tendency to eliminate vowel length as a distinctive phonological feature and the centralization and merger of the front and back high vowels */i/, */u/ and the mid front vowel [e] in Tocharian A [i̥] <ä>, and Tocharian B [i̥] <ä> with stressed allophone [ʌ] <a>. See 2.1.1.1 above. Diphthongal verbal roots, when in productive ablaut, tend to restore the resulting zero-grades *-i̥-* (< PIE **-i-*), and *-ä-* (< PIE **-u-*) as ^{AB}*i* and ^{AB}*u* (Adams 1978: 446–448; Hackstein 1995: 35; Ringe 1996: 135–138; Malzahn 2010: 24 f.).

- PIE **pik-* ‘color, paint, fashion’: ^B*pinkām* ‘paints, writes’, ^B*paikāmai* ‘I wrote’ (Malzahn 2010: 724).
- PIE **liθ₂-* ‘fall down, go (away)’: ^B*litau* ‘fallen off’, ^B*laitotār* ‘is falling off’ (Hackstein 2002: 8 f.; Malzahn 2010: 846–848); **lejith₂-* > Gk. λειθ- in **lojith-to-* > λοῖσθος ‘gone/left behind; last’.
- PIE **mith₂-* ‘let go, change’: ^B*mitentār* ‘they are setting out’, prt. ^B*maitar(e)* ‘they set out’ (Hackstein 1995: 27 f., Malzahn 2010: 769 f.)
- PIE **h₁lud^h-* ‘rise, go off, leave’: ^B*lutaṣṣeñca* ‘removing’, ^B*lyautwa* ‘I removed’ (Hackstein 1995: 245 ff., Malzahn 2010: 856–858).

- PIE **luk-* ‘shine, illuminate’: ^B*lukšäm* ‘shines, illuminates’, ^B*lyauksa* ‘illuminated’ (Hackstein 1995, Malzahn 2010: 854 f.).
- PIE **(s)prug-* ‘jump’: ^B*pruknam* ‘jumps’ (Hackstein 1995: 36, Malzahn 2010: 736 f.).

3.2. Short and long vowels

For the development of the vowel system from PIE to Common Tocharian see Jasanoff (1978: 33 f.), Ringe (1996: 7–37, 89–100, 124 ff. [loss of vowel length as a distinctive feature]), Kümmel (2007: 430), Pinault (2008: 420–444), and Malzahn (2010: 1–3).

3.2.1. Short vowels

3.2.1.1. PIE **(h₂)a* > CToch. **a* > ^B*a*, ^Á*á*, ^À*ā*, cf. above 2.1.1.1.

- PIE **amso-* > **ansæ-* > ^B*ántse* ‘shoulder’, **aysæ* > ^À*es* (Hilmarsson 1989: 127 f.); **omso-* > Gk. ὄμος (Hackstein 2002: 190 f.); **amso/omso-* > Hitt. *an(aš)ša-* ‘upper back’ (Puhvel 1984: 63).
- PIE **sal-* > **sal[-^jen]-^jēn* > CToch. **sal’äye* (cf. below 4.1.5.3.1) > ^B*salyiye* f., ^À*sāle* ‘salt’; cf. Gk. ἅλς, ἅλός, Arm. *al*.
- PIE **h₁ál̥io-* > CToch. **al̥iæ-* + *-kä* > ^B*alle-k*, ^À*āla-k* ‘(an)other’ : Lat. *alius*, Gk. ἄλλος.
- PIE **h₂ént-s*, **h₂nt-és*, **h₂ént* ‘at the front’ → **h₂ént-o-* ‘pertaining to the front’ > **h₂ánto-* > CToch. **ántæ-* > ^B*ánte* ‘forehead, surface’, ^À*ānt* : Ved. *ánta-* ‘border, edge’.
- PIE **h₂euh₂os* > ^B*áwe* ‘grandfather’ : Hitt. *ḫuḫḫaš*.

3.2.1.2. Pre-Toch. B **-a-* > *-o-* / *-y*

- PIE **sh₂-ie-* > CToch. **sajä-* > Pre-Toch. B **sojä-* > ^B*soyä-* ‘satisfy’.
- PIE **suh_{2,3}-ius* > CToch. **suaiä(s)* > Pre-Toch. B **suojä-* > ^B*soy* ‘son’ (Ringe 1996: 61–63).

Pre-Toch. B optative of verbal stems ending in *-a*: **k̥rs-a-ī-* > ^B*kársoym* ‘may I know’.

3.2.2. PIE **e*

3.2.2.1. PIE **e* word-initially > CToch. **jä*

- PIE collective **h₁ésh₂ōr* > CToch. **jásar* > ^B*yásar*, ^À*ysār* ‘blood’; cf. the abstract nouns Gk. ἔαρ (Hesych.), Hitt. *ēšhar* ‘id.’.

3.2.2.2. PIE *e word-internally > CToch. *jǎ

- PIE *gʷénh₂ > CToch. *k'äna > ^Bsána, ^Asäm f. ‘woman, wife’ : Gk. γυνή ‘woman’, Ved. jāni- ‘woman, wife’. See below 4.1.4 (g).
- PIE *jésetor > CToch. *jäs-jä-tär > ^Byaštär ‘lets boil, excites’ : Gr. ζέω ‘I boil, seethe’.

3.2.2.3. PIE *e after labials > CToch. *jǎ > ^Bi (cf. Peyrot 2008: 55–57)

- PIE *pénkʷe > CToch. *p'änk'jǎ > *p'änš > ^Bpiš, ^Apän ‘five’.
- PIE *méqʰu > CToch. *m'ätä > ^Bmit ‘honey’ : Ved. mádhū, OHG metu.
- PIE *mēms/*méms- (Ved. máh ‘flesh’), collective *mems-h₂ > CToch. *m'ämsa > ^Bmīsa ‘flesh’ (cf. Vine 1998: 95 f.).

3.2.3. PIE *o > CToch. *æ > ^Be, ^Aa

- PIE *kʷotos > ^Bkete, ket ‘whose’, with ablatival suffix PIE *-tos as in Lat. fundi-tus, Ved. kútaḥ ‘whence?’.
- PIE *ǵombʰo- > CToch. *kæmbæ- > ^Bkeme, ^Akam ‘tooth’ : Gk. γόμφος ‘peg’, OCS zǫbŭ ‘tooth’.
- PIE *h₂eǵomh₁no- > CToch. *akæmanæ > ^Bakémane, ^Aākāmāḡ ‘being led’ : Gk. ἀγόμενος, Ved. ājamānaḥ.

3.2.3.1. Word-final *-or > CToch. *-ār

- PIE *-tor, *-ntor > ^{AB}-tār, -ntār (cf. Ringe 1996: 86 f.).

3.2.3.2. Word-final *-o > CToch. *æ > ^Ba

- PIE aorist imperative *deh₃ ‘give!’ > *do(h₃) > *dō > *tsæ > affirmative particle ^Btša, ^A-ts (Hackstein 2001: 26–39).
- PIE *kʷis so > unstressed indefinite *kʷiso > ^B-ksa, in contrast to stressed proclitic *kʷisó > ^Bkᵤsé, e.g. ^Bkᵤsé=ksa ‘whoever’, ^Ballék=ksa ‘someone else’; cf. Lat. quisquis, aliquis (Pinault 1997b: 470–472, Hackstein 2001: 32 f.).

3.2.3.3. Weakening of PIE *o in proclitics to CToch. *æ > ^{AB}ǎ

- PIE *tod > CToch. *tæ → + -mV > proclitic article ^Atām in contrast to stressed demonstrative PIE *tod > ^Bte ‘this, it’.
- PIE *tod u > CToch. *tæ u > *tǎ-u > ^{AB}tu ‘this, it’ (Hackstein 2007: 135 n.7).

3.2.4.1. PIE *(H)i word-initially > CToch. *iä- > ^Byá, yä-, ^Ayä-

- PIE *h₁ités: ^Byacer, ^Ayäc ‘you [pl.] (will) go’; *h₁i-énti > *yämt(ä) > ^Byam, ^Ayiñc ‘they (will) go’.
- PIE collective *h₁éitōr → *h₁itōr > *itōr > CToch. *yätar-(yæ)- > ^Bytárye, ^Aytār ‘way’; cf. Lat. *iter*.

3.2.4.2. PIE *i between non-syllabics except after *u* > CToch. *(i)ä > ^Bá, ä/Ø, ^Aä

- PIE *lejimon-, *limn- > ^Blyam, ^Alyäm ‘lake’ : Gk. λεμῶν ‘meadow’, λίμνη ‘pool, (marshy) lake’ (Ringe 1996: 109, 125).
- PIE inanimate *ni-ǵ^hutó-m ‘godly power, libated on earth (cf. Gk. χυτή γαῖα) > CToch. *n’ä-kätæ- > ^Bñákte, ^Añákt ‘god’, whose inflectional behavior indicates a neuter noun (Hilmarsson 1989: 48 f.); likewise in Germanic, neuter PIE *ǵ^hutóm, collective *ǵ^hutéh₂ > Gmc. *guða- neuter sg./pl. > German *Gott* (Hackstein 2006: 100–103).

3.2.4.3. PIE *si- > CToch. *sä- > ^{BA}s(ä)

- PIE *osilo- ‘ash-tree’ > ^Basale, ^Aasäl ‘post’, cf. Lith. *úosis*, Russ. *jasen’*, Lat. *ornus*, OIr. *uinnius* (Pinault 2009: 241).
- PIE *luk-si- ‘shimmering fish’ > *läksä- > ^Blaks ‘fish’, cf. Gk. λεῦκος, OIcel. *lýr* (Pinault 2009: 241 n.74).

3.2.5.1. PIE *(H)u word-initially > CToch. *uä- > ^Bwá-, w(ä)-, ^Awä-

- PIE *h₁ur- > *ür > ^Bwár, ^Awär ‘water’, whose genus alternans points to an old neuter, cf. PIE *h₁ur- > *uh₁r- as in ON *ūr* ‘drizzle’, Lat. *ūrīnāre/-ī* ‘to dive, become submerged’, preserving *ūrīnā* in its archaic sense ‘water’. PIE *ueh₁-r > Luv. *wār*, Ved. *vár*, OIr. *fír* ‘milk’ (Watkins 1985: 403).
- PIE *h₂us- > *wäs- > ^Bwäs- ‘stay, dwell, rest’, Prs. IX 3sg. act. ^Bwsaššäm.
- PIE *uh₂g- > CToch. *wak- > ^{AB}wak- ‘to split, burst’; cf. Gk. ἄγνουμι ‘I break’, κυματωγή (*uoh₂g-) ‘place where the waves break, beach’, Hitt. *wáki* ‘bites (off)’.

3.2.5.2. PIE *(H)u between non-syllabics > CToch. *ä > ^Bá, ä/Ø, ^Aä/Ø

- PIE *h₁rud^hrós > CToch. *rätré > ^Bratre, ^Artär ‘red’ : Lat. *ruber*, Gk. ἐρυθρός, Ved. *rudhirá-*.
- PIE *d^hubrós > CToch. *täpré > *täpáre > ^Btäpre, ^Atpär ‘high’ ; cf. Lith. *dubùs* ‘deep, hollow’, *d^heubo- > Goth. *diups*.

3.2.6. Long vowels

3.2.6.1. PIE **eh₂* / *_C*, **ǎ* > CToch. **ǎ* > ^B*o*, ^A*a*

- PIE **b^hreh₂tēr* > **b^hrātēr* > CToch. **prāt^{er}* > ^B*procer*, ^A*pracar* ‘brother’.
- PIE **h₂ues-*, **h₂us-* with laryngeal metathesis (cf. below 3.2.6.5 ad ^B*musk-*) > **uh₂s-*, whence **ueh₂s-tu-* > CToch. **uǎstä* > ^B*ost*, ^A*wašt* ‘house’: Ved. *vástu* ‘house’.
- PIE acc. sg. fem. **h₂al-ia_h₂-m* > **alyām* > ^B**alyoñ-k* > *alyok*, *allok*, ^A*alyak-äm* ‘another (one)’.
- PIE acc. pl. fem. **h₁al-ia_h₂-ns* > **alyān* > ^B*alloñ-k=na*.
- PIE **d^hoHnah₂* > **d^hōnā* > CToch. **tanǎ* > ^B*tano* ‘corn’: Ved. pl. *dhānāḥ* ‘grain’, Lith. *dūona* ‘bread’ (Hilmarsson 1986: 18, Ringe 1996: 93).
- PIE collective **h₁rud^h-reh₂* > feminine plural ^B*rätro-na* ‘red things’, with additional neuter plural marker *-na* from enclitic neuter plural article **na(h₂)* < **neh₂* of the PIE demonstrative stem **eno-*, *ono-*, *no-* ‘that one’ (Pokorny 1959: 319–321).

Denominative, factitive suffix **-eh₂ie-* (as in Lat. *novāre*) > **-ǎiä-* > **-ǎ-* > ^B*-o₂-* (triggering *o₂*-umlaut, cf. 3.2.10.2 below), ^A*-a-*, e.g.:

- PIE **h₁arto-* → denominative **h₁art-eh₂ie-* > **art-o₂-* > ^B*orttotär* ‘approves’ (Malzahn 2010: 401n.57).
- PIE **spord^huo-* → denominative **spord^hu-eh₂ie-* > **spertw-o₂-* > ^B*sporttotär* ‘turns’.

Apparent exceptions to 3.2.6.1 are either due to analogy (a) or to the operation of sub-rules (b):

- a) PIE **meh₂tēr* > CToch. **mǎt^{er}* > ^B**mocer* → *mācer* ‘mother’ with *-ā-* after *pācer* ‘father’.
- b) PIE **-eh₂* > **-a(h₂)* in pausa with loss of final *h₂* (Kuiper 1955) > CToch **-a* > ^B*-a*:
 - PIE **séh₂* > **sá(h₂)* > **sá* > ^B*sǎ* ‘she’ (generalized pausa form, Ringe 1996: 5, 94, see below 4.1.7.1); the dichotomy between PIE pausa and non-pausa variants such as **sa* and **sah₂* among demonstratives is not without parallels, cf. pausa PIE **sos* ‘he, this one’ versus non-pausa PIE **so*.
 - PIE **h₂ál-ia(h₂)* > ^B**álya* > *alyǎ-k* f. ‘another (one)’.

3.2.6.2. PIE **eh₁* and **ē* before consonant > CToch. **ye* > ^B(*y*)*e*, ^A(*y*)*a*

- PIE **-tēr* in kinship terms > CToch. **-t^{er}* > ^B*-cer*, ^A*-car*.
- PIE **eh₁* and **ē* > ^B*a* in absolute final position in monosyllables, e.g. PIE **meh₁/*mē* > ^B*mǎ* (prohibitive particle and sentence negation).

3.2.6.3. PIE **eh₃*, **oH*, **ō* > **ō*

- a) > ^B*ǎ*, *a*, ^A*ā*, *a* word-internally between non-syllabics and finally before *_r*
 - PIE **d^hoHnah₂* > **d^hōnā* > CToch. **tanǎ* > ^B*tano* ‘corn’, see above 3.2.6.1.

- PIE **ĝneh₃-* > ^A*knā-* ‘recognize, know’, ^A*knānat* ‘you recognize’, ^B*nā-* in ^B*nanātār* ‘appears’ (< **‘becomes known’*).
 - PIE **ġneh₃-t-ih₂* ‘ignorance’ > CToch. **æn-knātsā* > ^B*aknātsa*, ^A*āknats* ‘ignorant one’ (abstract noun → adjective, cf. Hackstein 2012a: 156–158, second-syllable *ā* → *a* as per 2.1.2.1 above).
 - PIE **peh₂uōr* → **ph₂uōr* > **ph₂uuōr* > CToch. **pāuar* > ^B*pūwar* ‘fire’; CToch. **puwar* > **powar* > ^A*por* ‘fire’ (Hilmarsson 1989: 135).
- b) > ^{AB}*u* in absolute word-final position
- PIE **h₂ntb^hoh₁e/*h₂entb^hoh₁e* > **an(t)pō(u)* > ^A*āmpu-k* ‘both’ (fem., Hilmarsson 1989: 56–58) : Gr. ἄμφω (masc. = fem.).
 - PIE **-oh₂* > ^B 1sg. act. ending of the athematic subjunctive *-u* : ^B*āyu* ‘I will give’ from which thematic ^B*-a-u* is analogically derived (Lane 1976: 140–142; Ringe 1991: 100n.95).
 - PIE gen. pl. masc. **h₁al-īōm* > ^A*ālu* ‘to/of/by others’ = Skt. *pareśāṃ* A353b6; *ālu nankunt* = Skt. *paragarhitāṃ* ‘scolded by others’ A354b2; *ālu kāsu yatsi* ‘to do good for others’ YQ I.3a4.

3.2.6.4. PIE **ih₁*, **T* > **T* > CToch. **i* > ^{AB}*i*

- PIE nom.-acc. dual **duih₁* plus **dékmtih₁* ‘two tens’, reduced to **duih₁dkmtih₁* in univerbation (Hackstein 2010: 61 f.) > PIE **uih₁dkmtih₁* (by early dissimilation **du₁dk₁ → *Ø_udk₁*) > PIE **uih₁kmtih₁*, by PIE **dkm̄* > **k̄m̄* (Mayrhofer 1986: 152). This in turn yields CToch. **u₁ikānt¹ä* > **u₁ikān* > ^B*ikām* ‘twenty’; CToch. **u₁ikān* > Pre-Toch. A **uikāy* > ^A*wiki* ‘twenty’. For the phonological development of **-nti* in Tocharian, cf. Szemerényi (1960: 48), Cowgill (1985: 104, 2006: 548), and Ringe (1996: 77).
- PIE optative **n^h-ih₁-* in ^B*nś-i-tār* ‘may it perish’.
- PIE **uih₁-ro-* > CToch. **u₁iræ* > **u₁ira* > ^A*wir* ‘vigorous, young, fresh’; cf. Lith. *výras*, Av. *vīrō*, Hom. Gk. (F)ἵρος (personal name), Lat. *vīr*, etc., very likely from PIE **uei^hih₁-* ‘to strive for’, for which *-h₁-* was posited on the grounds of Hom. Gk. ἕμαι (Harðarson 1993: 159 f.).

3.2.6.5. PIE **uh₁*, **ū* > **ū* > CToch. **u* > ^{AB}*u*

- PIE **m(i)uh₁-skē/o-* > ^{AB}*musk-* ‘disappear’; root PIE **mieh₁u-* ‘move (away)’; zero grade **mih₁u-* by laryngeal metathesis > **m^huh₁-*; **i-* deletion root-initially after labials (Hackstein 1995: 191, 2012b: 112 f.); cf. **m(i)uh₁-to-* > Ved. *-mūta-* ‘moved’ and **m(i)ouh₁-éie/o-* in Lat. *movēre*; PIE root noun **m^huh₁-s* ‘the [quickly] moving [and gathering animal]’ > **mūs* > Engl. *mouse*; cf. **sp^huh₁-* (Gk. πύω) > **spuh₁-* in Lat. *spuere* ‘spit’.
- PIE **nu(n)* > **nū(n)* by monosyllabic lengthening > ^{AB}*nu*.

3.2.7. Diphthongs

3.2.7.1. *i*-diphthongs3.2.7.1.1. PIE $*ei >$ CToch. $*\acute{ai} >$ ^{AB}j

- PIE $*seik-$ ‘reach by foot’, $*seik-eh_2 >$ CToch. $*s^i\acute{a}iká >$ $^B\acute{s}iko$ ‘step’; Lith. *at-siekti* ‘reach’, Gk. *ικέσθαι* ‘arrive’.

3.2.7.1.2. PIE $*oi, *ai$

- before consonant, word-initially and internally in polysyllabics $>$ CToch. $*aj >$ $^{B}ai, ^{A}e$
 - PIE $*moiuo-$ $>$ Bmaiwe ‘young’ : Oic. *mjor* ‘soft, tender’ (Ringe 1996: 84).
- word-finally in polysyllabics $>$ CToch. $*\bar{e} >$ $^{B}i, ^{A}e$
 - PIE *o*-declension nom. pl. $*-\bar{o}s \rightarrow *-\bar{o}i$: PIE $*h_1ekūōs \rightarrow *h_1ekūoi >$ $^{B}yákwi\bar{e} >$ $^{B}yákwi$ ‘horses’.
- PIE $*-oih_1 >$ ^{AB}i
 - PIE $*duoih_1 >$ CToch. $*u\bar{e} >$ ^{B}wi ‘two (m.f.)’ (Hilmarsson 1989: 36).
 - PIE $*toih_1 >$ $*to-i >$ CToch. $*t\bar{e} >$ dem. pron. dual $^{A}ti-m/ti-m$ ‘these two’ (dual of *säs/säm*, Hilmarsson 1989: 38).
- $*-eh_2ih_1 >$ CToch. $*ai >$ $^{B}ai, ^{A}e$
 - PIE $*dueh_2ih_1 >$ CToch. $*uai >$ ^{A}we ‘two (f.)’ (*we snās* ‘two women [acc. pl.]’, Hilmarsson 1989: 37) and neuter (see 3.2.4.2) $^{A}we \acute{n}äktāñ$ ‘two gods’ (A299a7) : OCS dual f. *dvě*.
 - PIE $*teh_2ih_1 >$ CToch. $*tai >$ demonstrative nom. acc. dual ^{B}tai .
 - PIE stative-middle 1sg. $*-h_2ei >$ $*-ai >$ middle 1sg. secondary ($*perfect >$ preterite) ending $^{A}-e, ^{B}*-ai$ replaced by $^{B}-mai$, e.g. *s-prt. 1sg. middle* $^{A}präks-e, ^{B}parksamai$ ‘I asked’.

3.2.7.2. *u*-diphthongs3.2.7.2.1. PIE $*eu >$ CToch. $*\acute{äu} >$ $^{AB}ü$

- PIE $*h_2ieusōr >$ $*\acute{ä}usar >$ $^{A}yusār$ ‘season’ (Pinault 1998: 362).
- PIE $*h_1ara h_2iéu(i) >$ $^{A}ar-yú$ ‘long lasting’; cf. PIE $*ne h_1ara h_2iéu >$ Toch. A *mā aryu* ‘not well for long’, $*ne h_1ara h_2oiü$ ‘isn’t it right?’ $>$ Ancient Greek ($\acute{\eta}$ ἄρα οὐ) $>$ ἄρα οὐ ‘isn’t it?’; Hitt. *natta āra* ‘it’s not right’.
- PIE $*leuk-e- >$ CToch. $*l^i\acute{ä}ukj\bar{a}- >$ $^{B}lyus-$ in subj. II inf. *lyus-si* ‘illuminate’.

3.2.7.2.2. PIE $*\bar{e}ü >$ $^{B}äu, ^{A}ö$

- PIE $*h_2er dii\bar{e}ü$ ‘on this day’ $>$ $^{A}ārśo$ ‘today’; $*h_2er$ is an *er*-locative of the pronoun $*h_2e/o-$.
- PIE *s*-aor. $*l\bar{e}ük-s-$ in *s-prt. Blyauksa, Alyokäs* ‘illuminated’.
- PIE $*neh_1-ur >$ $^{B}ñor$ ‘beneath, under’ (Hilmarsson 1991: 135).

3.2.7.2.3. PIE $*a\bar{u}$ > ${}^B au$, ${}^A o$

- PIE $*h_2euk-s-$ > CToch. $*auk-s-$ > ${}^B auk-s-$, ${}^A ok-s-$ ‘grow’ (Hackstein 1995: 342 f.).

3.2.7.2.4. PIE $*o\bar{u}$ > ${}^B au/e_u$, ${}^A o$

- PIE nom. $*g^uou-s$ > ${}^B ke_u$ ‘cow’ (Katz 1997: 79) as in ${}^B kaur\bar{s}e$ ‘bull’ (cf. above 2.1.1.1 Note 1), with unequivocally attested short diphthong; PIE $*g^u\bar{o}u-s$ ought to have yielded ${}^B \dagger ku$.
- PIE $*h_1euk-$ ‘familiarize oneself, learn’, $*h_1ouk-m\bar{n}$ > ${}^A ok\bar{a}m$, dat. $*h_1ouk-(m)n-e\bar{i}$ > $*auk\bar{a}n\bar{a}$ > $*auk\bar{a}y$ > ${}^B auki$ ‘attention’, cf. A215b1 *okām pātstsār* ‘learn your lesson, be cautious!’; cf. OCS *učō* ‘I teach’, Arm. *owsanim* ‘I learn’, Ved. *ucyasi* ‘you are used to’.

3.2.7.3. PIE long diphthong reduction in absolute final position

A PIE diphthong $*-V:\bar{U}$ with acute long vowel drops $*\bar{U}$ in absolute final position, yielding $*-V:$, cf. Gk. $\pi\epsilon\theta\acute{\omega}$ ($*-\acute{o}\bar{i}$), Ved. *sákhā* ($*-\acute{o}\bar{i}$), Hitt. collective *udnē* (< PIE collective $*-n\acute{e}\bar{i}\bar{i}$ < $*ud-ne\bar{i}-h_2$, see Oettinger 2004: 168 f.). (By contrast, circumflex long vowels $*-\bar{V}:$ are exempt from this reduction rule, cf. dat. sg. $*-\acute{o}\bar{i}$ > Gk. $-\omega$.)

- PIE collective $*h_2es-te\bar{i}-h_2$ > $*h_2es-t\bar{e}\bar{i}$ ‘bony material, skull’ > ${}^B \acute{s}ce$ ‘head’; cf. Hitt. *haštāi* ‘bone’.
- PIE $*b^h\bar{h}_1-t\acute{e}\bar{u}$ > $*b^h\bar{h}_1-t\acute{e}$ > ${}^B pl\bar{a}ce$, ${}^A pl\bar{a}c$ ‘speech’, PIE $*b^h\bar{h}_1-tu-ns$ > acc. pl. ${}^B pla-t\bar{a}m$; cf. PIE $*b^h\bar{h}_1-eh_1-$ > Pruss. *billīt* ‘say, speak’ (Schaffner 2001: 490).

3.2.8. Laryngeals

3.2.8.1. Laryngeals between non-syllabics and postconsonantly in absolute word-final position > CToch. $*a$ > ${}^B \acute{a}$, a and ${}^A \bar{a}$, a , \bar{a} , \emptyset

- PIE middle participle $*-mh_1no-$ > CToch. $*-man\bar{c}e$ > ${}^B -mane$, ${}^A -m\bar{a}m$: Gk. $-\mu\epsilon\nu\acute{o}\varsigma$.
- PIE $*ph_2\bar{t}er$ > CToch. $*pat\bar{t}er$ > ${}^B p\acute{a}cer$, ${}^A p\bar{a}car$ ‘father’.

The interconsonantal a -vocalization of laryngeals resulted in the Tocharian dichotomy of A-roots and non-A-roots. Like Greek, Old Indic, and other Indo-European languages, Tocharian reflects the etymological contrast of roots ending in a laryngeal and roots not ending in a laryngeal.

Contrast	PIE	Greek	Old Indic	Tocharian
*-h ₁ #	*d ^h h ₁ -	θε-	<i>dhi-</i>	^B tā-, <i>tasém</i> ‘we put’
*-θ#	*h ₁ lud ^h -	τίθεμεν ἐλυθ- ἤλυθον	<i>dadhiṣvá</i> <i>rudh-</i> <i>ródhanti</i>	^B lät-, <i>lateṃ</i> ‘they leave’
*-h ₂ #	*ġerh ₂ -	γερα-	<i>jarⁱ-</i>	^B srā- ‘grow old’
*-θ #	*b ^h er-	γέρας φερ- φέρουσι	<i>jarimán-</i> <i>bhar-</i> <i>bháranti</i>	^B srāy ‘adult men’ ^B pār-, <i>páreṃ</i> ‘they carry’
*-h ₃ #	*ġneh ₃ -, *ġnh ₃ -	γνω-	<i>jñā-</i>	^B nā-, ^A knā- ‘know’
*-θ#	*men-, *mñ-	γινώσκει μεν- ἔμεινα	<i>jñāti</i> <i>man-</i> <i>mamandhi</i>	^B nānātār ‘appears’ ^B māsk- ‘become, be’ (< *mān-sk-)

While this distinction was faithfully preserved in the inherited lexicon, the *a*-character in particular tended to be morphologized as a stem vowel in productive verbal morphology and thus came to be generalized beyond its etymological source domain (cf. Hackstein 1995: 24 ff.; Peyrot 2013: 40 ff.).

3.2.8.2. Laryngeals word-initially before consonant

Word-initial laryngeals before nonsyllabics and /i, u/ are lost in Tocharian (Ringe 1996: 13-19):

- PIE *h₁s-ske-tor > copula ^Bstar ‘is’, cf. Gk. ἔσκετο ‘was’.
- PIE *h₂stēr → *h₂ster-jo- > *s^htār^hae- > ^Bsciryē ‘star’: Gk. ἀστήρ, Arm. *astl*.
- PIE *h₃reġ- > ^{AB}rāk- ‘stretch out’: Gk. ὀρέγω ‘I stretch out’.

Laryngeal loss occurs also in *Hi- (3.2.4.1) and *Hu- (3.2.5.1).

3.2.8.3. Initial shwa secundum [ə₂] yields *a*- in Tocharian

- PIE *h₂ə₂ġ-s- > AB *āk-s-* ‘proclaim; cf. PIE *h₂ġ-je/o- > *h₂ə₂ġ-je/o- > Lat. *aiiō* ‘I proclaim’ (Hackstein 2012b: 111 f.).
- PIE *h₂eh₁- ‘be hot’ (Pal. *hāri*, *hānta*) → inchoative *s*-present PIE *h₂eh₁-s- ‘dry out’, zero-grade PIE *h₂h₁s- > *h₂ə₂h₁s- > ^Bas- ‘dry’; cf. Hitt. *hašša-l/haššā-c*. ‘hearth’, Lat. *āra* ‘hearth, altar’, *ārēre* ‘be dry’.

3.2.8.4. PIE **ih₁*, **T* > **T* > CToch. **j* > ^{AB}*j*, cf. 3.2.6.4 above

3.2.8.5. PIE **-ih_{2,3}* > CToch. **-ja*

- PIE **g^uieh₃-*, **g^uih₃-* ‘live’ > Toch. ^{AB}*śā-* ‘live’, e.g. PIE **g^uih₃-ue-ti* > CToch. **k^uja-_uä-* > ^B*śai-m*; PIE **g^uih₃-uo-nti* > CToch. **k^uja-_uæ-* > ^B*śāwem* (> ^B**śā-*, cf. Ringe 1996: 156), ^A*śaweñc* (*śawa-*) ‘live’ : OCS *živŏ*, Lat. *vīvō*, Ved. *jīvati*.
- PIE **ĵeh₂-*, **ih₂-* > ^{AB}*yā-* ‘drive’, e.g. prs. V 3sg. act. ^B*iyam*.

3.2.8.6. PIE **uh₁*, **ū* > **ū* > CToch. **u* > ^{AB}*u*, cf. 3.2.6.5 above

3.2.8.7. PIE **-uh_{2,3}* > CToch. **-wa*

- PIE **geuh₂-*, **guh₂-* ‘call’: ^B*kwā-*, prs. V *kwātār* ‘calls, invites’; cf. Ved. *jōguve* ‘keeps invoking’; Gk. *γοῶν* ‘groan, weep’.
- PIE collective **luH-s-h₂* ‘little animals’ > ^B*lwāsa* ‘animals’ (Hilmarsson 1989: 119).
- PIE **mleuh_{2,3}-*, **mluh_{2,3}-* ‘speak’: ^B*pāl-wā-* ‘complain’, e.g., prs. V ^B*palwam*.
- PIE **suh_{2,3}-i_{us}* > CToch. **suaiä(s)* > Pre-Toch. B **suoiä* > *soy* ‘son’ (Hilmarsson 1989: 23; Ringe 1996: 61–63); > Pre-Toch. A **suaiä* > ^A*se* (cf. Doric Gk. nom. *υῖός*, Doric/Hom. gen. *υῖός*).

3.2.8.8. Laryngeal loss in PIE

- PIE **geh₂* + *d^heh₁-*, and **geh₂* + preverb *ui* + *d^heh₁-* ‘to set into beaming, make rejoice’; **gh₂d^h-ske/o-* > ^B*kātk-* ‘to rejoice’.
- PIE **dh₃e-dh₃ġ^h-uōn* > **de-dθġ^h-uōn-* > CToch. **tsātsk^uu-* > ^B*tsatku* ‘oblique, queer, erroneous’ (Pinault 2006: 148 f.); cf. Ved. *jihmá-* ‘oblique, athwart’ (EWAia I 590f); PIE *dh₃ġ^h-mó-* > Gk. *δομός* ‘across, aslant’.
- PIE **p^lth₂-s^ho-mh₁no-* > Prs. II ^B*plyetkemanē* to ^B*plātk-* ‘to surface, arise, swell’ (Hackstein 2002: 8).

3.2.8.9. Laryngeals word-finally after consonant

Word-final laryngeals after consonants are reflected as Toch. B *-a/ā* but are lost in Toch. A due to word-final apocope. The best example is the collective laryngeal suffix PIE **-h₂* in the consonantal declensions, whose phonology and function (number-indifferent use) represent a productive archaism of Tocharian. The examples fall into the following four categories:

a) PIE collective > Toch. singulative

- PIE **h₂eus-* ‘be bright, shimmer, gleam’, before collective *-h₂* resyllabified as state II root **h₂ues-h₂* ‘a shimmering, gleaming’ (cf. **h₂eug-* → **h₂uek-s-*; **derk-* →

- **drék-s-*) > CToch. **üäs-a* > nom. sg. ^B*yasa*, ^A*wäs* ‘gold’, gen. sg. ^B*ysāntse*; cf. PIE **h₂e-h₂us-om* > Lat. *aurum* (Driessen 2003; Oettinger 2012: 244).
- PIE **g^uen-h₂* collective ‘family association of women’ > ^B*śana* ‘woman as family member, wife’. For the palatalization see 4.1.4 (g) below and for the metonymic shift cf. PIE **uid^heueh₂* ‘family of slain hero’ → Lat. *vidua* ‘widow’.
- b) PIE collective > Toch. number-indifferent
- PIE collective **mems-h₂* > CToch. **m’ämsa* > ^B*mīsa* ‘flesh’; cf. 3.2.2.3 above.
 - PIE **még^h-h₂* in ^{AB}*māka* ‘a lot, much’, sometimes translating Skt. *bahu* ‘a lot, much’ (Hackstein 2012a: 154 f.).
- c) PIE collective > Toch. plural, when in productive paradigmatic number contrast:
- PIE collective **h₂ost-h₂* or **h₂est-h₂* ‘bony material, bones’ (Katz 1997: 72) > ^B*āsta* ‘bones’, singular ^B*āy* ‘bone’; cf. singular Ved. *āsthi*, plural Yav. *asti*.
 - PIE collective pl. (*-*h₂*) in *pwār-a*, singular ^B*pūwar* ‘fire’, dual (PIE *-*ih₁*) *pwār-i* ‘two fires’ (Hilmarsson 1989: 112 f.).
- d) Singulative use alongside coreferential singular/plural pronouns and despite paradigmatic contrast:
- PIE **ueh₂stu-h₂* (cf. 3.2.6.1) > ^B*ostwa*, ^A*waštu* ‘aggregate of houses, dwelling place, palace’, of individuals, e.g., *toṣ upādhyāy waštu* ‘this dwelling place of the teacher’ (YQI.5a2), *sās bhādhari brāṃne waštu* ‘this dwelling place of Bādhari the Brāhmin’ (YQI.7b2), cf. the parallel use in Myc. *wa-tu-wa /wastúw-ā/* ‘municipal district’, collective of *wa-tu /wastu/* (Leukart 1994: 155).

3.2.9. Syllabic liquids and nasals

3.2.9.1. PIE **r*, **l* between non-syllabics and in *auslaut* > CToch. **är*, **äl*

- PIE **k^ueturtós* (with dissimilatory loss of *-u-* in the second syllable) > CToch. **k^uät(w)ärtæ* > ^B*śarte*, ^A*śärt* ‘fourth’.
- PIE **b^hl^g-*, prs. **b^hlég^{ti}* → **b^hl^g-ti* > prs. I 3sg. act. ^B*palkām* ‘shines’ : Lat. *fulgeret /fulgēre*.

3.2.9.2. PIE **m*, **ŋ* between non-syllabics > CToch. **ām*, **än*

- PIE **dḱm-t-óm* > CToch. **kāntæ* > ^B*kante*, ^A*kānt* ‘hundred’.
- PIE **dn^gh^uah₂* → **d^hŋ^huh₂* > CToch. **tānk^uā* → (tabu deformation) **kānt^uā* > ^B*kantwo*, ^A*kāntu* ‘tongue’; cf. Osc. *fangvam* (Ringe 1996: 45–47).

3.2.9.3. **ŋ* word-initially > CToch. **æn-*

- PIE privative **ŋ-* and locative-comitative PIE **h₁ŋ-* yield CToch. **æn-* > a) B *e(n)-*, b) B **a(n)-* by *a*-umlaut (3.2.10.1), and c) B **o(n)-* by labial assimilation before

labial and B *-o-* in the following syllable (for the conditioning factors, see Hilmarsson 1991: 12–14, 192–198).

- a) PIE **ŋ-* before velar: ^B*eñ-klyaušätte* ‘unheard of’.
- PIE **h₁ŋ-pot-ih₂*, **h₁ŋ-pot-ieh₂m* > ^B*epetsa*, *epetso* ‘engaged to a husband’.
 - PIE **h₁ŋ-prot-* > ^B*eprete**, voc. *epréta* ‘resolute, bold’; cf. OHG *frad* ‘efficient’ (Schaffner 2001: 292).
- b) PIE **ŋ-ġneh₃-t-ih₂* ‘ignorance’ > CToch. **æñ-knātsā* > ^B*aknātsa*, ^A*āknats* ‘ignorant one’, see 3.2.6.3.a.
- c) ^B*omprotärtse* ‘in brotherly relation’ (Hilmarsson 1991: 134).

3.2.10. Umlaut

3.2.10.1. *a*-umlaut

CToch. **æ* and **je* (< PIE **o*, **ē*) undergo umlaut to **a* (> ^{AB}*ā*) before **a* in the following syllable (cf. Ringe 1996: 160–163).

- PIE **penk^ue=dekōm(t)* by reduction in univerbation → PIE **penk^ue=dekōm(t)* > **penk^ue=h₁kōm(t)* (as per Kortlandt 1983 or by compensatory lengthening) > **-ē-k̂-* > **pⁱānkⁱe-ka* > ^B*pišāka* ‘fifty’.
- PIE **próti_h3k^ueh₂* > CToch. **prétⁱak^uā* > **prátsakā* > ^B*pratsáko* ‘breast’.
- PIE **H₂ġ^u-* > ^{AB}*yok* ‘hair’ (cf. 3.2.10.4.b.), collective plural PIE **H₂ġ^u-h₂* > ^B*yák-wa* ‘hair’.

In Tocharian B, *a*-umlaut is independent of the word accent. In Tocharian A, however, *a*-umlaut seems to affect unaccented CToch. **æ*, **e* only. Therefore, no *a*-umlaut occurs in the first syllable of ^A*pratsak* ‘breast’ from PIE **próti_h3k^ueh₂* (*a*-umlaut would have yielded ^A**prátsak*), whereas the B-form shows *a*-umlaut (^B*pratsáko*).

Subjunctive I: stem in <i>-Ø</i> , hence no <i>a</i> -umlaut in the strong stem	Subjunctive V: stem in <i>*-a</i> , triggering <i>a</i> -umlaut in the strong stem
Strong stem <i>*prok̂-</i> > <i>preku</i> ‘I shall ask’ Weak stem <i>*prk̂-ih₁-</i> > <i>parši</i> ‘may (s)he ask’	<i>*(s)k̂odh₂-</i> > <i>kātaṃ</i> ‘(s)he, it will scatter’ <i>*(s)k̂dh₂-ih₁-</i> > <i>katoytār</i> ‘(s)he may scatter for herself/himself’ <i>*kors-a-</i> > <i>kārsau</i> ‘I shall know’ <i>*krs-a-ī-</i> > <i>kārsoym</i> ‘may I know’

3.2.10.2. Tocharian B *o₂*-Umlaut

The suffix of Prs. IV B *-o₂-* most likely reflects the contraction product of **-eh₂-je-* and causes a preceding root vowel to assimilate to ^B*o*, i.e., **a-o₂* > ^B*o-o₂*, which is a productive rule of Tocharian B (cf. Ringe 1996: 119–124):

- PIE **h₂h₁s-* (3.2.8.3) > **h₂ə₂h₁s-* > ^B**as-o₂tār* > ^B*osotār* ‘is drying out’.
- PIE **uh₂g-* (3.2.5.1) > CToch. **wak-*: ^B**wako₂tār* > ^B*wokotār* ‘is split, breaks (apart)’.

3.2.10.3. Pre-Tocharian B *o*₁-umlaut

Pre-Toch. B **u* and **æ* are assimilated to Pre-Toch. B **o*₁ (PIE **ōn* in final syllables) in the following syllable.

- a) **u-o* > ^B*o-o*: PIE **teuḗpōn*/**teub^hōn* > **cuwo* > **cowo*, obl. *cowai* ‘theft’; Gmc. **þiubaz* ‘thief’ (Hilmarsson 1991: 179).
- b) Pre-Toch. B **æ-o* > ^B*o-o* (Ringe 1996: 163), and **ä-o* > ^B*o-o* : PIE **g^hrud-n-iHōn* ‘grit’ > **kruñiyō(n)-* → **æn-kruñiyō(n)-* ‘made of grit’ > **onkrāñō* > **onkārño* > *onkorño* ‘sweet milk-and-rice porridge’, cf. OHG *gruzzi*, Germ. *Grütze* (Hilmarsson 1991: 137 f.).

3.2.10.4. *u/w*-umlaut

CToch (a) **æ* (< PIE **o*) and (b) **^je* (< PIE **ē*) undergo raising to **o* > ^{AB}*o* if followed by **u* or **ū* in the following syllable.

- a) PIE **dōru* ‘wood’ → **ōru* > CToch. **æru* > **oru* > **orä* > ^{AB}*or* ‘wood’ : Gk. *dórv*, Ved. *dāru*, Hitt. *tāru*.
- PIE **oktō(u)* > CToch. **æktū* > **oktu* → **ok(ä)t-än* (*-än* analogically after ‘9’ and ‘10’) > ^B*okt*, ^A*okät* ‘eight’ (cf. Rasmussen 1989: 123 f.).
 - PIE **mód^hu-* ‘sweetness’ > ^B*mot* ‘alcohol’ (Pinault 2003: 177 f.; on ‘sweetness’ → ‘sweet drink’, *ibid.* 179); **méd^hu-* (> ^B*mit* ‘honey’, see 3.2.2.3 above).
- u*-umlaut is triggered by the ending of the subj. class I, 1sg. act. ^B-*u*, e.g. ^B*yopu* ‘I shall enter’, *yoku* ‘I shall drink’.
- b) PIE **H₁ǵ^u-* > ^{AB}*yok* ‘body hair’; cf. Gk. *ἦβη* ‘youthful prime, youth’, Lith. *jėgà* ‘strength, vigor’ (Winter 2005: 225–228).
- PIE **seh₁-ntu-* > **sēntu-* > ^A*šont-* ‘road’; shortened by Osthoff’s Law in Celtic and Germanic: **sentu-* > OIr. *sét*, Welsh *hynt*, and Gmc. **sénþo-* > OHG *sind* ‘way’ (Hilmarsson 1986: 24–27).

4. Non-syllabics

4.1. Obstruents and Sonorants

As a result of the palatalization before mid and high front vowels, the Tocharian obstruents and sonorants exhibit palatalized allophones. The resulting allophony frequently persists synchronically in paradigmatic allomorphy (cf. Ringe 1996: xix–xxi; Winter 1998: 156 f.; Pinault 2008: 417–420; Malzahn 2010: 4 f.; Peyrot 2013: 69 ff.). The basic allophones are:

Overview of palatalization and resultant allophony of single consonants

	primary palatals	secondary palatals	Examples
<p> [p]	<p> [p ^j]	<py>	^B <i>spārtt-</i> : prt. II 1sg. act. <i>spyārta</i> ‘I turned’
<t> [t]	<c> [tʃ]	–	^B <i>pācer</i> , acc. <i>pātār</i>
<ts> [ts]	<ś> [ʃ]	<tsy>	^A <i>śām-</i> ‘grow’: prs. IV <i>śamantār</i> ‘they grow’ ^B <i>tsār-</i> : prt. II <i>tsyāra</i> ‘(s)he separated’
<k> [k]	<ś> [ʃ]	<ky>	^B <i>ākem</i> ‘they lead’ (PIE <i>*h₂eǵontī</i>): ^B <i>āsām</i> ‘(s)he leads’ (PIE <i>*h₂eǵeti</i>) ^B <i>kān-</i> : subj. V <i>kyānamar</i> ‘I shall fulfill’, ^B <i>ku-</i> : prt. II 3pl. act. <i>kyauware</i> ‘they let pour’ ^B <i>m’ütä</i> > ^B <i>mit</i> ‘honey’, see 3.2.2.3 ^B <i>mārs-</i> ‘forget’: prt. II 3sg. act. <i>myārsa</i> ^B <i>klin-</i> : opt. <i>kliñi</i> ‘(s)he ought to’ ^B <i>rātārya</i> f. ‘red’ ^B <i>-lle</i> , e.g., <i>āsalle</i> ‘to be led’, prt. II 3sg. act. <i>lyāma</i> ‘let sit’ ^B <i>yente</i> , ^A <i>want</i> ‘wind’
<m> [m]	<m> [m ^j]	<my>	^{AB} <i>sik-</i> : subst. ^B <i>siko</i> , ^A <i>şik</i> ‘step’
<n> [n]	<ñ> [ɲ]	–	
<r> [r]	<ry> [r ^j]	<ry>	
<l> [l]	<ly> [ɭ], <ll> [ɭɭ]	<ly>	
<w> [w]	B <y> [j], A <w> [w ^j]	–	
<s> [s]	<ś> [çʔ]	–	

The PIE system of stops underwent major reductions in Proto-Tocharian, first by the centum-type merger of PIE palatals and velars, and second by the collapse of the three PIE manners of articulation (unvoiced, voiced and voiced aspirated) into deaspirated voiceless articulation, which is unique in Indo-European. As regards the relative chronology of the deaspiration and devoicing of stops, the contrasting reflexes of PIE **t* and **d^h* as opposed to PIE **d* (see 4.1.3.1, 4.1.3.1.3) suggest that, at least in the case of the dentals, deaspiration and devoicing postdate the affrication of PIE **d* > CToch. **ts*.

4.1.1. Labials

- PIE **p*, *b*, *b^h* > ^{AB}*p* with palatalized allophone *p^j*.
- PIE **p* : PIE **pek^{u-}* ‘cook’: ^{AB}*pāk-* ‘ripen’; PIE **suepno-* > CToch. **šāpnæ* > **šāpānæ* > **špānæ* > ^B*spāne*, ^A*špām* ‘sleep’.
- PIE **b* : PIE **d^hubró-* > CToch. **täpræ* > **täpǎræ* > ^B*tāpre*, ^A*tpār* ‘high’.
- PIE **b^h*: PIE **b^hreh₂tēr* > **b^hrātēr* > CToch. **prāt’er* > ^B*procer*, ^A*pracar* ‘brother’.

4.1.2. Velars **k*, *g*, *g^h* and palatals **k̂*, *ġ*, *ġ^h* > CToch. **k* > ^{AB}*k* with palatalized allophone *ś*

- PIE **k* : PIE **leuk-* : ^{AB}*luk-* ‘shine, illuminate’; prs. ^B*luksām*, infinitive *lyuśsi* : Lat. *lūcēre*, Gk. *λεῦσσω* ‘I look’, Ved. *rócate* ‘shines’.
- PIE **g* : PIE **d^hugh₂tēr* > CToch. **tākat’er* > **tkat’er* > ^B*tkácer*, and with assimilated onset *c-* ^A*ckācar* ‘daughter’.

- PIE *g^h: PIE *serg^h- : ^Bsark, ^Asärk ‘illness’ : OHG *sorga* ‘care, sadness’, OIr. *serg* ‘sickness’, Church Slavic *sragŭ* ‘fearsome’, Lith. *sergù* ‘am sick’.
- PIE *k̂ : PIE *kléuos > CToch. *kl’äüæ > *käl’äüæ > ^B(ñem-)källywe, ^A(ñom-)klyu ‘glory’ : Gk. κλέος, Ved. *śrávas*.
- PIE *ǵ : PIE collective *h₂mélǵ-uōr, *h₂melǵ-uor-ŋ > ^Bmáلكwer ‘milk’; PGmc. *meluk- ‘milk’ (Griepentrog 1995: 301 f.); PIE *h₂melǵ- ‘squeeze (out), milk’ in OIr. *mligid*, Gk. ἀμέλω, Ved. *márjmi*. PIE *h₂molǵ- in ^Amalk-e ‘milk’; *máلكant kowi* ‘the cows were squeezed/milked’ (A63b5); Gk. (ἐν νυκτὸς) ἀμολγῶ ‘(in twilight’s) milk’ (Watkins 2009). PIE *ǵerh₂-on-es → *ǵerh₂-n-es > *śáráñä > *śáráy* (4.1.5.3.1) > ^Bśrāy ‘adult men’ (Carling 2003: 84 f., 92 f.; Pinault 2008: 484).
- PIE *ǵ^h: PIE *ǵ^heu- : ^{AB}ku- ‘pour’, prs. VIII ^Bkušäm ; ^A prt. III *śosā-ŋ* : Gk. χέω, Ved. *juhóti* ; PIE *b^hǵ^h-ro- > ^Bpärkäre, ^Apärkär ‘long’.

4.1.3. Dentals

4.1.3.1. PIE *t, d^h > ^{AB}t with palatalized allophone c

4.1.3.1.1. PIE *t > CToch. *t > ^{AB}t with palatalized allophone c

- PIE nom.-acc. neut. sg. *tod > ^Bte : Lat. *istud*, Engl. *that*.
- PIE *ph₂tēr > CToch. *pa^hter > ^B nom. sg. *pācer* ‘father’, obl. sg. *pātär*.
- PIE *kuH-ti-, oblique *kuH-tej- > CToch. *k^uā-^häy- > oblique case form ^Akāc ‘skin’ (instr. *kāc-yo*, A147b4): OHG *hūt*, Germ. *Haut*, Lat. *cūtis* (Hilmarsson 1986: 236).

4.1.3.1.2. PIE *d^h > CToch. *t > ^{AB}t with palatalized allophone c

- PIE *sperd^h-, *sprd^h-, *spord^h-u-, *spord^h-u-o-, *spord^h-u-eh₂- > ^Bspärttā-, ^Aspärtwā- ‘to turn’; cf. OHG *spurt* ‘competition, stadium’, Ved. *sprdh-* ‘competition’ (Griepentrog 1996: 373 f.).

Prs. VIII 3pl. act. ^Aspärtwseñc ‘they turn (tr.)’, prs. IV 3sg. ^Aspārcwš=äm ‘it turns up for him, appears’.

- PIE *h₁lud^het > *h₁lud^hed > CToch. *lät’ä > ^Blac ‘(s)he, it left’.
- PIE *h₁rud^hrós > CToch. *rät^hræ > ^Bratre ‘red’.
- PIE *d^hugh₂tēr > CToch. *täka^hter > *tka^hter > ^Btkācer, ^Ackācar ‘daughter’.

4.1.3.1.3. PIE *d before vowel > ^{AB}ts with palatalized allophone ś

- PIE *duh₂- ‘be apt, fitting’: B *tswetär* ‘it fits’; cf. Goth. *taujan* ‘make ready, do’, Gr. δύνασθαι ‘be able’ (Ringe 1996: 31).
- PIE *deuk- ‘pull’: prs. VI Inf. ^Atsäknätsi ‘pull’, impf. 3pl. act. *śākant* ‘they pulled out’.

- PIE **dem(h₂)-* > *ˆtsäm-* ‘grow’, prs. IV 3pl. mid. *ˆsasantär* ‘they grow’.
- PIE **der-* ‘split’ > prs. III 1pl. act. *ˆtsaramäs* ‘we separate’, *śralune* ‘separation’.
- PIE **dek̑m̑* > CToch. **ts jäkän* > **ts jäkä* > *ˆsāk*, *ˆsäk* ‘ten’.

4.1.3.1.4. Secondary **d* from PIE **d^h* by dissimilatory loss of aspiration > CToch.
**ts-* (Ringe 1996: 47)

- PIE **d^heg^{uh}-* > **deg^{uh}-* > CToch. **tsäky-* > *ˆBtsäk-* ‘burn’.
- PIE **d^hejġh-* > **dejġh-* > CToch. **tsäjġ-* → *ˆBtsik-* ‘shape, fashion’.

4.1.3.1.5. PIE **d* > *ˆBt* after dental nasal

Affrication of PIE **d* fails to occur, when **d* is adjacent to nasals.

- PIE **sked-h₂-* : **skd-n-h₂-* > *ˆBkatnaṃ* ‘strews, spreads’ (Ringe 1996: 147) : Gk. σκίδνῃμι ‘I spread’.
- PIE **spend-* > *ˆBspänt-* ‘trust’ (Malzahn 2010: 968) : Lat. *spondēre* ‘vow’.
- PIE **splend-* > *ˆBplānt-* ‘rejoice’ (Malzahn 2010: 742) : Lat. *splendēre* ‘shine’.
- PIE **tund-* > *ˆBtämts-* ‘scatter, disperse’ with *-ts-* analogically restored from **tud-* > *täts-* : Lat. *tundere* ‘beat, hammer, keep hammering, pound’, Ved. *-tudate* ‘strikes, wounds, hits’.

4.1.3.1.6. PIE **d* versus PIE **t, d^h* before *ġ*

- PIE *-dġ-* > *ˆB-y-*:
- PIE **suid-je/o-* > *ˆBsya-* ‘sweat’.
- PIE **-tġ-* > *ˆB-ts-*:

Nom. sg. fem. **uġh₂-nt-ih₂* by resyllabification → **uġh₂-nt-ih₂* > **ula-nt-ġa* > **wlāntsa* > *ˆBlāntsa*, *ˆAlānts* ‘queen’.

- PIE **-nt-*: collective *-nt-ih₂* > **-ntġa* > agent noun suffix *ˆB-ntsa* (*ˆBwapāntsa* ‘weaver’), *ˆA-nts*.
- PIE ptcp. prs. act. fem. **-o-nt-ih₂* > **-entsa* → *ˆB-eñca* analogically after agent noun suffix *ˆB-ca* (< **-tiġeh₂*).

Word-internally **-tiġ-* > *ˆB-c-*:

- PIE **-tiġo-* (Lat. *nuntius*, Ved. *śrūt, ġyam*): **-tiġeh₂* > agent noun suffix *ˆB-ca*, e.g. *ˆBkār-sau-ca* ‘knower’.
- PIE **tri-to-s* > *ˆBtrite* ‘third’: fem. **tri-tiġeh₂* > **tri-ṽġa(h₂)* > *ˆBtrica*.

4.1.3.1.7. Dental assibilation and palatalization of word-final PIE *-ti, -d^{hi}

- PIE **h₁id^{hi}* > **yät^{(h)j}ä* > **yät^{sj}ä* > **yäšä* > **pä* + *yäšä* > **päš* > ^A*piš*, ^B*paš* ‘go!’ (Jasanoff 1987: 109 f.).

	Dental affrication (4.1.3.1.3)	Weakening <i>a</i> > <i>ä</i> (2.1.1.5)	Syncope
PIE * <i>pó dh₃d^{hi}</i>	> * <i>pé-tsa-t^{sj}ä</i>	> * <i>pé-[tsä]-šä</i>	> * <i>pé-šä</i> > ^A <i>paš</i> ‘give!’ (2sg.)
PIE * <i>pó dh₃te</i>	> * <i>pé-tsa-t^vä</i>	> * <i>pé-[tsä]-cä</i>	> * <i>pé-cä</i> > ^A <i>pac</i> ‘give!’ (2pl.)

- PIE prs. 3sg. act. *-*eti* > ^A*-äš*.
 – PIE *(*h₁*)*eti* ‘thence’ > ^{Bš} ‘and’, PIE *(*h₁*)*eti épi* > ^{Bšp}, ^{šáp} ‘and’ (Hackstein 2003: 186, 2007: 134).

4.1.4. Labiovelars and umlaut next to labiovelars

The development of the PIE labiovelars into Tocharian is fairly complicated; for a detailed treatment see Kim (1999). Some basic developments are:

- a) PIE **k^u*, **g^u*, **g^{uh}* > CToch. **k^u*
 b) CToch. **k^uä* > ^A*ku*, ^B*ku-* (Kim 1999: 150 f.)
 – PIE **k^uers-ŕ* → **k^uŕs-ŕ* > pl. ^B*kursawa* (analogical sg. *kwars-är*), ^A*kursär*, pl. *kurtsru* ‘mile, vehicle’; PIE **pek^u-l̥* ‘(period of) ripening’ > ^B*pikul* ‘year’, collective pl. *pikwala* (Pinault 1998: 360 f.).
 c) Rounding of *-ä-* between labiovelars occurs in: PIE **k^ue-k^uh₁-o-* > CToch. **k^uäk^uälæ* > **k^uuk^uälæ* > ^B*kokále*, ^A*kukäl* ‘chariot’ : Engl. *wheel*, Gk. κύκλος, Ved. *cakrá-* m., n. ‘wheel’ (Ringe 1987: 258).
 d) PIE **k^ui* > ^{BA}*ku-*:
 – PIE **k^uis so* > **k^uiso* > **k^uäsæ* (cf. below 4.1.5.2) > ^B*k_use*, ^A*kus* ‘who, what’.
 e) Delabialization of CToch. **k^u* before PIE *-*o-* (Kim 1999: 149 f.):
 – PIE **k^uolh₁-o-* > ^B*kele* ‘navel’ : Gk. πῶλος ‘pole’.
 f) Delabialization of CToch. **k^u* before consonant and in *auslaut*:
 – PIE **penk^utó-* > CToch. **pⁱántæ-* > ^B*pínkte* ‘fifth’.
 – PIE **uok^u-s*, **uok^u-ŋ* > CToch. **uæk^u-s*, **uæk^uän* > ^B*wek*, ^A*wak* ‘voice’.
 g) Delabialization of palatalized CToch. **k^ui* > ^{ABš} (Ringe 1996: 105, 107):
 – PIE **g^uem-e-ti* in sbj. II ^A*šmäš*, ^B*ša-n-mäm* ‘(s)he, it will come’.

4.1.5. Resonants

4.1.5.1. PIE **u* > ^B*w* with palatalized allophone *y* ; ^A*w* with palatalized allophone *w* (<**w^j*)

- PIE **g^uih₃-ue-ti* > CToch. **k^uia-uä-tsi* > ^B*ša-i-* + *-m* ‘lives’; PIE **g^uih₃-uo-nti* > CToch. **k^uia-uæ-ntä* > ^B*šā-we-m* ‘they live’.

The alternation between *-i-* and *-we-* must have become synchronically opaque, as indicated by the levelling of both variants, e.g. 3rd pl. ^B*šā-ye-m* ‘they live’. The optative, however, continues to display palatalization synchronically:

- PIE optative 3sg. act. **g^uih₃-u-o-ih₁-t* > imperfect/optative ^{B-(*o*)i}, ^B*šāyi* ‘(s)he lived’.
- PIE **h₂ueh₁-ut-o-* → **h₂ueh₁-nt-o-* > **uēnto-* (consonantal onset of *nt-*suffix restored to achieve morphological transparency) > CToch. **u^lēnta-* > ^B*yente*, ^A*want* ‘wind’.
- PIE **ueǵ^hno-* > CToch. **u^läkna-* > ^B*yakne*, ^A*wkäm* ‘way’.
- PIE *s*-stem **ueiǵ-os* > CToch. **u^läk* > ^B*ike* (*genus alternans*, *neuter) ‘place’; cf. *s*-stem → *a*-stem, in Goth. *weihs* ‘town, village’ n. (Schaffner 2001: 593).

4.1.5.2. PIE **ui-* > CToch. **uü-* > **uä* (Ringe 1996: 66)

- PIE **uiso-* > **uäsæ-* > ^B*wase*, ^A*wäs* ‘poison’.
- PIE **duito-* > **duätä-* > **uätä-* > ^B*wate*, ^A*wät* ‘second’.
- PIE **k^uis so* > **k^uäsæ* > ^B*k_use*, ^A*kus* ‘who, what’.
- PIE nom. **h₂óuis*, gen. **h₂éuios* → **h₂éuis* > CToch. **auä* > nom. sg. fem. ^B*ā_uw* ‘sheep’, Luv. *hawis*, Gk. *óis*, Lat. *ovis*.
- PIE **uid-uo-s* > **uäuæ-* > ^B*uwe* ‘learned, capable’ (Klingenschmitt 2005: 401n.103; Schaffner 2001: 273n.24).

4.1.5.3. PIE **n* > ^{AB}*n* with palatalized allophone *ñ*(*ñ*)

- PIE **nu(n)* > **nū(n)* > ^{AB}*nu* (cf. 3.2.6.5). versus **ni-ǵ^hutó-m* > CToch. **nⁱä-kätæ-* > ^B*ñäkte*, ^A*ñäkt* ‘god’ (cf. 3.2.4.2)

4.1.5.3.1. **n-* > **ñ-* > *-y-*, if between *ǣ* and *ē*

- PIE **g^uen-h₂* collective ‘family association of women’ > ^B*šana* ‘woman as family member, wife’ as opposed to individualizing *n*-stem **g^un-en-ēn* ‘[the one] endowed with womanhood; feminine; single woman’, dissimilated to **k^läne* > ^B*kliye* ‘woman’ (Schmidt 1980: 410; Hilmarsson 1996: 159). For the dissimilation of *n...n* → *l...n*, cf. *n...n* → *n...l* in Tocharian **anknatsa* → **anklatsa* > Niya-Gāndhārī *aṃklatsa* [epithet of a camel] (Burrow 1935: 673).

- PIE nom. **uérh₁-ēn* → CToch. **uér-en-ēn* > **u'äräñen* > **yäräye* > ^B*yriye* ‘sheep’ : Gk. ἀρίην, Arm. *gaṙn* ; PIE acc. **urh₁-én-ṃ* → (-ñ- analogically with the nominative) CToch. **uer-än-äm* > **u'är'änän* > **yäräyü* > obl. **yäri* > ^B*yäri* (cf. 2.1.1.3).

4.1.6. *s > š

- PIE **selpos* ‘ointment’ > ^B*šalype*, ^A*šälyp* ‘ointment, fat’ : cf. OHG *salba*, Gk. ἔλπος ‘oil, fat’.
- PIE **sem-ēn* → **sem-en-ēn* > **šämäne* (4.1.5.3.1) > **šämäy* > **šäme* > ^A*šme* ‘summer’; cf. Ved. *sámā-*, OHG *sumar* (Pinault 1998: 362).

4.1.6.1. *su-

- PIE **sué* > **s'u'ä* > ^A*šu* ‘to oneself’ → ‘hither, away’ (Hackstein 2004b: 74, 82 f.). Cf. PIE acc. **túé* ‘you’ > **t'u'ä* > ^A*cu* (Ringe 1966: 113).

4.1.7. Obstruents in word-final position

4.1.7.1. PIE loss of final laryngeals in pausa (Kuiper’s law)

Like other Indo-European languages (Kuiper 1955), Tocharian offers traces of post-syllabic word-final laryngeal loss, that occurred in PIE pausa position.

- For PIE **-eh₂* > **-ah₂* > **-ǎ*, e.g. ^B*sā* ‘she’, cf. 3.2.6.1.b. above.
- For PIE **ih₁* > **-ǐ*, cf. B *ikäm* ‘twenty’, cf. 3.2.6.4 above.
- PIE vocative dual m. **-oh₁* > **-o* > ^B*-e* as in ^B*pacére* ‘parents’ (Malzahn 2000: 45–50).

4.1.7.2. Tocharian loss of word-final obstruents

With the exception of laryngeals, *-r*, and *-l*, all PIE word-final consonants were lost by Common Tocharian. Note that word-final consonants that arose later in the Common Tocharian period were exempt from this rule. Thus, in word-final position we find the 1pl. act. ending ^B*-m* (< PIE **-mes*), which still appears with a final vowel as *-mä* or in poetic texts as *-mo*, see 2.1.1.2 (a). above.

- PIE **-s* : **h₁lud^{hes}* > ^B*lac* ‘you left’.
- PIE **-ns* > CToch. **-äns* > obl. pl. ^B*-äm*, e.g. ^B*alyenkäm*; PIE **h₁ék^{mons}* > obl. pl. ^B*yákwem* ‘horses’.
- PIE **-t*: **h₁lud^{het}* > **h₁lud^{hed}* > ^B*lac* ‘(s)he, it left’ : Hom. Gk. ἦλυθε, OIr. *luid*.
- PIE **-nt*: **h₁lud^{hont}* > prt. VI 3pl. ^B*latem* ‘they left’ : Gk. Hom. ἦλυθον. Cf. PIE **-nti* > CToch. **-än^tä* > prs. 3pl. act. ^B*-äm*, ^A*-i* ; see above 3.2.6.4 on ^B*ikäm*, ^A*wiki* ‘twenty’.
- PIE **-m*: **dek^m* > CToch. **ts'äkän* > **ts'äkä* > ^B*šak*, ^A*šäk* ‘ten’; PIE **dk^mtóm* > CToch. **käntä* > ^B*kante*, ^A*känt* ‘hundred’.

- PIE **-n* : PIE **-ēn* > B *-ie* (B*yriye* ‘sheep’ : Gk. ἀρίην, 4.1.5.3.1); PIE **-mōn* > B *-mo* (B *klyomo* ‘renowned’, cf. Rau 2009: 55 f.).

4.1.7.3. Tocharian A and B preserve PIE vibrant **r* and lateral **l* in absolute final position, and additionally Tocharian B final laryngeals.

- PIE **-h₂* : **g^henh₂* > B*śana* ‘woman, wife’, but apocope in *^śām* (cf. above 4.1.4 (g) for the delabialization of palatalized CToch. **k^{wj}* to *ś*).
- PIE **-r* : **-(n)tor* > *^AB-(n)tär*.
- PIE **-l* : PIE **pek^h-l* > B*pikul* ‘year’, pl. *pikwala*.

5. Consonant clusters

Note in the chart below that the Tocharian B form *twe* arises either by syncope from an inherited disyllabic nom. CToch. **tuwæ* (Ringe 1996: 12, 113) or from monosyllabic PT CToch. **twæ* with analogical reshaping of its onset as **tw-* after the inherited acc. **tué* ‘you’ > CToch. **t^wwä* > *^cu* (4.1.6.1), cf. Gk. Ion. Att. nom. σ-onset σύ extended from acc. σέ (< **tué*).

Overview of palatalization and resultant allophony of consonant clusters

	palatalized	Examples
<i>tt</i> <i>tk</i> , cf. 5.1 below	<i>cc</i> <i>^tk</i> > <i>śś</i>	B <i>tattam</i> : opt. <i>täccimar</i> ‘may I put’ CToch. <i>*-tk-</i> before <i>^i</i> (< PIE <i>*-ih₂-</i>) : CToch. <i>*nātk-i-</i> > <i>^nāsi</i> ‘mistress’, pl. <i>nāśśāñ</i> PIE <i>*at=k^hih₁</i> > interrog. ptcle <i>^aśśi</i> , Lat. <i>at=quī</i> , Pol. interrog. ptcle <i>czy</i> (Hackstein 2004a: 173–175)
<i>tk</i>	<i>^ck</i> , B <i>cc</i>	before PIE <i>*e</i> : <i>^käckäš</i> , B <i>käckäm</i> ‘rejoices’ before PIE <i>*o_i</i> > <i>*ē</i> : nom. pl. m. <i>^nācki</i> ‘masters’
<i>tw</i>	<i>^cw</i> , B <i>c</i>	nom. <i>^tu</i> ‘you’ (2sg.): abl. <i>cwäš</i> nom. B <i>twe</i> , <i>^tu</i> ‘you’ (2sg.): acc. <i>^cu</i> (< <i>*tué</i> , cf. 4.1.6.1)
<i>kt</i>	<i>^ABkc</i>	PIE <i>*nok^h-t-ey-jo-</i> > B <i>nekcīye</i> , <i>^nakcu</i> ‘at night’ (Pinault 1990: 189)
<i>kst</i>	<i>śc</i> (B), <i>śś</i> (A)	PIE <i>*pek-s-tih₁</i> > CToch. <i>*päckē-</i> > <i>*päścä-</i> > B <i>päśca(-ne)</i> , <i>päś(-ne)</i> , <i>^päśśä(-m)</i> ‘breasts’ (Hilmarsson 1989: 99 f.), cf. PIE <i>*pek-tu-</i> > Lat. <i>pectus</i> , OIr. <i>ucht</i> (<i>*puktu-</i> ← <i>*pektu-</i>), PIE <i>*pek-so-</i> > Ved. <i>pakṣá-</i> ‘side, flank, wing’ acc. pl. B <i>lāntām</i> , : nom. pl. <i>lāñc</i> ‘kings’ B <i>tāñk-</i> : <i>cemštär</i> ‘hinders’ Gk. ἀστήρ: B <i>ściryē</i> ‘star’ sbj. 3sg. <i>^štamaš</i> ‘(s)he will stand’ : caus. impv. 2sg. <i>päśśäm</i> ‘put!’ sbj. 3sg. B <i>stāmañ</i> : prt. 3sg. B <i>śama</i> ‘(s)he stepped’
<i>nt</i> <i>ñk</i> <i>st</i> , <i>št</i> (A)	<i>ñc</i> , <i>ñś</i> , <i>mś</i> , <i>ś</i> <i>mś</i> , <i>ñś</i> , <i>ñc</i> <i>śc</i> > <i>śś</i> , <i>ś</i>	

Overview of palatalization and resultant allophony of consonant clusters (continued)

	palatalized	Examples
<i>sk</i>	> ^B š / <u>t</u> (prs. II)	prs. II 1sg. ^B páskau ‘I guard’ : 2sg. ^B pášt ‘you guard’
<i>sk</i>	> ^B s / <u>t</u> (prs. IX–XI)	prs. IXa 1sg. ^B känmaskau ‘I come’ : 2sg. ^B känmast ‘you come’
<i>sk</i>	> ^B šš (prs. II, IX–XI)	prs. II 3sg. ^B páššäm, prs. IXa 3sg. ^B känmaššäm ‘(s)he comes’
<i>sw</i>	^A šu, ^B š	PIE * <i>sue-</i> : ^A šu ‘to oneself, hither, away’, ^B šän ‘(one)self’

5.1. Historical developments

CToch. *-s- was lost between dental and velar: *-T-sk̂- > *-T-k-. This rule accounts for the verbal roots in -tk-, which arose by regular sound change where roots ending in a dental were suffixed with *-s̄ke/o- (Melchert 1978; Hackstein 2002: 3–5; Pinault 2006). – PIE **klud-s̄ke/o-* → *klutk-*, with restored *u*-vocalism as per 3.1 above, ^Bklautk- ‘turn, become’; cf. Oic. *hljótask* ‘happen, become’ (Hilmarsson 1996: 145).

This sound change also occurred in nominal formations, cf. CToch. **tsätsk̄u-* > ^Btsatku, see 3.2.8.8.

(Many thanks to Jared Klein, Dieter Gunkel, Christoph Bross, and Benedikt Peschl for the suggestion of numerous improvements in the contents of the present overview. Needless to say, the responsibility for any remaining factual errors or inadequacies is solely my own.)

6. References

- Adams, Douglas Q.
1978 Ablaut and umlaut in the Tocharian vowel system. *Journal of the American Oriental Society* 98: 446–450.
- Adams, Douglas Q.
1999 *A Dictionary of Tocharian B*. (Leiden Studies in Indo-European 10). Amsterdam: Rodopi.
- Adams, Douglas Q.
2013 *A Dictionary of Tocharian B*. Revised and greatly enlarged edition. Amsterdam: Rodopi.
- Burrow, Thomas
1935 Tokharian elements in the Kharoṣṭhi documents from Chinese Turkestan. *Journal of the Royal Asiatic Society* 1935: 667–675.
- Carling, Gerd
2003 New look at the Tocharian B medical manuscript IOL Toch. 306 (Stein Ch.00316a2) at the British Library – Oriental and India Office Collections. *Historische Sprachforschung* 116: 75–95.
- Cowgill, Warren
1967 Ablaut, accent, and umlaut in the Tocharian subjunctive. In: Walter W. Arndt, Paul W. Brosnan, Jr., Frederic E. Coenen, and Werner P. Friedrich (eds.), *Studies in historical*

- linguistics in honor of George Sherman Lane*. Chapel Hill: University of North Carolina Press, 171–181. [Reprinted in Klein (ed.), 445–450.]
- Cowgill, Warren
1985 The Personal Endings of Thematic Verbs in Indo-European. In: Bernfried Schlerath (ed.), *Grammatische Kategorien. Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft Berlin, 20.–25. Februar 1983*. Wiesbaden: Reichert, 99–108.
- Cowgill, Warren
2006 The Personal Endings of Thematic Verbs in Indo-European (longer version of preceding entry). In: Klein (ed.), 535–567.
- Driessen, Michiel
2003 **h₂é-h₂us-o-*, the Proto-Indo-European term for ‘gold’. *The Journal of Indo-European Studies* 31: 347–362.
- EWAia = Manfred Mayrhofer
1992–2001 *Etymologisches Wörterbuch des Altindoarischen*. 3 Bde. Heidelberg: Winter.
- Griepentrog, Wolfgang
1995 *Die Wurzelnomina des Germanischen und ihre Vorgeschichte*. (Innsbrucker Beiträge zur Sprachwissenschaft 82). Innsbruck: Institut für Sprachwissenschaft der Universität.
- Hackstein, Olav
1995 *Untersuchungen zu den sigmatischen Präsensstambildungen des Tocharischen*. Göttingen: Vandenhoeck & Ruprecht.
- Hackstein, Olav
2000 Review of Hilmarsson 1996. *Kratylos* 45: 96–104.
- Hackstein, Olav
2002 Uridg. *CH.CC > *C.CC. *Historische Sprachforschung* 115: 1–22.
- Hackstein, Olav
2002 *Die Sprachform der homerischen Epen*. Wiesbaden: Reichert.
- Hackstein, Olav
2003 Review of Adams 1999. *Indo-Iranian Journal* 46: 177–189.
- Hackstein, Olav
2004a Rhetorical questions and the grammaticalization of interrogative pronouns as conjunctions in Indo-European. In: Adam Hyllested, Anders Richard Jørgensen, Jenny Helena Larsson, and Thomas Olander (eds.), *Per Aspera ad Asteriscos. Studia Indogermanica in honorem Jens Elmegård Rasmussen*. Innsbruck: Institut für Sprachwissenschaft der Universität, 167–186.
- Hackstein, Olav
2004b Reflexivpronomina, Präverbien und Lokalpartikel in indogermanischen Sprachen. *Tocharian and Indo-European Studies* 10: 69–95.
- Hackstein, Olav
2006 La langue poétique indo-européenne: archaïsme et renouvellement dans les théonymes. In: Georges-Jean Pinault and Daniel Petit (eds.), *La langue poétique indo-européenne. Actes du Colloque de travail de la Société des Études Indo-Européennes. Paris, 22–24 octobre 2003*. Leuven: Peeters, 95–108.
- Hackstein, Olav
2007 Ablative Formations. In: Alan J. Nussbaum (ed.), *Verba Docenti. Studies in historical and Indo-European linguistics presented to Jay H. Jasanoff by students, colleagues, and friends*. Ann Arbor: Beechstave, 131–153.
- Hackstein, Olav
2011 Proklise und Subordination im Indogermanischen. In: Thomas Krisch und Thomas Lindner (eds.), *Indogermanistik und Linguistik im Dialog. Akten der 13. Fachtagung der Indogermanischen Gesellschaft in Salzburg*. Wiesbaden: Reichert, 192–202.

- Hackstein, Olav
2012a Collective and Feminine in Tocharian. In: Olav Hackstein and Ronald I. Kim (eds.), *Multilingualism and History of Knowledge*. Vol. II: *Linguistic Developments along the Silkroad. Archaism and Innovation in Tocharian*. Vienna: Österreichische Akademie der Wissenschaften, 143–177.
- Hackstein, Olav
2012b Review of Weiss 2009. *Kratylos* 57: 109–115.
- Harðarson, Jón Axel
1993 Griechisch (F)εματ. In: Gerhard Meiser (ed.), *Indogermanica et Italica. Festschrift für Helmut Rix zum 65. Geburtstag*. Innsbruck: Institut für Sprachwissenschaft der Universität, 159–168.
- Hilmarsson, Jörundur
1986 *Studies in Tocharian Phonology, Morphology and Etymology. With Special Emphasis on the o-Vocalism*. Reykjavik: Málvísindastofnun Háskóla Íslands.
- Hilmarsson, Jörundur
1989 *The dual forms of nouns and pronouns in Tocharian*. (Tocharian and Indo-European Studies Supplementary Series 1). Reykjavik: Málvísindastofnun Háskóla Íslands.
- Hilmarsson, Jörundur
1991 *The Nasal Prefixes in Tocharian. A Study in Word Formation*. (Tocharian and Indo-European Studies Supplementary Series 3). Reykjavík: Málvísindastofnun Háskóla Íslands.
- Hilmarsson, Jörundur
1996 *Materials for a Tocharian Historical and Etymological Dictionary*. Edited by Alexander Lubotsky and Guðrún Þórhallsdóttir with the assistance of Sigurður H. Pálsson. Reykjavik: Málvísindastofnun Háskóla Íslands.
- Jasanoff, Jay
1978 *Stative and Middle in Indo-European*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Jasanoff, Jay
1987 Some irregular imperatives in Tocharian. In: Watkins (ed.), 92–112.
- Katz, Joshua
1997 Ein tocharisches Lautgesetz für Monosyllaba. *Tocharian and Indo-European Studies* 7: 61–87.
- Kim, Ronald
1999 The development of labiovelars in Tocharian: a closer look. *Tocharian and Indo-European Studies* 8: 139–187.
- Klein, Jared S. (ed.)
2006 *The Collected Writings of Warren Cowgill*. Ann Arbor: Beech Stave.
- Klingenschmitt, Gert
2005 *Aufsätze zur Indogermanistik*. Ed. by Michael Janda, Rosemarie Lühr, Joachim Matzinger, and Stefan Schaffner. Hamburg: Kovač.
- Kortlandt, Frederik
1983 Greek numerals and PIE glottalic consonants. *Münchener Studien zur Sprachwissenschaft* 42: 97–104.
- Krause, Wolfgang
1952 *Westtocharische Grammatik*. Band 1: *Das Verbum*. Heidelberg: Winter.
- Krause, Wolfgang and Werner Thomas
1960 *Tocharisches Elementarbuch*. Band I. *Grammatik*. Heidelberg: Winter.
- Kuiper, Franciscus Bernardus Jacobus
1955 *Shortening of final vowels in the Rigveda*. Amsterdam: Noord-Hollandsche Uitg. Maatschappij. (= Kuiper 1997, *Selected Writings*, 284–320.)

- Kuiper, Franciscus Bernardus Jacobus
 1997 *Selected Writings on Indian Linguistics and Philology*. Ed. by Alexander Lubotsky, Marianne S. Oort, and Michael Witzel. Amsterdam: Rodopi.
- Kümmel, Martin Joachim
 2007 *Konsonantenwandel*. Wiesbaden: Reichert.
- Lane, George Sherman
 1976 Notes sur le sort des syllabes finales i.e. en tokharien. *Bulletin de la Société de Linguistique* 71: 133–161.
- Leukart, Alex
 1994 *Die frühgriechischen Nomina auf -tās und -ās*. Vienna: Österreichische Akademie der Wissenschaften.
- LIV² see Rix (ed.).
- Malzahn, Melanie
 2000 Tocharisch *ñāktene* ‚Götterpaar‘ und Verwandtes. *Tocharian and Indo-European Studies* 9: 45–52.
- Malzahn, Melanie
 2007 A Tocharian Brahmi chart. In: Melanie Malzahn (ed.), *Instrumenta Tocharica*. Heidelberg: Winter, 223–254.
- Malzahn, Melanie
 2010 *The Tocharian Verbal System*. Leiden: Brill.
- Mayrhofer, Manfred
 1986 *Lautlehre (Segmentale Phonologie des Indogermanischen)*. *Indogermanische Grammatik*. Ed. by Jerzy Kuryłowicz. Band I. 2. Halbband. Heidelberg: Winter.
- Melchert, H. Craig
 1978 Tocharian verb stems in *-tk-*. *Zeitschrift für vergleichende Sprachforschung* 91 [1977]: 93–130.
- Oettinger, Norbert
 2004 Griechisch γυνή Frau und der Ursprung des *i*-Kollektivums. In: Thomas Krisch, Thomas Lindner, and Ulrich Müller (eds.), *ANALECTI HOMINI UNIVERSALI DICATA. Arbeiten zur Indogermanistik, Linguistik, Philologie, Politik, Musik and Dichtung. Festschrift für Oswald Panagl zum 65. Geburtstag*. Band I. Stuttgart: Heinz, 164–170.
- Oettinger, Norbert
 2012 Nominal- und verbale Reduplikation im Indogermanischen und Anatolischen. In: Craig Melchert (ed.), *The Indo-European Verb. Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles 13–15 September 2010*. Wiesbaden: Reichert, 240–246.
- Peyrot, Michaël
 2008 *Variation and Change in Tocharian B*. (Leiden Studies in Indo-European 15). Amsterdam: Rodopi.
- Peyrot, Michaël
 2013 *The Tocharian Subjunctive*. Leiden: Brill.
- Pinault, Georges-Jean
 1990 Notes sur les manuscrits de Maitreyasamiti. *Tocharian and Indo-European Studies* 4: 119–202.
- Pinault, Georges-Jean
 1995 The Rendering of Buddhist Terminology in Tocharian. *Journal of the Dunhuang and Turfan Studies* 1: 9–35.
- Pinault, Georges-Jean
 1997 Sur l’assemblage des phrases (« Satzgefüge ») en tokharien. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft Madrid, 21–24 de septiembre de 1994*. Wiesbaden: Reichert, 449–500.

- Pinault, Georges-Jean
 1998 Tocharian Language and Pre-Buddhist Culture. In: Victor H. Mair (ed.), *The Bronze Age and Early Iron Age Peoples of Eastern Central Asia*. Volume I. Washington: Institute for the Study of Man Inc., 358–371.
- Pinault, Georges-Jean
 2003 Sur les thèmes indo-européens en *-u-: dérivation et étymologie. In: Eva Tichy, Dagmar S. Wodtko, and Britta Irslinger (eds.), *Indogermanisches Nomen. Akten der Arbeitstagung der Indogermanischen Gesellschaft Freiburg, 19. bis 22. September 2001*. Bremen: Hempen, 153–188.
- Pinault, Georges-Jean
 2006 Sur l'évolution phonétique **tsk* > *tk* en tokharien commun. *Münchener Studien zur Sprachwissenschaft* 62 [2002]: 103–156.
- Pinault, Georges-Jean
 2008 *Chrestomathie tokharienne. Textes et grammaire*. (Collection linguistique publiée par la Société de linguistique de Paris 95). Leuven: Peeters.
- Pinault, Georges-Jean
 2009 On the formation of the Tocharian demonstratives. In: Elisabeth Rieken and Paul Widmer (eds.), *Pragmatische Kategorien, Funktion und Diachronie*. Wiesbaden: Reichert, 221–245.
- Pinault, Georges-Jean
 2011 Sur l'histoire des cas en tocharien. In: Michèle Fruyt, Michel Mazoyer, and Dennis Pardee (eds.), *Grammatical Case in the Languages of the Middle East and Europe*. Chicago: The Oriental Institute of the University of Chicago, 383–398.
- Pokorny, Julius
 1959 *Indogermanisches Etymologisches Wörterbuch*. Bern: Francke.
- Puhvel, Jaan
 1984 *Hittite Etymological Dictionary*. Vol. 1,1. Berlin: De Gruyter.
- Rasmussen, Jens Elmegård
 1989 *Studien zur Morphophonemik der indogermanischen Grundsprache*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Rau, Jeremy
 2009 *Indo-European Nominal Morphology: The Decads and the Caland System*. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Ringe, Don
 1987 On the prehistory of Tocharian B accent. In: Watkins (ed.), 254–269.
- Ringe, Don
 1989 Tocharian B *ausu*, *ausu*, *aulsu*. *Tocharian and Indo-European Studies* 3: 35–50.
- Ringe, Don
 1991 Evidence for the position of Tocharian in the Indo-European family? *Die Sprache* 34 [1988–1990]: 59–123.
- Ringe, Don
 1996 *On the Chronology of Sound Changes in Tocharian*. Volume I. *From PIE to Proto-Tocharian*. (American Oriental Series 80). New Haven: American Oriental Society.
- Rix, Helmut (ed.)
 2001 *Lexikon der indogermanischen Verben. Die Wurzeln und Ihre Primärstammbildungen*. Unter der Leitung von Helmut Rix und der Mitarbeit vieler anderer bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp und Brigitte Schirmer. Zweite erweiterte und verbesserte Auflage bearbeitet von Martin Kümmel und Helmut Rix. Wiesbaden: Reichert.
- Schaffner, Stefan
 2001 *Das Vernersche Gesetz und der innerparadigmatische grammatische Wechsel des Urgermanischen im Nominalbereich*. (Innsbrucker Beiträge zur Sprachwissenschaft 103). Innsbruck: Institut für Sprachwissenschaft der Universität.

- Schmidt, Klaus Totila
 1980 Zu Stand und Aufgaben der etymologischen Forschung auf dem Gebiete des Tocharischen. In: Manfred Mayrhofer, Martin Peters, and Oskar E. Pfeiffer (eds.), *Lautgeschichte und Etymologie. Akten der VI. Fachtagung der Indogermanischen Gesellschaft Wien, 24.–29. September 1978*. Wiesbaden: Reichert, 394–411.
- Schmidt, Klaus T.
 1997 Problems of recording Tocharian in central Asian brāhmī. In: Shirin Akiner and Nicholas Sims-Williams (eds.), *Languages and Scripts of Central Asia*. London: School of Oriental and African Studies, 16–24.
- Sieg, Emil, Wilhelm Siegling, and Wilhelm Schulze
 1931 *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.
- Szemerényi, Oswald
 1960 *Studies in the Indo-European system of numerals*. Heidelberg: Winter.
- Thomas, Werner and Wolfgang Krause
 1964 *Tocharisches Elementarbuch*. Band II. *Texte und Glossar*. Heidelberg: Winter.
- Vine, Brent
 1998 Indo-European and nostratic: some further comments. In: Joseph C. Salmons and Brian D. Joseph (eds.), *Nostratic. Sifting the evidence*. Amsterdam: Benjamins, 85–105.
- Watkins, Calvert
 1985 Two Anatolian forms: Palaic *aškmāuwa-*, Cuneiform Luvian *wa-a-ar-ša*. In: George Cardona and Norman H. Zide (eds.), *Festschrift for Henry Hoenigswald On the Occasion of his Seventieth Birthday*. Tübingen: Narr, 399–404.
- Watkins, Calvert
 2009 The Milk of the Dawn Cows Revisited. In: Kazuhiko Yoshida and Brent Vine (eds.), *East and West. Papers in Indo-European studies*. Bremen: Hempen, 225–240.
- Watkins, Calvert. (ed.)
 1987 *Studies in Memory of Warren Cowgill (1929–1985). Papers from the Fourth East Coast Indo-European Conference Cornell University, June 6–9, 1985*. Berlin: De Gruyter.
- Weiss, Michael
 2009 *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave.
- van Windekens, Albert-Joris
 1976 *Le tocharien confronté avec les autres langues indo-européennes*. Vol. I. *La Phonétique et le Vocabulaire*. Louvain: Centre International de Dialectologie Générale.
- Winter, Werner
 1967 A note on cases. In: *To Honor Roman Jakobson. Essays on the occasion of his seventieth birthday*. The Hague: Mouton, 2250–2253.
- Winter, Werner
 1990 The importance of fine points in spelling. In: Jacek Fisiak (ed.), *Historical Linguistics and Philology*. Berlin: De Gruyter, 371–391.
- Winter, Werner
 1994 Syncope in Tocharisch A. In: Jens E. Rasmussen (ed.), *In honorem Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Unter Mitwirkung von Benedicte Nielsen. Wiesbaden: Reichert, 401–415 (= Winter 2005, *Kleine Schriften* I, 450–464).
- Winter, Werner
 1998 Tocharian. In: Anna Giacalone Ramat and Paolo Ramat (eds.), *The Indo-European Languages*. London: Routledge, 154–168.
- Winter, Werner
 2001 Adams 1999. *Central Asiatic Journal* 45: 128–138.
- Winter, Werner
 2003 Review of Adams 1999. *Journal of the American Oriental Society* 123: 202–208.

Winter, Werner

2005 *Kleine Schriften. Selected Writings*. 2 Bände. Ausgewählt und herausgegeben von Olav Hackstein. Bremen: Hempen.

Olav Hackstein, Munich (Germany)

76. The morphology of Tocharian

- | | |
|---------------------------------|--------------------------------|
| 0. Introduction | 4. Pronouns and determinatives |
| 1. Nominal morphology | 5. Numerals |
| 2. Types of declension of nouns | 6. Verb morphology |
| 3. Adjectives | 7. References |

O. Introduction

Tocharian (henceforth Toch.) A and Toch. B reflect divergent evolutions from the morphological system of Common Tocharian (henceforth CToch.) that can be reconstructed in great detail despite the discrepancies between the two languages. On the whole, Toch. A shows a more advanced stage of evolution towards simplification of paradigms and reduction of the quite numerous allomorphs that are still retained in Toch. B, although in some instances Toch. B has been the innovative language. The morphology of nouns and verbs can be described along the same lines as in the other IE languages and offers as expected both innovation and preservation of archaic features. In addition to the reshaping of many categories, CToch. has seen the emergence of new morphological properties that do not belong to the IE mainstream: the semantic categorization of some endings according to animacy in the noun and the signalling of valence in the verb stems. In addition, the nominal system is characterized by an ongoing, albeit still incomplete, shift towards an agglutinative type.

1. Nominal morphology

For reasons of space, the present survey will not consider word formation (on which see Adams, this handbook), and only the most productive adjectival suffixes will be mentioned. Although nominal compounding is frequent in the texts, this is due largely to the imitation of Sanskrit, for nominal composition does not seem to have been that common as a genuine Tocharian feature. A peculiarity of CToch. is the existence of compounds based on binominal phrases referring to complementary or similar concepts.

Winter, Werner

2005 *Kleine Schriften. Selected Writings*. 2 Bände. Ausgewählt und herausgegeben von Olav Hackstein. Bremen: Hempen.

Olav Hackstein, Munich (Germany)

76. The morphology of Tocharian

- | | |
|---------------------------------|--------------------------------|
| 0. Introduction | 4. Pronouns and determinatives |
| 1. Nominal morphology | 5. Numerals |
| 2. Types of declension of nouns | 6. Verb morphology |
| 3. Adjectives | 7. References |

O. Introduction

Tocharian (henceforth Toch.) A and Toch. B reflect divergent evolutions from the morphological system of Common Tocharian (henceforth CToch.) that can be reconstructed in great detail despite the discrepancies between the two languages. On the whole, Toch. A shows a more advanced stage of evolution towards simplification of paradigms and reduction of the quite numerous allomorphs that are still retained in Toch. B, although in some instances Toch. B has been the innovative language. The morphology of nouns and verbs can be described along the same lines as in the other IE languages and offers as expected both innovation and preservation of archaic features. In addition to the reshaping of many categories, CToch. has seen the emergence of new morphological properties that do not belong to the IE mainstream: the semantic categorization of some endings according to animacy in the noun and the signalling of valence in the verb stems. In addition, the nominal system is characterized by an ongoing, albeit still incomplete, shift towards an agglutinative type.

1. Nominal morphology

For reasons of space, the present survey will not consider word formation (on which see Adams, this handbook), and only the most productive adjectival suffixes will be mentioned. Although nominal compounding is frequent in the texts, this is due largely to the imitation of Sanskrit, for nominal composition does not seem to have been that common as a genuine Tocharian feature. A peculiarity of CToch. is the existence of compounds based on binominal phrases referring to complementary or similar concepts.

1.1. Categories

1.1.1. Gender

Masculine and feminine are opposed in nouns and adjectives. The neuter proper is found only in the singular of demonstrative pronouns, when they refer to a nominal clause. Besides nouns with fixed gender, several classes of nouns feature a so-called “alternating” gender (*genus alternans*), taking masculine agreement in the singular and feminine agreement in the plural. Since the plural endings of these nouns go back ultimately to the PIE neuter plural, they indirectly reflect original neuters. In synchrony the situation is blurred by the fact that some former neuter endings have been extended to the feminine in the plural of the adjectives and of some nouns. The extinction of the neuter is a consequence of the merger of its nom.-acc. endings with the endings of the masculine in the singular, and with those of the feminine in the plural in the prehistory of Tocharian.

1.1.2. Number

In addition to singular and plural, one finds the dual, without any formal distinction between paired and unpaired items. But this category was already falling into decay. In the plural, one can note a trend to form individualizing plurals contrasting with collectives. Toch. B has a specific formation with suffix *-aiwenta*: *misaiwenta* ‘single pieces of meat’ vs. *mīsa*, pl. *tantum*, ‘meat’; the singulative plural goes back to a phrase with attribute **aiwe-nta*, plural of **aiwe* < **(h)oi-uo-* ‘one’ (cf. Av. *aēuua-*).

1.1.3. Case

CToch. has a two-storied system. The so-called primary cases are nominative, accusative (or oblique = Germ. *Obliquus*), genitive, and (in Toch. B only) vocative. These follow the PIE inflectional type. The genitive covers partly the function of the dative as well. The second story comprises the so-called secondary cases that are made by adding to the accusative (singular, plural, or dual) postpositions that are indifferent to number: perlativ, comitativ, allativ, ablativ, locativ, instrumentiv (only in A), and causal (only in B). The relative lateness of the univerbation of these affixes is shown by several features: 1. they are generally (except for the ablativ and causal) ignored by the Toch. B stress assignment rule, e.g., B *yákwe* ‘horse’ < **yákwe*, perl. sg. *yákwe-sa* vs. gen. sg. *yákwéntse*, 2. they are still felt as independent, since they can be separated from the governed noun, 3. in a coordinative or attributive clause these affixes are normally omitted from all but the last or head term, the preceding terms showing the accusative form (so-called “group inflection”), 4. on the surface they differ in the two Toch. languages: perl. B *-sa*, A *-ā* com. B *-mpa*, A *-aśśäl* all. B *-śc(ä)*, A *-ac* abl. B *-mem*, A *-äs* (*-as*) loc. B *-ne*, A *-am* instr. A *-yo* and causal B *-ñ(ä)*. In some cases (locativ, allativ, perlativ) the discrepancies can be explained by differing results of the decomposition that affected sequences of accusative singular or plural followed by these postpositions. A few postpositions are obviously related to independent adverbs, pre- or postpositions,

e.g., A com. *-aśśäl* and A *śla*, B *śale/śle* ‘with, together’, B abl. *-meṃ* and B *mante*, adv. ‘upwards’.

2. Types of declension of nouns

Due to the blurring of PIE final syllables and the emergence of new classes, the inflectional class of a noun cannot be predicted from the form of the nominative singular.

2.1. Two major declensional types can be distinguished: a) nominative and accusative plural are identical, hence the plural has a distinct suffix to which the genitive ending and the secondary affixes are added, b) nominative and accusative plural are different. The nouns of type a) reflect ancient neuters (hence of “alternating” gender), but comprise also some animates (classes I–III), the nouns of type b) are prevaillingly animates and correspond to different PIE formations of masculine and feminine gender (class IV–VII). In the singular, many nouns do not show any distinction between nominative and accusative forms. This is expected for nouns of type a), but it is found also for classes of nouns of type b), with two exceptions: 1. the noun has inherited an acc. sg. form that differs by itself from the nom. sg., 2. the nouns signalled as [+human] are provided with an overt ending of acc. sg. *-ṃ* (anusvāra in Brāhmī script), noting the dental nasal /-n/. In Toch. B one finds this ending also in some types of adjectives, whereas Toch. A has extended it to all types of adjectives and to their feminines (*-ā-ṃ*) as well. For instance, Toch. B *yakwe* (A *yuk*) ‘horse’ and *eṅkwe* (A *oṅk*) ‘man’ have exactly the same declension, except for acc. sg. B *yakwe* (A *yuk*) vs. *eṅkweṃ* (A *oṅkaṃ*). The ending *-ṃ* has its origins in the accusative sg. of nasal (*-on-) stems that provided an enlarged declension with individualizing value to several types, and especially to stems ending in a vowel, including the thematic vowel: *-on-ṃ > *-aenā(n) > *-aenā, decomposed as *-aenā and specialized as a marker of the feature [+human]. One may leave aside the various forms of the vocative sg., which are attested only in Toch. B, since they cannot be traced to any ancient declensional class.

2.2. The normal ending of the genitive singular of nouns is B *-ntse* (*-āntse* after a consonant), A *-s* (*-is* after a consonant from earlier *-āⁱnsa, *-es* in thematic stems from *-aⁱnsa < *-aensæ) < CToch. *-nsæ. Its origin lies certainly in *-n-stems, starting from the recharacterization of the final gen. sg. *-én-s as *-ensos, through extension to the animate *-on-stems, hence *-on-os remade as *-onsos > *-aensæ, decomposed as *-aensæ. The ancient athematic ending *-os > CToch. *-æ is retained in nouns and adjectives based on *-nt-stems, e.g., B *lānte* (A *lānt*) gen. of *walo* (A *wäl*) ‘king’, acc. sg. AB *lānt*, and B *pernente* gen. of *perne_v*, acc.sg.masc. *pernent*. The standard masculine gen. sg. ending in most adjectives in Toch. is B *-epi*, A *-(y)āp*, and this has been extended in Toch. A to nouns marked as [+human]. The constituent *-p-i* of this ending recalls the gen. sg. *cwi/cpi* of the demonstratives (see below 4.2), and one may assume this to be the source of this ending, which has been added in Toch. A to a former thematic ending. There are other gen. sg. endings, but of quite restricted extension: AB *-i* (< dative sg. *-ei) for nouns marked as [+ person], such as kinship terms (e.g., B *pātri*, to *pācer* ‘father’), feminine institutional terms (e.g., *lantsoy*, to *lāntsa* ‘queen’), and proper names B *-ñ* and A *-y* in nouns borrowed from Sanskrit.

2.3. The genitive plural shows less diversity: B *-nts* (*-nts*) < **-nsä* < **-nsu* < **-nsōm*, through remaking of **-n-ōm* from **-n*-stems after the model of the genitive sg. **-ns-os* > **-nsæ* A *-śśi* in nouns of type (b) and in adjectives, possibly from **-s+śi*, that is **-s* < gen. pl. **-nsä*, which merged with the result of the acc. pl. **-ns* and was then reinforced by an adjectival suffix. In the great majority of nouns of type (a) Toch. A uses an ending *-is* which is actually identical to the singular ending (see above). This fact is a further indication of the growing agglutinative trend of the declension.

2.4. Classes of nouns of types a): I. Plurals B *-ā*, *-wa* (A *-ā*, *-u*, recharacterized as *-wā*) are derived from the neuter pl. in **-ā* < **-a* < **h₂* of consonant and **-u*-stems, e.g., B *pwāra* of *puwar* ‘fire’, B *ostuwa*, A *waštu* < CToch. **wāstuwā*, of B *ost* (< **wost*), A *wašt* ‘house’. This final has been enlarged to *-u-nt* in Toch. A with another plural ending (see below under III.). II. Plurals B *-na*, *-nma* (after metathesis), A *-(ā)ṃ* < CToch. **-nā* and **-mnā* have their nucleus in the neuter pl. of **-en-* and **-men-*stems, which has been extended to other neuters. It has been used also to renew ancient collectives, as in B *akrūna* ‘tears’, cf. A *ākrunṭ* (sg. *ākār* < **ākru*). The final **-nā* came to be added to enlarge a feminine pl. after the merger in **-ā* < **-ās* of the ancient nom. pl. and acc. pl., hence B *śnona* (< **śāno-nā*) of *śana* ‘wife’; the same process is seen in adjectives in *-re* (B *-rona*, A *-raṃ*), gerunds (B *-llona*, A *-laṃ*), privatives in B *-tte*, fem. pl. *-ttona*, etc. III. The plural B *-nta*, A *-nt* < CToch. **-ntā* has been quite productive, it has been renewed in Toch. A as *-ntu* < **-nt-wā* after **-wā* competing with final **-ā*. The starting point for this formation lies in the plural of individualizing derivatives in **-nt-* of former collectives, cf. B *śānta* ‘items of small cattle’. The suffix came to be added to former **-s*-stems, cf. B *cake* ‘river’ (< **cākæ* < **ték^u-os*), pl. *ckenta*, to various other neuters, e.g., B *war* ‘water’, pl. *wranta* (A *wār*, pl. *wrāntu*), to abstracts, e.g., B *palsko* ‘thought’, pl. *pālskonta* (A *pāłtsāk*, pl. *pāłskant*), and even to loan words.

2.5. Classes of nouns of type b): IV. The kinship terms with *-r*-stems form a small class filled with relics: B *pācer* (A *pācar*) ‘father’, B *mācer* (A *mācar*) ‘mother’, B *tkācer* (A *ckācar*) ‘daughter’, B *procer* (A *pracar*) ‘brother’, B *ṣer* (A *ṣar*) ‘sister’. They keep a frozen acc. sg. differing from the nom. sg.: e.g., B *pātār*, *protār*. The expected pl. forms, e.g., nom. B **pācārā*, obl. **pātrāṃ* have been remade after the pattern of other classes based on **-i*-stems and **-n*-stems, hence nom. pl. B *pātārñ*, A *pācri* (acc. pl. *pācrās* with levelling of the palatalization). The collective forms of type *pacera* (nom. = acc.) originated probably from the plural of feminine kinship terms. V. Plural B nom. *-i*, acc. *-(ā)ṃ*, A nom. *-i* or *-ñ*, acc. *-(ā)s*. The nom. pl. B *-i* < CToch. **-āi* has two PIE origins: **-oi* of thematic stems (ultimately coming from the demonstratives), and **-ei-es* of **-i*-stems. The latter ought to trigger palatalization of the preceding consonant, but this has been partly undone under the influence of the former ending. CToch. **-āi* becomes normally Toch. A *-e*, which is kept in thematic adjectives but has been replaced in many nouns by *-i* and mostly by *-añ* (taken from nasal stems) in thematic nouns, cf. B *yakwi* ‘horses’, acc. *yakweṃ* (< **-āi*, **-æns* < **-o-ns*), but A *yukañ*, acc. *yukas*. Apart from thematic nouns, this type has been extended from **-i*-stems to older consonant stems, due to the ambivalence of acc. pl. **-āns*, cf. B *kaum* ‘sun, day’, pl. nom. *kaunī*, acc. *kaunāṃ*, A *koṃ*, *koñi*, *konäs*, A *kowi* ‘cows’, acc. *kos* < **kowäs* (B *kewāṃ* < **kæw-āns*), B *meñe* ‘month, moon’, pl. nom. *meñi*, acc. *meñāṃ*, A *mañ*, *mañi*, *mañäs*. In this class one finds acc. sg. forms of Toch. B that descriptively drop the final *-e* of the nom.

sg.: e.g., *meñ* of *meñe*, *maś(c)* of *maśce* ‘fist’. This is the result of the older final **-i-m* > CToch. **-ä*. VI. Plural AB *-ñ*, acc. B *-ṃ*, A *-s* < **-ns*. This productive class, that started from nasal stems, falls into different subtypes according to the vowels that precede the endings: nom. pl. B *-eñ*, A *-añ* < **-æñä* < **-on-es*, AB *-iñ* < **-iñä* < **-en-es*, etc. It has been further extended to remodelled nasal stems, nom. pl. **-añä* < **-ōn-es* (alternatively < **-oHones*), hence Toch. A *-añ*, B *-aiñ* (or *-añ*, depending on the acc. sg.) under the accent, *-añ* elsewhere, e.g., *okso* ‘ox’, nom. pl. *oksaiñ*, but *oñkolmo* (A *oñkalām*) ‘elephant’, pl. *oñkolmañ* (A *oñkalmāñ*). Older feminine **-ā*-stems have joined this class due to the coincidence of the nom. sg. B *-o*, e.g., B *kantwo* ‘tongue’, acc. sg. *kantwa*, pl. nom. *kāntwāñ* (A *kāntu*, pl. *kāntwāñ*, *kāntwās*). The older acc. sg. **-ān(ä)* is the source of B *-ai* (A *-e*), which has spread extensively in nouns (masc. and fem.) and adjectives, e.g., B *pyāpyo*: *pyāpyai* ‘flower’, B *yoko/yokiye*: *yokai* (A *yoke*) ‘thirst’, B *yolo*: *yolai* ‘bad’, B *kauşēnca*: *kauşēncai* ‘killing’. VII. Plural B *-ñc*, A *-ṃś*, acc. *-ntām*, A **-ntäs* levelled to *-ñcäs*, go back to **-nt-* stems, pl. nom. **-ñcä* < **-nt-es*, acc. **-ntäns* < **-nt-ṃs*, cf. B *walo* (A *wäl*) ‘king’, pl. nom. B *lāñc* (A *lāṃś*, *lās*), acc. *lāntām* (A *lāñcäs*). One should note that the acc. sg. AB *lānt* reflects also the expected outcome **-ntä(n)* of **-nt-ṃ*.

2.6. One may trace back many of the multifarious dual forms to the PIE endings of nom.-acc. dual, while others have been enlarged by the dual suffix B *-ne* A *-(ä)ṃ* < **-næ*, which is parallel to the plural suffix **-nā*. In thematic stems masc. **-o-h₁(e)* > **-ō* > **-u* > **-ä*, e.g., in **ānsä-næ* ‘shoulders’ > B *antsane*, *āntsne*, A *esām*, **-o-(h₁)* through “lex Kuiper” > **-æ*, in B *ñakte-ne* ‘god [and] goddess’, extended to *pacere* ‘father [and] mother’, nt. **-o-h₁*, merging with **-oi* > **-äy* > B *-i*, cf. adj. B *kartsi* of *kartse* ‘good’. In consonant stems, masc. **-h₁(e)* > **-yä*, triggering palatalization of the preceding consonant or yielding AB *-i* when the preceding consonant is not palatalized. This is identical to the expected outcome of the nt. ending **-ih₁* > **-yä* > AB *-i*, e.g., from B *ek* A *ak* ‘eye’ < **æk*, dual **æśä-næ* > B *es(a)ne*, A *asām*, B *pwāri* ‘two fires’ (pl. *pwāra*), A *pratri* ‘two brothers’. The past history of the rare genitive dual forms remains largely a matter of conjecture.

3. Adjectives

3.1. Thematic types, ending in B *-e*. Two subtypes can be distinguished

3.1.1. In adjectives ending in B *-re* (A *-r*) < **-ræ* (< **-ro-*), the declension shows the influence of a weak inflection with nasal extension, e.g., B *tāpre* ‘high’, acc. *tāprēm* (< **-ænä[n]* < **-on-ṃ*), pl. nom. *tāprēñ* (< **-æñä* < **-on-es*), acc. *tāpré(nä)ṃ*, vs. A *tpār*, acc. *tāprām*, pl. nom. *tāpre*, acc. *tāpres*.

3.1.2. The other thematic types show the plural of class V nouns, but the acc. sg. is marked, when possible, by the palatalization of the suffixal consonant. This is seen in ordinals, B *-te*, A *-t*, acc. sg. *-ce* (A *-cām*), privatives, B *-tte* (A *-t*), acc. sg. *-cce*, gerunds B *-lle* (A *-l*), acc. sg. *-lye* (A *-lām*) and adjectives with suffix *-tstse*, acc. sg. *-cce*. Adjectives with suffix B *-(i)ye*, A *-i* < **-iyæ*, B *-ṣṣe*, A *-ṣi* < **-ṣiyæ*, B *-ññe*, A *-ñi*

(< *-ñiyæ) are not affected. This palatalization has been extended to the gen. sg. masc. and to the masc. plural. It has been certainly modeled after the demonstratives, cf. B masc. *se*, acc. *ce*, gen. *cwi/cpi*, nt. *te*, pl. nom. *cai*, obl. *ceṃ*.

3.1.3. In the feminine singular, the extension of the *devī*-type at the expense of the thematic type should be noted, hence sg. nom. *-yā (< *-ih₂), e.g., in an adjective in B *-re* (see above), B *tparya*, acc. *tparyai*, A *tāpri*, acc. *tpäryām*, and the final *-ā of the productive adjective types with palatalized suffixal consonant, cf. B *-šša*, *-ñña*, *-ca*, *-cca*, *-lya*, *-ttsa*.

3.2. The most productive athematic types show the outcomes of animate *-n-stems and *-nt-stems, with ongoing mutual influence of the two, and their endings correspond accordingly to the endings of noun classes VI and VII. These adjectives possess a feminine in *-yā, hence B *klyomo*, A *klyom* ‘noble’ (reflecting a *-mon-stem), fem. nom. B *klyomña*, A *klyomiṃ*. The type of adjectives in *-uont- is represented by *perne*_u ‘glorious’ (derived from B *perne*), acc. *pernent* (A *parno*, *parnont*), pl. nom. *perneñc*, acc. *pernentäm*, with contraction of *-æ-wænt- to *-ænt- in this type, the feminine plural nom.-acc. in *-ntā, B *pernenta* (A *parnont*) reflects the old neuter *-nt-h₂. The same characteristics are shared by adjectives featuring the nom. sg. masc. -u (with subtypes in *-ssu*, *-nu*) < *-uō(n). The preterite participle, while possessing different subtypes, has a nom. sg. masc. suffix -u < *-uōs, the palatalization of the final consonant of the suffix is used also to mark the acc. sg. (in B) and the nom. pl., e.g., B *kekesu*, A *kaksu* ‘extinguished’ (verb *kās-*), acc. sg. B *kekesoṣ*, A *kaksunt* (modeled after the -u-adjectives), masc. pl. nom. *kekesoṣ*, A *kaksuṣ*. The nom. pl. -oṣ (-uweṣ in another subtype) reflects indirectly the allomorph *-uos-es > *-wæšä. The feminine (CToch. *-usā) goes back ultimately to the PIE type *-us-ih₂.

4. Pronouns and determinatives

4.1. Personal pronouns preserve many familiar features of archaic IE languages: 1. the number distinction is signalled by suppletion, 2. the 2nd person sg. pronoun shows a contrast of stems between nominative and non-nominative forms, 3. besides the autonomous pronouns, there is a set of enclitic pronouns, that are suffixed to verbs or nominal predicates: they express direct object as well as genitive-dative case. There is, however, a striking innovation in Toch. A, which distinguishes in the 1st person sg. a feminine form from the masculine, whereas PIE did not have any gender distinction in personal pronouns. Singular 1st person: nom.=acc. B *ñās* (*ñis*) < CToch. **ñāsā*, A masc. *nāš*, probably from CToch. **ñāš*, through depalatalization, A fem. nom.=acc. *ñuk* < CToch. **ñāku*, gen. B *ñi* (A masc. *ñi*) < CToch. **ñāñi*, fem. A *nāñi*, probably remade from a demonstrative stem. Enclitic B *-ñ* < CToch. **-ñā* (ultimately from PIE *-me), remade as A *-ñi* with genitive ending -i. The form of B *ñās* (*ñis* through secondary palatalization), though originally unstressed, has been generalized also to stressed position. Singular 2nd person: nom. B *twe* (*tuwe*), A *tu* < CToch. **tuwæ*, acc. B *ci*, A *cu* < CToch. **cwā* (< PIE **tue*), gen. B *tañ* < **tāñ* < **tā-ñä*, recharacterized in A as **tāñ-i* with genitive ending -i, hence *tñi*. Enclitic B *-c* < CToch. *-cā* (ultimately from PIE *-te), remade as A

-*ci* parallel to the 1st sg. enclitic. Plural 1st person: nom.=acc. B *wes*, A *was* < CToch. **wæs*, gen. archaic B *wesäm*, A *wasäm* < CToch. **wæsän*, besides analogical forms, B *wesän*, *wesi*. Plural 2nd person: nom.=acc. B *yes*, A *yas* < CToch. **yæs*, gen. archaic B *yesäm*, A *yasäm* < CToch. **yæsän*, besides analogical forms, B *yesän*, *yesi*. The dual forms, nom.=acc., occur only in Toch. B, 1st pers. *we-ne*, 2nd pers. *ye-ne* they are formed with the dual suffix *-*næ* added to previous dual forms **wæ* and **yæ*. In addition, there is a reflexive possessive pronoun (undifferentiated for number and gender), reflecting a genitive form, parallel to the 2nd sg. pronoun: B *šan* < **šän* < **šä-ñä*, recharacterized in A as **šän-i*, hence *šni*. The enclitic paradigm also has forms for the 3rd person: singular B *-ne*, A *-ṃ* < *CToch. *-*næ* (ultimately from a demonstrative stem); plural, actually common to the three persons, B *-me*, A *-m* < CToch. *-*mæ*. The reflexive pronoun is expressed by a phrase that combines the inflected noun meaning originally 'self' (B *änme*, A *äncäm* < **äncmæ*) preceded by the possessive reflexive.

4.2. Demonstrative pronouns have four paradigms in Toch. B and three in Toch. A. These are mostly differentiated in synchrony by final vocalic or consonantal markers, viz. *-u*, *-s/-š*, *-m*, *-n* (noted with *-ṃ*), *-m(p)*. The last of these reflects probably the notation of a single consonant, the bilabial fricative /-β/. Three main functions can be ascertained, see the forms of the nom. sg.: 1. anaphoric: A masc. *säm*, fem. *sām*, nt. *täm*, B masc. *su*, fem. *sāu*, nt. *tu*; 2. proximal deictic: A *säs*, *sās*, *täs*, B *se*, *sā*, *te*; 3. distal deictic: A *saṃ*, *sām*, *taṃ*, B *saṃ(p)*, *som(p)*, *taṃ(p)*. Toch. B has a further set, *seṃ*, *sām*, *teṃ*, which is used, like the second set, as a discourse deictic but has the specific value of activating information shared by the speaker and the hearer. It is the formal match of the set Toch. A *saṃ*, *sām*, *taṃ*, but it has been replaced in Toch. B by a new set for distal deixis. As determiners, these demonstratives have only masculine and feminine gender. The common structure can be observed in the simplest pronoun, B *se*. The paradigm contrasts *sV* in the nom. sg. animate (masc. and fem.) vs. *tV* in neuter singular, *sV* for nom. sg. vs. *tV* elsewhere in the feminine, *sV* for nom. sg. vs. *cV* elsewhere in the masculine. The starting point for all this is clearly the PIE paradigm that had two stems: sg. nom. masc. **so*, fem. **seh₂*, nt. **to-d*. The last of these is reflected by CToch. **te* > B *te*, A *ta-* (with further substitutions). The evolution can be sketched as follows: for the masculine, sg. nom. **so* > **sæ* > B *se*, A *sa-* (with further substitutions) acc. **to-m* > **tæ*, replaced by the palatalized form **caē* > B *ce*, A *ca-*, pl. nom. **toj* > **tæi*, replaced by **caēi* > B *cai*, A *ce(-)*, acc. **to-ns* > **tæns*, replaced by **caens* > B *ceṃ*, A *ces*. The replacement of *tV* by *cV* served to avoid the merger of masculine and neuter in acc. sg., and it has been extended to distinguish masculine and feminine; the allomorph *cV* with palatalization can be sought out in an inherited form **te-* > CToch. **cä-*. A likely candidate is the gen. sg. masc. *cwi/cpi* (the latter form being the older) that is common to most paradigms: it reflects a CToch. form **cäβi*, possibly going back to a dative form. In the feminine, sg. nom. **seh₂* > **sa(H)* ("lex Kuiper") > **sā* > B *sā*, A *sā-*, acc. **tah₂-m* > **tā* > **tā*, replaced by **tā* after the nom. > B *tā*, A *tā-*, pl. nom. **teh₂-es* and acc. **teh₂-ns* merge to **tās* > **tā* > B *to(-)*, A *to(-)* nom.=acc., with various extensions (e.g., B *toy* after masc. *cai* in the paradigm of B *se*, fem. *sā*; A *tom* in the paradigm of A *sām*, fem. *sām*, etc.). The gen. sg. of B *sā* is *tāy* < **tā+i* with addition of the ending signalling [+person], which occurs also in the masculine B *cwi/cpi*. In the paradigm of Toch. A *säs*, the nom. sg., masc. *säs* and fem. *sās*, contrasts with the rest of the paradigm owing to the assimilation (depalatalization) of the final marker

-š by the preceding dental sibilant. The agglutinative pattern has asserted itself more strongly in Toch. A, where the genitive ending *-i* has become neutral to number and is added to the acc. form bearing the marker of the paradigm: for A *säm*, sg. acc. *ca-m*, gen. *ca-m-i*, pl. nom. *ce-m*, acc. *ces-äm*, gen. *cesmi* (< **ces-äm-i*); for A *säs*, sg. acc. *ca-š*, gen. *ca-š-i*, pl. nom. *ce-š*, acc. *ces-äs*, gen. *cessi* (< **ces-äs-i*); for A *saṃ*, sg. acc. *ca-m*, gen. *ca-n-i*, pl. nom. *ce-m*, acc. *ces-äm*, gen. *cesni* (< **ces-än-i*). Beyond the basic structure, CToch. did not yet possess an entirely fixed system of the different demonstratives, since the markers differ in both languages. The stems of the demonstratives have been used to form adverbs and conjunctions, e.g., B *tumem*, A *tmäs* (< **täm-äs*) ‘thereupon, then’ (frozen ablative), B *tusa*, A *tämyo* ‘therefore’ (frozen perlativ or instrumental), B *tane/tne*, A *tṣaṃ* ‘here’ (frozen locative), etc.

4.3. Interrogative and relative pronouns are not distinguished in Toch. B, whereas Toch. A differentiates the relative paradigm through the addition of the particle *-ne*: nom. *k^use* (A *kus*, rel. *kus-ne*), acc. *k^uce* (A *kuc*, rel. *kuc-ne*), gen. *kete* (A *ke*, rel. *ke-ne*). These forms, which are undifferentiated for number and gender, go back to the univertation of the indefinite-interrogative pronoun and a demonstrative pronoun, nom. sg. **kwäsæ* (with accent on the second syllable, hence B **kwäsé* > *k^use*) < **k^ui-s+so* the genitive B *kete* (A *ke*) goes back to a form reflecting an allomorph **kæ-* < PIE **k^uo-*. In addition, there are complex (and defective) pronouns, which are also based on univertations with demonstratives: B *intsu*, obl. *iñcew* (*iñcau*), A *äntsam*, obl. *äñcam*, B masc. *mäksu*, fem. *maksā_u*, nt. *mäktu*, etc. Interrogative and relative adverbs, as well as conjunctions are based ultimately on the stems CToch. **kwä-*, **kæ-*, through addition of various morphemes, e.g., B *k^uce*, A *kucne* ‘since’, B *kwri*, A *kupre* ‘when, if’, AB *kos* ‘how much?’, B *kā* ‘why?’, A *k^uyal* ‘why’, *k^uyalte* ‘because’.

4.4. The pronominal system comprises also an indefinite pronoun, undifferentiated for number and gender: B nom. *ksa* (A *saṃ*), acc. *kca* (A *caṃ*), gen. *ket(a)ra* (A *ke*). These forms are based on the inherited paradigm of the interrogative pronoun, through generalization of a clitic form **kw(ä)sā*, in addition to the genitive **kæ*. In Toch. A the corresponding forms **sa* (and **ca*) have been extended by a sentence particle; A **sa* goes back probably to a phrase with the pronoun ‘other’: **älak ksa* (matching lat. *aliquis*) > **älaksa*, hence **älaksa+n*.

4.5. The pronoun and determinative ‘other’ is nom. masc. sg. B *alyék*, A *älak* < CToch. **älyæ-kä*, pl. B *alyaik*, A *älyek* < CToch. **älyai-kä*; it is based entirely on the paradigm of the PIE stem **aljo-* with addition of the particle CToch. **-kä* to the inflected forms. An exception to this is the gen. pl. masc. A *älu*, which may go back to an adjectival form derived from **äl* < gen. pl. **aljōm*. In addition, there is a reflexive intensive pronoun, having the same value as Lat. *ipse*, which is based on the univertation of particles preceding a demonstrative pronoun, with the same final particle as in ‘other’ in Toch. A: B masc. nom. *makte* (A *mättak*), acc. *makce* (A *mäccakäm*), fem. nom. *mäkcīya* (A *mäccäk*), obl. *mäkcīyai* (A *mäccäkīyäm*), etc. The morpheme **mä-* (< CToch. **mæ-* < PIE **mo-*) is also found in other adverbs and conjunctions, cf. B *mant* ‘so, thus’, B *mäkte* ‘how?, as, in order that’, A *mänt* ‘how?’, A *mäm̄tne* ‘as, in order that’, etc.

5. Numerals

5.1. The cardinals reflect the PIE forms with some remodelling. The cardinals ‘one’ through ‘four’ are inflected for gender (except for ‘two’ in Toch. B and ‘four’ in Toch. A) ‘one’ has also plural forms. Several forms go back directly to PIE, e.g., ‘four’, B *štwer* (masc.), A *štwar* < CToch. **šätwærä* < **k^uet^uor-es*; ‘ten’, B *šak* (A *šäk*) < CToch. **šäkä(n)* < **dék^m*; ‘hundred’, B *kante* (A *känt*) < CToch. **käntæ* < PIE **k^mtóm* < **dk^mtóm*; ‘thousand’ is CToch. **wⁱälsæ* > B *yaltse*, A *wälts*. The paradigm of CToch. ‘one’ is based on the combination of forms of PIE **sém-* and **sió-* (cf. Hitt. *šia-*) with the adjective **somh₂-ó-*, fem. **somh₂-éh₂-*, hence nom. masc. B *še* (< **ššæ*), A *sas* (through assimilation < **šæs*), bound form A *ša-*, adv. A *ši* (< **šyæ*), fem. B *sana*, A *säm* (< **sänā*), acc. masc. B *šeme*, A *šom* < CToch. **šæmæ*, fem. B *somo* (A *šom*) < CToch. **šæmä*, etc. For the additive cardinals, ‘eleven’ through ‘nineteen’, an asyndetic construction *decad+digit* is used in Toch. B, while in A the digit is followed by a particle *-pi*, e.g., ‘eleven’ B *šak še*, A *šäk ša-pi*. The decads, with the exception of ‘twenty’, use a suffix CToch. **-kā*, which is possibly a remaking of **-kän* with substitution of the nominal suffix **-ā* of the plural (collective). A good representative of an archaic form is ‘thirty’, B *täryäka* (A *taryäk*, instead of **täryäk*, influenced by *štwaräk* ‘forty’) < **täryäkā* ← **triya-kän* < **trih₂-k^mt*; cf. Lat. *trīginta*, Ved. *triṁśát*. ‘Twenty’ is B *ikäm* < **wⁱikän*, based on an item differing from B *wi* ‘two’; this CToch. form ends up as A *wiki* after refashioning based on the model of the dual of nouns. The multiples of ‘hundred’ and ‘thousand’ are made through univerbation, featuring in Toch. B the bound forms of the respective numbers.

5.2. The ordinals from ‘second’ to ‘sixth’ contain a suffix CToch. **-tæ*, e.g., ‘second’, B *wate*, A *wät* < **wätæ* < **dui-to-s*; so do B *ikante* ‘twentieth’, and ‘tenth’, B *škante*, A *škänt* < B **šäkäntæ* < **dék^m-to-s*. Partly from this latter form, CToch. **-ntæ* > B *-nte*, A *-(ä)nt* has been extended to ‘seventh’ through ‘ninth’. As expected, the term for ‘first’ is independent from that for ‘one’: B *pärwešše* is an adjective based on the adverb *parwe*, related to Toch. A *pärwat* ‘first-born (son)’; A *maltowinu* is based on a compound with adv. *malto* ‘first’ and a participle form based on the allomorph **yän-* of the verb *yä-/i-* ‘go’, cf. B *ynūca* ‘going’. Ordinals based on decads have in Toch. A the suffix *-ñci* < **-ñcyæ*, that goes back ultimately to a doublet of **-ntæ*. The additive ordinals are formed through combination of the simple ordinals in Toch. B, whereas in Toch. A the suffix *-nt* occurs not after the digit but after the complex form with particle *-pi*, e.g., A *šäk-šapint* ‘eleventh’.

5.3. Distributive expressions are made through repetition, e.g., B *šeme šeme* (A *šom šom*) ‘one by one’, or by addition of the suffix CToch. **-ärä* > B *-är(ä)*, still with two syllables in archaic Toch. B, e.g., B *somār*, *wyār*, *škäsār*, *ñ_uwār*, *škār* ‘in groups of one (fem.), two, six, nine, ten’. The shape of this suffix precludes its equation with an adverbial suffix based on PIE collective **-ōr*.

6. Verb morphology

6.1. The following categories are expressed by finite forms

6.1.1. Person

Three persons, except for the imperative, which has only 2nd pers. forms.

6.1.2. Number

Singular, plural, and dual. But only a few dual forms are attested: 3rd du. pres. act. *-teṃ*, but all other dual endings (*-aitār*, *-ais* [A *-es*], *-ait*) are based on **-ai-* plus the plural ending of present or preterite.

6.1.3. Voice

Active vs. mediopassive. Special intransitive/passive formations are known for some verbs. Non-finite forms are normally indifferent to voice: infinitive, gerunds, preterite participle. The value of the so-called “mediopassive” present participle (with suffix B *-mane*, A *-mām*), which can be found alongside active finite forms of the same verb, does not fall under the scope of diathesis, and its usage is close to that of a converb (as in Turkic languages), referring to a secondary action or state. The contrast with the so-called “active” present participle (with suffixes B *-ñca*, A *-nt*, cognate with agent noun suffixes) is based on valence: the latter is the marked category, since it is based usually on transitive verbs.

6.1.4. Tense

In the indicative, non-past vs. past, the latter being expressed by the imperfect or the preterite.

6.1.5. Aspect

Imperfective (durational) vs. perfective (non-durational), the latter being expressed by the preterite in the indicative, in contrast to the imperfect. The stative meaning of the perfect, which has disappeared as a finite category, can be expressed by periphrasis with the preterite participle.

6.1.6. Mood

Indicative, subjunctive, optative, and imperative. The so-called subjunctive has both tense value, as a future, and modal values, in main as well as in independent clauses. Whereas an order is expressed by the imperative, and alternatively by the optative or the subjunctive, prohibition is expressed by the present or by the subjunctive, depending on its aspect: with negation (B *mā*, undifferentiated for every type of sentence, A *mar* contrasting with the basic negation *mā*), the former bears on the current situation, whereas the latter is preventive.

6.1.7. Transitivity

There is a general tendency, which is carried through most thoroughly in Toch. B, to signal by formal means the transitive character of a verb. The most common markers signalling transitivity are accent (stress) on the initial syllable (which can be observed in Toch. B), palatalization of the initial consonant of the root, and the thematic suffixes *-s-* and *-sk-* in the present and subjunctive.

6.2. A Tocharian verb has normally three stems: present, subjunctive and preterite. 1. The present stem is used to form the present tense, the imperfect with suffix B *-i-* (*-oy-* after stems ending in **-ā-*), the participles or agent nouns B act. *-ñca*, A *-nt* (cf. the agent noun suffix B *-nta*, A *-nt*) and “mediopassive” B *-mane*, A *-māṃ* (< **-mānæ*), the gerund (I) of obligation or necessity: B *-lye/-lle*, A *-l* (< **-lylyæ* < **-[i]jo-*), and the Toch. A infinitive with suffix *-tsi*. 2. The subjunctive stem is used to form the so-called subjunctive, the optative with suffix AB *-i-* (B *-oy-* after stems ending in **-ā-*), the gerund (II) of possibility, with the same *l*-suffix as above, the Toch. B infinitive with the suffix *-tsi*, and the privative (productive in Toch. B) with prefix CToch. **æ-* (< PIE **ḡ-*) and suffix B *-tte*, A *-t* < **-ta* < **-to-*, type Gk. ἄσχετος ‘irresistible’, Ved. *amṛta-*, Gk. ἄμβροτος ‘immortal’. 3. The preterite stem is used to form the preterite tense and the preterite participle, with the suffix AB *-u*, B *-au* : A *-o*, reflecting the suffix **-uos-*, nom. sg. masc. **-uōs* of the PIE perfect participle. Some of the formations of this participle keep the reduplication of the PIE perfect, while others are modelled after the CToch. preterite stems without reduplication.

6.3. Other important morphological features are the following: the imperative is built with a prefix CToch. **pā-* added to the preterite stem, less commonly to the subjunctive stem. The imperfect was made with the suffix of the optative added to the present stem. Except for the verbs ‘be’ and ‘go’, Toch. A has replaced this suffix **-i-* (< PIE **-ih₁-*) with the suffix **-ā-* of the preterite. The verbal adjectives served as basis for productive abstract formations, including the gerund-based abstracts in B *-llāññe* (*-lñe*), A *-lune*, and the abstract in *-r* based on the preterite participles (B *-uwer*, *-or*, *-ar*, A *-or*, *-ur*), from which is formed the absolutive (matching in function the Skt. absolutive in *-tvā* or *-ya*) by the addition of the ablative case affix.

6.4. The whole conjugation is governed by two fundamental distinctions

6.4.1. In both languages each verb has a double paradigm in non-finite as well as in finite forms, with causative stems contrasting with non-causative stems. The non-causative verb can be intransitive or transitive: in the former case, the corresponding causative verb is usually transitive, e.g., AB *tsälp-* ‘be free of’, caus. ‘free of, redeem’, AB *täm-* ‘be born’, caus. ‘engender, produce’; in the latter case, it may have the same value as the transitive basic verb, e.g., *käl-* ‘bring’, caus. ‘id.’, or it has ditransitive (or factitive) value, e.g., AB *kärs-* ‘know’: caus. ‘make known to someone’. Several stems are specialized as belonging to the causative paradigm.

6.4.2. Verbs with final **-ā-* vs. verbs without final **-ā-*. This vowel surfaces in the subjunctive stem and typically in the infinitive of Toch. B. It reflects originally the final laryngeal (> **-a-* between consonants, yielding CToch. **-ā-*) of PIE roots, but the contrast is far from reflecting the opposition between verbs with and without final laryngeal. Some verbs have lost, for various reasons, any reflex of this laryngeal, and conversely its CToch. reflex **-ā-* has been extended to numerous verbs, so as to become the commonest marker of the preterite. Nevertheless, the presence of this final vowel determines to a large extent the structure of the paradigm and the combination of the different stems. In Toch. B *-ā-*verbs build presents of classes III to VI and feature *-ā-* before the suffix *-sk-* of classes IX to XI; they build preterites (classes I, II) and subjunctives (class V) that are characterized by the same vowel. In transitive verbs, the present is often formed by insertion of a nasal before this vowel, producing a suffix *-nā-* on the surface (class VI). Intransitive verbs have presents belonging to classes III and IV. By contrast, non-*-ā-*verbs build presents of classes I, II, VII and VIII and alternatively presents of classes IX to XI, but without *-ā-* before the suffix *-sk-*. Accordingly the preterite participle has a form nom. sg. masc. B *-au* (A *-o*) < **-ā-wu* in the *-ā-*verbs vs. AB *-u* in the non-*-ā-*verbs.

6.5. The quite numerous verbal stems go back with relative clarity to familiar PIE types and keep root ablaut as well as ablaut of the thematic vowel **-e/o-*. The **-e-* grade of the thematic vowel is reflected by the palatalization of a preceding consonant, either of the root or of the suffix in present classes. One can recognize in the various preterite classes the athematic root aorist (class I), the thematic root aorist (class VI), the sigmatic aorist or what appears to be an outcome (or an ancestor?) of it (class III), the athematic reduplicated aorist (class II in A), etc. With some exceptions, subjunctive and present classes share the same formations: thematic (class II, relatively infrequent) or athematic root present (class I), present in **-se/o-*, in **-skē/o-* (class VIII, IX, frequent and normal in the causative paradigms, with further extension in classes X and XI), denominative or deverbative presents in **-je/o-* (class XII), nasal presents (class VII with infix, class VI with suffix *-nā-*). Classes III and IV are characterized by predominant mediopassive inflection and a non-ablauting thematic vowel, respectively B *-e-*, A *-a-* (< **-æ-*) and B *-o-*, A *-a-* (< **-ā-*), which may result from contractions (stems in **-e/o-je/o-*, or **-i-je/o-*, and **-eh₂-je/o-*, respectively). In the so-called subjunctive two classes (I and V) show descriptively a contrast between the reflex of PIE **o-* grade in the strong stem and zero grade in the weak stem, carried through to different degrees in the two languages.

Despite its apparent resemblance to the PIE perfect, the origin of this category is still much debated.

6.6. Suppletion is a well-attested feature of the Tocharian verbal system. Many verbs fill out their paradigms with stems based on two (or more) different roots, one of which, with lexically durative value, gives the present stem, and the other, with lexically non-durative value, the preterite and subjunctive stems. As often in Indo-European languages, the verbs ‘be’ and ‘go’ are diachronically suppletive and synchronically irregular. Thus, the present stem proper of the verb ‘be’ is B *nes-*, A *nas-* < **næs-* < **nos-* ‘be back home’ from the root **nes-* ‘return safely’, whereas the preterite and subjunctive stem **tākā-* is based on a suffixed form of the root *(*s*)*teh₂-* ‘to stand’. However, Toch. B possesses a specific copula based on forms (3rd sg. *ste*, pl. *skente*, *stār-* and *skentār-* before enclitic pronouns, dialectal and late *stare*) that go back ultimately to the iterative present **h₁s-ske/o-* of the root **h₁es-*.

6.7. Verbal endings

Four sets of endings are distinguished: present, imperfect and optative (uniquely for B), preterite, and imperative.

6.7.1. Present endings are used for the present and the so-called subjunctive in both languages. In Toch. A they are used also for the optative, and for the two imperfects act. of the verbs ‘be’ (A 1st sg. *šem*) and ‘go’ (A 1st sg. *yem*), which go back to earlier optatives. We will take as examples the paradigm showing thematic inflection.

6.7.1.1. Active: verb AB *āk-* ‘lead’, pres. II, stem CToch. **ākæ-/āśä-* < **h₂ágê-o-/h₂ágê-e-*.

	B		A	
sg. 1.	<i>ākau</i>	< * <i>āk-aw</i>	<i>ākam</i>	< * <i>āk-æ-m(ä)</i>
2.	<i>āšt(o)</i>	< * <i>āś-ä-tV</i>	<i>āšt</i>	< * <i>āś-ä-tV</i>
3.	<i>āšäm</i>	< * <i>āś-ä(-ž)+nā</i>	<i>āšāš</i>	< * <i>āś-ä-š</i>
pl. 1.	<i>akem(o)</i>	< * <i>āk-æ-mV</i>	<i>ākamäs</i>	< * <i>āk-æ-mäs(ä)</i>
2.	<i>āšcer</i>	< * <i>āś-ä-caer</i>	<i>āšac</i>	< * <i>āś-ä-cä</i>
3.	<i>ākem</i>	< * <i>āk-æ-n</i>	<i>ākeñc</i>	< * <i>āk-æ-ñc(ä)</i>

Athematic inflection has special endings only for the 1st sg. (B *-u* in class I presents and subjunctives, < PIE **-ō* ?) and 3rd pl. (B *-äm* A *-iñc* < **-āñcä*). An exception to this pattern is the present of the verb *yä-/i-* ‘go’, which goes back ultimately to the athematic root present of the PIE root **h₁ej-*. Whereas Toch. A has the expected forms based on a stem *yä-*, the forms of Toch. B are based on a stem **yä-* with allomorph **yän-* in 1st and 3rd pl, as well as in the participles: sg. 1. *yam*, 2. *yat*, 3. *yam*, pl. 1. *ynem(o)*, 2. *yacer*, 3. *yanem*, part. act. *yneñca*, medpass. *ynemane*. Otherwise, athematic inflection shows the same endings as the thematic inflection, minus the palatalization of the final

consonant of the root in 2nd and 3rd sg. and 2nd pl. Toch. A has an (at face value) truncated variant of the 3rd pl., e.g., *tränki* instead of *tränkiñc* ‘they speak’, *sälpe* instead of *sälpeñc* ‘they burn’. The CToch. endings reflect a mixture of indicative and injunctive endings, hence 1st sg. **-æ-mä*, which has been lenited (except in B **yä-mä* ‘I go’) to B **-ävä* > archaic B *-e_u*, standard *-au*; 3rd pl. **-nt* > **-n* (in B) vs. **-nti* > **-ñcä* (in A); 2nd sg. **-e-s* > **-ä-ø* remade with the addition of the personal pronoun **-tä* < **tu*; 3rd sg. **-e-d* > CToch. **-äz* > A *-äs*, remade in B with the addition of a particle **-nä* < PIE **nu* ‘now’. As for the 2nd pl., by contrast with Toch. A *-c(ä)*, which goes back to the ending **-te* or **-tes*, Toch. B *-cer* < **-cæ-r* may go back to **-tē*, with addition of a particle of address; 1st pl. A *-mäs* can reflect PIE primary **-mesi*.

6.7.1.2. Mediopassive: verb B *klyaus-*, A *klyos-* ‘hear’, pres. II, stem CToch. **klyausæ-/klyausä-*: **klēus-o-/klēus-e-*.

	B	A	CToch.	
sg. 1.	<i>klyausemar</i>	<i>klyosmār</i>	<i>*^os-æ-mār</i>	< <i>*-o-[m-]h₂e+r</i>
2.	<i>klyoštār</i>	<i>klyoštār</i>	<i>*^oš-ä-tār</i>	< <i>*-e-th₂e+r</i>
3.	<i>klyauštār</i>	<i>klyoštār</i>	<i>*^oš-ä-tār</i>	← <i>*-e-to-r</i>
pl. 1.	<i>klyausemt(t)är</i>	<i>klyosamtär</i>	<i>*^os-æ-mätär</i>	← <i>*-o-med^hh₂+r</i>
2.	<i>klyauštär</i>	<i>klyoššär</i>	<i>*^oš-ä-cär</i>	< <i>*-e-d^{he} (-d^{hu}e)+r</i>
3.	<i>klyausentär</i>	<i>klyosantär</i>	<i>*^os-æ-ntär</i>	< <i>*-o-ntro+r</i> ← <i>*-o-nto-r</i>

The CToch. endings are based on the PIE middle endings with generalization of the final *-r*, which is not found in any other paradigm. As in the active, one may surmise a mixture of injunctive and indicative forms. PIE **(n)to-r* ought to yield CToch. **(n)taer*, so one should assume some intermediary step in order to get the actual endings, and an interference with the expected result of the secondary ending **-ntræ* < **-ntro*. Except for the final *-r*, the endings B *-tär* and A *-cär* of 2nd pl. are incompatible, but compare B *-t* and A *-c* in the preterite medpass.: A **-cä* < PIE **-d^{he}*. One may note also that the expected 1st sg. **-h₂e-r* > **-är* has been recharacterized by insertion of the nasal of the active, a phenomenon that is paralleled in Gk. $-\mu\alpha\iota$. The 1st pl. has received the final **-är* of the 3rd sg., 3rd pl., and 2nd pl. endings.

6.7.2. Special endings of the optative and imperfect are found only in Toch. B for the active singular: 1. *-m*, 3. $-\emptyset$ the final B 3rd sg. *-i* (*-oy* for stems in *-ä-*) goes back to **-yä* < **-ih₁-t*, with secondary ending. Toch. A uses instead the preterite endings for the imperfect and the present/subjunctive endings for the optative. The imperfect of the verbs ‘be’ and ‘go’ contradicts this pattern only in synchrony, since these forms go back to an ancient optative: sg. 1. *šem*, 2. *šet*, 3. *šeš*, pl. 3. *šeñc* (B sg. *šaim*, *šait*, *šai*, pl. *šeyem*), and sg. 1. *yem*, 2. *yet*, 3. *yeš*, pl. 3 *yeñc* (B sg. *yaim*, *yait*, *yai*, pl. *yeyem*). These two verbs conceal the strong allomorph of the optative suffix **-jēh₁-* > CToch. **-yæ-*, while the weak allomorph has been generalized in the regular optative (and imperfect) formations with suffix *-i-* < *-yä-* < **-ih₁-*.

6.7.3. Preterite endings

6.7.3.1. Active: verb B *prek-*, A *prak-* ‘ask’, pret. III with stem **præk(-s)-*, **præks-ā-* in 3rd sg., cognate with the sigmatic aorist of the root **prēk-*, having lengthened grade **prēk-s-*.

	B	A	CToch.
sg. 1.	<i>prekuwa, prekwa</i>	<i>prakwā</i>	<i>*prækuwā</i>
2.	<i>prekasta</i>	<i>prakāšt</i>	<i>*prækāstā</i>
3.	<i>preksa</i>	<i>prakās</i>	<i>*præksā</i>
pl. 1.	<i>prekam(o)</i>	<i>prakmās</i>	<i>*prækāmā(s)</i>
2.	<i>prekas(o)</i>	<i>prakās</i>	<i>*prækāsā</i>
3.	<i>prekar</i>	<i>prakār</i>	<i>*prækār</i>

The zero ending of the 3rd sg. goes back to final **-t*, accordingly **-H-t* (> **-ā-ø*) in roots ending with a laryngeal. In pret. VI, which reflects the thematic aorist, the final has triggered the palatalization of the preceding consonant, cf. B *lac* ‘he went out’ < **lācā* < **h₁lud^h-e-t*. The 1st pl. shows the same ending as in the present and subjunctive. The normal 3rd pl. ending is B *-re* (A *-r*) < **-ræ*, (*-āre/-are* [A *-ār/-ar*] in *ā*-classes, depending on the place of the accent). Loss of final *-e* occurs in late and colloquial Toch. B, and especially before clitic pronouns. The short ending *-ār* (< **-ārā*) is found originally in pret. III, and always bears the accent in Toch. B, hence surfacing as *-ar*. The reflex of the aorist ending **-nt* is still found in the thematic aorist, *latem* ‘they went out’ < **lātæn* < **h₁lud^h-o-nt*, *kamem* ‘they came’ < **kāmæn* < **g^wṇm-o-nt*. The ending **(-ā)r*, by itself as well as part of the ending **-ræ*, can go back to the perfect ending **-r* or **-rs*. In the 2nd sg., **-stā* replaces secondary **-s* which had been lost without any trace in final position, and it may go back to the perfect ending **-th₂e* with epenthetic sibilant in roots ending with dental stop. The 1st sg. has in some Toch. A verbs a variant *-u*, which is the expected outcome of CToch. **(u)wā* > B *-wa*; the ending *-wā*, which is found in pret. II, III, IV and in the imperfect, has been recharacterized with the normal ending *-ā*. The latter reflects the contraction of **-ā-wā*, based on preterite stems with final **-ā-*. One may surmise that this ending **(u)wā* reflects also the perfect ending **-h₂e* > **-h₂a* > **-ā*, combined with the reflex of the previous aorist ending. The 2nd pl. ending **-sā* cannot be explained from any PIE source, and is probably analogical after the 2nd sg. ending.

6.7.3.2. Mediopassive: same verb, stem **pärksā-*, remade in Toch. A as *präksā-* on the model of *prak-* in the active, from the zero grade **prk-s-* of the aorist stem.

	B	A	CToch.
sg. 1.	<i>parksamai</i>	<i>präkse</i>	<i>*pärksai</i>
2.	<i>parkсатаi</i>	<i>präksāte</i>	<i>*pärksā-tai</i>
3.	<i>parkсate</i>	<i>präksāt</i>	<i>*pärksā-te</i>
pl. 1.	<i>parkсамt(t)e</i>	<i>präksāmät</i>	<i>*pärksā-mätē</i>
2.	<i>parksat</i>	<i>präksāc</i>	<i>*pärksā-tā</i>
3.	<i>parkсante</i>	<i>präksānt</i>	<i>*pärksā-ntē</i>

The 3rd sg. **-tæ* and pl. **-ntæ* go back to the familiar PIE endings **-to* and **-nto*; their CToch. final vowel has triggered the replacement of 1st pl. **-mätā* (< **-med^hh₂*) by **-mätæ*. In the 2nd pl. the discrepancy between the two languages can be explained by divergent generalization of primary ending **-cä* (< **-d^he*) and secondary **-tä* (< **-d^huue*). As for the 1st sg., one finds a variant A *-we*, remade after act. *-wā*, in pret. III *yāmwe* (*yām-* ‘make’) and in the imperfect; the final *-e* < **-ai* is based on the contraction of **-ā(w)ai*. This ending has been made more transparent in Toch. B by insertion of *-m-* before **-ai*, parallel to present *-m-ār*. The final **-ai* has been extended to the 2nd sg., replacing expected **-tā* < **-th₂e*. This final diphthong is enigmatic. One may consider the ending as a direct reflex of PIE **-h₂ai*, comparable to Lat. *-ī* < **-ai*, etc., but a primary ending (marked with the near deictic particle **-i*) is not expected in a preterite paradigm. Alternatively, one can trace back CToch. **-ai* by sound law to **-ān(ā)* or **-āñ(ā)* (the source of the acc. sg. ending B *-ai*, cf. 2.5 above). This may have started from the reflexive middle, by the addition of the 1st person clitic **-ñā* to the 1sg. medpass. **-ā* (< **-h₂a*), which had merged at some stage with the 1sg. perfect (active) ending **-ā* (< **-h₂a*).

6.7.4. Imperative endings

The imperative is inflected only in the 2nd sg., pl., and dual. The plural endings, active and middle, are actually the endings of the preterite. The sole specific ending is 2nd sg. *zero* in the active, with the addition of *-r* in the mediopassive: e.g., to the stem **tākā-* of the pret. and subj. of the verb ‘be’, act. B *ptāka*, pl. *ptākas* (A *pāštāk*, pl. *pāštākās*) from the verb *yām-* ‘make’, to the pret. III stem **yām-sā-*, medpass. B *pyāmtsar*, pl. *pyāmtsar*, du. *pyāmtsait* (A *pyāmtsār*, pl. *pyāmtsāc*). The situation is made complicated by the absence of the final **-ā* of the stem in the 2nd sg. act. of some imperatives (e.g., AB act. *pyām*, pl. B *pyāmtso*, A *pyāmās*) and by various deviant 2nd sg. act. forms in common verb types: in *-s*, cf. B *paş*, A *piş* ‘go!’ (verb *i-/yā-*), A *paş* ‘give!’ (verb *e-*, B *ai-*), B *päklyauş*, A *päklyoş* ‘listen!’; in *-e*, cf. B *pokse* ‘explain!’ (verb *āks-*), hence 2nd pl. act. in *-es*, replacing less transparent forms, e.g., *petes* instead of older *petso*, pl. of *pete* ‘give!’.

7. References

- Adams, Douglas Q.
 1988 *Tocharian Historical Phonology and Morphology*. (American Oriental Series 71). New Haven: American Oriental Society.
- Carling, Gerd
 2000 *Die Funktionen der lokalen Kasus im Tocharischen*. Berlin: Mouton de Gruyter.
- Hackstein, Olav
 1995 *Untersuchungen zu den sigmatischen Präsensstambildungen des Tocharischen*. Göttingen: Vandenhoeck & Ruprecht.
- Hilmarsson, Jörundur
 1989 *The dual forms of nouns and pronouns in Tocharian*. (Tocharian and Indo-European Studies Supplementary Series 1). Reykjavik: Málvísindastofnun Háskóla Íslands.

Hilmarsson, Jörundur

- 1991 *The nasal prefixes in Tocharian. A study in word formation.* (Tocharian and Indo-European Studies Supplementary Series 3). Reykjavik: Málvísindastofnun Háskóla Íslands.

Jasanoff, Jay H.

- 2003 *Hittite and the Indo-European Verb.* Oxford: Oxford University Press.

Klingenschmitt, Gert

- 1994 Das Tocharische in indogermanistischer Sicht. In: Bernfried Schlerath (ed.), *Tocharisch. Akten der Arbeitstagung der Indogermanischen Gesellschaft (Berlin, September 1990).* (Tocharian and Indo-European Studies Supplementary Series 4). Reykjavik: Málvísindastofnun Háskóla Íslands, 310–411.

Krause, Wolfgang

- 1952 *Westtocharische Grammatik.* Band I: *Das Verbum.* Heidelberg: Winter.

Krause, Wolfgang and Werner Thomas

- 1960–1964 *Tocharisches Elementarbuch.* Band I: *Grammatik.* Band II: *Texte und Glossar.* Heidelberg: Winter.

Malzahn, Melanie

- 2010 *The Tocharian Verbal System.* Leiden: Brill.

Melchert, H. Craig

- 2000 Tocharian plurals in *-nt-* and related phenomena. *Tocharian and Indo-European Studies* 9: 53–75.

Peyrot, Michaël

- 2008 *Variation and Change in Tocharian B.* (Leiden Studies in Indo-European 15). Amsterdam: Rodopi.

Peyrot, Michaël

- 2013 *The Tocharian Subjunctive. A study in syntax and verbal stem formation.* Leiden: Brill.

Pinault, Georges-Jean

- 2008 *Chrestomathie tokharienne. Textes et grammaire.* (Collection linguistique publiée par la Société de Linguistique de Paris 95). Leuven: Peeters.

Pinault, Georges-Jean

- 2009 On the formation of the Tocharian demonstratives. In: Elisabeth Rieken and Paul Widmer (eds.), *Pragmatische Kategorien. Form, Funktion und Diachronie. Akten der Arbeitstagung der Indogermanischen Gesellschaft vom 24. bis 26. September 2007 in Marburg.* Wiesbaden: Reichert, 221–245.

Pinault, Georges-Jean

- 2012 Tocharian *-nt-* participles and agent nouns. In: Olav Hackstein and Ronald I. Kim (eds.), *Linguistic developments along the Silkroad: Archaism and innovation in Tocharian.* (Multilingualism and history of knowledge II). Vienna: Österreichische Akademie der Wissenschaften, 179–204.

SSS = Emil Sieg, Wilhelm Siegling, and Wilhelm Schulze

- 1931 *Tocharische Grammatik.* Göttingen: Vandenhoeck & Ruprecht.

Stumpf, Peter

- 1971 *Der Gebrauch der Demonstrativ-Pronomina im Tocharischen.* Wiesbaden: Harrassowitz.

TEB see Krause and Thomas.

Van Windekens, Albert-Joris

- 1979 *Le tokharien confronté avec les autres langues indo-européennes.* Vol. II, 1: *La morphologie nominale.* Louvain: Centre International de Dialectologie Générale.

Van Windekens, Albert-Joris

- 1982 *Le tokharien confronté avec les autres langues indo-européennes.* Vol. II, 2: *La morphologie verbale.* Louvain: Centre International de Dialectologie Générale.

Winter, Werner

1992 Tocharian. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. (Trends in Linguistics. Studies and Monographs 57). Berlin: De Gruyter, 97–161.

Winter, Werner

2005 *Kleine Schriften. Selected Writings*. Band 1: *Tocharisch und Indogermanisch*. Ausgewählt und herausgegeben von Olav Hackstein. Bremen: Hempen.

Georges-Jean Pinault, Paris (France)

77. The syntax of Tocharian

- | | |
|---|--|
| 1. Word classes: a typological survey | 6. Verbal morphosyntax and periphrastic formations |
| 2. Typologically distinctive features | 7. Adverbials |
| 3. Independence of the syntax in translated texts | 8. Word order |
| 4. Stylistics and metrics | 9. Sentence syntax |
| 5. Nominal morphosyntax and adpositional phrases | 10. References |

1. Word classes: a typological survey

Fundamentally, Tocharian syntax corresponds to that of most other Indo-European languages, such as Old Indo-Aryan or Greek. The basic system is accusative and the backbone of the grammar is an inflectional system which in its rough outlines can be traced back to the Indo-European system (cf. Pinault, this handbook).

To begin with the open word classes, nouns, verbs, adjectives, and adverbs (cf. Schachter 1985: 5 f.), Tocharian noun inflection shows in most instances an important distinction between human and non-human (see Pinault 1989: 77; TEB I §141). The former shows a special oblique marker that allows a distinction between Agent and Patient in most paradigms. Furthermore, the inflectional-agglutinative structure of the nominal system, especially the division into primary and secondary cases, has had consequences for the syntax (see 5.1). The verbal system, like that of the noun, is partially based on Indo-European inflection, and partly on a system of suffixes used mainly to mark valency (see 6.1). Adjectives are basically inflected, but a large number of them are uninflected, which often blurs the boundary between adjectives and adverbs. Adverbs show no distinctive morphological marking. A number of adverbs are used also as adpositions (pre-/postpositions).

Tocharian is a heavily synthetic language, with rich morphemic systems in both the nominal and the verbal system. This gives the closed word classes of pronouns, noun adjuncts (adpositions), verb adjuncts (auxiliaries), particles, conjunctions, clitics, copula, interjections, and negations a less prominent role in the syntax, as is typical for synthetic languages (cf. Schachter 1985: 23 f.). Among the pronouns, the three-fold deictic stems

Winter, Werner

1992 Tocharian. In: Jadranka Gvozdanović (ed.), *Indo-European Numerals*. (Trends in Linguistics. Studies and Monographs 57). Berlin: De Gruyter, 97–161.

Winter, Werner

2005 *Kleine Schriften. Selected Writings*. Band 1: *Tocharisch und Indogermanisch*. Ausgewählt und herausgegeben von Olav Hackstein. Bremen: Hempen.

Georges-Jean Pinault, Paris (France)

77. The syntax of Tocharian

- | | |
|---|--|
| 1. Word classes: a typological survey | 6. Verbal morphosyntax and periphrastic formations |
| 2. Typologically distinctive features | 7. Adverbials |
| 3. Independence of the syntax in translated texts | 8. Word order |
| 4. Stylistics and metrics | 9. Sentence syntax |
| 5. Nominal morphosyntax and adpositional phrases | 10. References |

1. Word classes: a typological survey

Fundamentally, Tocharian syntax corresponds to that of most other Indo-European languages, such as Old Indo-Aryan or Greek. The basic system is accusative and the backbone of the grammar is an inflectional system which in its rough outlines can be traced back to the Indo-European system (cf. Pinault, this handbook).

To begin with the open word classes, nouns, verbs, adjectives, and adverbs (cf. Schachter 1985: 5 f.), Tocharian noun inflection shows in most instances an important distinction between human and non-human (see Pinault 1989: 77; TEB I §141). The former shows a special oblique marker that allows a distinction between Agent and Patient in most paradigms. Furthermore, the inflectional-agglutinative structure of the nominal system, especially the division into primary and secondary cases, has had consequences for the syntax (see 5.1). The verbal system, like that of the noun, is partially based on Indo-European inflection, and partly on a system of suffixes used mainly to mark valency (see 6.1). Adjectives are basically inflected, but a large number of them are uninflected, which often blurs the boundary between adjectives and adverbs. Adverbs show no distinctive morphological marking. A number of adverbs are used also as adpositions (pre-/postpositions).

Tocharian is a heavily synthetic language, with rich morphemic systems in both the nominal and the verbal system. This gives the closed word classes of pronouns, noun adjuncts (adpositions), verb adjuncts (auxiliaries), particles, conjunctions, clitics, copula, interjections, and negations a less prominent role in the syntax, as is typical for synthetic languages (cf. Schachter 1985: 23 f.). Among the pronouns, the three-fold deictic stems

of the demonstratives (cf. Morphology, Stumpf 1971: 90 ff.; TEB I § 266–269; Pinault 1989: 116–117) are of particular interest from a semantic and typological perspective (cf. Anderson and Keenan 1985: 286 f.). Personal pronouns have independent as well as cliticized forms, although the use of the latter is restricted to certain syntactic positions (see 5.2). Adpositions include both pre- and postpositions, though postpositions are more common. However, because of the rich case system (see 5.1) adpositional phrases are not as common as case constructions in discourse. The pre- and postpositions are syntactically connected with certain cases (see 5.3). Likewise, verbal particles and auxiliaries occur, but they are not numerous in discourse. Most semantic and syntactic nuances are expressed by the rich verbal system (see 6).

2. Typologically distinctive features

Tocharian has a number of typologically distinctive features that differ from other Indo-European languages and can be explained as a result of the restructuring of pre-Tocharian. They are:

1. agglutinative technique in the case paradigm (see K. H. Schmidt 1975, 1982)
2. group inflection (see 5.1) and lack of concordance of head and modifier in Noun Phrases (cf. K. H. Schmidt 1974)
3. case-inflected adpositions and adverbials (cf. Carling 1999)

These innovations have been ascribed to substratum or areal influences, possibly from Caucasian, Uralic-Altaic, Turkic, and Finno-Ugric (K. H. Schmidt 1990; Thomas 1994), or internal evolution. The almost parallel developments in the Indo-Aryan languages may serve as an argument for internal evolution (cf. Carling 2005: 49–52).

3. Independence of the syntax in translated texts

Tocharian literature is heavily dependent on Buddhist Sanskrit literature, even though there was also an independent literary tradition (cf. Pinault 1989: 12 f., 2008: 7–8; Winter 1955). Because most texts, with the exception of some fragments in Tocharian B that are mostly of economic and administrative character, have originals in Sanskrit, one might think that Sanskrit would have influenced aspects of Tocharian syntax. However, the large typological differences between the languages, for instance the inability of Tocharian to create lengthy compounds (for the principles of compounding in Tocharian, see Bernhard 1958), speak against any deeper influence. The independence of the Tocharian language in texts directly translated from Sanskrit has been demonstrated in a number of studies by Werner Thomas (for an overview, see Thomas 1985: 91–93). The presence of parallel categories, such as the absolutive (Toch. A *-äṣ*, B *-mem* with a verbal noun in *-r*), corresponding to the Sanskrit absolutives in *-tvā*, or correlative constructions, like Toch. A *āntan(n)ene ... tāmne*, Skt. *yena ... tena*, do not provide enough basis to infer direct syntactic influence from Sanskrit (see Pinault 1997: 467–470).

4. Stylistics and metrics

Typical for Tocharian syntax are fixed formulas, consisting of word pairs, either nouns, like A *tkaṃ ākās* ‘earth [and] atmosphere’, or verbs, like A *ārtant pālant* ‘they honoured and praised’, combinations of verb + noun, as A *ākālac kāka-* ‘request (somebody) for a desire’, verb + adjective, adjective + noun or fixed phrases with a case construction, e.g., A *ākārmunt aśān-yo* ‘with tearful eyes’. Formulas are also found in the presentation of names or titles and in colophons.

Many Tocharian texts are of a metrical nature, and metrics has a relatively strong influence on syntax, in particular inversion of the normal word order (see 8.). Another phenomenon typical of metrical texts is chiasmus, which means a mirrored repetition of either parts of Noun Phrases, e.g., adjective – substantive : substantive – adjective or substantive – adjective : adjective – substantive, or of sentence arguments, e.g., Subject – Predicate : Predicate – Subject, Predicate – Subject : Subject – Predicate, Object – Predicate : Predicate – Object, Predicate – Object : Object – Predicate (see Zimmer 1976: 84 f.; Thomas 1985: 102).

5. Nominal morphosyntax and adpositional phrases

5.1. Case functions

In Tocharian, grammatical relations are encoded almost entirely in the Noun Phrase, either by case marking or by adpositional phrases.

Morphologically, the Tocharian case paradigm is divided into two segments: inflectional and inflectional/agglutinative. This corresponds roughly to the syntactic usages of the paradigm: the primary (inflectional) cases, i.e., nominative, oblique, and genitive, basically encode syntactic core functions (Subject, Agent, Object, Indirect Object), whereas the secondary (inflectional/agglutinative) cases are basically used for all kinds of non-core functions, such as location in space and so forth. The S/A is marked by the nominative and the O by the oblique. The case of the IO is genitive, or with particular verbs, the oblique or exceptionally the allative. The genitive is also used in the role of Logical Subject in inverse constructions, which occur for instance with the verbs A *nes-*, B *nes-* ‘be’, or AB *māsk-* ‘become’. Clitic pronouns also occur frequently in syntactic core positions (see 5.2). A particular feature of Tocharian noun phrases, especially those showing the same secondary cases, is group inflection, whereby, as an alternative to marking such cases on each individual noun, it is possible to mark only the last noun in the group with the agglutinative case marking. And the same procedure is often found with the more basic genitive case (Carling 2000, 2005, 2008).

Turning to non-core functions, we find the oblique, which can express Goal with a certain set of verbs, and Extension and Distribution in space. The genitive can denote Inalienable and Alienable Possession. It is also used as Agentive with non-finite verb forms (preterite participle, gerundive, infinitive) and finite verbs in the passive. In this function, the genitive marks the Logical Subject (see Zimmer 1985: 562). The genitive’s role in modifying nouns is shared with the productive denominal adjective formation (see Adams, this handbook).

Among the secondary cases (Carling 2000) we consider first the perlocative (A *-ā*, B *-sa*), a case with a wide range of functions. As a local case it denotes Path as well as Incoherent Location and extended to time it denotes a non-limited time period, e.g., B *piś kṣumntsa* ‘during the fifth year of reign’. Further, the perlocative can be used to designate the Agent with the preterite participle, gerundive, and infinitive. In Tocharian B the perlocative denotes Instrumental, a function that is expressed by a separate instrumental case in A. Instrumental functions of the perlocative can occur in Tocharian A as well. In Tocharian B comparative constructions, the perlocative is used with nouns to denote the standard followed by the comparative (*cisa lāre* ‘dearer than you’), which is itself not marked by any special morpheme. In addition, the perlocative denotes Cause in both Tocharian A and B, even though there is a separate case for that in Tocharian B.

The locative (A *-am*, B *-ne*) expresses basically Location, but also Motion within a limited space, and Goal Attainment. The local functions of the locative and perlocative overlap, but minimal pairs reveal several important distinctions (locative ~ perlocative): in ~ on, coherent ~ incoherent, permanent ~ non-permanent. In temporal expressions the locative marks a defined period of time during which an action takes place, e.g., A *aṣṭam konam* ‘on the eighth day’. The locative is also used with nouns in superlative constructions to indicate the class over which preeminence obtains (e.g., B *rṣākeṃne śpālmeṃ* ‘the most excellent among the Ṛṣis’). As was the case with the comparative, the superlative shows no special morphological marking.

The allative (A *-ac*, B *-ś[c]*) as a local case denotes Direction (without Goal Attainment) and is therefore basically used with motion verbs. However, it occurs also as second argument with some individual verbs, such as AB *kārk-* ‘bind (against)’, AB *lānk-* ‘hang (against)’. The ablative (A *-āṣ*, B *-meṃ*) denotes Source. Like the perlocative, it is used to mark the standard of comparatives. A frequent function is the absolute ablative with verbal nouns in *-r*, corresponding to the Sanskrit formations in *-tvā*, e.g., A *yāmurāṣ* B *yāmormēṃ* ‘having done, doing’ (instrumental and perlocative can be used in this construction as well). The function of the comitative (A *-śśāl* B *-mpa*) is to denote accompaniment. The instrumental (A *-yo*) occurs only in Tocharian A and is used both concretely, e.g., A 91 a3 *wsā-yokās poken-yo añcāl pañwāṣ* ‘with his gold-coloured arms he is drawing the bow’, and abstractly, e.g., A *tsopatsām kācke-yo* ‘with great joy’. The instrumental is also used in temporal expressions denoting time extension. As a marker of Agentive, the instrumental is used with inanimate objects only, beside the perlocative (in Tocharian A). The causal (*-ñ*) occurs in Tocharian B only. Like the comitative, it has a relatively simple function: it denotes the emotional basis of an action (e.g., *treme[m]ñ* ‘out of anger’).

The perlocative, locative, ablative, causal, and instrumental are also used in adverbs/adpositions.

5.2. Clitic pronouns

Clitic pronouns (Carling 2006), found in 1–3 singular and plural, are used generally in syntactic core positions. They occur as Direct Object of mono- and ditransitive verbs, e.g., B H.149.X.5 a5 *kāryorttau ksa lyāka-ne* ‘some merchant saw her’, Indirect Object of ditransitive verbs, e.g., B 492 a3 *parso lyāwa-ś* ‘I have sent you a letter’. They are also

used as Logical Subject in inverse constructions, e.g., B Monastery letter *sātre lauke mänketär-ne* ‘and grain is lacking far away for us’. Furthermore they can denote Alienable and Inalienable Possession, replacing a genitive, e.g., A 253 b6 *wināsamās-ši tosām krañt pñintu* ‘we honour these your good merits’ and be used to signal the Agent in passive constructions, e.g., B 11 a4 *laute mā kätkojtär-me* ‘the right moment should not be passed over by you’.

5.3. Adpositional phrases

Compared to case constructions, adpositional phrases are less common in discourse. Often, an adpositional phrase is used to further define location already expressed by the verb/case. Thus, an adpositional phrase can mark degree of coherence, e.g., A 144 a1 *mām̄tne kokāśśi lām̄ś [tsopatsām wrā eṣāk ... ymām̄]* ‘like the king of the ruddy geese [moving over a great water]’, or degree of penetration into a Reference Object, e.g., B 46 b5 *troñne eneñka* ‘(they hide) deep inside a cave’. Otherwise, special situations that are not clearly enough expressed by a case only, are marked by an adpositional phrase. Adpositions themselves are often, like adverbs, frozen case forms. Normally they govern the genitive, especially if the form is transparent. However, other cases may occur with adpositions as well (cf. 7).

6. Verbal morphosyntax and periphrastic formations

The verbal system, like that of the noun, shows a rich inflection and a large number of morphological variants. A verbal root with a full set of derivations can be expected to have the following stems: present, subjunctive, preterite, preterite participle, and imperative. Further, these stems normally (but not for all verbs) occur two times or more, forming the base verb (Grundverb) and causative paradigms. This chapter will not deal further with the morphological structure of the system, except to summarize its semantic-syntactic realizations.

6.1. Valency

The main purpose of the base verb ~ causative system is to denote change in valency (valency change [decrease] can be marked by the mediopassive endings as well, cf. 6.3). By means of the base verb ~ causative system the lexical valence of a verb can be increased. The valency of the lexical root corresponds to the active form of the present tense of the base verb (exceptions are *media tantum*). A very common pattern is base verb = intransitive ~ causative = transitive, e.g., A *ritwa-* B *ritte-* ‘connect’ ~ A *ritws-* B *rittäsk-* ‘join, put together’. Other patterns are base verb = transitive ~ causative = ditransitive, e.g., A *kālpnā-* B *kālpāsk-* ‘find, get, achieve’ ~ A *kālpāṣ-* B *kalpāsk-* ‘bestow’. A particular type with three paradigms occurs mostly in Tocharian B (roots ending in *-k*). Morphologically, we would here expect base verb = intransitive, causa-

tive 1 = transitive, causative 2 = ditransitive, but the semantic patterns of these verbs have been blurred and there is no real example of this structure (for details see TEB I § 376; Carling 2003: 69–73).

6.2. Number and person

Verbs are inflected in 1–3 person singular and 1–3 plural. Dual forms occur, but only with certain verbs and categories, and in certain persons. There are imperative endings in 2nd person singular, plural, and dual. The forms of 2nd plural/dual can also be used with Hortative function in the 1st plural/dual (see Pinault 2005).

6.3. Voice

Besides a set of active endings, Tocharian also has a set of mediopassive endings (cf. Pinault, this handbook). These can be divided into three structural types (cf. K. T. Schmidt 1974): 1. *media tantum*, i.e. verbs inflected in mediopassive only, 2. *medio-actives*, i.e. verbs inflected as mediopassive in the present, active in the preterite, and mediopassive/active in the subjunctive, and 3. *mediopassives* showing matching active forms throughout their entire paradigm.

The semantic types encompassed by mediopassive inflection include first of all typical Reflexive middles, such as A *yäštär* B *yatäštär** ‘adorns (oneself)’ and Reciprocal middles, such as A *kroptär* ‘gathers together’, e.g., B PK.AS.6B (=A1 Lévi) a6 *sonopitär likšitär wätsanma krenta yäššitär* ‘he anointed himself, washed himself and put on good clothes’. It is interesting to observe that these middles can also be used in non-reflexive and non-reciprocal function, e.g., A *äyantu kropamt* ‘they collected bones’. But there are cases where the non-reciprocal form is marked by an active, e.g., B 42 b7 *näskäsyem lyikš(y)e(m-ne)* ‘they bathed and washed him’. We find other typical middle functions as well, such as those denoting Body actions, A *šewimtär* ‘yawns’, Emotions, A *šerttär** ‘cries’, Speech actions, B *kwätär* ‘cries’, Spontaneous events, both of animate subjects, B *kwretär* ‘grows old’, A *šalpatär* B *tsälpetär* ‘crosses, is delivered’, and inanimate subjects, denoting Disruption, A *aratär* B *orotär* ‘ceases’, Motion, A *klawatär** B *kloyotär* ‘falls’ (also of animates), Position, A *tränkäštär* B *treñkastär* ‘is attached’, Physico-chemical change, B *pälketär* ‘burns’, Disruption of material integrity, A *wikatär* B *wiketär* ‘disappears’, Change of shape or internal structure, A *asatär* B *osotär* ‘becomes dry’. Here, an intransitive mediopassive can occur beside a transitive active of the same lexical root, e.g., A *nkäš* B *nakšäm* (active) ‘destroys’, A *näknäštär* B *nakštär* (mediopassive) ‘disappears, perishes’. Middles indicating Emotion can be either intransitive, A *planta-tär** B *plontotär* ‘is satisfied, enjoys’, or transitive, A *yärksatär** ‘reverses’.

Another function of the mediopassive is to signal outright Passive with transitive verbs, either with or without an Agent (see K. T. Schmidt 1974: 203 ff.); for the case of the Agent see 5.1–5.2). Passive values may also occur with verbs that are *media tantum*, but in these instances the passive function must be determined by the construction.

Turning to verbs with both active and mediopassive forms, we find cases where the difference in diathesis signals transitivity change, as in A *äkl-* (active) ‘teach’, (medio-

passive) ‘learn’ (for this verb, in Tocharian B this difference is marked by change of verbal stem, cf. 6.1) and other instances where it signals change of semantic content but not of transitivity, e.g., B *aun-* A *on-* (active) ‘meet’, (middle-passive) ‘begin’. In other cases, there is no difference in either semantics or transitivity, as in A *āk-/wā-*, B *āk-/wāya-* (active/mediopassive) ‘lead, bring’.

6.4. Tense and aspect

The Tocharian tense system has three dimensions: Past, Present, and Future. Present is denoted by the present tense, which can have other functions as well (see below). Past is denoted by the imperfect and preterite and occasionally by the historical present (for perfect and pluperfect see 6.7). Future is marked by present, subjunctive, or the periphrastic future (6.7). As in several Indo-European languages Aspect dimensions are coded in the Past tense (cf. Sasse 2002; Dahl 1985: 81 ff.). Basically, the imperfect marks Imperfective and the preterite Perfective, as in B H.149.add.8 b7 *pañākte phalmaguti cakesa mäskiträ* (impf.) *wārtoṣṣe ikene aśup śamānents ākṣa* (pret.) ‘Buddha was near the river Phalgumafī in a woody place. He told the monks about the body in decay’.

However, the situation is more complicated in reality and a reading of Thomas (1957) reveals a number of unexpected usages of the imperfect and the preterite, two forms which continuously interchange in discourse. Here are a few examples: events seen as more important for the story can be marked by a preterite, whereas less important events are marked by an imperfect. Embedded stories, e.g., in direct speech, are marked by imperfect, whereas other events in past tense are marked by preterite. Repeated actions or events are normally marked by imperfect, but, if they are seen as individual and momentary, i.e. as taking place one after another, preterite is used. Preterite is the usual form for narrative texts, often interrupted by absolutes, and sometimes by a historical present (cf. above). Preterite is also the usual form of verbs in economic and administrative documents referring to events in the past tense, as well as in remarks and colophons.

Another important function of the preterite, like the perfect of Indo-European and the classical languages, is to express events seen in Perfect aspect (cf. Sasse 2002: 209), e.g., A 253 a5 f. *maitreyasamiti nā(ṭkaṃ guru)darśaṃ ñomā śākṣapint nipānt ār* [pret.] ‘in the Maitreyasamiti-nāṭaka the 11th act, named Gurudarśana, has come to an end’. An alternative construction is here the periphrastic perfect (see 6.7).

6.5. Mood

The modal system consists of the indicative, the subjunctive, the optative, and the imperative. As in the classical languages, subjunctive and optative are more frequently used in subordinate than in main clauses.

In main clauses, the subjunctive denotes Obligative, e.g., A 221 b4 *štwar-wkām kārmet-sunentu kārseñc* [subj.] *ākṣiñlūne-yo* ‘they will understand through the teaching of the four-fold truths’, whereas the optative denotes Volitive and Speculative, e.g., A 23 a5 *tārkor kälpmār ṣñi-tsar ptāñkāt yrāṣimār* ‘I would like to get permission, I would wash Buddha with my own hand’. For Imperative and Jussive the imperative, subjunctive, or

optative can be used. The imperative is here the “strongest form”, used normally, but not exclusively, in the 2nd Sg./Pl./Du. The subjunctive as Obligative and the optative as Volitive represent “weaker” requests (Jussives), e.g., A 342 a5-b1 *oñi-cmolši rākšats (sām) kalkaš* [subj.] *lo ymār caš ypeyāš* ‘The Rākšasa related to human birth should immediately leave this country!’ Jussives, normally in the 3rd person, are impersonal and can also be marked by a verbal noun (type I) or an infinitive, as is typical in medical texts, e.g., B PK.AS.9 a4 *te pākšälle* [verbal noun] *šälype lipātsi* [inf.] ‘this should be cooked and the oil should be left over’.

In subordinate clauses we find the subjunctive in relative, temporal, modal, local, and conditional clauses, and the optative in local, temporal, modal, final, and conditional clauses. Unfortunately, sentence syntax (cf. below) and subordination is a poorly investigated field within Tocharian syntax, which makes it difficult to summarize the various usages of the moods in subordinate clauses.

6.6. Non-finite forms

Non-finite forms of verbs can be used as verbal complements (cf. 6.7), adverbials, Noun Phrases, or predicates.

The infinitive, which is not bound to tense-mood-aspect, is used in subordinate clauses (infinitive clauses), when the subject of the sub- and superordinate sentences is the same (see Thomas 1954). It can also be inflected as a noun with secondary case affixes, e.g., B 45 b3–4 *toṃ ykenta wnołmeṃts nestsine* [loc.] ‘the places of beings are in existence’. The infinitive can also serve as the predicate under certain circumstances. There are three types of participles: active present, mediopassive present, and preterite (for their formations see Pinault, this handbook). The active present participle is inflected and serves normally as a complement to a noun, but it can also be used as a separate Noun Phrase (see Dietz 1981).

Mediopassive participles are uninflected and represent normally an event parallel or simultaneous with the predicate, e.g., B 81 b2 *tumeṃ cai brāhmaṇi tot ike postām ynemane* [medpass. part.] *araṇemiñ lānte yapoyne kameṃ* ‘thereupon the Brahmins, walking one behind the other, came into the kingdom of king Araṇemi’ (cf. Dietz 1981: 74 f.; TEB I § 315).

The preterite participle can have both active and passive values and can be used as an independent Noun Phrase, a complement to a noun, or with or without copula in the value of a predicate (see Saito 2006: 64–68).

There are two types of gerundives, I and II, of which the basic difference is to code Necessity (I) or Possibility (II) (see Thomas 1952; TEB I § 318; for the modality cf. Palmer 2001: 72–74). Like active present participles, gerundives can be nominalized, in which case their original modality is often weakened. Gerundives are used in periphrastic constructions together with past participles (see 6.7).

6.7. Periphrastic constructions

Periphrastic constructions are less common in Tocharian syntax than independent forms (cf. 1.) All periphrastic constructions are based on a form of the copula A *nes-* B *nes-* with forms of the preterite participle or gerundive.

Various forms of the copula can be used with the preterite participle. With the present of the copula the syntagm is used in the sense of a perfect in other Indo-European languages and with the imperfect/preterite in the sense of a pluperfect. The difference between the imperfect and preterite of the copula is not completely clear. The copula can also be in the subjunctive or optative, denoting various degrees of Irrealis, e.g., A 130 b5 *kusne mā walu tās* [subj.] *mā pat wlatär* [subj.] ‘who would not be dead or die’.

Gerundive II with the present of the copula (mainly in negative clauses) denotes Future, and gerundive II with the imperfect of the copula is used to denote past Irrealis, normally in subordinate clauses, but exceptionally also in main clauses. With the subjunctive of the copula, gerundive II appears in subordinate clauses to signal a hypothetical action.

7. Adverbials

Under this heading, adverbs and adverbial phrases will be dealt with, whereas adverbial clauses will be dealt with under 9.

As expected, adverbs are more common than adverbial phrases (cf. 1.). Furthermore, many adverbs also serve as adpositions, and in some instances it is difficult to distinguish the two. There are three types of adverbs (see SSS § 386; Carling 1999: 101): 1. adjectives used as adverbs without any formal marking, 2. (frozen) case forms of nouns, 3. adverbs that appear only as such.

As adpositions (type 2), case inflected nouns, whether oblique, instrumental, locative, perlative, or allative, can govern oblique and genitive, but also perlative, allative, locative, instrumental, and comitative. All these cases, except for the genitive and to a certain extent the oblique (cf. 5.1), can function as independent case-inflected Noun Phrases, which blurs the distinction between adverbs and adpositions.

8. Word order

Tocharian word order (Zimmer 1976; Knoll 1994) is a complicated issue, in particular because of the large number of metrical texts in which an inversion of normal word order is frequent. However, even in prose texts the word order is not fully fixed, and inversion can be used as a stylistic device.

Basic word order is SOV, postpositions are more common than prepositions, adjectives normally precede nouns, and adverbs normally precede verbs. Thus, Tocharian follows expected universal patterns (see K. H. Schmidt 1982).

Initial position of the verb is used to mark the importance of an event or to focus it. Jussive and Imperative verb forms are normally in initial position. In questions the order

is normally reversed to VSO; in Tocharian A the verb is preceded by the particle *te*. Verb-initial position is also found in formulas in dramatic literature, e.g., A *lcār poñś* ‘all went out [from the stage]’, or in Buddhastotras, e.g., B *wināsam-ci ... puṇḍarikām* ‘I honour you ... Lotus flower’.

The normal position of the verb is sentence-final, sometimes followed by a particle, i.e. conjunction, subjunction, postposition, clitic pronoun; an adverbial; or a non-core Noun Phrase. We find also examples of medial position, SVO, OVS, and VOS, all of which seem to be caused by rhetorical and metrical demands.

Likewise, an inversion of the pattern adjective – noun is frequently found in metrical texts, where we also find other peculiarities, such as separation of adjective – noun or adjective – verb syntagms, or circumposition of modifiers: adjective – noun – adjective.

9. Sentence syntax

The copula is not obligatory in the present tense with periphrastic constructions (cf. 6.7). Both Tocharian A and B have the ability to construct so-called nominal sentences without a finite verb.

9.1. Coordination

Tocharian has a number of coordinating particles, e.g., Conjunctive B *spä* A *škam*, which is sentence-coordinating, and B *wai* A *yo*, which is word/phrase-coordinating, both meaning ‘and’, Disjunctive B *ra* A *pe* ‘even, also’, B *no* A *nu* ‘but, then’ (in Wackernagel position), B *wat*, *wat no* and A *pat*, *pat nu* ‘or’ (in Wackernagel position). A *yo* ‘and’ is in metrical texts normally positioned after the coordinated parts (words/phrases) and this is sometimes, but perhaps not normally, the case also with the other coordinating particles, both word/phrase and sentence-coordinating (cf. 8.)

9.2. Subordination

There are several types of subordinate clauses: Relative clauses, characterized by inflected relative pronouns, various adverbial clauses expressing Time (with A *äntāne* B *ente* ‘when’), Location (with A *äntā*, *äntāne* B *ente* ‘where’), Cause (with A *kucne* B *kuce* ‘because’), Purpose (with A *mämtne* B *mäkte* ‘in order to’), and Condition (with A *ku-prene* B *kwri*, *kru* ‘if’). In Conditional clauses the form of the predicate is often a subjunctive or an optative (cf. 6.5). The subordinating particle is often in sentence-final or sentence-medial position, a pattern which does not seem to be dependent on metrics, and the subordinating particle can be omitted if the predicate is a subjunctive.

(I thank Werner Winter, Georges-Jean Pinault, Folke Josephson, and Douglas Adams for valuable remarks. I thank Judith Josephson for correcting my English. N. B. Categories with initial capital letter, e.g., Jussive, refer to general terms, whereas those with

non-capital letter refer to Tocharian terms, e.g., ablative, middle-passive. An asterisk* after the word indicates that it has been restored by internal reconstruction.)

10. References

- TEB I = Krause, Wolfgang and Werner Thomas
 1960 *Tocharisches Elementarbuch*. Band I: *Grammatik*. Heidelberg: Winter.
- TEB II = Thomas, Werner and Wolfgang Krause
 1964 *Tocharisches Elementarbuch*. Band II: *Texte und Glossar*. Heidelberg: Winter.
- SSS = Sieg, Emil, Wilhelm Siegling, and Wilhelm Schultze
 1931 *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.
- Anderson, Stephen R. and Edward L. Keenan
 1985 Deixis. In: Shopen (ed.), Vol III: 259–307.
- Bernhard, Franz
 1958 *Die Nominalkomposition im Tocharischen*. Diss., Göttingen.
- Blake, Barry, J.
 1994 *Case*. Cambridge: Cambridge University Press.
- Carling, Gerd
 1999 The Tocharian inflected adverbials and adpositions in relation to the case system. *Tocharian and Indo-European Studies* 8: 95–110.
- Carling, Gerd
 2000 *Die Funktionen der lokalen Kasus im Tocharischen*. Berlin: Mouton de Gruyter.
- Carling, Gerd
 2003 Middle-passive and causative: valence-change in the Tocharian B -e-presents without initial palatalization. In: Georges-Jean Pinault and Brigitte Bauer (eds.), *Language in Time and Space. A Festschrift for Werner Winter on the Occasion of his 80th Birthday*. Berlin: Mouton de Gruyter, 63–76.
- Carling, Gerd
 2005 Proto-Tocharian, Common Tocharian, and Tocharian – on the value of linguistic connections in a reconstructed language. [Appendix to: Mair, Victor, Genes, Geography, and Glottochronology: The Tarim Basin during Late Prehistory and History.] In: Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe, and Miriam Robbins Dexter (eds.), *Proceedings of the Sixteenth Annual UCLA Indo-European Conference. Los Angeles November 5–6, 2004*. (Journal of Indo-European Studies Monograph Series, No. 50). Washington DC: Institute for the Study of Man, 47–71.
- Carling, Gerd
 2006 Notes on the syntactic functions of the Tocharian clitic pronouns. In: Gerd Carling (ed.), *GIS^{HUR}gul-za-at-ta-ra. A Festschrift for Folke Josephson on the Occasion of his 70th Birthday*. (Meijerbergs Arkiv för Svensk Ordforskning 32). Gothenberg: Meijerbergs institut för svensk etymologisk forskning, 39–49.
- Carling, Gerd
 2008 Reconsidering the system: verbal categorization and the coding of valence in Tocharian. In: Jens Elmegård Rasmussen and Thomas Olander (ed.), *Internal Reconstruction in Indo-European. Methods, results, and problems. Section papers from the XVIth International Conference on Historical Linguistics, University of Copenhagen, 11–14th August, 2003*. Copenhagen: Museum Tusulanum Press, 49–63.
- Dahl, Östen
 1985 *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Dietz, Rudolf

- 1981 *Der Gebrauch der Partizipia Präsens im Tocharischen. Eine syntaktische Untersuchung.* Diss., Frankfurt-am-Main.

Kemmer, Suzanne

- 1993 *The Middle Voice.* Amsterdam: Benjamins.

Knoll, Gabriele

- 1996 *Die Verwendungsweisen der Adjektive im Tocharischen.* Diss., Frankfurt-am-Main.

Kölver, Bernhard

- 1965 *Der Gebrauch der sekundären Kasus im Tocharischen.* Diss., Frankfurt-am-Main.

Palmer, F.R.

- 2001 *Mood and Modality.* Cambridge: Cambridge University Press.

Payne, John R.

- 1985 Complex phrases and complex sentences. In: Shopen (ed.), Vol. II: 3–41.

Pinault, Georges-Jean

- 1989 Introduction au tokharien. *Lalies* 7: 5–224.

Pinault, Georges-Jean

- 1997 Sur l'assemblage des phrases (« Satzgefüge ») en tokharien. In: Emilio Crespo and José Luis García Ramón (eds.), *Berthold Delbrück y la sintaxis indoeuropea hoy. Actas del Coloquio de la Indogermanische Gesellschaft Madrid, 21–24 de septiembre 1994.* Wiesbaden: Reichert, 449–500.

Pinault, Georges-Jean

- 2003 Contacts linguistiques en Asie Centrale à la lumière des textes tokhariens. In: Sven Bretfeld and Jens Wilkens (eds.), *Indien und Zentralasien. Sprach- und Kulturkontakt. Vorträge des Göttinger Symposions vom 7. bis 10. Mai 2001.* Wiesbaden: Harrassowitz, 45–83.

Pinault, Georges-Jean

- 2005 Impératif et exhortation en tokharien. In: Gerhard Meiser and Olav Hackstein (eds.), *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle-an-der-Saale.* Wiesbaden: Reichert, 495–523.

Pinault, Georges-Jean

- 2008 *Chrestomathie tokharienne. Textes et grammaire.* (Collection linguistique publiée par la Société de Linguistique de Paris 95). Leuven: Peeters.

Saito, Haruyuki

- 2006 *Das Partizipium Präteriti im Tocharischen.* Wiesbaden: Harrassowitz.

Sasse, Hans-Jürgen

- 2002 Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6: 199–271.

Schachter, Paul

- 1985 Parts-of-speech systems. In: Shopen (ed.), Vol. I: 3–61.

Schmidt, Karl Horst

- 1974 Zur formalen Verkettung von Determinans und Determinatum. In: Luigi Heilman (ed.), *Proceedings of the Eleventh International Congress of Linguistics, Bologna-Florence, Aug. 28–Sept. 2, 1972.* Bologna: Società editrice il Mulino, 445–453.

Schmidt, Karl Horst

- 1975 Das indogermanische Kasusmorphem und seine Substituenten. In: Helmut Rix (ed.), *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft Regensburg, 9.–14. September 1973.* Wiesbaden: Reichert, 268–286.

Schmidt, Karl Horst

- 1982 Typusrelevanter Sprachwandel flektierend zu agglutinierend und seine Korrelationen. *Études Finno-Ougriennes* 15: 335–346.

- Schmidt, Karl Horst
 1990 The postulated Pre-Indo-European substrates in Insular Celtic and Tocharian. In: Thomas L. Markey and John A. C. Greppin (eds.), *When Worlds Collide*. Ann Arbor: Karoma, 179–202.
- Schmidt, Klaus Totila
 1974 *Die Gebrauchsweisen des Mediums im Tocharischen*. Diss., Göttingen.
- Shopen, Timothy (ed.)
 2007 [1985] *Language typology and syntactic description*. Vol. I–III. 2nd edn. Cambridge: Cambridge University Press.
- Stumpf, Peter
 1971 *Der Gebrauch der Demonstrativ-Pronomina im Tocharischen*. Wiesbaden: Harrassowitz.
- Thomas, Werner
 1952 *Die tocharischen Verbaladjektive auf -l. Eine syntaktische Untersuchung*. Deutsche Akademie der Wissenschaften zu Berlin. Institut für Orientforschung, Veröffentlichung Nr. 9. Berlin: Akademie-Verlag.
- Thomas, Werner
 1954 Die Infinitive im Tocharischen. In: Johannes Schubert and Ulrich Schneider (eds.), *Asiatica. Festschrift Friedrich Weller*. Leipzig: Harrassowitz, 703–764.
- Thomas, Werner
 1957 *Der Gebrauch der Vergangenheitstempora im Tocharischen*. Wiesbaden: Harrassowitz.
- Thomas, Werner
 1981 Indogermanisches in der Syntax des Tocharischen: zum Ausdruck eines Gebotes und Verbotes. In: Johann Wolfgang Goethe-Universität Frankfurt. Wissenschaftliche Gesellschaft (ed.), *Festschrift der Wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main*. Wiesbaden: Steiner, 481–498.
- Thomas, Werner
 1985 *Die Erforschung des Tocharischen (1960–1984)*. Stuttgart: Steiner.
- Thomas, Werner
 1994 Zur Frage der nichtindogermanischen Komponente im Tocharischen. In: Roland Bielmeier and Reinhard Stempel (eds.), *Indogermanica et Caucasia. Festschrift für Karl Horst Schmidt zum 65. Geburtstag*. Berlin, 222–233.
- Thompson, Sandra, Robert Longacre, and Shin Ja J. Hwang
 1985 Adverbial clauses. In: Shopen (ed.), Vol. II: 237–269.
- Winter, Werner
 1955 Some aspects of “Tocharian” drama: Form and techniques. *Journal of the American Oriental Society* 75: 26–35.
- Winter, Werner
 1980 Morphological signalling of selection properties: transitivity in Tocharian B and A verbs. In: Jacek Fisiak (ed.), *Historical morphology*. The Hague: Mouton de Gruyter, 421–442.
- Zimmer, Stefan
 1976 *Die Satzstellung des finiten Verbs im Tocharischen*. The Hague: Mouton.
- Zimmer, Stefan
 1985 Die Gebrauchsweisen des adnominalen Genetivs im Tocharischen. In: Bernfried Schlerath (ed.), *Grammatische Kategorien. Funktion und Geschichte. Akten der VII. Fachtagung der Indogermanischen Gesellschaft Berlin, 20.–25. Februar 1983*. Wiesbaden: Reichert, 560–574.

78. The lexicon of Tocharian

1. Introduction
2. Inherited words
3. Borrowed words
4. Special words
5. Word formation
6. References

1. Introduction

Because of the fullness of its attestation, Tocharian is one of the “major” branches of Indo-European. However, the texts in Tocharian are meager when compared to those of the other ten “major” branches of the family (excluding the similarly meager Anatolian branch as well). A modern lexicographer, working with a language used for all purposes in a medium-sized state and with access to both a wide range of written documents and native speakers, could expect to produce a dictionary of some 50,000 to 60,000 words, though a significant portion of those are going to be modern scientific and administrative words. As another kind of comparison, we may note that for Old English, partially contemporary with Tocharian, the array of recoverable vocabulary has allowed the creation of dictionaries of approximately 35,000 words (Hall 1960). A different comparandum is the number of words in the mono-generic Homeric vocabulary; Cunliffe’s dictionary (1924) has some 8,200. However, for Tocharian B, far the better attested of the two Tocharian languages, we have somewhat more than 5,400 words attested once proper names have been subtracted (Adams 2013). Thus the attested Tocharian B lexicon is some 15% of that of Old English and 65% of Homer’s. Another constraint on our knowledge of the Tocharian lexicon is the nature of the surviving texts. The bulk of these texts are Buddhist in content and Indian in orientation. Thus they have a high proportion of Indic loanwords and do not represent the overall situation of the Tocharian lexicon as well as they might.

We should also note, with regard to Tocharian B, that we have texts attested over half a millennium. Naturally there was change over this period of time, and some words and derivational processes common in the earlier period disappeared or became less productive over time while, conversely, other words were added and other derivational processes developed and/or became more common. This discussion focuses on Classical Tocharian B, roughly the 6th and 7th centuries CE. The occasional use of “Classical Tocharian B”, rather than just “Tocharian B”, in the text is meant to suggest that the phenomena thus characterized may well be restricted to the Classical period. Exceptionally, the section on names (5.1.7.) draws its data primarily from later sources, as texts from this later era provide the overwhelming bulk of our knowledge of Tocharian B names. Tocharian A texts are attested from a much more restricted period of time (the 7th century CE) and may, therefore, be treated as representing a single, temporally undifferentiated language.

The Tocharian languages have a rich set of derivational processes by which one part of speech may be derived from another. The bulk of these processes involve suffixation. Tocharian makes very little use of prefixes, having only three (exemplifying with Tocharian B forms): *e(n)-* ~ *(a)n-* ‘not’, the homophonous *e(n)-* ~ *a(n)-* ‘in’ (or often only an intensifier), and *y(n)-* ‘in’. On the border between compound members and prefixes are

śle- ‘with’ and *snai-* ‘without’. There are widespread examples of Indo-European ablaut in verbs, as well as in nominal derivation. A verb showing typical interplay of ablaut grades (and with concomitant palatalization of consonants before original front vowels) is seen in *kārs-* ‘know’: present *kārsānā-* (PIE zero-grade), modal *kārsā-* (PIE *o*-grade, umlauted to *-ā-* by the following *-ā-*), preterite *śārsā-* (PIE *e*-grade), causative present and modal *śārsāsk-*, preterite *śārsā-* (PIE *ē*-grade and palatalization). Nominal derivatives with *o*-grade are common (see 5.1.2); *ē*-grade can be seen in B *lyauto* ‘hole, perforation’ (cf. A *lot* ‘hole’ with *o*-grade) and Hittite *luttāi-* ‘window’ with zero-grade), *śerk(w)* ‘cord’ (from *kārk-* ‘bind’); *ō*-grade in B *wašamo* ‘friend’ (from *wās-* ‘dwell’), B *pār* ‘plumage’ (cf. *parwa* ‘feathers’), A *wārtsāk* ‘having neighbors’ (A *wartsī*, B *wertsīya* ‘retinue’). Reduplication is productive in the formation of past participles (e.g., B *tetriku*, A *tatriku* ‘confused, deceived’), and in Tocharian A in the formation of the causative preterite (e.g., *kakāl* ‘brought’), but exceedingly rare elsewhere: B *pyāpyo* (A *pyāpi*) ‘flower’ and B *ckācko* ‘thigh’ have been held to be examples of reduplication; B *sāsuwa* ‘sons’ (the suppletive plural of *soy* ‘son’) probably represents **su-suh₃-eh₂*, a reduplicated participle ‘*those engendered’. Productive verbal formation has as its vowel B *-e-*, A *-a-* (Proto-Tocharian **-e-*) which may irregularly reflect PIE **-e-* preserved in this particular morphological configuration, just as it is, for example, in Gothic.

2. Inherited words

A very rough count of attested Tocharian B words suggests that some 35–40 % of dictionary entries are borrowings from Buddhist Hybrid Sanskrit. This raw statistic is misleading, however, since most of the attested borrowings are highly technical terms (e.g., names of medicinal plants, Buddhist technical terminology). Almost all the “ordinary” vocabulary is inherited. This point can be illustrated by any number of semantic fields, provided we have sufficient attestation of lexical items. Consider, for example, agriculture, particularly grains and domesticated herd animals. In both areas we have substantial lexical attestations, though even here there are some obvious gaps in our knowledge (etymological discussions are from Adams [2013]).

Grains:

- B *āka* [pl.] ‘barley’ or ‘millet’(?) (< **h₂eĕ-* ‘sharp, pointy’, cf. Gothic *ahs* [gen. *ahsis*] ‘ear of grain’, Old English *ēar* [< **ahuz-*] ‘id.’; formally B *āka* looks like it might be a PIE **h₂eĕs* [nt.pl.] corresponding to the **h₂eĕs* [nt.sg.] of Gothic).
- AB *klu* ‘rice’ (an early borrowing from Old Chinese **gləw?* ‘rice, rice-paddy’ [in New Chinese *dào*]) (Schuessler 1987: 116).
- B *klese* ‘barley(-meal)’ (< **kolsos* ‘[ear of] grain’, cf. 16th century Albanian *kall* ‘ear of grain’, OCS *klasъ* ‘ear of grain’, Persian *kaška* ‘barley’ [< Proto-Iranian **karšaka-*]) (Blažek, 1999 79–82).
- B *traksiñ* [pl.] ‘grains of millet’ (cf. Sanskrit *drākṣā* ~ *dhrākṣā* ‘grape’, Late Khotanese *drrāmśā-* ‘millet’, Old Irish *derc* ‘berry’ [an *s*-stem]).
- B *yap* ‘barley’ (< PIE **yeb^(h)om*, by oral obstruency assimilation from **yewom* ‘grain, particularly barley’, cf. Sanskrit *yáva-* ‘bread; grain, particularly barley’, Greek. *zeiai* [pl.] ‘an inferior sort of wheat [einkorn or emmer wheat]’, Hittite *ewan* ‘≈ barley’).

- B *ysāre* ‘wheat’ (A *wsār* [‘heap of] grain’) (< **wesōro-* as a derivative of **wesʔ* ‘spring’, a reference to a seasonal variety of wheat) (Huld 1990: 420, fn. 15; Ivanov 2003: 194).
- B *śātre* ‘grain’ (< **g^wyeh₃wo-tro-* ‘Lebensmittel’).
- B *tāno* ‘grain’ (< **d^hoh_xneh₂-*, cf. Sanskrit *dhānāḥ* [f. pl.] ‘grain’, Khotanese *dāna-* [f.] ‘grain’, Lithuanian *dūona* ‘bread’).
- B *proksa* [pl.] ‘millet’ (< **prókseh₂* [pl.], cf. Russian *próso* ‘millet (*Panicum milia-ceum*)’, Old Prussian *prassan* ‘millet’ from **próksom* [sg.]) (Ivanov 2003: 197).

Domestic animals:

Sheep and goats:

- B *ariwe* ‘ram’ (< **h₁ōreiwo-*, cf. Sanskrit *āreya-* ‘ram’, Latin *aries* [gen. *arietis*] ‘ram’, Umbrian *eriet-* ‘ram’).
- B *alā_u* ‘wether (sheep or goat)’ (etymology uncertain).
- B *ā_u* ‘ewe’ (< **h₂ówis* [gen. **h₂áwis*] ‘sheep’, cf. Latin *ovis* ‘sheep’, Old English *ēow* ‘sheep’, Lithuanian *avis* ‘sheep’, Greek *ó(w)is* ‘sheep’, Luvian *hāwali-* ‘sheep’, Lycian *xawa-* ‘sheep’, Sanskrit *ávi-* ‘sheep’; Tocharian shows the generalization of **h₂áw-* of the weak forms).
- B *āl* ‘ram, he-goat’ (< **h₂eli-* ‘male [of animals]’, cf. Hittite *aliyan[a]-* ‘roebuck’ (with apparent early dissimilation of *h-*) and Macedonian *aliē* ‘boar’ [< **h₂elih₁ēn*]).
- AB *ās* ‘(she-)goat’ (usually assumed to be a borrowing from some Middle Iranian source; one might compare Middle Persian *azak* ‘goat’ [also Sanskrit *ája-* (m.) ‘buck’, *ajā-* (f.) ‘goat’]; however, it is hard to see why the putative Iranian source would not have given a TchB **ese[k]* or the like).
- B *eye* ‘sheep’ (< **h₁oyós*, an agentive derivative of **h₁ei-* ‘go’; the semantic development is like that of Hittite *iyant-* ‘sheep’ which is, in origin, the present participle of Hittite *i-* [PIE **h₁ei-*] ‘go’, or Oscan *eitiuvam* ‘pecuniam’, another derivative of PIE **h₁ei-* ‘go’. [From a different verb but with a similar metaphor (i.e. ‘moveable chattels’) we have Greek *próbata* ‘cattle; sheep’.]).
- B *yrīye* ‘lamb’ (< either **werh₁en-*, cf. Sanskrit *úraṇa-* ‘sheep, ram’, Avestan *varən-* ‘lamb’, Greek *warén* (Cret.) ‘lamb’, Armenian *garñ* ‘lamb’, or from PIE **h₁er[i]-* ‘lamb, kid’, cf. Greek *éripfos* ‘kid’, Armenian *oroj* [< **eroj*] ‘lamb’, Old Prussian *eristian* ‘lamb’, Lithuanian [*j*]*éras* ‘lamb’, Latvian *jērs* ‘lamb’).
- B *śantālya* ‘breeding ewe/she-goat’ (a nominalized gerundive from /śāntā-/ , a denominative verb built to the same participial formation from PIE **g^wyeh₃-* ‘live’ which, when nominalized, gives the plural *śānta* ‘sheep/goats’).
- B *śari* ‘kid’ (< **steruh₁en-*, cf. Albanian *shtjerrë* ‘lamb’, *shtjerri* ‘flock of lambs, kids, and/or heifers’ and, a little more distantly, Sanskrit *starīḥ* ‘cow that neither gives milk nor is pregnant, heifer; barren’, Greek *steira* ‘barren [of animals or women]’, etc.).
- B *śaiyye* ‘sheep/goat’ [pl. *śānta*] (< **g^wyeh₃w-yo-*, and exactly equivalent to Greek *zô(i)on* ‘animal’; the same form gives TchA *śāyu* [unknown animal species] regularly, though with different semantic specialization. TchB *śānta* reflects a neuter plural participle from **g^wyeh₃w-*).
- A *śos* [pl.] ‘sheep and goats’ (< the same **g^wyeh₃w-* of the preceding with an uncertain nominal suffix).

Cattle:

- *okso* (A *opsi* [pl.]) ‘ox’ (< **uk^wse/on-*, cf. Sanskrit *ukṣán-* ‘ox, bull’, Avestan *uxšan-* ‘id.’, Welsh *ych* ‘ox’, Gothic **auhsa* ‘ox’, Old English *oxa* ‘id.’, OHG *ohso* ‘id.’).
- B *ke_u* (A *ko*) ‘cow’ (< **g^wou-* ‘cow’, cf. Sanskrit *gáuh* [m./f.] ‘cow’, Avestan *gauš* [m./f.], Armenian *kov*, Greek *boûs* [m./f.] [Doric *bôs*], Latin *bōs* [m./f.], Old Irish *bó* [f.], OHG *chuo* [f.], Old English *cū* [f.], Latvian *gūovs*, Hieroglyphic Luvian *wawa-*, Lycian *wawa-* ~ *uwa-* ‘cow’).
- *kau_urše* (A *kayurs*) ‘bull’ (< **g^wou-wršen-* ‘cow-male’ as in Sanskrit *go-vr̥ṣa-* ‘bull’, Old Norse *kursi* ‘bull-calf’).
- B *paitar* ‘calf’ (etymology uncertain).
- A *mahirš* ‘buffalo’ (< Sanskrit *mahiṣá-*).

Equids:

- B *kercao* ‘ass’ (the exact equivalent of Sanskrit *gardabhá-* ‘donkey, ass’ [< **gordeb^ho-*]; if, as has so often been suggested, *kercao* is a borrowing from Indic *gardabhá-*, the borrowing must be very early, before the merger of the non-high vowels in Indo-Iranian; more likely a common inheritance from PIE).
- B *khare* ‘ass’ (< Sanskrit *khara-*).
- B *yakwe* (A *yuk*) ‘horse’ (< **h₁ekwo-*, cf. Sanskrit *ásva-*, Avestan *aspa-*, Latin *equus*, Greek *hippos*, Old English *eoh*, Old Irish *ech*, Hieroglyphic Luvian *azu(wa)-*, Lycian *esbe*).
- B *haye* ‘horse’ (< Sanskrit *háya-*).
- B *etswe* ‘mule’ (< Proto-Iranian **aćwa-* ‘horse’ [Avestan *aspa-*]).

Miscellaneous:

- AB *ku* ‘dog’ (< **kúwō* [nom. sg.] ‘dog’, cf. Sanskrit *ś(u)vá*, Greek *kúōn*, Latin *canis*, Old Irish *cú*, Gothic *hunds*, Lithuanian *šuo*, Hittite *kuwas* [acc. *kuwanan*] [Melchert 1989], Hieroglyphic Luvian *zuwana/i-*).
- B *partáкто** ‘camel’ (?) (cf. adj. *partákaññe* ‘pertaining to a camel’ [?]) (both meaning and etymology uncertain).
- B *wästarye* ‘pertaining to camels’ (implying **wästär*, **wästare*, or **wastre* ‘camel’ [Chen Ruixuan, p.c., 2013]), cf. Avestan *uštra-*, Sanskrit *uṣtra-* ‘camel’ (Tocharian [and Indic?] presumably borrowed the word from Iranian, but the Tocharian phonology is not exactly what one would expect [why *-s-* and not *-š-*?], and a borrowing in the other direction would be phonologically easier).
- B *mārijāre* ‘cat’ (< Sanskrit *mārijāra-*).
- B *suwo* ‘hog, pig’ [known only from its use in the twelve-year calendrical cycle] (< **sū-*, cf. Avestan [gen. sg.] *hū*, Latvian *suvēns* ‘young pig, piglet’, Greek *hūs* [m.] ‘boar’, [f.] ‘sow’, Albanian *thi* ‘pig’, Latin *sūs* ‘id.’, Old English *sū* ‘sow’).

Of the words referring to grains only *klu* ‘rice’ is a borrowing; of the words referring to domestic animals, only *mahirš* ‘buffalo’, *khare* ‘ass’, *haye* ‘horse’, *etswe* ‘mule’, and *mārijāre* ‘cat’ are certain borrowings (and these words, except *etswe*, appear only with reference to India), and *ās* ‘she-goat’, *kercao* ‘ass’, and *partáкто* ‘camel’ are possible borrowings. Other “homely” semantic spheres, e.g., kin, body parts, are even more monochromatically indigenous in origin. Thus, to give just one more example, the words

for ‘father’, ‘mother’, ‘brother’, ‘sister’, ‘son’, and ‘daughter’ ([A/B] *pācer/pācar*, *mācer/mācar*, *procer/pracar*, *šer/šar*, *soy/se*, *tkācer/ckācar*) are all patently Proto-Indo-European in origin (cf. English *father*, *mother*, *brother*, *sister*, *son* [Greek *huyús*, *daughter*]).

3. Borrowed words

Tocharian shows a host of borrowings from Indo-Iranian and several significant ones from Chinese but no sure examples from Tibetan and only one likely example from Altaic, namely B *kenek* (A *kanak*) ‘cotton cloth’, probably from Proto-Turkic *köjnek* ‘shirt’ (Lubotsky and Starostin 2003: 260).

3.1. Borrowings from Indo-Iranian

The oldest Indo-Iranian lexical connection is represented by TchB *kercao*, Sanskrit *gardabhā-* ‘ass’, though whether the correspondence represents a late PIE dialect word or the Tocharian word is borrowed from pre-Proto-Indo-Iranian, or vice versa, or both from a third source is unknown. The common ancestor would have been **gordhebho(h₃on)-*. There are a small number of words which are clearly borrowed from Iranian which show very early, even pre-Proto-Iranian phonology. Thus we have B *tsain* ‘arrow’, pl. *tsainwa* (< **dzainu-*, cf. Avestan *zaēnu-* ‘arrow’), B *tsaiññe* ‘ornament’ (a derivative of pre-Proto-Iranian **dzay-* ‘equip’), B *etswe* (< Proto-Iranian **ačwa-* ‘horse’) and B *waipecce* ‘possession’ (< Proto-Iranian **hwaipapya-* ‘one’s own’ [cf. Avestan *h³aēpaiṣya-*] or perhaps even earlier pre-Proto-Iranian **hwaipatya-*), *iscem* ‘brick, tile’ (< Proto-Iranian **ištyām*). B *sāte* ‘rich’ (Avestan *šyāta-* ‘joyful’) shows at least an Old Iranian phonology (*šy-*) since the Middle Iranian *š-* would have become Tocharian *š-*. Borrowed words for irrigation, *ārte* ‘raised irrigation ditch’ (cf. Khufi *wurđ* < earlier Iranian **ārda-*), *newiya* ‘canal’ (cf. Sarikoli *wanew* ‘irrigation ditch’ < earlier Iranian **wi-nāwīya-*), *murye* ‘irrigation ditch’ (cf. Sogdian *mwry’y* < earlier Iranian **mūrya-*) and commerce (broadly defined), *nip-* ‘≈ pledge’ (cf. Manichaean Sogdian *np’k* ‘pledge’), A *pāšim* ‘treasury, treasure-house’ (cf. Khotanese *pārgyiñi-* ~ *pājiñi-*), B *pīto* ‘price’ (Khotanese *pīha-*), B *peri*, A *pare* ‘debt’ (Khotanese *pīra-* < **parya-*) suggest that Tocharian speakers modeled themselves on Iranian speakers in certain culturally important ways. It may be significant that most of the commercial terms would seem to come from neighboring Khotanese while the irrigation terminology is not matched in Khotanese but only in languages west of the Tarim Basin. Other borrowings from Iranian include *akālk* (A *ākāl*) ‘wish’ (Sogdian *āyal[ak]-*), *ašām* (A *āšām*) ‘worthy’ (Khotanese *āšana-*), *ekšinkāññe* ‘pigeon-’ (Ossetic *äxsinäg*, Khotanese *aššānaka-*), B *epyac* (A *opy-āc*) ‘thing remembered’ (Middle Iranian **abyāt*), AB *krāke* ‘filth’ (Khotanese *khārgga-* ‘manure’), A *twantam* ‘reverently’ (Khotanese *tvamdanu*), AB *perāk* ‘believing’ (cf. Buddhist Sogdian *pyr’k*), B *peret* (A *porat*) ‘ax’, B *perne* (A *parām*) ‘rank, station’, B *maññe* (A *māññe*) ‘servant, slave’ (cf. Old Persian *māñiya-* ‘domestic servant’), B *mišši* (A *miši*) ‘community’ (cf. Buddhist Sogdian *mydry*), AB *menāk* ‘comparison’ (Sogdian *myñ’k*), B *ršake* (A *rišak*) ‘seer’ (Khotanese *rišaya-* < **rišaka-*, itself ultimately

from Sanskrit), B *wrākai* (A *wrok*) ‘pearl’ (Khotanese *mrāha-*), *špakiye* ‘pill’ (Khotanese *svak[y]e-*), B *sanāp-* ‘rub in/on’ (pre-Khotanese **zənāf-* ‘wash’ < Proto-Iranian **snāp-*), B *sām* ‘enemy’ (Khotanese *sāna-*, Sogdian *s’n*), AB *senik* ‘under the care of’ (cf. Khotanese *ysīnīta-* < **zīnīka-*, Sogdian *zynyh*, Kroraina *jhenīga-*), B *tserēññ-* ‘trick, deceive, lead astray’ (cf. Khotanese *jsīr-* ‘deceive’). The homophony of B *yok* ‘animal hair’ and *yok* ‘color’ probably reflects the same homophony of Iranian, e.g., Avestan *gaona-*, with both meanings.

The cultural influence was not all from Iranian to Tocharian. TchB *eñcuwo* ‘iron’ (cf. A *añcwaši* ‘pertaining to iron’) appears to reflect earlier **en-šu-* ‘≈ something poured into’ (like German *Guß-eisern* ‘cast iron’) and have been borrowed into Iranian languages as 1. **ančuwān-* (Khwarezmian *hnču* [with some irregularities]), 2. **áčwana-* (Khotanese *hiššana-*, Ossetic *æfsæn*), 3. **áčwanya-* (Sogdian *spen*, Waxi *[y]išn*, Shughni *sipin*), 4. **áčuna-* (Turfan Middle Persian *’hwn*, New Persian *āhan*, Turfan Middle Parthian *’swn*, Zoroastrian Pahlavi *’syn*, Balochi *āsin*). Similarly B *yolo* ‘evil’ (< PIE **h₁edwolu-*, cf. Hittite *idālu-* ‘evil’, CLuvian *adduwal-*) has been borrowed into Khotanese as *yola-* ‘falsehood’. Early TchB *tmāne* (A *tmām*) ‘10,000’ yields later TchB *t₄māne* which, in turn, was borrowed into Middle Iranian, cf. Modern Persian *tumān*. Borrowings from Iranian (and Indic) have obviously strengthened the status of inherited *-ke/k* as both a noun- and adjective-forming suffix. Whether the Tocharian B diminutive suffix *-ške*/*-śke* is a borrowing from an Iranian source is not clear. At least the *-śke* variant would be phonologically unlikely to be a borrowing.

There are a host of loanwords from Buddhist Hybrid Sanskrit that deal with religion and religious doctrine, philosophy, and medicine which are little changed from their Sanskrit originals and which belong probably only to specialists’ vocabularies (e.g., A *ārši* ‘Indic/Aryan’ [a Prakrit descendant of Sanskrit *aryá-*], AB *cākkār* ‘wheel’ [as mystical symbol, < *cakrá-*], B *lekhāke* ‘scribe’ [< *lekhaka-*], B *hom* ‘oblation’ [< *hóma-*], B *divvrāni* ‘ulcer or wound’). However, there are other, more popular, words, many of which are borrowed from some Prakrit, e.g., *kwrakār* (A *k₄rekār*) ‘upper room, apartment at the top of the house’ (< Sanskrit *kūṭāgāra-*), B *šamāne* (A *šamām*) ‘monk’ (< Prakrit equivalent of Sanskrit *śramaṇá-*), B *sakw* (A *suk*) ‘good fortune’ (< Sanskrit *sukhá-*), B *sāñ* (A *šāñ*) ‘plan, skill’ (< Northwestern Prakrit *sañā* or via Khotanese intermediary), B *sāñk* (A *sāñk*) ‘community; monastery’ (< *saṃgha-*), B *sāmte* (A *sāmte*) ‘medicine’ (< Prakrit equivalent of BHS *sāntaka-*), A *sāñce* ‘doubt’ (< Sanskrit *saśaya-*). An interesting group are the TchB euphemisms *asāñcne* ‘buttocks’ (lit, ‘the two seats’ < Sanskrit *āsana-* ‘seat’), *indri* ‘penis’ (< Sanskrit *indriyá-* ‘power, virile power, organ of generation’), *strīndri-* ‘female genitals’ (< *strīndriya-* ‘id.’), *kukši* ‘womb’ (< Sanskrit *kukši-* ‘id.’).

3.2. Borrowings from Chinese

The most obvious borrowings from Chinese are those of the Tang period, and reflect the Chinese commercial orientation of the Tarim Basin. Thus we have words for measurement: B *šañk* ‘liter’ (= modern Mandarin *shēng*), *tau* ‘10 litres’ (= *du*), *cāk* ‘100 liters’ (both dry and wet measure of volume; = *shì*), and *tsum* ‘inch’ (= *cùn*), *cak* ‘foot’ (= ten *tsum*; = *chǐ*); Tocharian A adds *šipākiñc* ‘abacus’ (= *shū-pán* +?). Other words from

probably the same period are B *cok* ‘lamp’ (*zhú*), *šitsok* ‘millet-alcohol’ (= *shǔ-jiǔ*), and AB *tsem* ‘blue’ (= *qīng*). Two other words are of particular interest because their phonological shape demands that they have been borrowed during the Han-period, i.e., AB *klu* ‘rice’ (= *dào* ‘rice plant’) and *rāp* ‘twelfth month’ (= *là* ‘winter sacrifice’). Also very interesting are Chinese *mì* ‘honey’, an Old Chinese borrowing from the ancestor of B *mit* ‘honey’ and possibly similarly for *lǐ* ‘village’ from the ancestor of B *riye* (A *ri*) ‘city’. As in the case of Iranian, linguistic borrowing between Tocharian and Chinese is both very old and two-sided (see also Lubotsky and Starostin 2003).

3.3. Integration of borrowed words

Early borrowings of verbs from Iranian result in full Tocharian verbs: B *tsereññ-* ‘deceive’ (Khotanese *jsīr-*), B *nipā-* ‘pledge’ (no verb attested in Iranian). There are few examples of borrowed adjectives in Tocharian, no doubt because of the productivity of denominal adjectives within Tocharian – see 5.3.3). Earlier borrowings were fully integrated into Tocharian morphology (e.g., TchB *šāte* ‘rich’ from Iranian [m.n.pl. *šateñ*] and AB *tsem* ‘blue’ from Chinese [B f.n.sg. *tseñña*, f.pl. *tseññāna*, A f.n.pl. *tseññān*]). Later, technical borrowings from Buddhist Hybrid Sanskrit may either be fully integrated (e.g., TchB *kāyike* ‘corporeal’ from BHS *kāyika-*) or used indeclinably (e.g., *aupacayik* ‘based on accumulation’ from BHS *aupacayika-*). Borrowed nouns denoting masculine beings are normally provided with an *-e* in Tocharian B and, in both languages, declined like old PIE *o*-stems (nom. Pl. *-i* in B, *-añ* in A). Similarly, nouns denoting feminine beings are normally provided with an *-a* in Tocharian B and, in both languages, declined like old *ā*-stems. Inanimate nouns are taken in as Tocharian neuters, usually without any special ending in Tocharian B. Typically they show up as plurals with the addition of B *-nma* or *-nta*, or A *-ntu*, e.g., B *sām̄tke* ‘medicine’, pl. *sañ̄tkenta* (A *sām̄tāk*, pl. *sām̄tākantu*), AB *dhātu* ‘element’, pl. (B *dhatunma*, A *dhātuntu*).

4. Special words

As with every language, the Tocharian languages have many words which are neither clearly inherited nor clearly borrowed. A classic example from Germanic would be the family of English *drink*, whose morphology is purely Indo-European but which has no cognates outside of Germanic. Of the words discussed above (2), *alā_u*, *ās*, *paitar*, and *partāktō* have neither a good inherited pedigree nor any obvious extra-Tocharian source. There is no trouble in identifying others of the same type: B *ausw-* ‘cry out (for help)’, B *karep* (A *kāryap*) ‘damage, harm’, B *tuñe* ‘flower, blossom’, etc. Special in a different way is A *akmal* ‘face’. Literally it is ‘eye-nose’, a very un-Indo-European formation. B *sār_wāna* [pl.] ‘face’ is also obscure in origin but morphologically more expected. Finally we might note the typological oddity, otherwise unknown in Indo-European and very rare elsewhere, of the lack of separate words for ‘animal’ and ‘bird’. They are both (B) *luwo* (A *lu*). In B birds are *lwāsa šlyamñāna* ‘flying animals’ and animals, *sensu stricto*, are *lwāsa ynamñāna* ‘walking animals’.

5. Word formation

5.1. Nouns

5.1.1. Introduction

Both Tocharian languages have three noun gender classes: masculine, feminine, and neuter (or “alternating”). The last group takes masculine concord in the singular and feminine concord in the plural, e.g., B *luwo* (neuter singular) ‘animal’ in *salamo luwo* ‘a flying animal’ (= ‘bird’), where *salamo* is masculine singular, but *lwāsa* (neuter plural) in *kowān lwāsa šlyamñana ynamñana* ‘he will kill flying animals and walking animals’ (= ‘birds and animals’), where both *šlyamñana* and *ynamñana* are feminine plural. Nouns also have three numbers: singular, dual, and plural, e.g., B *ñakte* ‘god’, *ñaktene* ‘the two gods’, *ñakti* ‘gods’.

5.1.2. Stem-classes

Most Tocharian nouns are morphologically descendants of PIE *o*-stems, *ā*-stems, *n*-stems, and (neuter) *nt*-stems. Thus we have (B) *yakwe* ‘horse’, pl. *yakwi*, from PIE **ekwos/*ekwoi*; *sana* ‘woman’, acc. sg. *sano*, from PIE **g^weneh₂/*g^weneh₂m*; *kantwo* ‘tongue’, pl. *kantwañ*, from PIE **d^ṅg^wheh₂ōn/*d^ṅg^wheh₂(o)nes*; and *cake* ‘river’, pl. *ckenta*, from PIE **tekontl/*tekonteh₂*. There are remnants of animate *i*-stems (*laks* ‘fish’, pl. *läkši*, from PIE **loksis/*lokseyes* with analogical root vowel), neuter *u*-stems (*yok* ‘hair’, pl. *yākwa*, from PIE **yēku/*yēkweh_a*), animate *r*-stems (*tkācer* ‘daughter’, gen. sg. *tkātre*, from PIE **dhuǵ₂tēr/*dhuǵ₂tros*), or obstruent stems (*wek* ‘voice’ from PIE **wok^ws*). All of the Tocharian A declensional types show considerable mutual influence, e.g., unlike B *yakwe/yakwi*, A has *yuk/yukañ*. Tocharian B largely preserves the Proto-Tocharian situation, while the loss of final vowels in Tocharian A has led to much restructuring.

5.1.3. Action nouns

The Tocharian languages have several formations by which they can create action nouns or verbal abstracts. Rarely is there a root noun, e.g., B *wek* (A *wak*) ‘voice, noise’ from PIE **wok^ws* or a minimally derived, “quasi-root noun”, B *lyuke* ‘light’ representing PIE **leukos*. Old in Indo-European are more obvious derivatives in **-r* and **-l* (abundant still in Hittite). Tocharian contains numerous examples of both, e.g., B *milar* ‘≈ damage’ [the further derived A *milārts* ‘damaged’] (*mil-* ‘≈ wound, damage’), AB *l(y)ipär* ‘remainder’ (*lip-* ‘remain, be left over’), B *tsänkär* [A *tsänkär*] ‘summit’ (*tsänk-* ‘raise up’), and the productive TchB abstracts and concrete nouns built from past participles, e.g., *karyor* ‘buying, commerce’ (*kāryau* ‘having bought’), *kekamor* ‘(the) past’ (*kekamu* ‘having come’), *srukor* ‘dying, death’ (*srukau* ‘having died’), *yāmor* ‘deed’ (*yāmu* ‘having done’), *āyor* ‘gift’ (*āyu* ‘having given’); probably also *malkwer* ‘milk’ if it reflects a past participle whose oblique stem was *-weš-*. In Tocharian A this latter type is very

rare as a pure noun (only *kärsor* ‘knowledge’, *kuryar* ‘buying, commerce’ [no underlying verb attested in Tocharian A], and *tärkor* ‘permission’), but in both languages they occur freely in their respective ablatives (B *-mem*, A *-äš*) forming absolutes, e.g., TchB *kekteſē* *kärsormem* ‘having known the body’, A *nākās āsuk kätkorāš* ‘the nāgas having passed’.

Tocharian also retains numerous traces of PIE verbal nouns in **-l* (cf. Anatolian verbal nouns in *-al* and Armenian infinitives in *-l*): TchA *el* ‘gift’ (*e-* ‘give’), AB *eñkäl* ‘suffering’ (*eñk-* ‘take’, TchA word borrowed from B), B *camel* [A *cmol*] ‘birth’ (*täm-* ‘be born’), B *treñkäl* ‘clinging; (worldly) attachment’ (*treñk-* ‘adhere, cling’), A *rkäl* ‘blanket’ (*räk-* ‘cover’), B *watkal* ‘decision’ [further derived in A *wätkäłts* ‘distinctive’] (*wätk-* ‘distinguish’), B *warkšäl* [A *wärkšäl*] ‘power, strength’ [no underlying verb], A *wsäl* ‘garment’ (*wäs-* ‘be clothed’), B *šaul* [A *šol*] ‘life’ (*šāw-~šāy-* ‘live’), B *šaṃšäl* ‘number’ (*ās-* ‘count’), A *šwāl* ‘meat’ (*šwā-* ‘eat’; cf. B *šwāl* ‘corpse’). Surely belonging here is B *pikol* [A *p_ukäl*] ‘year’, but the underlying verb is uncertain. Insofar as these originally abstract nouns had plurals, they were probably in **-ā* (thus *cmela*, *pikwala*, *puklā*), but as concrete nouns they show up with a variety of plural suffixes (e.g., *elā*, *cmolu*, *treñkalwa*, *räklunt*, *wsälu*, *šaulanma*, *šöläntu*). Historically at least these verbal nouns in *-l* underlie the fully productive gerunds in B *-lle* [A *-l*] (< PIE **-l-yo-*) and they, in turn, form the basis for the productive abstracts in B *-lñe* [A *-lune*], e.g., B *camel* > *cmelle* > *cmelñe*. Likewise they lie behind the small group of nouns in (B) *-elme*, i.e., *onolme* ‘creature’ (cf. *anāsk-* ‘breathe’), *yšelme* [A *yšäläm**] ‘[sexual] pleasure’ [underlying verb not attested], *wpelme* ‘spider’s web’ (*wäp-* ‘weave’), *syelme* ‘sweat’ (*si-* ‘to sweat’), *tsrerme* (with assimilation of *-l-* to *-r-*) ‘fortification ditch’ (*tsär-* ‘separate’). Another small set of action nouns come from PIE neuters in **-men* [sic; and not the **-mñ* predicted by Greek and Sanskrit], e.g., A *nākām*, B *nāki* ‘blame’, A *wākām*, B *wāki* ‘distinction’. The unusual B ending is actually regular (**-Cmen* > **-Cämⁱän* > **-Cäwⁱä(n)* > **-Cäyā* > *-Ci*).

A very common derived noun throughout the history of Tocharian is the *tómos*-type. Tocharian B has at least fifty exemplars and Tocharian A at least thirty-five: A *kark* ‘a part of a bow’ (*kärk-* ‘bind’), B *kene* (A *kaṃ*) ‘song, melody’ [no underlying verb], B *kerke* ‘≈ fetter’ (*kärk-* ‘bind’), B *keme* (A *kam*) ‘tooth’ [no underlying verb], B *kraupe* (A *krop*) ‘group’ (*kraup-* ‘gather’), B *klepe* ‘≈ theft’ (*kälyp-* ‘steal’), A *klop* ‘suffering’ (B *klup-* ‘squeeze’), B *klautke* (A *lotäk*) ‘way, manner’ (*klutk-* ‘turn’), B *tenke* ‘(sheep)fold; checkpoint’ (*täñk-* ‘check, stop’), B *treñke* ‘clinging; (worldly) attachment’ (*treñk-* ‘adhere, cling’), B *traike* (cf. A *treke*) ‘confusion, delusion’ (*trik-* ‘go astray, be confused’), B *tweye* (A *twe*) ‘dust’ [no underlying verb], B *nete* (cf. A *nati*) ‘≈ power’ [no underlying verb], B *netke* ‘urging, prompting’ (*nätk-* ‘urge’), B *nerke* ‘hesitation’ (*närk-* ‘keep away’), B *newe* (cf. A *nawem*) ‘roar’ (*nu-* ‘roar’), B *petwe* ‘bank (of a river)’ (*pätt-* ‘climb up’), B *pautke* (A *potäk*) ‘≈ share, tribute’ (*putk-* ‘divide’), *preñke* (A *prañk*) ‘island’ [*präñk-* ‘keep away’], B *prautke* (A *protäk*) ‘≈ enclosure’ (*prutk-* ‘shut up’), B *plewe* ‘raft’ [*plu-* ‘fly’], A *mañk* ‘lack’ (cf. B *meñki*) ‘lack’ (*mäñk-* ‘be lacking’), B *meli* [pl.] (A *malañ* [pl.]) ‘nose’ [no underlying verb], B *meske* (A *masäk*) ‘joint, portion’ [no underlying verb attested], B *yenme* ‘portal’ (*yäp-* ‘enter’), B *leke* (cf. A *lake*) ‘bed’ (*lyäk-* ‘lie’), B *laitke* (cf. A *letke*) ‘vine’ [no underlying verb], B *laupe* ‘ointment’ (*lup-* ‘smear’), B *wente* ‘covering’ (*wänt-* ‘cover’), B *wepe* ‘≈ corral, pad-dock’ (*wäp-* ‘weave’), B *were* (A *war*) ‘smell’ (*wär[-sk]-* ‘smell’), B *werpe* (A *warp*) ‘enclosure’ (*wärp-* ‘surround’), B *waike* ‘lie’ (TchA *wek-* ‘lie’), B *waipe* ‘banner’ (*wip-* ‘shake’), B *sekwe* (A *saku*) ‘pus’ [no underlying verb], B *skeye* (A *ske*) ‘zeal, effort’

(*skai-* ‘strive, attempt’), B *sleme* (A *slam*) ‘flame’ [no underlying verb], B *tsaipe* ‘dance’ (*tsip-* ‘dance’). It is not certain that this was a productive formation in attested Tocharian, but it is certain that it remained productive relatively late into Tocharian prehistory, since we find it built to verbs which have incorporated the present formative *-sk-* into the root (e.g., *netke*, *pautke*, *laitke*) or, at least in one case, the present-forming suffix *-n-* (i.e., *yenme*). In many cases the underlying verb is not attested in Tocharian: either the underlying verb has died out in the prehistoric period or it is simply accidentally unattested.

There are a significant number of derived verbal nouns ending in TchB *-o*, often alternating with *-iye* in the nominative singular (and sharing *-ai* in the accusative singular): e.g., B *kāko* ‘invitation’ (*kāk-* ‘invite’), AB *krāso* ‘≈ vexation, torment’ (*krās-* ‘torment’), B *tran̄ko* ‘sin’, B *trik̄šo* ‘≈ error, mistake’ (*trik-* ‘go astray, be confused’), B *palsko* (A *pālt̄sāk*) ‘thought’ (*pālsk-* ‘think’), B *pauto* (A *poto*) ‘honor, flattery’ (*paut-* ‘honor, flatter’), B *prosko* (~ *proskiye*) (A *praski*) ‘fear’ (*pārsk-* ‘fear’), B *yoko* (~ *yokiye*) (A *yoke*) ‘thirst; desire’ (*yok-* ‘drink’), B *yošiye* ‘≈ irrigation’ (*yok-* ‘drink’), B *raso* ‘span’ (*rās-* ‘stretch’), B *laiko* ‘bath’ (*lik-* ‘wash’), B *šart̄to* ‘encouragement’ (*šart̄t-* ‘incite’), B *šiko* (A *šik*) ‘(foot)step’ (*šik-* ‘set foot’), B *tsamo* ‘growth’ (*tsām-* ‘grow’), B *tsāro* ‘cloister’ (*tsār-* ‘separate’). Some of these clearly have, in PIE terms, an *e*-grade (*trik̄šo*, *misko*, *yarpo*, *šiko*), while others have an *o*-grade (*prosko*, *laiko*, *tsāro*), and at least one an *ē*-grade (*šart̄to*); the others merely repeat the vowel grade of the underlying verb. In origin this is a heterogeneous group. The nucleus was presumably provided by verbal nouns in (PIE) **-eh₂-* (extended either as **-eh₂-n-* or **-eh₂-h₁en-*) but there are also certainly **-u-h₁en-* (e.g., B *proskiye*, A *praski* ‘fear’ < **prosk-u-h₁en-*) and **-i-h₁en-* (B *trik̄šo*, *yošiye*) as well.

There are many examples of Tocharian A derivatives ending in *-e* without a match in Tocharian B: *-krase* in *pār-ra-krase* ‘distance of an arrow-shot’ [*kārs-* ‘shoot, propel’], *tampe* ‘power’, *pate* ‘agriculture; plowing’, *peke* ‘painting’ [*pik-* ‘paint’], *malke* ‘milk’, *rape* ‘music’, *lame* ‘place’ [*lām-* ‘sit’], *wampe* ‘jewel’, *smale* ‘lie’ [no underlying verb], *tseke* ‘construction’ [*tsik-* ‘fashion’]. Others have Tocharian B equivalents in *-e*: *treke* (B *traike*) ‘confusion, delusion’ (*trik-* ‘go astray, be confused’), *retwe* (B *raitwe*) ‘≈ application, means’ [*ritt-* ‘apply’], *lake* (B *leke*) ‘bed’ (*lyäk-* ‘lie’), *letke* (B *laitke*) ‘vine’ [no underlying verb], *swase* (B *swese*) ‘rain’ (*su-* ~ *swāsā-* ‘rain’). Tocharian A *-e* has a mixed origin, PIE **-eh₂-h₁en-* and Proto-Tocharian **-äi-* of whatever origin.

The most common abstracts in Tocharian B are those formed with *-än̄ne* to which Tocharian A corresponds with *-une* ~ *-one*, and occasionally *-ñe*, e.g., B *astarñe* ‘purity’ from *astare* ‘pure’ (A *āstrone* from *āštār*), B *aśamñe* ‘wisdom’ from *aśamo* ‘wise’, A *tsopt̄sune* ‘size’ from *tsopats* ‘large’, *kāswone* ‘virtue’ from *kāsu* ‘good.’ Exceptionally B has *-uñne* in *lant̄uñne* ‘kingdom’ (A *lānt̄une*) and a few others. In both languages this abstract ending can be productively added to verbal adjectives in B *-lle*, A *-l* to form verbal abstracts, e.g., B *lkālñe* ‘sight, insight’ (< *lāk-* ‘see’), A *kālpālune* ‘achievement’ (< *kālp-* ‘achieve’), B *nesalñe* (A *naslune*) ‘being’ (< *nes-/nas-* ‘be’). When formed from transitive verbs these verbal abstracts can be constructed with either a direct object, e.g., B *lam̄tuñne ike källālñe* ‘the achievement of a royal position’ or a genitive, e.g., B *waipeccent̄am̄ts källālñe* ‘the achievement of possessions’. All other abstracts noted above take only the genitive.

5.1.4. Agent nouns

Tocharian B shows two examples of the Proto-Indo-European agent-noun formation **-tor-*: *s•ltre* (vowel of first syllable unknown) ‘a type of artisan’ (derived from an unknown verbal root) and *yerter* ‘felloe’ from PIE **h₂wērg-tor-* (cf. B *yerkwanto* ‘wheel’, A *wärkänt* ‘wheel’, and Hittite *hurki-* ‘wheel’). Tocharian B also shows two examples of the derived instrumental suffix **-tro-*: *šātre* ‘grain’ (< **‘Lebensmittel’*) (*šāy- ~ šāw-* ‘live’) and *tsarātre* ‘≈ extract’ (*tsär-* ‘separate’, modal stem *tsārā-*). Again found only in Tocharian B are root nouns used as agents: *yape* ‘spider’ (< **webhos*) and *lyak* ‘thief’ (< **leghs* ‘one who lies in wait’). A more recent layer of agent nouns is derived from prehistoric participles in **-nt-*: A *āsant* ‘leader’ (*āk-* ‘lead’), B *kausenta* ‘killer, murderer’ (*kau-* ‘kill’), A *košant* ‘executioner’ (*ko-* ‘kill’), A *pekant* ‘painter’, B *prekšenta* ‘judge’ (*pärk-* ‘ask’), B *yokānta* ‘drinker’ (*yok-* ‘drink’), B *weñenta* ‘speaker, intercessor’ (*we-* ‘speak’), A *märtkant* ‘shearer’ (B *märtk-* ‘shear’), A *rāpant* ‘musician’ (cf. *rape* ‘music’ from an unattested **rāp-* ‘play music’), *tsepant* ‘dancer’ (*tsip-* ‘dance’), and the further derived adjective *spärkšantik* ‘destroying’. Derivatives in **-ntyeh₂-* are almost exclusive to Tocharian B: *tarkāntsa* ‘carpenter’ (*tärk-* ‘work, fashion’), *mallāntsa* ‘vintner’ (*mäl-* ‘crush’), *wapāntsa* ‘weaver’ (*wāp-* ‘weave’), *walāntsa* ‘hindrance (to religious life)’ (*wāl-* ‘surround, hem in’). Where one can tell, this second type is derived only from the modal stem. Tocharian A has no attested agent nouns of this second sort but does have the derived abstract *wāpāntsune* ‘work of a weaver’ which presupposes a **wāpants* [= B *wapāntsa*] ‘weaver’ in its prehistory. Also, in Tocharian B participles in *-eñca* may be at least semiproductively nominalized as agent nouns, e.g., *trikšeñca* ‘sinner’.

Only Tocharian B has productive agent nouns from verbs. They are of two types: 1. *-uca* added to the modal stem and 2. *-uki* added to the non-modal stem. In both cases these formations preserve full verbal rection; that is, they preserve any direct objects in the accusative rather than change them to the genitive, e.g., *pilko palsko kärkauca* ‘a stealer of insight and thought’ (and not ***pilkontse pälskontse kärkauca*). Thus we have 1. *yāšsuca* ‘beggar’ (*yāšk-* ‘beg’ [A *yāšsuce* is borrowed from Tocharian B]), *ynuca* ‘one who goes’ (*i-* ‘go’), *kärkauca* ‘stealer’ (*kärk-* ‘rob, steal’, modal stem *kärkā-*), *kärsaUCA* ‘one who knows’ (*kärs-* ‘know’, modal stem *kärsā-*), *kärstauca* ‘destroyer’ (*kärst-* ‘cut off’, modal stem *kärstā-*), *pälkaucäkka* ‘(female) fortune-teller’ (*pälk-* ‘see, look at’, modal stem *pälkā-*), *pälskauca* ‘thinker, philosopher’ (*pälsk-* ‘think’, modal stem *pälskā-*) and 2. *aksāšsuki* ‘announcer’ (*āks-* ‘announce’, present stem *āksäsk-*), *kälpāšsuki* ‘achiever’ (*kälp-* ‘achieve’), *klāšsuki* [*epyāc klāšsuki* ‘rememberer’] (*käl-* ‘bring’), *tanmašsuki* ‘engenderer’ (*tām-* ‘engender’), *yamašsuki* ‘maker’ (*yām-* ‘do, make’), *yirpšuki* ‘inspector’ (*yärp-* ‘pay attention’), *wešsuki* ‘talker’ (*we-* ‘speak’), *šärpšuki** ‘one who shows the way’ (*šärp-* ‘point out’), *sparkāšsuki* ‘≈ one who puts to flight’ (*spärk-* ‘put to flight’). Neither *-uca* nor *-uki* have any convincing extra-Tocharian parallels.

Both languages allow the creation of agent nouns from nouns, though again, Tocharian B shows the bulk of the examples. Thus we have B *olyitau* ‘boatman’ (AB *olyi* ‘boat’), B *karyorttau* (A *kuryart*) ‘merchant’ (*karyor* [A *kuryar*] ‘buying and selling’), B *yirmakka* ‘≈ treasurer, steward’ (*yarm* ‘measure’), B *yenme_u* ‘gatekeeper’ (*yenme* ‘portal’), B *yotkolau* ‘≈ foreman, director’ (presupposing **yotkol* ‘≈ order’ and ultimately from *wätk-* ‘separate; command’), B *laukito* (A *laukit*) ‘stranger, guest’ (*lauke* [A *lok*]

‘far off; away’), A *waco* ‘warrior’ (*wac* ‘strife’), B *wetāu* ‘warrior’ (*weta* ‘strife’), B *saṃtkinau* (A *sāmtkenū*) ‘doctor’ (B *sāmtke* [A *sāmtāk*] ‘medicine’). B *pālkostau* ‘spy’ presumably belongs here, but the underlying noun is not attested. In Tocharian B there is obviously a tendency to make all these agent nouns *w*-stems in the nominative singular, but AB *lokit/laukito* and A *kuryart* show that in some cases at least this is secondary. Traces of PIE **-ā-* (cf. Latin *agricola* ‘farmer’) and **-tā-* are apparent here, though in PIE these suffixes would normally be added to verbs, not nouns. There are a few examples of agent nouns from nouns resulting from nominalizations of derived adjectives in *-tstse* (A *-ts*) ‘having-X’, e.g., AB *amok* ‘art’ and B *amoktse* (A *amokāts*) ‘artist’, B *-yamortstse* in *yolo-yamortstse* ‘evil-doer’ (B *rinatsstse* ‘renouncer’ is exceptional in being derived from a verb, *rin-* ‘renounce’), or B *werpiške* (borrowed in A *warpiške*) ‘garden’ and B *werpiškatsstse* ‘gardener’. Similar is B *prekšātsstse* ‘questioner’ though the underlying noun (**prekša?*, **prekšo?* ‘question’) is not attested.

5.1.5. Derived feminines

Both Tocharian languages have processes whereby a specifically feminine form may be derived from masculine nouns denoting animate beings. The processes vary. In Tocharian B the simplest is a change of final *-e* or *-o* to *-a*. In addition we find examples of *-ñña*, or *-āñca* (originally a borrowing from Iranian), or, once, *-kka*. The many fewer examples in Tocharian A show the A equivalents of *-ñña* (i.e. *-im*) and *-āñca* (*-āñc*); there are also a couple of instances of (PIE) **-yā-*. Examples include: A *ākālšāl* ‘student’, *āklāšālyi* ‘female student’, B *upāsake* ‘(male) lay follower’, B *upāsakāñca* (A *wāskāñc*) ‘female lay follower’, B *oñkolmo/oñkolma* ‘(she-)elephant’, B *ostaññe/ostañña* ‘male/female householder’, B *ñakte* (A *ñkāt*) ‘god’, B *ñakteñña* (A *ñakteñña*, probably borrowed from B) ‘goddess’, A *nātāk* ‘lord’, A *nāsi* ‘lady’, B *paše* ‘hare’, B *paša** [attested in the derived adjective *pšāññe*] ‘doe’, B *pālkaucākka* ‘(female) fortune-teller’ (*pālk-* ‘see, look at’), B *plaktukāñña* ‘female door-keeper’, B *plāñkši/plāñkšiñña* ‘seller’, B *mañiye/mañiya* ‘(female) servant’, B *māñcuške* ‘prince’, B *māñcuška* ‘princess’, A *māskit* ‘prince’, A *māskitāñc* ‘princess’, B *mewiyo/mewiya* ‘(female) tiger’, B *mokomške/mokomška* ‘(female) monkey, ape’, B *rakšatse/rakšatsa* ‘(female) rakšasa’, B *ršāke* ‘seer’, B *ršakāñca* ‘seeress’, B *walo* (A *wāl*) ‘king’, B *lāntsa* (A *lānts*) ‘queen’ (this pair reflects a PIE **wlānt-/*wlāntyā-*), A *wāsak* ‘lay-brother’, A *wāskāñc* ‘lay-sister’, B *šanmire/šānmirāñca* ‘(female) novice’, B *šamaške* ‘child, boy’, B *šamñāmška* ‘girl’, A *šom* ‘boy’, A *šomiṃ* ‘girl’. The Tocharian A pair *nātāk* ‘lord’ and *nāsi* ‘lady’ are particularly interesting as they would appear to be the exact equivalents of Greek *ánaks* ‘king’ and *ánassa* ‘queen’ from PIE **w(ṛ)nakts* and **w(ṛ)naktyā-* respectively (Winter 1970). At least in Tocharian B (no data from A) large herd animals (cows, sheep, goats) do things a bit differently. The basic species-designating words may also be used specifically of the female, and there are separate words, usually not etymologically related, for the male. Thus *ke_u/kau_urše* ‘cow/bull’, *ā_u/ariwe ~ āl* ‘sheep (ewe)/ram’, and *ās/āl* ‘(she-)goat/he-goat’. The sex of young sheep and goats at least is denoted by collocations with words for ‘female’ and ‘male’, e.g., *klaina šroñ* ‘female kids’ and *āl yriye* ‘ram-lamb’. In adjectival formations *ala-*, *klai(m)-* are used as parts of compounds, e.g., *alam-šrotāññe* ‘pertaining to a male kid’, *klaiṃ-šrotāññe* ‘pertaining to a female kid’, *klai-yritāññe* ‘pertaining to a ewe-lamb’.

5.1.6. Diminutives

Only Tocharian B shows a productive class of diminutives in *-ške* or (more rarely) *-ške*, e.g., *kapyāriške* ‘≈ little worker’ (from *ka(l)pyāre* ‘worker, laborer’), *kālyške* ‘boy, youth’ (no non-diminutive attested), *kāntwāške* ‘little tongue’ (*kantwo* ‘tongue’), *kuntiške* ‘little pot’ (*kunti* ‘pot’), *tapākiška* [f.] ‘little mirror’ (*tapākye* [f.] ‘mirror’), *tarske* ‘potsherd’ (no non-diminutive attested), *pšaška* [f.] ‘female bunny/leveret’ (*paše* ‘hare’), *mokomške/mokomška* [m./f.] ‘monkey, ape’ (< *‘little old man’), *ylaške* ‘young gazelle’ (*yal* ‘gazelle’), *werpiške* ‘garden’ (*werpiye* ~ *werwiye* ‘garden’; cf. A *warpiške* ‘garden’, borrowed and adapted from B), *šarmir(i)ške* ‘young novice’ (*šarmire* ~ *šanmire* ‘novice monk’), *šerška* ‘little sister’ (*šer* ‘sister’), *soške* ‘(dear) son’ (*soy* ‘son’); *māncuške* ‘prince’, *māncuška* ‘princess’, *yākwaške* ‘young horse’ (*yakwe* ‘horse’), *tsaške* ‘elder, senior monk’ (no non-diminutive attested); with *-kk-*: *appakke* ‘little father’, *ammakki* ‘little mother’, *tanākkō* ‘grain, spot’ (*tāno* ‘seed, grain’), *tsarākkañ* ‘weapons [of some sort]’.

One can use the same suffixes to form adjectives with affective meaning, e.g., *añmālaške* ‘pitying, sympathetic’, *amiške* ‘bad-tempered’, *tallānciške* ‘miserable’ (*tallāw* ‘miserable, unfortunate’), *tparške* ‘shallow’ (*tapre* ‘high, fat’), *lykaške* ‘small’ (cf. A *lykāly* ‘small’ with a diminutive suffix *-āly* seen only here), *naumikke* ‘≈ jewel-like, shining’ (*naumiye* ‘jewel, pearl’), *malyakke* ‘youthful, puerile’.

5.1.7. Names

Most of the attested texts in Tocharian B and virtually all of them in Tocharian A reflect Indian situations and the personal names contained in them are transliterations of Indian names. Many of the administrative documents in Tocharian B were generated by various Buddhist monasteries and the names recorded therein are also usually Indian in origin, presumably because the monks had taken specifically Buddhist names on ordination (e.g., *Candratewe* [*Candradeva*], *Indrasene*, *Jñānavīre*, *Somacandre*). A number of names end in *-ai(y)še* (e.g., *Mitraiše*, *Puttaiše*, *Yataiš*). This formation reflects Sanskrit compounds of the type **mitra-yaśa-* ± ‘Possessing the glory of Mitra’. So too members of the royal house normally appear with Sanskrit names, e.g., *Swarnatepe* (king of Kuca, CE 624–646), reflecting *Suvaradeva*, or his father *Swarnabūspe* (died CE 624), reflecting *Suvarapuṣpa*. However, the latter is also to be found in our texts under a Tocharian translation, *Ysāṣṣa Pyāpyo* ‘Golden Flower’. *Suvara-* is a recurring element in the names of the Kuchean (Tocharian B) royal house and *-arjuna-* in the Agnean (Tocharian A) royal house. It has been suggested (Pinault 1987) that real Tocharian B compound names might occur in *Kleñkarako*, if *Kleñka-rako* ‘Wagon-Director’ and *Kweṃtoko*, if *Kweṃ-toko* ‘One who sets dogs running’ (like Greek *kunēgós* ‘hunter?’); *Citrapaške* might be another such, *Citre* with an obscure second part (‘singer’? cf. A *rapant* ‘musician’), with the whole made into a diminutive. (More on compounds below, 5.2) There are, however, administrative texts in Tocharian B that list people not in monastic orders and these lists presumably give us a truer picture of ordinary naming conventions in Tocharian B society. There are some 175 “native” names (probably including a few Turkic names as well). Statistically they are distinguished from non-names by their much higher incidence of

diminutives in *-š/ške/a* (17%, e.g., *Puluške, Mitraške, Klyotiška, Muštiške*) and another set, probably diminutive, in *-le* (12%, e.g., *Catile, Caukile, Ptompile, Šaṅkale*). Ten percent end in *-k(k)e*, which again is likely to be diminutive, at least in part (e.g., *Capiśakke, Pällentakke* ‘Little Full-Moon’ [*pällent* (acc.) ‘full moon’], *Wiśikke*). Other phonetic peculiarities are the large number of names with geminate consonants (e.g., *Koppesale, Kwappale, Ńatti, Pissure, Ptāmparre, Mossotti*), the larger than expected number ending in *-i* (e.g., *Cauši, Cati, Ńatti, Tuši, Turki*) or *-āne* (e.g., *Mikisāne, Mukusāne*), and the presence of the initial cluster *pt-* (e.g., *Ptompile, Ptāmparre, Ptepār*). Where we have diminutives they are not always of the sort the English speaker would expect. Though we have *Lariška* ‘Little Dear’ (*lāre* ‘dear’), we also have both *Kercapiške* and *Kercipile* ‘Little Ass’ (*kercapo* ‘ass, donkey’) or *Cowaške* which looks like ‘Little Theft’ (*cowai*) or the like. *Indriške* is delightfully ambiguous as to whether it is ‘Little Indra’ (*Indre*) or ‘Little Penis’ (*indri*). Occasionally we have non-diminutive descriptors: *Šimprāye* ‘Wintry’ or *Yūšo* ‘Dullard’ (*yūše* ‘dull [of senses]’).

We have even less knowledge about Tocharian place-names. There is the obvious *Ńuwa Riye* ‘New-Town’, and the toponymic adjective *lestāññe* may conceal a place-name *Lesto* ‘The Nest’. We also know of a *Laikar*, ‘The Place of the Baths’(?), if related to *laiko* ‘bath’.

5.2. Compounds

The Tocharian languages are rich in nominal compounds. Essentially all that are familiar in other Indo-European languages can be found. There are copulative compounds (the *dvandvas* of Sanskrit grammatical tradition), e.g., A *akmal* ‘face’ (‘eye-nose’), B *ñem-kālywe* (A *ñom-klyu*) ‘fame’ (‘name-fame’), A *pācar-mācar* ‘parents’ (‘father-mother’), AB *śwātsi-yoktsi* ‘food and drink’.

There are also numerous examples of determinative compounds acting as adjectives (the Sanskrit grammarians’ *bahuvrīhis*): A *atra-tampe* ‘having heroic power’, B *astre-pālsko* ‘with pure spirit’, A *kāswa-pāłtsāk* ‘having a good thought’, B *tāryā-yākne* ‘three-fold’, B *totkā-yārm* ‘of small measure’, B *treya-meskeṃ* ‘having three joints’, B *pārkre-śaul* ‘long-lived’, B *pārkre-klauts* ‘having long ears’, A *ptāñkāt-ñom* ‘having the name of a buddha’, B *pyapyai-yok* ‘flower-colored’, A *prākra-pratim* ‘having a firm decision’, B *ywarca-meñ* ‘bimonthly’, B *śka-maiyya* ‘having ten powers’, B *šeme-yārm* ‘having a single measure’, B *še-śuke* ‘having a single taste’, A *šoma-pācar* ‘having the same father’. Sometimes the *bahuvrīhis* are provided with an adjectival derivational suffix (*-tstse* in B, *-um* in A) as well, e.g., A *ākār-aśnum* ‘with tear-filled eyes’, B *orotstse-cāmpamñetstse* ‘having great capabilities’, B *yolo-pilkotstse* ‘with bad insight’, A *wla-lune-ākatsum* ‘with death at the end’ (with both *-ts-* and *-um*), A *śka-tampeyum* ‘having ten powers’ (cf. B *śka-maiyya* without a suffix).

Finally we find substantial numbers of endocentric determinative compounds (*tatpuruṣas*), although these may be less numerous than, say, in Germanic, because they have heavy competition in the form of denominal adjectives plus nouns (see below, 5.3.3). With a regular verbal derivative (participle, verbal noun, agent noun) as the second member we have, for instance, B *aśca-sanāpalle* ‘head-salve’, A *ākāl-kāmše* ‘wish-fulfilling’, B *āñm-nākalñe* ‘self-reproach’, B *osta-šmeñca* ‘householder’, B *kārtse-aqšu*

‘well instructed’, B *kärtse-rita* ‘seeking good’, B *telki-yāmor* ‘sacrifice’, A *ri-pāše* ‘city-watchman’, B *lwaksā-tsaika* (A *kuntis-tsek*) ‘potter’, B *tsain-yamāššeñca* ‘arrow-making’. With nouns as second members we have B *amäkṣ-pānte* ‘≈ wagon-master’, A *ārāntiś-parām* ‘dignity of an arhat’, B *orotstse-pācer* ‘grandfather’, B *kašār-wāstsi* ‘*kāšāya*-clothing’, B *kuñi-mot* ‘grape-alcohol’ (i.e. ‘wine’), B *kurār-lūwo* ‘osprey’, B *kaum-ñäkte* (A *koṃ-ñkät*) ‘sun-god’, B *kaucū-wār* ‘upstream’, B *kwipe-ike* ‘shame-place’ (i.e. ‘penis’), B *pañäkte* ~ *pūd-ñäkte* (A *ptā-ñkät*) ‘Buddha’, B *poyši* ‘Buddha’ (i.e. ‘all-knowing’), B *ywärt-taś* ‘commander of the middle’, A *lāntune-abhišek* ‘royal consecration’, B *war-katso* ‘dropsy’, A *wsi-yats* ‘gold-skin’, B *šaṅ-añm* ‘oneself’, B *šaṅ-šaumo* ‘relative, retainer, follower’, A *se-ākāl* ‘wish for a son’, A *tsärk-rape* ‘lute-music’.

On the border of composition and derivation are privative compounds whose first member is *e(n)-* ~ *a(n)-*, e.g., verbal privatives, common in Tocharian B. Examples are B *atākatte* ‘untrue, unfounded’, *amplākätte* ‘without asking permission’, *ekamätte* ‘what has not come, future’, *etañkätte* ‘unhindered’, *ešuwatte* ‘uneaten’, *ontsoyotte* ‘insatiable’ (in Tocharian A there are only two attestations of this formation: *atāñkät* ‘unhindered’ and *asinät* ‘insatiable’), and such isolated formations as B *atāmo* ‘unfertile ground’ and *aswāre* ‘unsweetly’. More clearly compositional are semantically identical compounds with *snai* ‘without’ or *mā* ‘not’ as their first member (e.g., B *mā-yśelme* ‘imperceptible to the senses’, A *mā-wrāṣlune* ‘absence of judgment’, B *mā-šekāñne* ‘inconstancy’, A *sne-kaś* ‘numberless’, A *sne-yärm* ‘without measure, numberless’, A *sne-wars* ‘without spot’, B *snai-trāñko* ‘sinless’, B *snai-märkär* ‘unturbid, clear’ [cf. the semantically identical *snai-märkartse*]).

The canonical form of Tocharian compounds shows the first member in its “base” form. For adjectives, that is the masculine nominative singular (e.g., B *orotstse-pācer* ‘grandfather’, *pärkre-šaul* ‘long-lived’; *šeme-yärm* and *šoma-pācar* would be exceptions) and for nouns, the accusative singular (B *kau_urṣa-pkai* ‘having a chowrie’ [nom. *kau_urṣe*], B *pyapyai-yok* ‘flower-colored’ [nom. *pyāpyo*], B *meñ-yok* ‘moonlike’ [nom. *meñe*]). Where the resultant compound is a noun, the second member inflects normally for case and number (e.g., all the compounds in B *-ñakte*, A *-ñkät* ‘god’ [B *pañäkte*, gen. *pañäktentse*], B *saim-wāsta* ‘O protector’ [vocative of *saim-wāste*], B *ška-maiyya* [nom.], *ška-maiyyantse* [gen.], *ška-maiyya* [acc.] ‘ten-powered one’ [cf. nom. sg. *maiyyo* ‘power’]). However, most adjective *bahuvrīhis* have as their second member a noun frozen in the accusative case and are thus indeclinable. Usually it is the accusative singular, but there are cases where we find appropriately either the dual (*pärkre-klauts* ‘having long ears’) or plural (*yuše-yndrinta* ‘having dulled senses’, *treya-meskeṃ* ‘consisting of three parts’). Exceptionally, some *bahuvrīhis* whose second member is *-yok* ‘color, likeness’ are inflected in Tocharian B for number and case (masculine accusative singular *ysā-yokām* ‘golden’, feminine plural *meñ-yokāññana* ‘moon-like’). Of course, in such cases *-yok* may have become more of an adjective-forming suffix than the second member of a compound.

Sometimes, when the first member is a single syllable, it is extended by *-a-* so as to form a second syllable. The *-a-* of Tocharian A represents PIE compositional **-o-*; while the *-a-* of Tocharian B is stressed *-ā-* and represents a number of PIE antecedents (**-e-*, **-i-*, and **-u-*). The pattern of stress is only determinable in Tocharian B. In that language, when both members of the compound are disyllabic, the most common pattern is to have a single stress on the last syllable of the first member, including those cases where *-a-* has been added to make a second syllable (*pärkré-klauts* ‘having long ears’,

yušé-yndrinta ‘having dulled senses’, *treyá-meskeṃ* ‘consisting of three parts’, *ská-maiy-ya* ‘having ten powers’). There are, however, other patterns as well: *āñm-nākalñe* ‘self-reproach’, *kurár-lúwo* ‘osprey’, *pákri-nesályñe* ‘obviousness’ (with both members of the compound independently stressed), or *māntrákka-yäkne* ~ *māntrakká-yäkne* ‘so formed, so shaped’ (with an unstressed second member but variation as to the placement of stress on the first member). Tocharian A *ārāntiś-parām* ‘dignity of an arhat’, A *puttiś-parā* ‘rank of a buddha’, and *kuntis-tsek* ‘potter’ have a ‘joining element’ *-iś-* of obscure origin; in *kuntis-tsek* ‘potter’ the *-ś-* has become *-s-* by assimilation to the following consonant. Most compounds have only two members, but three-member compounds are attested: A *wsā-yok-yats* ‘having golden skin’, B *wartto-wšeññai-saim* ‘forest-dwelling-refuge’.

5.3. Adjectives

5.3.1. Introduction

Adjectives in the Tocharian languages agree with their head noun in number and gender and in case in the primary cases of nominative, accusative, vocative, and genitive. For the secondary cases of the noun the agreement is with the accusative of the adjective, as it is, at least half the time, with the genitive as well. Indeed, there appear to be no attested instances of a feminine genitive singular in Tocharian B and probably only one, or at most two, examples of the feminine genitive plural; examples of the feminine genitive are more plentiful in Tocharian A.

5.3.2. Adjectival stem-classes

In Indo-European terms the Tocharian languages show adjectives which reflect *yo*-stems and *n*-stems (both very common) and *o*-stems, though the latter are mixed with both *yo*-stems and *n*-stems (e.g., TchB *lāre* ‘dear’ from PIE **leh₂dros* has feminine nominative singular *lariya*, and masculine nominative plural *lareñ*, B *astare* ‘pure’, f.n.sg. *astarya*, m.n.pl. *astari* [A *āštār*, f.n.sg. *āštri*, m.n.pl. *āštre*]). Most commonly PIE *o*-stems are found in B *-šše* (A *-ši*), B *-ññe* (A *-ñi*) from **-syo-* and **-nyo-* respectively (see 5.3.3). TchB has *-tstse* (A *-ts*) from PIE **-tyo-* where other Indo-European languages have **-to-* in adjectives of possession, cf. B *kokaletstse* ‘having a wagon’, *stanātstse* ‘having trees’, *eśanetstse* ‘having eyes’ (such adjectives can be made from singular, plural, or dual nouns as appropriate), A *wākmats* ‘outstanding’ (from *wākām* ‘distinction’). TchB also has adjectives derived from verbs in *-ārške*, which look like diminutives (see below) but do not have discernible diminutive meaning, e.g., *takarške* ‘faithful; clear; gracious’ from *tāk-*, modal and preterite stem of ‘be’. However, both languages show fully productive deverbal adjectives in B *-lle* (A *-l*). They are formed from both the non-modal (“present”) stem and the modal (“subjunctive”) stem, e.g., B non-modal *tsrelle* ‘what is to be separated’ and (modal) *tsrälle* ‘separable’. As these examples show, the derivative of the non-modal stem is an adjective of necessity while the modal derivative is an

adjective of possibility. (The modal adjective forms the basis for the fully productive abstract verbal nouns of the languages, e.g., B *tsrālñe* ‘separation’.)

Adjectives also show productive remnants, often very much rebuilt, of *w(e)nt*-stems (e.g. TchB *perneu* [masculine accusative singular *pernent*], A *parno* ‘glorious’ [*parnont*]) and particularly of *w(e)nt*-stems originally built to *s*-stems (e.g., TchAB adjectives in *-ssu*; TchB *ymassu* [m.n.pl. *ymassoñc*], A *ymassu* [m.n.pl. *ymassuṣ*] ‘mindful’), past active participles in **-w(o)s-* (e.g., Tchb *yāmu* ‘done’, masculine nominative plural *yāmoṣ*, A *yāmu* and *yāmuṣ*), and, at least semi-productive, verbal adjectives/agent nouns in **-mon-* (e.g., TchB *weñmo* ‘speaker’, *salamo* ‘flying’). Non-productive are remnants of *nt*-stems (i.e., TchB *krent-* ‘good’ and *erkent-* ‘black’, A *krant-*, *arkant-*) and *m(e)nt*-stems (only AB *klyomo* ‘famous’ from PIE **k̑leumont-*). More innovative than Tocharian B, Tocharian A has extended *n*-stems secondarily to *yo*-stems (thus masculine accusative singular *-ṣim* [B *-ṣṣe*] and feminine accusative singular *-ṣinām* [B *-ṣṣai*]) and merges very evenly participles in **-nt-* and those in **-w(o)s-* (m.n.sg. *yāmu*, m.a.sg. *yāmunt*, m.n.pl. *yāmuṣ*, f.sg. *yāmuṣ*, f.pl. *yāmunt*). Likewise, Tocharian A has extended adjectives in *-ā* and *-u* to *-ām* and *-um* respectively (with an *-m* not yet well-explained). On the boundary between agent-nouns and derived verbal adjectives are the rare formations in *-a* (e.g., B *-rita* ‘seeking’, from *rit-*); all such formations in Tocharian A have been remade as adjectives in *-ām* (e.g., *-ritām*); outside of Tocharian one can compare the type seen in Latin *agricola* ‘farmer’. Much more common are formations with similar meaning in B *-i*, A *-e* (e.g., B *ākṣi* ‘announcing, announcer’, B *poṣi* ‘Buddha’ [from *po aik-* ‘know all’], B *yāmi* ‘doer’ [*kärtse-yami* ‘doer of good’], A *ākāl-kāmṣe* ‘wish-fulfilling’, *ri-pāṣe* ‘city-guard’). This formation reflects PIE **-ih₂en-*, the agent equivalent of the action noun forming **-ih₂on-* so common in Italic (e.g., Latin *opiniō* ‘opinion’).

5.3.3. Denominal adjectives

One of the more striking characteristics of the Tocharian languages is the widespread use of derived denominal adjectives where in other Indo-European languages one would expect a nominal genitive, e.g., TchB *śaiṣṣeṣṣem* [adj.] *skwanma* ‘the fortunes of the world’ or *po eśaneṣana* [adj.] *te[ka]nmane kartse* ‘[it] is good for all diseases of the eyes’. Certainly at times the two formations are semantically identical, e.g., TchB *ṣlentse* [gen.] *troñkne lyam=ompalskoññe* ‘he sat [preterite] in meditation in a hollow of the mountain’ but *ṣl[i]ye* [adj.] *gune cau ṣamy ompolskoññe* ‘he sat [imperfect] in meditation in that mountain cave’, or TchB *oñkolmamṣt* [gen.] *walo* ‘king of the elephants’, A *oñkāl-mem* [adj.] *wäl* ‘king of the elephants’. Denominal adjectives can be derived from singulars (e.g., B *ekaṣṣe* ‘of an eye’ from *ek* ‘eye’, B *cmelṣe*, A *cmolṣi* ‘pertaining to [a] birth’ from *camel* and *cmol* ‘birth’), duals (e.g., B *eśaneṣṣe* ‘of the two eyes’ from *eśane* ‘the two eyes’), and plurals (e.g., B *cmelaṣṣe*, A *cmolwāṣi* ‘pertaining to births’ from *cmela* and *cmolu* ‘births’), though, naturally, the number of derivatives of duals and plurals is small.

If the noun represented by the genitive or denominal adjective is low in animacy, the adjective is overwhelmingly chosen; if the noun is high in animacy, the genitive is overwhelmingly chosen. Thus, while the overall incidence of denominal adjectives is 56% in Tocharian B and genitives are correspondingly 44% of the total, for abstract

nouns the ratio is 76 to 24, for concrete inanimate nouns the ratio is 69 to 31, for concrete animate but non-human nouns the ratio is 51 to 48, for common human nouns the ratio is 45 to 55, and for proper human nouns the ratio is 4 to 96. The ratios for Tocharian A appear to be quite similar.

Genitives and denominal adjectives may be conjoined as modifiers of the same noun, e.g., B *sankantse* [gen.] *pelaiykneṣṣe* [adj.] *wāntare* ‘a legal affair of the community’. Moreover, the following patterns of noun modification also occur: a noun may be modified by a genitive which is itself modified by an adjective, e.g., B *kroksām̐ts* [gen.] *weṣeñña māka* [adj.] ‘the sound of many bees’; by a genitive which, in turn, is modified by another genitive, B *Airawantam̐tse* [gen.] *oñkolmām̐ts* [gen.] *lānte* [gen.] *seyi* [gen.] ... *ṣuñc* ‘the trunk of Airawanta, the son of the king of elephants’; by a denominal adjective in turn modified by a genitive, *akālk seyi* [gen.] *emelñeṣṣe* [adj.] ‘the wish for the birth of a son’; by a denominal adjective which, in turn, is modified by another denominal adjective, B *laksañai* [adj.] *klautsaiṣṣe* [adj.] *ṣpel* [nom. sg. noun] ‘poultice of fish ears’ (here the more deeply embedded denominal adjective [*laksañai*] agrees in the genitive [formally accusative, see above] case appropriate to the genitive noun underlying the denominal adjective *klautsaiṣṣe*); and finally, by a genitive which, in turn, is modified by a denominal adjective, TchB *jambudvipṣem* [adj.] *śāmnantsä* [gen.] *naumye ysāṣṣe* [adj.] ‘the golden jewel of the peoples of India’.

Denominal adjectives in *-(i)ye* [A *-i*], *-ññe* [A *-ñi* and *-em*], and *-ṣṣe* [A *-ṣi*] would appear to all serve the same range of functions and thus to be synonymous. Occasionally there are indeed synonymous doublets: (all B) *riññe* ~ *riṣṣe* ‘pertaining to a city’, (common) *ñäkcye* ~ (rare) *ñäkteññe* ‘divine’, *ṣliye* ~ *ṣleṣṣe* ‘pertaining to a mountain’, *keñiye* ~ *kenäṣṣe* ‘pertaining to the ground, land’. Their distribution, however, is not random. The suffix *-(i)ye* appears rather infrequently in all semantic groups except those at the ends of the animacy scale (abstract inanimates and human proper nouns): it is of course the ubiquitous PIE suffix **-iyo-* and the nouns from which it derives denominal adjectives all appear to be inherited from PIE or at least are very old in pre-Tocharian. It does not appear to be productive. The Tocharian B suffix *-ññe*, does not normally appear with inanimate abstract nouns at all; it appears rarely with inanimate concrete nouns (e.g., *eñcuwaññe* ‘of iron’, *ysārñe* ‘wheaten’), but appears, beside *-(i)ye*, just about two-thirds of the time (68% *-ññe*, 32% *-[i]ye*) with words designating animals and 79% of the time with human common nouns (otherwise *-ṣṣe*). Given this connection with animacy, it is surprising that there are no adjectives in *-ññe* from human proper nouns. In PIE terms *-ññe* represents **-n-yo-*. Very rarely we have *-uññe* rather than *-ññe*, i.e., *kotruññe* ‘pertaining to the family’, *lantuññe* ‘royal’, *lykuññe* ‘pertaining to a thief’; Tocharian A *-ñi* is rare and does not show the same animacy distribution as TchB *-ññe* (A *oñi* ‘human’ from *oñk* ‘human being’ [cf. B *eñkuwaññe*], but *praskañi* ‘frightful’, and *yokañi* ‘thirsty’). The suffix *-em* (= B *-aññe* and *-eññe*), however, does show the expected predilection for animate nouns (e.g., *oñkalem* ‘pertaining to an elephant’ [B *oñkolmaññe*], *pättāmñktem* ‘pertaining to the Buddha’ [B *paññaktāññe*], *lwem* ‘pertaining to an animal’ [B *lwaññe*]).

In Classical Tocharian B the suffix *-ṣṣe* [A *-ṣi*] occurs everywhere else and is the only suffix that can form denominal adjectives from duals and plurals, no matter what semantic group they may belong to. (The one exception is B *-pikwalaññe* used in forming compound adjectives denoting age, e.g., *ikaṃ-pikwalaññe* ‘twenty years old’; in later Tocharian B *-ññe* begins to encroach on *-ṣṣe* in the singular but never in the plural.) It

is universal in forming denominal adjectives from abstract inanimates, and human proper nouns at the other end of the animacy scale. It is hard not to take *-sše* as arising from *-āsše* by redivision into *-ā-sše* with subsequent spread of *-sše* to other noun types (just as the productive suffix *-tstse* ‘having X’ is redivided from *-ātstse* from PIE **-eh₂tyo-*). Proto-Tocharian **-āsše* is from PIE **-eh₂syo-* as in Latin *-ārius* (e.g., *agrārius* ‘agrarian’) and, without the **-y-*, in Anatolian *-assa-*. Though it is not customary to do so, I would add Greek adjectives in *-aios* (< **-eh₂syo-*) as well. If a genitive singular in **-osyo* survived into Proto-Tocharian, it would have appeared as **-esše* and helped the redivision of *-āsše* into *-ā-sše*.

5.4. Adverbs

As expected, Tocharian adverbs modify verbs, adjectives, and other adverbs. On a semantic basis we can divide them into two groups: qualifying adverbs (i.e., adverbs of manner or quantifying adverbs) or circumstantial adverbs (i.e., temporal, local, or causal adverbs). Thus qualifying adverbs answer the questions ‘how?’ or ‘how much?’, while circumstantial adverbs answer the questions ‘when?’, ‘where?’, or ‘why?’. An example of a qualifying adverb occurs in B *twe pitka wes ām* [adv.] *lamam* ‘command thou [that] we sit quietly’ or *kokaletstse* [adv.] *īyoy sū Prasenacī walo ot* ‘then King Prasenajit was traveling by wagon’. An example of a circumstantial adverb is B *śaumo ks-allek komt tsonkaik* [adv.] *tsaṅkoy* ‘may someone else arise daily at dawn’. On a formal basis we can divide adverbs into four groups: (synchronically) underived, derived from adjectives, derived from nouns, and pronominal adverbs. There is no strong correlation between semantic function and form. Examples of underived adverbs include B *at(e)* [A *atas*] ‘≈ away’, A *ārt* ‘over a distance’, B *ot* ‘then’, B *kauc* [A *koc*] ‘(on) high’, B *ñerwe* [A *ārśo*] ‘today’, B *naus* [A *neš*] ‘earlier, before’, and A *tsiṅk* ‘hard, tightly’. Adverbs derived from adjectives appear in a form identical to the nominative masculine singular, which is probably historically the nominative-accusative neuter singular. This formation appears to be productive in both languages. Examples include B *arwäre* [A *ārwar*] ‘ready’, B *kartse* ‘well’, A *kāsu* ~ *kāswe* ‘well’, B *kokaletstse* ‘by wagon’ (as an adjective it means ‘possessing a wagon’), A *tpär* ‘high’, B *läre* ‘dearly’, B *lykaške* ‘finely’, and A *salu* [B *solme*] ‘completely’. Sometimes we find a prefixed version of the adjective used as an adverb, e.g., B *emparkre* [A *apärkär*] ‘widely, long, expansively’, B *enkätkre* ‘deeply’, B *eñwetstse* ‘ever anew’. Deriving adverbs from nouns may not be a productive process in attested Tocharian but there is good evidence that it once was, the adverb being identical with the accusative singular of the noun. The evidence is best of course in those rare cases where the noun and adverb exist side by side, e.g., B *tsamo* ‘very’/‘growth’, B *ām* ‘silently’/‘silence’, B *kaum-yaši* ‘day and night’. More often we have an adverb which looks to be the accusative singular of a noun but the noun itself is not attested, e.g., A *letäk* ‘separately’, B *wetke* ‘away’, A *spānte* ≈ B *spantai* ‘trustingly’. We also find accusatives of nouns prefixed with *e(n)-* ~ *a(n)-*, *y(n)-* (both meaning either ‘in’ or serving simply as an intensive) (see above), or suffixed by *-k* (intensive), e.g., B *akek* ‘finally’, A *ākak* ‘constantly’ (both from a noun meaning ‘end’, B *āke*, A *āk*), B *tsonkaik* ‘in the morning’ (no attested noun), B *anaišai* ‘attentively’ (noun *aišai* ‘≈ attention, notice’), B *aplāc* ‘in conversation’ (noun *plāce*, acc. *plāc* ‘conversation’),

A *oklop* ‘in danger’, A *opärkā* ‘in the morning’, A *ykoṃ*, B *iñkaṃ* ‘by day’, AB *ynāñm* ‘worthily’, B *yneśne* ‘really, manifestly, obviously’ (lit. ‘in the eyes’), B *yśāmna* ‘among men’. Occasionally, at least in B, we find *e(n)-* ~ *a(n)-* used in its second etymological meaning, ‘not’, e.g., *aswāre* ‘unsweetly’, *enersānk* ‘inadvertently’. Not surprisingly, we find many examples of nominal cases other than the accusative used as adverbs, e.g., A *elā* ‘under cover’ (no noun otherwise attested), B *aurtsesa* (A *wārtsyo*) ‘fully’, B *aultsorsa* (A *waltsurā*) ‘in short’, A *koṛpā* ‘against’ (no noun otherwise attested), B *poyknesa* (A *puk-wāknā*) ‘in whatever way’, B *watesa* ~ *wtentse* ‘again’ (*wate* ‘second’), B *špālmem* ‘best’ (*špāl* ‘head’). Both languages have a host of pronominal adverbs, almost exclusively with circumstantial meaning, e.g., B *ālām* ‘otherwise’, B *tane* ‘here,’ A *tanne* ‘in this way, so’, B *tumem* ‘thereupon’, B *tusa* ‘thus’, A *tmā* ‘over there’, A *tmāš* ‘thereupon, then’. Finally, both languages productively form distributive (and other) adverbs by iteration, e.g., B *kālymi-kālymi* ‘region by region’, A *kālyme-kālyme* ‘everywhere’, A *kumpa-kump* ‘pot by pot’, A *koṃ-koṃ* ~ *kona-koṃ* ‘day by day’, A *ksānk-ksānk* ‘chuckling’, B *nāno-nāno* (A *n_nnak-n_nnak*) ‘again and again’, B *pālyca-pālyc* ‘fleetingly’, A *pkānt-pkānt* ‘separately, each individually’, B *somo-somo* ‘one by one’, B *šek-šek* ‘forever’, A *šñi-šñi* ‘each his own’. (These collocations are in every way like Vedic *āmreḍitas* [Klein, 2003].)

5.5. Verbs

5.5.1. Introduction

Tocharian verbs are inflected for number (singular, dual, plural), person (first, second, third), tense (past, non-past), mood (indicative, imperative, subjunctive, optative), aspect (imperfective, perfective – but only in the past indicative), and voice (active, mediopassive). There are approximately 450 verb roots attested in Tocharian B. Some 150 of them are basically intransitive and the rest basically transitive (the roughly one-third to two-thirds ratio of intransitives to transitives is found in many languages). A quarter or more of these 450 verbs have derived causatives (see below, 5.5.3). In addition there are some sixty periphrastic verbs known from Tocharian B, e.g., TchB *yoktsi ai-* ‘give to drink’, *postām i-* ‘follow’, *ñem tā-* ‘name’, *ñwecce klutk-* ‘renew’, *parra spāw-* ‘disburse’, and particularly the substantial number composed of noun + *yām-* ‘do, make’ (see further below, 5.5.4). With periphrastic verbs the ratio of intransitives to transitives is much lower, on the order of one-to-nine. Counting the basic verb and any derived causative independently, and adding the periphrastic verbs, we arrive at a grand total of about 625 attested Tocharian B verbs. Tocharian A has a smaller attested corpus and thus fewer attested verbs (some 330), but still there are some 110 verb roots attested in Tocharian A that are not attested in Tocharian B (of which there are about fifty whose meaning is unknown). If the Old English corpus were the size of Tocharian B’s, Old English would have some 860 verbs attested (again basing our estimate on the number of entries in Hall 1960). Thus there is a demonstrably thinner “verb density” in Tocharian than in Old English. The “average verb”, if one can think in such terms, in Tocharian must have a more generic meaning than the “average verb” in Old English.

5.5.2. Verbal stem-classes

Tocharian indicative stems and many modal stems (modal stems are often old presents that have been relegated to modal status by the creation of new indicative stems) preserve recognizable traces of many Proto-Indo-European present-stem formations. We find *e*-grade athematics with a first person singular in **-mi* (e.g., AB *i-* ‘go’ from PIE **h₁ei-*), *o*-grade athematics with a first person singular in **-h₂ei* (e.g., B *neku* ‘I will destroy’ from PIE **nok* – these are the equivalent of Hittite *hi*-verbs and in Tocharian have been almost always relegated to the modal stem), and *e*-grade thematics in **-e/o-* (e.g., B *parām* ‘carries’ from PIE **bher-*, *klyaušām* [A *klyošās*] ‘hears’ from PIE **klēus-e/o-*). More complex present-stem formations are those in **-ye/o-* (by itself largely relegated to modal stems [but there remain a few presents like *keri-/kerye-* ‘laugh’] and then only in Tocharian B, but more commonly as part of the conglutinate suffix **-n-ye/o-*), **-se/o-*, **-ske/o-*, **-eh₂-*, **-n(e)w-*, **-n(e)h₂-*, **-eye/o-* (i.e. B *auk-* ‘give to drink’ from PIE **h₁og^wheyel/o-*; these have fallen together with the *o*-grade athematics [Adams 2003]), and **-eh₁-* (only in the derived middle **-h₁-ó-*).

5.5.3. Causatives

Roughly one-third of all Tocharian verbal roots are basically intransitive, the rest basically transitive. Unlike English, very few verbs in Tocharian are indifferent as to transitivity. Clear examples of the same form of the verb being used both intransitively and transitively are limited to *iyā-*, both ‘travel’ (intr.) and ‘lead’ (tr.), *kätk-*, both ‘proceed; pass’ (intr.) and ‘cross’; ‘make cross’ (tr.), and *pälk-*, both ‘shine’ (intr.) and ‘illuminate’ (tr.). On the other hand, Tocharian verbs rarely make explicit any difference between stative and eventive verbs (on the order of English *stand* vs. *stand up*, *sit* vs. *sit down*), though Tocharian B would appear to make a distinction between stative (traditionally so-called, but perhaps better, “eventive durative”) *täl-* ‘keep upraised’ and its eventive “causative” *täläsk-* ‘raise upward’ (however, in this case they share a single preterite, *cālā-*). Another example is B *länkš’ä/e-* ‘hang’ [tr.] and *länkäsk’ä/e-* ‘let dangle’.

In Proto-Indo-European it appears that the most common way by which a basically transitive verb root might lose an argument and be made intransitive would be by becoming mediopassive, and the most common way a basically intransitive verb would gain an argument and become transitive would be by the addition of a causative suffix. The same is true of Tocharian, though the forms these conversions take, particularly with regard to the causative, have changed. Just as in Proto-Indo-European, intransitives are made from transitives by changing active into mediopassive endings, e.g., B *nakšām* ‘destroys’ beside *nakštār* ‘is destroyed, perishes’. In this instance A shows an additional stem difference: *nkäš* ‘destroys’ but *näknäštār* ‘is destroyed, perishes’.

Of the basically intransitive roots, some 55% or more have attested derived transitives or “causatives”, e.g., using Tocharian B examples, *tsälpetār* ‘is saved’ vs. *tsalpästār* ‘redeems’, *lyuketār* ‘shines’, vs. *lukšām* ‘illuminates’, *tmetār* ‘grows’ (intr.) vs. *tsamšām* ‘makes grow’. Given the vagaries of attestation and the smallness of the overall sample, deriving a causative from an intransitive must have been extremely productive in Tocharian. Of the roughly 270 basically transitive roots, only about twenty-five (roughly 9%)

have derived causatives. The small number of causative derivatives from transitive roots makes it doubtful that this is a fully productive process in Tocharian. Only two-thirds of this latter group are actually semantically causatives ('make/cause s. o. do-X') of the root they are derived from (meanings of the transitive and derived causative separated by a slash): B *aik-* 'know/cause to know', AB *kätk-* 'traverse/allow passage', AB *kärs-* 'know/make known', AB *kälp-* 'obtain/bestow', B *ku-* 'pour/make pour', B *kau-* 'kill/let kill', B *yäm-* 'achieve/obtain', B *ränk-* 'rise above/ascend, take control of', AB *läk-* 'see/show', B *tsuk-* 'suck/cause to suck'.

While the two most important causative suffixes in Proto-Indo-European were **-new-* and **-eye/o-* (together with *o*-grade of the root), only the second has left any trace in Tocharian (B *auk-* 'give to drink', see above, 5.5.2). The causative-forming suffixes in Tocharian are **-selo-* and **-skelo-* (in Tocharian A the two have fallen together under the form of the first). In Proto-Indo-European these suffixes were iterative or intensive in meaning (cf. the situation in Hittite); the shift to causative meaning presumably parallels the same shift seen in the *o*-grade verbs in **-eye/o-* (which shift, if indeed that is what occurred here, had happened by Proto-Indo-European times). Thus we find B *prutketär* 'is confined, shut up; is filled to overflowing' beside *prutkäššäm* 'shuts up, fills up', B *orotär* 'ceases, comes to an end' beside *äršäm* 'leaves behind, renounces' (A *aratär* beside *äräs*), B *plänketär* 'comes up for sale' beside *plänkšäm* 'sells', A *kärsnäs* 'knows' beside *särsäs* 'teaches' (B *kärsanam* beside *särsäššäm*; root-initial palatalization and root-stress are other hallmarks of a causative verb), or, in a case like B *täm-* 'be born', where both the intransitive present ('be born') and transitive present ('engender') end in *-äsk-*, the intransitive is stressed on the stem-formative and has mediopassive endings while the transitive is stressed on the root and has active endings: *tänmastär* 'is born' beside *tanmäššäm* 'engenders'. The non-causative and causative preterite and past participle stems also are characteristically different. In B the causative preterite ends in *-ššā-*, e.g., *tanmäšša* 'he engendered' (cf. *temtsate* 'he was born') or has the root vowel *-ā-* typically with root-initial palatalization, e.g., *šārsa* 'he taught' (cf. *šarsa* 'knows'). In A the causative preterite is usually reduplicated, e.g., *šasärs* 'he taught' (cf. *särs* 'he knew'). In both languages the preterite participle of causative verbs is normally reduplicated, e.g., B *tetanmäššu* 'engendered', *šeššarsu* 'taught'. Only in Tocharian B are representatives of a periphrastic causative attested: *klyaušsi ai-* 'let hear' (lit. 'give to hear'), *yoktsi ai-* 'give to drink' (semantically equivalent to *auk-* 'give to drink'), and *šwätsi ai-* 'feed'. On the border between true causatives and verbs of command with sentential objects (e.g., 'order someone to do something') are such sentences as B *se šamāne šaṅ šarsa kem rapanam räptasi wat watkäššäm* 'whatever monk should dig in the earth with his own hand or have someone else dig it, it is forbidden', where the Tocharian B *räptasi watkäššäm* translates the straight Sanskrit causative *khānayet*. Therefore, derived causatives may be very common in Tocharian, but they are not nearly as universal as in Sanskrit; thus the necessity of creating these periphrastic equivalents.

5.5.4. Denominative verbs

While Tocharian has a rich system for deriving nouns and adjectives from verbs, the creation of verbs from nouns or adjectives is much more limited. Tocharian shows sub-

stantial but unproductive traces of two PIE denominative verb formations, **-ye/o-* and **-eh₂-*. The bulk of Tocharian denominatives in **-ye/o-* are those formed originally to *n*-stems, hence the double suffix **-n-ye/o-* (except where noted, all examples are from Tocharian B): *añmaññ-* ‘wish, desire’ (*añme* ‘wish’), *kwipeññ-* ‘be ashamed’ (*kwipe* ‘shame’), *tāñkwaññ-* ‘love’ (*tāñkw* ‘love’), TchA *tuñkiññ-* ‘love’ (*tuñk* ‘love’), *lareññ-* ‘love’ (*lāre* ‘dear’), *skwaññ-* ‘be lucky, fortunate, happy’ (*sakw* ‘good fortune’), *tsereññ-* ‘deceive’ (*tserekwa* [pl.] ‘deception’). Tocharian B shows one probable example of a **-ye/o-* denominative not to an *n*-stem in *šeri-* ‘hunt’ (*šerwe* ‘hunter’; in PIE terms **g^wh₂er-wo-* and **g^wh₂er-w-ye/o-*). The denominatives in **-eh₂-* have mostly been relegated to modal stems by the creation of new indicative stems in either *-nā-* or *-o-*. Having become modals in *-ā-*, they tend to lose their identity among *-ā-* modals of different origins. Possible examples of such cases are those modal stems in *-ā-* that show no ablaut and have beside them nouns in *-e* or *-o* which show PIE *o*-grade: *klautkā-* (modal) ‘turn, become’ (*klautke* ‘way, manner; behavior’ < *klutk-* ‘turn’ [A *lutk-*, *lotäk*, (modal) *lotkā-*]), *laikā-* (modal) ‘wash (away) [intr.]’ (*laiko* ‘bath, washing’ < *lik-* ‘wash [another person]’ [A *lik-*, (modal) *lekā-*]), *wāpā-* (modal) ‘weave’ (*wepe* ‘≈ corral, paddock’ [A only *wāp-*]), *skāyā-* (modal) ‘strive, attempt’ (*skeye* ‘zeal, effort; temptation’ [A *ske*, (modal) *skāyā-*]), AB *spālkā-* (modal) ‘move forcefully’ (*spel[t]ke* ‘zeal’ [A *spaltäk*]), *swārā-* (modal) ‘please’ (*swāre* ‘sweet’ [A *swār*]), *swāsā-* (modal) ‘rain’ (*swese* ‘rain’ < *su-* ‘rain’ [A *su-*, *swase*, (modal) *swāsā-*). There is also in Tocharian B one verb which at least looks like it might possibly be a denominative in *-sk-*, *wināsk-* ‘honor’ [A *winās-*] to *wīna* ‘pleasure’; but its formal isolation and the loose semantic connection between (supposed) underlying noun and verbal derivative invite caution. Both the formations in *-ññ-* and *-ā-* seem clearly unproductive. The combination of noun + *yām-* is the primary (and probably only productive) way of creating denominative verbs in attested Tocharian. The resultant verb may be (rarely) intransitive (all examples are Tocharian B unless noted): *añkaiṃ yām-* ‘vomit’, *apākārtse yām-* [A *pākār yām-*] ‘be visible, be manifest’ (= BHS *prakās-*), AB *onmiṃ yām-* ‘repent’, *ārwer yām-* [A *ārwar yām-*] ‘prepare oneself’ (this verb may also be transitive), *wīna yām-* ‘be pleased, amused, gratified’ (= BHS *ram-*). Overwhelmingly more common are transitive verbs. Usually the direct object of the compound verb is in the accusative, e.g., *ate yām-* ‘take off [clothes]’ (= B *rutk-*), *anumodit yām-* ‘give approbation to’, *appamāt yām-* ‘mistreat’, A *abhišek yām-* ‘ordain’, *arwer yām-* ‘prepare’, A *āštār yām-* ‘purify’, *aiśai yām-* ‘take care of, tend’, *ite yām-* ‘fill’, *keś yām-* ‘count’ (= B *šāms-*), *kṣānti yām-* ‘forgive’, A *cārit yām-* ‘hand out’, *cotit yām-* ‘accuse’ (= BHS *codaya-*), A *tuñk yām-* ‘love’ (= A *tuñkiññ-*), *telki yām-* [A *talke yām-*] ‘sacrifice’ (possibly intransitive), *pākri yām-* ‘make public’, *ynes yām-* ‘realize’, *lāre yām-* ‘love’ (= BHS *sev-*, B *lareññ-*), A *šotre yām-* ‘mark, label’, *sarit yām-* ‘memorize’, and *saiṃ yām-* ‘take refuge in’. Less frequently the direct object is in a case other than the accusative, usually the genitive, e.g., *āke yām-* ‘end’ (+ gen.), *ākli yām-* ‘teach’ (+ locative), *pāke yām-* [A *pāk yām-*] ‘share’ (+ gen.), *paucčiṃ yām-* ‘renounce’ (+ ablative), *pkante yām-* [A *pkānt yām-*] ‘hinder’ (+ gen.), *yarke yām-* [A *yār k yām-*] ‘honor’ (+ gen.), *wasampaṃ yām-* ‘ordain’ (+ gen.), AB *spaktām yām-* ‘serve’ (+ gen.).

6. References

- Adams, Douglas Q.
1988 *Tocharian Historical Phonology and Morphology*. (American Oriental Series 71). New Haven: American Oriental Society.
- Adams, Douglas Q.
2003 'Give to Drink' in Tocharian B and the Reflexes of the PIE Causative. *Tocharian and Indo-European Studies* 10: 1–10.
- Adams, Douglas Q.
2013 *A Dictionary of Tocharian B*. 2nd edn. (Leiden Studies on Indo-European 10). Amsterdam: Rodopi.
- Bauer, Brigitte L. M. and Georges-Jean Pinault (eds.)
2003 *Languages in Time and Space: A Festschrift for Werner Winter on the Occasion of his 80th Birthday*. Berlin: Mouton de Gruyter.
- Blažek, Václav
1999 Alimenta Tocharica (1–3). *Journal of Indo-European Studies* 8: 79–82.
- Cunliffe, Richard John
1924 *A Lexicon of the Homeric Dialect*. London: Blackie and Son.
- Hall, J. R. Clark
1960 [1894] *A Concise Anglo-Saxon Dictionary*. 4th edn. with a supplement by Herbert D. Meritt. Toronto: University of Toronto Press. [Cambridge: Cambridge University Press.]
- Huld, Martin E.
1990 The Linguistic Typology of the Old European Substrate in North Central Europe. *Journal of Indo-European Studies* 18: 389–424.
- Ivanov, Vyacheslav V.
2003 On the origin of Tocharian terms for GRAIN. In: Bauer and Pinault (eds.), 189–210.
- Klein, Jared S.
2003 Āmreḍitas and Related Constellations in the *Rigveda*. *Journal of the American Oriental Society* 123: 773–802.
- Krause, Wolfgang
1952 *Westtocharische Grammatik*. Band 1: *Das Verbum*. Heidelberg: Winter.
- Krause, Wolfgang and Werner Thomas
1960 *Tocharisches Elementarbuch*. Band I: *Grammatik*. Heidelberg: Winter.
- Lubotsky, Alexander and Sergei Starostin
2003 Turkic and Chinese loan words in Tocharian. In: Bauer and Pinault (eds.), 257–270.
- Melchert, H. Craig
1989 PIE 'dog' in Hittite. *Münchener Studien zur Sprachwissenschaft* 50 [1990]: 97–101.
- Pinault, Georges-Jean
1987 Notes d'onomastique koutchéenne. *Tocharian and Indo-European Studies* 1: 77–97.
- Schuessler, Axel
1987 *A Dictionary of Early Zhou Chinese*. Honolulu: University of Hawaii Press.
- Sieg, Emil, Wilhelm Siegling, and Wilhelm Schulze
1931 *Tocharische Grammatik*. Göttingen: Vandenhoeck & Ruprecht.
- Winter, Werner
1970 Some Widespread Indo-European Titles. In: George Cardona, Henry M. Hoenigswald, and Alfred Senn (eds.), *Indo-European and Indo-Europeans: Papers Presented at the Third Indo-European Conference at the University of Pennsylvania*. Philadelphia: University of Pennsylvania Press, 49–54.

Douglas Q. Adams, Moscow, ID (USA)

79. The dialectology of Tocharian

1. Preliminary remarks
2. Geographical approach
3. Diachronic approach
4. Palaeographical evidence
5. Absolute chronology
6. Summary
7. References

1. Preliminary remarks

First of all, it should be pointed out that Tocharian A and Tocharian B are considered separate languages here, deriving from a common ancestor usually called Common Tocharian (German *Gemeintocharisch*, French *tokharien commun*). In his contribution on the interrelationship between Tocharian A and B, Lane calls them dialects (Lane 1966), but he himself holds that the differences between Tocharian A and Tocharian B are greater than between any two languages of, e.g., the Slavic or Germanic branch (Lane 1966: 213). So far, however, no general agreement has been reached concerning this matter (cf. Thomas 1984: 126–127 with references). The least that can be said is that the differences observable within the Tocharian B corpus are much less substantial than those between Tocharian A and Tocharian B. All in all, the two idioms differ from each other too much phonologically, lexically, and esp. morphologically to be treated as dialects of a single language (cf., e.g., Winter 1998: 155).

There is virtually no phonological or morphological diversity within the Tocharian A corpus of texts that could be described in terms of dialectology; thus, the dialectological description of Tocharian is first and foremost concerned with Tocharian B.

2. Geographical approach

The first linguistic features to attract the attention of scholars were the so-called “*MQ-Schreibungen*” found in texts from western sites around Kucha. These particular spellings clearly diverge from what is usually found in standard Tocharian B texts unearthed in the central region of Tocharian settlement in and around Šorčuq. In his thorough synchronic treatment of the Tocharian B verbal system, Krause listed a series of features that occur systematically in texts found in the western part of the Tocharian B area, especially in the caves of Ming-oi Qizil (MQ) near Kucha and, to a lesser extent, in Kucha itself (Krause 1952: 1 ff.). Krause suggested cautiously that the texts containing MQ-spellings basically belong to an older layer of Tocharian B. According to his interpretation, it would be incorrect to speak of Tocharian dialects at all.

Based on Krause’s description, Winter (1955) gave a systematic account of the distribution of some 25 phonological features in the texts of the Berlin collection that set the MQ texts apart from the texts written in the standard language. Additionally, he isolated a third distinct group of texts, all of which were discovered in the area around Turfan. Since Winter was unable to devise a chronological scenario for the corpus of Tocharian B, he pursued the matter from a purely synchronic point of view. Winter succeeded in

Tab. 79.1: Shibboleth Features of the Western Dialect as Opposed to the Central Dialect

Western Dialect	Central Dialect
occurrence of < \ddot{a} , a , \bar{a} > not linked to stress rules < \ddot{o}_i > < \ddot{e}_u , ew > < \ddot{o}_u , ow , au , auw >	occurrence of < \ddot{a} , a , \bar{a} > linked to stress rules < oy > < au > < au >

Tab. 79.2: Shibboleth Features of the Eastern Dialect as Opposed to the Central Dialect

Eastern Dialect	Central Dialect
/ \ddot{a} / spelled < i > in palatal environment - \acute{s} - w - $\acute{s}w$ -	/ \ddot{a} / spelled < \ddot{a} > in palatal environment - c - p - $\acute{s}tw$ -

isolating three major dialectal groups: the Central Dialect (area of Šorčuq), the Western Dialect (area of Kucha; esp. MQ), and the Eastern Dialect (area of Turfan). The shibboleth features given by Winter (1955: 224) are provided in tables 79.1. and 79.2.

3. Diachronic approach

As mentioned above, Winter's dialectological categorization is based on the premise that all differentiation within Tocharian B should be ascribed to geography rather than chronology. This premise was not challenged until 1977, when Peter Stumpf in his habilitation thesis (published posthumously in 1990) returned to the position upheld by Krause (1952) and argued convincingly that the diversity in Tocharian B can best be accounted for by positing chronological layers within one single language. Methodologically, he extended the scope of the investigation by taking into account morphological features more comprehensively than Winter.

Stumpf isolated two major linguistic layers, I and II, in the manuscripts, with I being the older layer and II the younger layer (Stumpf 1990: 74, 108).

In layer I, Stumpf identified three sublayers on the basis of the spelling rules for the vowel triad \ddot{a} , a , \bar{a} (Stumpf 1990: 76–79) and the spelling of the u -diphthongs (Stumpf 1990: 79–82; cf. table 79.4).

At this point, Stumpf faced the problem that, in the absolute chronology he established, the oldest manuscripts of Tocharian B date to the middle of the 7th century CE, while the linguistically definitely younger layer II type documents occur as early as the 7th century as well (Stumpf 1990: 154 f.). This overlap in written attestation contradicted Stumpf's concept of two chronologically layered language varieties. He solved the problem by assuming that the older layer I type language represents the standard that was coined during the heyday of the Kuchean kingdom in the 4th century CE and that contin-

Tab. 79.3: Principal Features Used by Stumpf (1990: 64–107)

Feature	Ex. Layer I	Ex. Layer II	
consonant cluster simplification	<i>nest</i>	<i>nes</i>	‘you are’ (2 nd sg.)
assimilation	<i>šarm</i>	<i>šaräm</i>	‘reason’
articulatory simplification	<i>parna</i>	<i>parra</i>	‘outside’
fricativization	<i>plāc</i> [-c]	<i>plās</i> [-ç]	‘speech’ (obl. sg.)
morphological/lexical variation	<i>šalype</i> [-p-]	<i>šalywe</i> [-β-]	‘salve’
	<i>skente</i>	<i>stare</i>	‘they are’
	<i>wes</i>	<i>wesäm</i>	‘us’

Tab. 79.4: Principal Features of Stumpf’s Sublayers

Feature	I A (≈ MQ)	I B	I C (≈ Š)	II (≈ Turfan)	
accented /ä/	<i>pälsko</i>	<i>pälsko</i>	<i>pälsko</i>	<i>pälsko</i>	‘thought’
unaccented /ā/	<i>šärmänmä</i>	<i>šärmänma</i>	<i>šärmánma</i>	<i>šärmánma</i>	‘reasons’ (nom./obl.)
*æu-diphthong	<i>ēurtse(/aurtse)</i>	<i>ēurtse/aurtse</i>	<i>aurtse</i>	<i>aurtse</i>	‘broad’

Tab. 79.5: Correspondences between Winter (1955) and Stumpf (1990) According to Stumpf (1990: 149)

Winter’s Dialects	Stumpf’s Layers
Western Dialect	I A & I B
Central Dialect	I C
Eastern Dialect	II

ued to be used as a literary language for more than 300 years (Stumpf 1990: 156). According to Stumpf, this standardized literary form underwent, in the course of time, successive influences from non-literary registers that continued to evolve. At the stage of layer II documents, the non-standard variety had finally replaced the old formal style (Stumpf 1990: 144–147, 157). Stumpf’s scenario thus combines historical and sociolinguistic arguments.

4. Palaeographical evidence

Until recently, the study of palaeography did not play a decisive role in settling the question. Although the editors of the Tocharian B texts stored in Berlin (Sieg and Siegling 1949, 1953) made sufficiently reliable remarks about the writing ductus of each fragment, their judgment had to remain somewhat vague since no investigation of Tocharian palaeography had been undertaken at the time of their editorial work. Even today, a detailed investigation is still lacking, and it is to be hoped that the ongoing digitalization will at last make available scans of all texts necessary for this task.

Tab. 79.6: Palaeographical Periodization According to Malzahn (2007: 259, 296–297)

Period	Text with Shabboleth Signs
Most Archaic	THT 1520
Middle Archaic	B 273 ff.
Early Common Archaic	THT 1661 ff.
Late Common Archaic	B 133 ff.

Winter, who did not yet have the opportunity to check the palaeographical labeling by Sieg and Siegling, already noticed (Winter 1955: 220–221) that ligature writing is much more frequent in texts of western origin (i.e. from Kucha), whereas *virāma* writing clearly predominates in standard Tocharian B texts from Šorčuq and in texts from the eastern sites around Turfan. This line of orthographic argumentation was integrated into his scenario by Stumpf as well (Stumpf 1990: 105–107). He expanded on it by adding the spelling of older *k_wu_i* vs. younger *kwri* ‘when(ever), if’ and older *šai* vs. younger *sey* ‘was’ (3 sg. imperfect). Incidentally, *šai* vs. younger *sey* (probably *še* [*< šai*] + *y*) belongs to the phonological level, while *ō_i* vs. *oy* (Winter 1955: 217) is an orthographical matter.

Winter must undoubtedly have been aware of the fact that scribal habits do not carry great weight in establishing synchronically defined dialectal areas. Indeed, the parallel development of specific scribal habits with linguistic change as, e.g., consonant cluster simplification, suggests that the orthographic norm developed alongside the language. In this light, it would seem that Stumpf’s chronological approach is superior to the geographical one proposed by Winter.

Additional evidence in support of Stumpf’s diachronic classification has been adduced by Malzahn in a pioneering palaeographical analysis of the Tocharian B texts, which was facilitated by the high resolution scans of the Berlin and London texts now publicly accessible on the world wide web (<http://titus.fkidg1.uni-frankfurt.de/texte/tocharic/thtcat.htm>; <http://idp.bl.uk>).

Malzahn (2007) provides a solid framework for the palaeographical development of the Tocharian B texts by taking as a starting point the oldest type of the local *brāhmī* script used in Sanskrit manuscripts, viz. the Early Turkestan *Brāhmī* (alphabet r), and tracing it down to the standard Northern Turkestan *Brāhmī* in a series of steps.

The palaeographical layering established by Malzahn fits nicely with the linguistic one by Stumpf, so that, at this point, it can hardly be doubted that the linguistic diversity observable in the Tocharian B texts has to be ascribed, first and foremost, to language development over time.

5. Absolute chronology

With regard to absolute chronology, carbon-14 datings of Tocharian manuscripts offer corroborating evidence for a chronological layering. Tamai (2005) has made available the results of a carbon-14 analysis of nine Tocharian B manuscripts (see Table 79.7). The implications of Tamai’s findings are, first of all, that Tocharian B is attested in

Tab. 79.7: C-14 Dating of Tocharian B documents (Tamai 2005)

MS	Location	C-14 Dating
B 333	Ming-oi Kizil (R)	CE 394–473
B 240	Ming-oi Kizil	CE 428–524
B 601	Kucha	CE 669–780
B 352	Qumtura	CE 679–776
B 178	Sāngim	CE 697–716
B 367	Murtuk	CE 737–773
B 483	Ming-oi Kizil	CE 770–888
B 296	Xocho	CE 1178–1255

writing roughly from 400 CE to 1200 CE, thus supplying additional reasons to assume that the diversity of Tocharian B is due at least partly to language change over time.

From this sample of dated texts a clear pattern emerges: The oldest manuscripts (B 333 [MQ], B 240 [MQ]) are from the western area around Kucha, i.e. Winter's western dialect, and all manuscripts from either central Šorčuq or eastern Turfan areas are much younger than the oldest manuscripts. Furthermore, most of the texts written in Winter's central dialect date from the second half of the 7th century onwards.

The c-14 datings also resolve some of the chronological difficulties Stumpf encountered: They show that not all of the manuscripts that display layer I A or I B features necessarily have to have been written (or copied) in the 7th century CE or later in a literary form that was coined three centuries earlier (Stumpf 1990: 156); at least some of them are originals written in the 4th or 5th century CE.

On the basis of the c-14 datings, Adams (2006: 386) convincingly proposes four chronological/geographical stages:

1. Early Tocharian B (ca. 400–600 CE) only in Kucha and environs
2. Middle Tocharian B (ca. 600–900 CE) everywhere in “attestation area”
3. Late Tocharian B (ca. 900–1100 CE) in Turfan and Kucha and environs
4. Very Late Tocharian B (ca. 1100–1300 CE) in Turfan

This scenario obviously relies on a small sample of dated texts, and not every text is necessarily contemporary with the medium it is written on (cf. the remarks on B 558 in Tamai 2005: 4). Additionally, an older literary style can be imitated by a skilled writer, thus obscuring the picture further. However, it can hardly be due to mere coincidence that the absolute dating squares perfectly with the results of the palaeographical analysis performed by Malzahn and the linguistic statements by Stumpf.

Concerning the younger characteristics of the most eastern variety of Tocharian B in Turfan, Stumpf suggests that they were introduced to the written language through strong interference from more informal styles. He reached this conclusion mostly because some features such as consonant cluster simplification are not confined to texts of eastern provenance, but also occur in documents containing profane texts from Kucha and Šorčuq. In Malzahn's assessment (Malzahn 2007: 289 f.), the peculiarities of the eastern texts are such that one should rather treat them as representing a dialect of their own. In the same breath, she admits that the formal language of the Turfan area population may have been influenced by more informal registers to a higher degree than elsewhere and that both scenarios need not be mutually exclusive.

It should be emphasized that the purely chronological scenario as designed by Adams harmonizes perfectly with Stumpf's interference concept. The assumption would be that low register features (and some peculiarities developed through contact with Tocharian A) had already entered the language of the normative literary standard by the time literary text production was being established in Turfan.

6. Summary

The diversity of the Tocharian B corpus may thus well be described in terms of chronology, implying an eastward migration of the Tocharian B scriptorial activities and a successive modernization of the literary standard over time.

Nevertheless, it is by no means ruled out that Tocharian B had different dialects, and a subset of their constitutive features may be identical to the features ascribed to different registers and chronological stages. However, on the basis of the available evidence and the research done so far, their existence is difficult to prove. It would be essential to find more morphological and syntactic evidence in addition to the phonological arguments predominating so far. A relevant syntactic phenomenon may be seen in the use of a PP with the preposition *y* 'in' (otherwise only known as the first member *y[n]/-* of compounds) instead of the usual inflected case form to express locational relationship in three bilingual (Tocharian B and Prākṛit) documents containing commercial records (Schmidt 2001). The contents of these texts present some difficulties, but the correspondences in table 79.8. (Schmidt 2001: 18 f.) are sufficiently clear.

Note that the Prākṛit version of THT 4062 seems to have taken over the PP construction from Tocharian.

Evidence for dialectal diversity in morphology is present in *śemare*, 3 plural preterite of the verb *kām* 'to come'. It is glossed, and thereby clearly labeled as unusual, in the graffito G-Su 35 as *kameṃ*, the standard form (cf. Widmer 2001: 186 f.). But again, it is difficult to ascertain whether these differences should be ascribed to dialectal diversity or to level and/or style of speech.

In summary, the facts suggest that the linguistic diversity in the Tocharian B documents is due to language change over time and to register phenomena, although the existence of dialects, however defined, cannot be ruled out.

Tab. 79.8: Preposition *y* in Three Bilinguals

MS	Toch.	Kucha-Prākṛit	
THT 4059	a1 <i>y lāpar</i> ₂	b1 <i>laparamiṃ</i>	'in Lapar' (<i>Lapār</i> , cf. Stein 1928: 830) SI P/141
THT 4059	a2 [y]/[aika] ₂ r	b2 <i>laisāraṃ[mī]</i>	'in Laikar'
THT 4062	a1 <i>y tākkaiṃ</i>	b1 f. i [t]ākkai	'in Tākkain'

7. References

- Adams, Douglas Q.
 2006 Some Implications of the Carbon-14 Dating of Tocharian Manuscripts. *Journal of Indo-European Studies* 34: 381–389.
- Krause, Wolfgang
 1952 *Westtocharische Grammatik*. Band I. *Das Verbum*. Heidelberg: Winter.
- Lane, George Sherman
 1966 On the Interrelationship of the Tocharian Dialects. In: Henrik Birnbaum and Jaan Puhvel (eds.), *Ancient Indo-European Dialects*. Berkeley: University of California Press, 213–233.
- Malzahn, Melanie
 2007 The Most Archaic Manuscripts of Tocharian B and the Varieties of the Tocharian B Language. In: Melanie Malzahn (ed.), *Instrumenta Tocharica*. Heidelberg: Winter, 255–297.
- Schmidt, Klaus Totila
 2001 Entzifferung verschollener Schriften und Sprachen, dargestellt am Beispiel der Kučā-Karoṣṭhī und des Kučā-Präkritis. *Göttinger Beiträge zur Asienforschung* 1: 7–35.
- Sieg, Emil and Wilhelm Siegling
 1949 *Tocharische Sprachreste, Sprache B, Heft 1: Die Udānalankāra- Fragmente, Text, Übersetzung und Glossar*. Göttingen: Vandenhoeck & Ruprecht.
- Sieg, Emil and Wilhelm Siegling
 1953 *Tocharische Sprachreste, Sprache B, Heft 2: Fragmente Nr. 71–633*, aus dem Nachlaß herausgegeben von Werner Thomas. Göttingen: Vandenhoeck & Ruprecht.
- Stein, M. Aurel
 1928 *Innermost Asia, Detailed Report of Explorations in Central Asia, Kan-su and Eastern Īrān*, Vol. II. Text. Oxford: Clarendon Press.
- Stumpf, Peter
 1990 *Die Erscheinungsformen des Westtocharischen. Ihre Beziehungen zueinander und ihre Funktionen*. (TIES Supplementary Series 2). Reykjavík: Málvísindastofnun Háskóla Íslands.
- Tamai, Tatsushi
 2005 Paläographische Untersuchung und ¹⁴C-Prüfung. Digitalisierung der chinesischen, tibetischen, syrischen und Sanskrit-Texte der Berliner Turfansammlung, 2. 6. 2005 – Berlin. <http://www.bbaw.de/bbaw/Forschung/Forschungsprojekte/turfanforschung/bilder/Tamai.pdf> [Last accessed 21 October 2006].
- Thomas, Werner
 1984 *Die Erforschung des Tocharischen (1960–1984)*. Stuttgart: Steiner.
- Widmer, Paul
 2001 Ererbtes und Geneuertes: Die Präteritumklasse VI des Tocharischen. *Studia Etymologica Cracoviensia* 6: 181–191.
- Winter, Werner
 1955 A Linguistic Classification of “Tocharian” B Texts. *Journal of the American Oriental Society* 75: 216–225.
- Winter, Werner
 1998 Tocharian. In: Anna Giacalone Ramat and Paolo Ramat (eds.), *The Indo-European Languages*. London: Routledge.

Paul Widmer, Zurich (Switzerland)

