The Indo-European Languages

Anna Giacalone Ramat, Paolo Ramat

Proto-Indo-European: Comparison and Reconstruction

Publication details
Calvert Watkins
Published online on: 20 Nov 1997

How to cite :- Calvert Watkins. 20 Nov 1997, Proto-Indo-European: Comparison and Reconstruction from: The Indo-European Languages Routledge
Accessed on: 23 Aug 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Proto-Indo-European: Comparison and Reconstruction

Calvert Watkins

Indo-European is the name given for geographic reasons to the large and well-defined genetic family including most of the languages of Europe, past and present, and extending across Iran and Afghanistan to the northern half of the Indian subcontinent.

A curious by-product of the age of colonialism and mercantilism was the introduction of Sanskrit in the eighteenth century to European intellectual and scholars long familiar with Latin and Greek, and with the European languages of culture, Romance, Germanic and Slavic. This new third member of the comparison, in addition to the two classical languages, revolutionized the perception of linguistic relationships. The English jurist and orientalist Sir William Jones (1746–94), speaking to the Asiatick Society in Calcutta on 2 February 1786, uttered his now famous pronouncement:

The Sanscrit language, whatever be its antiquity, is of a wonderful structure, more perfect than the Greek, more copious than the Latin, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of verbs and in the forms of grammar, than could possibly have been produced by accident; so strong indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists. There is a similar reason, though not quite so forcible, for supposing that both the Gothick and the Celtick, though blended with a very different idiom, had the same origin with the Sanscrit, and the old Persian might be added to the same family, if this were the place for discussing any question concerning the antiquities of Persia.

Jones was content with the assertion of the common origin of these languages; it remained for others to explore systematically the true nature of their linguistic relationship in the first decades of the nineteenth century. The new
science of comparative grammar was founded in 1816 by the young German Franz Bopp (1791–1867), after four years of study of oriental languages in Paris, with the appearance of his work *Ueber das Conjugationssystem der Sanskritsprache, in Vergleichung mit jenem der griechischen, lateinischen, persischen, und germanischen Sprachen, nebst Episoden des Ramajan und Mahabharat in genauen metrischen Uebersetzungen aus dem Originaltexte und einigen Abschnitten aus den Veda’s*. To Bopp and his near contemporary the Dane Rasmus Rask (1787–1832) must go the principal credit for first correctly seeing the relationships and systematically evaluating the similarities of the Indo-European family. The publication of the manuals of Lithuanian, with its strikingly archaic appearance, and the development of Slavic philology enabled Bopp to add the Baltic and Slavic branch to the growing Indo-European family. The Celtic languages, with their apparently idiosyncratic initial consonant mutations were the first real challenge to the emergent comparative method. That they too were Indo-European was asserted by Rask and others, and conclusively demonstrated by Bopp in 1838, with an elegant historical explanation of the troublesome initial mutations. Albanian was later added to the family, and in 1875 Hübschmann showed that Armenian was an independent branch rather than an aberrant form of Iranian. Representatives of all eight of these branches survive to the present day; Celtic is the only one threatened with extinction. Two new branches, both now extinct, were added early in the twentieth century, thanks to the discovery of new documents: Anatolian (including Hittite and others) and Tocharian.

The similarities among these languages, attested over nearly four millennia, require us to assume they are the continuation of a single prehistoric common language, spoken perhaps some seven thousand years ago, called *Indo-European* or *Proto-Indo-European*. The systematic investigation of the resemblances among these languages, by the comparative method, enables us to reconstruct the principal features of the grammar and lexicon of this proto-language. The reconstruction in turn (as with any proto-language) provides us with an initial stage starting from which we can describe the history of the individually attested daughter languages, which is the ultimate goal of historical linguistics.

**Principal Branches**

The principal branches of the Indo-European family are given below, in the order of their earliest historical attestation, focusing on the principal languages and the character of their documentation in the earlier periods.

Three branches are attested in the second millennium BC.

**Anatolian.** Excavations in central Turkey at Hattusas, the capital city of the Hittite Empire (near the village of Boğazköy, now Boğazkale), have unearthed extensive documents in *Hittite* written on clay tablets in a cuneiform script. Philologically we can distinguish *Old Hittite* (c. 1700–1500 BC), *Middle Hittite*
and Neo-Hittite (1350–1200 BC). Fragmentary remains of two other related languages are found in the same cuneiform Hittite sources: Palaic, in texts contemporary with Old Hittite and spoken to the north-west of Hattusas, and Cuneiform Luvian, in texts contemporary with Old and Middle Hittite and spoken over much of southern and western Anatolia (a form of Luvian in the north-west may have been the language of the Trojans). The preponderance of Luvian personal names and the loan words in neo-Hittite texts would indicate widespread use of the Luvian language in Hittite context as well. A very closely related dialect is Hieroglyphic Luvian (formerly called Hieroglyphic Hittite), written in an autochthonous pictographic syllabary, attested on seals and isolated rock inscriptions from Middle and neo-Hittite times, and from a number of monumental and other inscriptions from the region of northern Syria, 1000–750 BC. In classical times in south-western Anatolia we have sepulchral and administrative inscriptions (some quite extensive) in Lycian (5th to 4th century BC), and further north in the west, short inscriptions in Lydian (6th to 4th century BC), both written in epichoric alphabets. Lycian is clearly developed from a variety of Luvian; the other Anatolian languages cannot yet be organized in subgroups.

Archaeological excavations in Turkey are ongoing, and continue to yield new texts and fragments of texts in Hittite and the other cuneiform languages, in Hieroglyphic Luvian, and in the alphabetic languages of the first millennium. Each discovery brings new additions and new precision to our understanding of these languages, and the process may be expected to continue.

Indo-Iranian consists of two large and ancient groups, Indic (or Indo-Aryan) and Iranian, and a third, Nuristani (formerly called Kafiri, and sometimes improperly Dardic), attested from modern times in remotest North Afghanistan and neighbouring Pakistan and India. The earliest Indic consists of words and names in Anatolian texts (fifteenth century BC).

Extended Indic texts in Vedic Sanskrit begin with the RigVeda, whose earliest parts were probably composed in the Punjab in the second half of the second millennium BC, and continue through the other Vedas, Brāhmaṇas, Sūtras, etc. By c. 500 BC the language was codified in the grammar of Pāṇini as Classical Sanskrit, used to the present day as a learned literary language. From the fifth century BC on we have extensive Middle Indic documents (Pali, Prakrits; the very numerous Modern Indo-Aryan languages begin to be attested around AD 1000 (see Map 2.2).

Iranian, once spoken over vast stretches of south-eastern Eurasia, is first attested in Old (Gathic) Avestan, the hymns (gāthās) composed by Zarathustra, of uncertain date, but significantly older than the Younger Avestan of the middle of the first millennium BC. Old Persian is known from the monumental inscriptions of the Achaemenid kings of the sixth to fourth centuries BC; it is the ancestor of Middle (Pahlavi) and Modern Persian (Farsi). Several other Middle Iranian languages, known only from twentieth-
century discoveries are attested from Seleucid, Arsacid and Sassanian times, like Parthian, Sogdian and Saka, descended from the language of the ancient Scythians. By Islamic times we find the first attestations of some of the numerous Modern Iranian languages (see Figure 2.2).

Greek. First attested in documents is Mycenaean Greek, on the mainland and in Crete from the thirteenth century BC, written in the Linear B syllabary, and deciphered only in 1952. Greek was written on the island of Cyprus in the Cypriot syllabary, clearly of common origin with the Linear B syllabary. The oldest inscription is a single name from the eleventh century BC; the rest date from the eighth century BC to Hellenistic times. Alphabetic Greek is attested continuously from c. 800 BC, beginning with the Homeric poems and continuing through the Classical and Hellenistic (koiné) periods, to medieval (Byzantine) and modern times. A striking historical fact about Greek is that over this long period of virtually continuous documentation this branch maintained its identity as a single language (with different dialects at different times) down to the present day. Greek shares this feature with Armenian, and it is perhaps significant that some scholars posit an intermediary common Graeco-Armenian language.

Two substantial branches and several fragmentary languages are attested in the first millennium BC.

Italic. Old Latin and the closely related Faliscan are attested in short inscriptions from the sixth to the third century BC; from then on we have extensive documentation of Classical Latin. The main other Italic dialects South Picene, Oscan and Umbrian (together constituting the Sabellic group) are attested in inscriptions from the seventh or sixth to the first century BC. The affinity of the Venetic language to the Italic branch is controversial; the very recently reported Venetic inscriptions said to be found in southern Hungary seem to be imaginary. The Italic affinity of the various ancient languages of Sicily, Siculan (Sikel), Sicanian and Elymian is likewise uncertain due to the scarcity of evidence. Latin spread by conquest, replacing the other Italic dialects, and ultimately prevailed over large areas of southern and central Europe; the descendants of the spoken language are the medieval and modern Romance languages from Portugal to Romania (see Map 2.2), whose differentiation can be documented from the seventh and eighth century AD.

Celtic languages were in the first millennium BC spoken over large areas of Europe from the Iberian peninsula through southern Germany, the Po valley and Austria to the Danube plains, and as far as Galatia in central Anatolia. In our documentation we distinguish geographically between Continental Celtic (third century BC to third century AD, extinct; inscriptions in Gaulish, Celtiberian, Lepontic and others) and Insular Celtic, the languages spoken now or formerly in the British Isles. These form two groups, Goidelic (Gaelic) in Ireland and Brythonic (British) in Britain. The former comprise Irish (Primitive or Ogam AD 400–600, Old Irish AD 600–900,
Middle AD 900–1200, and Modern 1200 +); Scottish Gaelic (1200 +) and the extinct Manx. Brythonic includes Welsh (Old eighth to twelfth century AD; Middle twelfth to fifteenth century, Modern), Breton (Old and Modern) and the extinct Cornish.

Other fragmentarily attested languages, clearly Indo-European but of disputed origin, are Phrygian (western central Anatolia, short inscriptions eighth to fifth century BC, first to second century AD), and Messapic (the ‘heel’ of Italy, short inscriptions fifth to first century BC). Both are sometimes, but for geographic reasons only, grouped with the (poorly understood) ancient Balkan languages. (See Map 2.1.)

The remaining five branches of the family are first attested in the Christian era.

Germanic. The earliest extensive representative is Gothic (extinct), known from the Bible translation of the fourth century, which together with the other language remnants (Vandalic, Burgundian, etc.) form East Germanic. North Germanic is attested from a few runic inscriptions (third century AD +) and principally from Old Norse (ninth to sixteenth century) and the later West (Norwegian, Icelandic) and East (Danish, Swedish) Scandinavian languages. The principal earliest West Germanic monuments are in Old English (c. AD 700 +), Old Saxon (c. AD 850 +), and Old High German (c. AD 750 +), with the later medieval and modern forms of English, Frisian, Dutch, Low German and High German.

Armenian. Known from the Bible translation of the fifth century and subsequent literature is Classical Armenian, with its medieval and modern descendants spoken in several dialects, notably the Eastern (Russian) and the Western (Turkish and post-diaspora).

Tocharian. Two languages, now extinct, found in documents (mostly Buddhist translation literature) from the eastern (Toch. A) and western (Toch. B) parts of the Tarim basin in Chinese Turkestan (Xinjiang), dating from sixth to eighth century AD.

Balto-Slavic. The Slavic and Baltic languages appear to form a single subgroup within Indo-European, though some scholars would keep them apart. Slavic is first attested in the Bible translation of the ninth century in Old Church Slavonic. The dialect division into East Slavic (Russian, Ukrainian, Belorussian), West Slavic (Polish, Czech, Slovak, etc.), and South Slavic (Slovene, Serbo-Croatian, Macedonian, Bulgarian) is probably not much older than the middle of the first millennium AD. Old Church Slavonic is basically Common Slavic with some South Slavic dialect features.

Of the Baltic languages, the earliest attested is the extinct Old Prussian (fourteenth to seventeenth century) followed by the two flourishing East Baltic languages Lithuanian and Latvian (sixteenth century). Despite this late attestation Baltic and Slavic languages are in phonology and morphology remarkably conservative.

Albanian is known only from the fifteenth century on, in two dialects, a
Map 2.1 Indo-European languages in the first millennium BC
Used with permission of C. Watkins
northern (Geg) and a southern (Tosk). Some scholars are now inclined to see Albanian as the descendant of Daco-Mysian, a reconstructed language of the Eastern Balkan area.

**Subgrouping**
The ten or eleven attested branches can be arranged in larger subgroups, constituting dialectal divisions within Proto-Indo-European which go back to a period long before the speakers arrived in their historical areas of attestation. It is clear from the archaeological record that the Indo-European languages were brought into the areas they historically occupied in Europe and south-western Asia by a series of movements of people, small or large, over many millennia. A number of archaic features in morphology and phonology set Anatolian apart from the other branches, and indicate that it was the earliest to hive off. But Anatolian remains derivable from Proto-Indo-European, and periodic efforts to situate Anatolian as a sister language to Indo-European, both deriving from a putative ‘Indo-Hittite’, have not found a following.

On the basis of a number of shared innovations and other common features we may schematically display the dialectal affinities among the remaining ten branches in four quadrants; corresponding to the points of the compass (Figure 2.1). Each branch shares certain features with the nearest branch in the adjacent quadrants, the closest affinities of Anatolian are with the western group.

**The Homeland of the Indo-Europeans**
Many scholars hold that the Siberian steppe zone north and east of the Black Sea was, if not the ultimate ‘cradle’ of the Indo-Europeans, at least a significant staging area for movements to the west into the Balkans and beyond, to Anatolia, and to the south and then east into Iran and India, beginning in the mid-fifth millennium BC. This is what is termed by archaeologists the Kurgan culture, after the Russian word for its characteristic grave monuments or barrows; see Gimbutas (1980); Mallory (1989).

Still other areas have been proposed on sometimes skimpy evidence, ranging from central Europe and the Balkans to northern Europe, and even the circumpolar far North at the end of the last Ice Age. Most recently (Gamkrelidze and Ivanov 1995) an area in eastern Anatolia south of the Caucasus has been proposed, to account for various alleged contact phenomena with neighbouring Semitic, Kartvelian and other language families. This area would have been the starting point for a circular northward movement east of the Caspian Sea, then turning westward into the Kurgan staging area. But the linguistic evidence for the cross-family contact remains controversial. It would appear that the break-up of the Indo-European dialects had already
Map 2.2  Indo-European languages in modern times
Used with permission of C. Watkins
taken place in the transpontic Kurgan or a similar staging area. Proto-IndoEuropean was doubtless spoken over a fairly wide geographical area in Eurasia, and what we reconstruct and term Proto-Indo-European almost certainly refers to a language spoken over a temporal continuum of more than one millennium.

A Grammatical Sketch of Reconstructed Proto-Indo-European

Phonology
In the Indo-European phonological system we distinguish four groups of sounds on the basis of clustering and other morphophonemic rules. These are, in ascending order of syllabicity, *obstruents* (stops and sibilants) (symbol T); *laryngeals* (symbol H); *resonants* (symbol R); and *vowels* (symbol E).

*Obstruents*
A fairly rich system of stop consonants has been traditionally reconstructed, with five points of articulation (labial, apico-dental, palatal, velar and labiovelar) and three manners of articulation (unvoiced, voiced and voiced aspirated):
For series III, equivalent notations are \(b^h\ d^h\), etc. For the palatal stops, equivalent notations are \(\kappa'\ g'\), etc. For the question of the three tectals (‘gutturals’) see pp. 38f. below. These are illustrated by the following correspondence charts.

Compare


*\(k\kappa\) - ‘excrement’ in Ved. śákr, gen. śáknás, Gk. kópros; *(h)έku-ο- ‘horse’ in HLuv. azúwa-, Ved. ásvas, Arm. aspō, Myc. Gk. iqo (Gk hippos), Lat. equus. Goth. aihwa-, OIr. ech (Gaul. eпо-, dial. equus), Lith. ašvā ‘mare’, Toch. B yakwe


The voiced labial \(b\) is disproportionately rare, and correspondence sets are typically limited to few languages, like Hitt. lip- [lib-] ‘lick’, Luv. lapan-[laban-] ‘saltlick’, Eng. lip, Lat. labium, IE *leb- (probably imitative); Gk steibō ‘trample’, Arm. stipem ‘press’. Such equations may be posterior to the common period; it is possible, but not demonstrable, that PIE *m reflected an earlier merger of and *m. For the others compare:

IE *dek\(m\) 10: Skt dāśa, Avest. dasa, Gk déka, Lat. decem, Osc. deket-, Goth. taihun, Arm. tasn, OIr. deich n-, OCS desetb, Lith. dėšimts, Toch. B šak

IE *гenu-/гоνу- ‘knee’: Hitt. genu, Skt jānu, Avest. zānu, Gk gónu, Lat. genu, Goth. kniu, Arm. cunr, OIr. glún, Toch. B keni-, Alb. gju

IE *g\(w\)ow- ‘bovine’: HLuv. waw(a)~ (< *g\(w\)-), Skt gaus, Avest. gaoš, Gk boūs, Lat. bōs (Sabine dialect borrowing); OEng. cū, Arm. kov, Fr. bó, OCS *gov-edo.

Compare:

IE *b\(h\)er\(g\)h- ‘high’: Hitt. park- [parg-] ‘be high’, Skt brʰant- ‘lofty’, Avest. br̥ZX- and br̥XZant- ‘high’, OLat. ethnic name Forcti, Gmc ethnic name Burgund-, Arm. barjr ‘high’, Ir. name Brigit, Toch. B pärk- ‘rise’ parkare ‘long’.

IE *d\(h\)(e)uh₂- ‘smoke’: Hitt. tuh₂- [tux-] ‘smoke’, Skt dhūmás, Gk thūmós ‘spirit’, Lat. fūmus, OCS dymu, Lith. dūmai.
Table 2.1 Unvoiced stops

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Osca</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Irish</th>
<th>Old Church</th>
<th>Slavic</th>
<th>Lithuanian</th>
<th>Tocharian</th>
</tr>
</thead>
<tbody>
<tr>
<td>*p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>f(b)</td>
<td>h(w)</td>
<td>---</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p(п)</td>
</tr>
<tr>
<td>*t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>p(d)</td>
<td>t(c)</td>
<td>t(th)</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t(c)</td>
</tr>
<tr>
<td>*k</td>
<td>k</td>
<td>k</td>
<td>s</td>
<td>k</td>
<td>c</td>
<td>k</td>
<td>h(g)</td>
<td>s</td>
<td>c(ch)</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>k(š)</td>
</tr>
<tr>
<td>*kw</td>
<td>ku</td>
<td>k(c)</td>
<td>k(c)</td>
<td>k(c)</td>
<td>k</td>
<td>p(t)</td>
<td>qu</td>
<td>p</td>
<td>hw</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>k</td>
</tr>
</tbody>
</table>

1. c before *e, i, j in Indo-Iran. by the Law of Palatals.
2. ĉ before *e, i, j in Slavic, c before *oi, ai (which became ě, i).
3. t before *e, i in most dialects. Myc. Q = [kʷ].
6. h-initially, w between vowels.
7. Inner-Toch. palatalization.
8. Irish unvoiced t etc. only initially.
9. Ogam Ir. Q = [kʷ].
11. Luv. z < *k, k < *k, ku < *kʷ.
Table 2.2  Voiced stops

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Oscan</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Irish</th>
<th>Old Church Slavic</th>
<th>Lithuanian</th>
<th>Tocharian</th>
</tr>
</thead>
<tbody>
<tr>
<td>*b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>p</td>
<td>p</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>p</td>
</tr>
<tr>
<td>*d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>t</td>
<td>t</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>t</td>
</tr>
<tr>
<td>*g</td>
<td>g</td>
<td>j</td>
<td>z</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>k</td>
<td>c</td>
<td>g</td>
<td>z</td>
<td>z</td>
<td>k</td>
</tr>
<tr>
<td>*gw</td>
<td>gu</td>
<td>g(j)</td>
<td>g(j)</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>k</td>
<td>k</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>k</td>
</tr>
</tbody>
</table>

2. OCS  ż by Slavic Law of Palatals, dz like c above; see Chapter 15, p. 429.
3. d before e in most dialects. Myc. ‘q’ = [gʷ].
4. Lat. gu after n.
5. Word initially these are stops [b d g], word internally they are voiced continuants [fr g it].
6. With palatalized variants as with unvoiced stops.
7. Probably devoiced word-initially.
Table 2.3 Voiced aspirated stops

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite 8</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin 4</th>
<th>Oscan</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Irish 5</th>
<th>Old Church Slavic</th>
<th>Lithuanian</th>
<th>Tocharian 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>*bʰ</td>
<td>b</td>
<td>bh</td>
<td>b</td>
<td>ph</td>
<td>f(b)</td>
<td>f</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>*dʰ</td>
<td>d</td>
<td>dh</td>
<td>d</td>
<td>th</td>
<td>f(d)</td>
<td>f</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>*gʰ</td>
<td>g</td>
<td>h</td>
<td>z</td>
<td>kh</td>
<td>h(g)</td>
<td>h</td>
<td>g</td>
<td>j</td>
<td>g</td>
<td>z</td>
<td>ž</td>
<td>k</td>
</tr>
<tr>
<td>*gʰw</td>
<td>gu</td>
<td>gh(h) 1</td>
<td>g(j) 1</td>
<td>kh</td>
<td>h</td>
<td>h</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>g</td>
<td>k</td>
</tr>
</tbody>
</table>

1 Skt h, Avest. j by the Law of Palatals.
2 ċ by the Slavic Law of Palatals, dz like c.
3 th before e, i in most dialects. Myc. Ҁ = [kʰw].
4 Sounds in parentheses are word-internal.
5 As for voiced stops.
6 b- is controversial.
7 As for unvoiced stops.
8 Probably devoiced word-initially.

The Glottalic Theory
The typological rarity (‘unnaturalness’) of the three series (I) unvoiced: (II) voiced: (III) voiced aspirated (murmured, breathy voiced), together with the rarity of b, has led a number of scholars recently to reinterpret the traditional voiced series (II) as an unvoiced ejective (glottalized) one, and the unvoiced (I) as aspirated (with unaspirated allophones):

\[
\begin{align*}
\text{I} & : p^h, t^h, k^h, k^{wh} \\
\text{II} & : p', t', k', k^{wh} \\
\text{III} & : b^h, d^h, ģ^h, g^h, g^{wh}
\end{align*}
\]


While the new ‘typological interpretation’ may account for some problematic features, it introduces considerable complexities (and ‘unnaturalness’) elsewhere in the system, and the question still awaits a generally acceptable solution.

The traditional three series of stops (see I—III of chart on p. 34 above) are nowhere preserved intact, but everywhere in the dialects somehow altered. We find the following developments:

1 development of a new series IV, unvoiced aspirates (Indic, pre-Iranian, perhaps pre-Greek and pre-Armenian);
2 series III is devoiced (Greek, possibly pre-Italic);
3 de-aspiration, merger of series II and III as II (Anatolian, Iranian, Balto-Slavic, Phrygian, Celtic, Albanian, Messapic);
4 merger of series I, II, III as I (Tocharian);
5 spirantization (Iranian I, IV → [+continuant], Italian III → [+continuant]);
6 ‘mutation’ (Germanic I, III → ‘+continuant, a voiced], II → [−voiced]; Armenian I → [+aspirated], II → [−voiced]).

A ‘glottalic’ restatement (see I—III of chart above) would simplify item (6) but considerably complicate (1) through (5).

Indo-European forms in this chapter, as in most current handbooks, appear in their traditional shape.

The Tectals
Nearly all the languages show good evidence only for a front/back contrast in the tectals or dorsals (‘gutturals’), k : kʰ in the western and ģ : k in the eastern dialects: Lat. centum ‘100’: quattuor ‘4’, Lith. šimtas ‘100’: keturi ‘4’. Such dialect variation can be typologically paralleled (e.g. in the Athabascan language Kayukon of Alaska). Some scholars therefore assume only four
points of articulation in the PIE stops, \( p, t, k, kw \), despite the classical three-way contrast of reflexes in identical environments, for example, initially before the liquid \( r \) in Skt śrad-: Lat. crēd- (IE *kr-); Skt krāvis: Gk krēas (IE *kr-); Skt (ā-vi-)krītas: Gk (a-)prīatos (IE *kwr-). New evidence from Luvian as well as other arguments supports the traditional reconstruction of an original three-way contrast \( k : kw : k' \). Compare Luv. za-/zi- ‘this’ < *ko-/ki- (Lith. sīs, Arm. sī); zīyar(i) ‘lies’ < *kej-or (Ved. sāye < *kej-aj); kīša(i) ‘com’ < *kes- (OCS česoq), kars- ‘cut’ (Gk a-kerse-kōmēs ‘with unshorn locks’); kui- ‘who’ < *kwi- (Lat. quis), -kuva < *kw'e (Lat. que). See Melchert (1987).

Most western dialects merged \( k, k > k \) (the ‘centum’ solution); the eastern group merged \( k, kw > k \) (the ‘satem’ [Avest. satm 100] solution). In the latter case, the front palatals further developed to affricates, then sibilants (see Chapter 4, p. 105).

**Sibilants**
Beside the wealth of stops, Proto-Indo-European had only a single continuant, the sibilant \( s \), allophonically voiced to [z] when followed by distinctively voiced segments, that is, voiced stops. In several languages (independently) \( s \) is weakened to [h] before vowel: Old Irish, Greek, Armenian, Brythonic.

IE *sekʷ- ‘follow’: Lat. sequitur, Skt sācate, Avest. haicaitē, Gk hépetai ‘follows’
IE *misēh- ‘recompense’: Gk misthós, Avest. mižda-, Goth. mizdō, OCS mźzda
IE *₃sēh- ‘star’: Hitt. hašter-, Gk astér, Arm. astt, Avest. star-, OIr. ser

**Cluster Phenomena**
Certain phonological effects are caused by the contact of obstruents in a string. These are context-sensitive rules of IE date. A selection is given below.

**Voicing Assimilation**
Non-aspirated obstruents assimilate in voicing to a following stop. With the loss of the vowel of IE *ped- ‘foot’, the resultant *pd- appears as *bd-: Avest. fra-bd-a- ‘forefoot’, Gk epì-bd-ai ‘day after [‘at the foot of’] a festival’. (The voicing of \( s \) above is simply an instance of this rule.) IE *nigʷ- ‘wash’ + -to- > *nikwtō- ‘washed’: Skt nikta-, Gk ἀνιπτό ‘unwashed’.

**Bartholomae’s Law**
If the first member of an obstruent cluster is (voiced) aspirated, the assimilation is progressive: \( D^h + T \rightarrow DD^h \). This rule is most clearly seen in early Indo-Iranian: Skt budh- ‘awaken’ + -tā- > buddhā; Indo-Iranian *augh- ‘speak solemnly, aver’ +3 sg. mid. -ta → Gāthic Avest. aogdā ‘he spoke’. The rule also applied to \( s \): *awgh- + 2 sg. mid. -sa → Indo-Iran. *augzʰa, Gāthic Avest. aogxā. Outside of Indic the effects of Bartholomae’s Law have been largely undone by paradigmatic analogy: Young Avest. aoxta ‘speak’ like Gk eïkto, leaving only isolated relics. But that it was an IE rule...
is shown by the suffix doublets like *-tro-l *-dho-ro-, *tlo-l *-dho-lo- (Lat. -trum, -brum, -colum, -bulum), whose existence is scarcely to be explained otherwise.

Dental plus Dental
In IE clusters of root-final dental stop followed by morpheme-initial dental stop (like the common third-person singular ending *-ti, -t, the verbal adjective *-to-, agent suffix *-ter-), an s was inserted between the two dentals. From the root *h₁ed- 'eat' + 2 pl. ending *-te(-), Hittite e-ez-te-en 'eat!' [etsten] < *edsten (voicing assimilation). Common Indo-Iranian had also *stt- (and *-zdth-) where the second dental was voiced), whence Indic -tt- (and -dd-), but Avest. -st- (and zd-): Ved. attana 'eat!'. In other branches the result is either -st- (Gk, Balto-Slavic) or -ss- (Italic, Celtic, Germanic); IE *gwhedh-to- 'in-exöräbilis' (*gwhedh- 'entreat, pray') in Lat. infestus, YAV ajasta- (remade from *ajazda-), with *-dh-t- > *zdth- in both forms according to Bartholomae’s Law (see p. 39).

Dental plus Tectal (‘Thorn Clusters’)
In a small but important group of words we observe the correspondences Skt ks : Gk kt (rkṣa- 'bear': ārktos). Skt ks : Gk khθ (ksám- 'earth': khθn); Skt ks : Gk phθ (ā-kṣi-ta- 'imperishable' ā-phθitos). The evidence of Hitt. tekan [teqgan], gen. taknaš 'earth', hart(a)kkas 'bear(?)', as well as Toch. A tkam, Toch. B kem made it clear that the correspondences result from metathesized tautosyllabic clusters of dental plus tectal: *tk, *d(h)g, *d(h)gwh, with differing reflexes dependent on phonetic context.

Geminates
The sequence s+s at morpheme boundary was simplified to single s: root *h₁es- 'be' + 2 sg. *-si yielded *h₁esi, whence Skt āsi, Gathic Avest. ahi, Gk eis(s), Lith. esi. Compare the dental plus dental rule above, which also eliminated geminates. Geminated or long consonants in general were avoided in IE ‘ordinary’ language, while they played an important role in the onomastic system of hypocoristics (cf. Ger. Fried-rich [friːd-] but Fritz [*frit-]), appellatives like Gk átta, Goth. atta, Hitt. attaš, ‘poppa’, and expressive or iconic deformations in poetic language like Gk ópphin 'snake!', synnékhés ‘continuously’.

Laryngeals
The term laryngeal is applied (loosely) to a set of h-like sounds in the proto-language, of not entirely specifiable phonetic value. It is probable that they belonged to the natural class of ‘gutturals’ now recognized by phonologists working with Semitic languages. The PIE laryngeals are here noted h₁ h₂ h₃; equivalent notations are ɾ₁ ɾ₂ ɾ₃, E A O, ɾ x x̆, and other symbols as well. (Some scholars assume only one laryngeal, others four or more. See Kuryłowicz and Mayrhofer 1.121 ff. and Eichner 1988.)
The system of the IE laryngeals was in large part worked out by F. de Saussure in his brilliant *Mémoire* of 1878, from internal structural evidence, corrected by the work of H. Møller (1879), A. Cuny (1912) and J. Kuryłowicz (1927), also on structural grounds. The last further identified Saussure’s A = a, by then recognized as a consonantal element, with Hittite h₂. While this material confirmation of the theory was as welcome as it was dramatic, evidence for the laryngeals comes from a variety of sources in many languages, and, for example, despite the absence of a single phonetic continuator, the laryngeals were probably better preserved in Greek than in Hittite.

The basic rules affecting the laryngeals are two: (1) H-coloration, and (2) H-loss.

(1) H-coloration. Laryngeals h₂ and h₃ had the property of colouring (lowering, backing or rounding) an adjacent vowel e (and only that vowel) to a and o respectively; h₁ had no colouring effect. The results were

\[
\begin{align*}
    h₁e & > h₁e \\
    h₂e & > h₂a \\
    h₃e & > h₃o
\end{align*}
\]

The new vowels a and o merged with the previous /a/ and /o/, and increased their frequency and distribution. This change took place already in the proto-language.

(2) H-loss. At a later period, beginning in the proto-language and continuing into the dialects, the laryngeals tended to become lost, with different phonological consequences in different environments and according to dialect. Only a selection of these can be illustrated. When preceded by a short vowel (e, a, o and the high vowels i and u) and followed by a non-syllabic, the loss of the laryngeal resulted in the corresponding long vowel, by compensatory lengthening. Using the cover symbol H for laryngeal and C for consonant,

\[
\begin{align*}
    Hc & > c \\
    Ha & > a \\
    Ho & > o \\
    Hi & > i \\
    Hu & > u
\end{align*}
\]

\[
\begin{align*}
    eHC & > ēC \\
    aHC & > āC \\
    oHC & > ōC \\
    iHC & > ĩC \\
    uHC & > ũC
\end{align*}
\]

Compare IE *peh₂- ‘protect’, extended form *peh₂-s- *pah₂-s- in Lat. pās-tor ‘herdsman’, Hitt. pahš-. The o-grade of the same root is not coloured, but can be lengthened: *pōh₂ju- in Gk pōu ‘flock (of sheep)’ (< *‘protected’), *poh₂yi- in Ved. Skt pōyā- ‘protector’. When H is lost before a syllabic, compensatory lengthening does not take place: *poh₂i-mén- in Gk poimēν ‘herdsman’, Lith. piemuō, gen. piemeĩš ‘shepherd boy’.
H-loss in the dialects as stated above applied only in cases where H was adjacent to a vowel. Where H was not adjacent to a vowel and not in word-initial position, it became syllabic, with different vowel reflexes in the dialects: Skt i, most others a, but Greek keeps the three distinct as e, a, o. These are the reflexes of the older reconstructed ‘schwa’ o of traditional pre-laryngeal handbooks. Thus:

*ph₂tér- ‘father’ in Skt pitár-, Gk patér, Lat. pater, Goth. fadar, Olr. athair, Arm. haya
*dh₂-nó-/l-tó- ‘given’ in Skt diná-, Gk dotós, Lat. datus
*dʰh₁s- in Lat. fānum ‘piece of consecrated ground, temple’ < *fús-nom, Gk thēs-phatos ‘destined by a god, bounded, fatal’, the-ós ‘god’, beside e-grade of the following
*dʰeh₁s- in Osc. fišnū ‘piece of consecrated ground, temple’ < *fēs-nā, Lat. fēs-tus (dies) ‘(religious) holiday’, Arm. di-k’ ‘(pagan) gods’.

/h₁/ for which a phonetic value [h] or [ʃ] has been suggested, has no colouring effect on an adjacent vowel. Its former presence can be inferred, however, from ablaut and vocalization phenomena (pp. 51f) and lengthening effects. Thus:

*h₁é-s-ti ‘is’ in Hitt, e-es-zi/’esši/ Gk étis ‘there is’, Skt ásti, Lat. est
*h₁s-énti ‘are’ in Myc. Gk e-e-si/lehensi/, Skt sánti, Lat. sunt, Goth. sind
negated participle *h₁s-ont- in Skt ásat- ‘false’.

/h₂/, for which a velar [x] or pharyngeal [h] have been suggested, colours e to a, but has no colouring effect on o or long vowels. It is directly attested in Hitt. and Luv. h-, -h(h)- and Lyc. x-. Compare uncoloured *h₂ow-i- ‘sheep’ in Luv. ḫawīš, Lyc. xawā- (acc.), Lat. ovis, Gk ᴧ(w)is, beside coloured *h₂ent- > h₂ant- ‘front, forehead’ in Hitt. ḫanza ‘forehead’, ḫantezzi(ya)s ‘first’, Lyc. xītawati- ‘ruler, king’, Gk antí ‘before’, Lat. ante. Initially before consonant, *h₂-nér- ‘man, hero’ with initial vowel in Gk anér, Arm. ayr and Phrygian anar, lengthening in Skt sūnāra- ‘handsome’ < *su-h₂ner-o-, and loss in Osc. niīr ‘magistrate’, Lat. Ner-ō, Skt nāram (acc.), Avest. nā, Olr. nert ‘strength’.

/h₃/, for which a labialized velar [xʷ], [ɣʷ] or (labialized?) voiced pharyngeal [G], [Gʷ] has been suggested, colours e to o, but has no colouring effect on a or long vowels. It is directly attested in Hitt. h- corresponding to Lyc. zero: IE *h₃ep- > *h₃op- in Hitt. ḫāppar ‘transaction, deal’, ḫāppinānt- ‘wealthy’, Lyc. epiρijei- ‘sell’, Lat. ops ‘power, wealth’, op-ulent- ‘wealthy’. IE *gneh₂ > *gnoh₃ ‘know’ in Lat. (g)nō-scō, but uncoloured long vowel (lengthened grade, pp. 52f.) in *gnēh₃-s- Hitt. ganes-zi ‘recognizes’, Toch. A knas-āst ‘du kennst dich aus’.

Under the heading of cluster phenomena we may note some important
laryngeal reflexes; only a selection is possible.

**RH.** Clusters of syllabic liquid or nasal plus laryngeal when followed by a non-syllabic have special reflexes. These are the traditional ‘long sonants’ ḫ etc. as reconstructed in pre-laryngealists' handbooks. The basic correspondences in the branches that preserve these distinct are

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Sanskrit</th>
<th>Italic, Celtic</th>
<th>Lithuanian (under the accent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rH(C)</td>
<td>ḫ/̛r</td>
<td>rā</td>
<td>ḫr</td>
</tr>
<tr>
<td>lH(C)</td>
<td>̛l/̛r</td>
<td>là</td>
<td>̛l</td>
</tr>
<tr>
<td>mH(C)</td>
<td>ān</td>
<td>mā(?l)</td>
<td>āmen</td>
</tr>
<tr>
<td>nH(C)</td>
<td>ā</td>
<td>nā</td>
<td>nā</td>
</tr>
</tbody>
</table>

[Here as on page 42 Greek has a threefold reflex, with further complications (type -ne- / -ene-, -nā- / -ana-).] The essence of the system is the preservation of a bimoric reflex of the bimoric sequence RH, typically by transfer of one mora to the vowel. Lith. īr etc. with ‘acute’ (falling) intonation reflects an earlier *īr < *iɾ < *r(C); contrast īr with ‘circumflex’ (rising) intonation <*īr < *r(C), consisting of one mora (pp. 44f. below).

IE *ph₁-h₁-nō- ‘full’: in Skt pūrnās, Olr. lūn, Lith. pilnas
IE *gn̄h₁-tő- ‘born’: in Skt jātās, Avest. zātō, Lat. (g)nātus, Gaul. personal name Cintu-gnātus ‘First-born’.

Germanic shows no distinction between ṛ and ḫ: Goth. fulls < *fulnaz and -kunds < *-kundaz.

**Metathesis.** The sequences CHi, CHu as zero grades (p. 52) of CeHi, CeHu tend to undergo metathesis:

IE *peh₂-wr ‘fire’: in Hitt. paḥhur, zero grade *ph₂ur → *puh₂r in Gk pyr, Umbr. pir, beside unmetathesized *ph₂ur- before vowel in Gk gen. pyrōs, Umbr. abl. pure
IE *seh₂-wel > *sah₂-wel ‘sun’ in Dor. Gk aēlios, Goth. sauil, beside metathesized *suḥ₂-el (*suwel) in Ved. sūvar, antevocalic gen. sūrās = Avest. huuaro, hūrō.

Saussure was the first to note regular cases in Greek where in roots of the structure CeRH the laryngeal disappeared in the o-grade: péra-ssa ‘I sold’ but pór-nē ‘prostitute’, -bremé-tēs ‘thunderer’ but bron-tē ‘thunder’; tór-nos ‘carpenter’s tool for drawing a circle’ and tór-mos ‘hole, mortice, tenon’ but tēre-tron ‘drill’. The rule appears to operate in Hittite as well (tar-mas ‘peg, nail’ beside Gk tór-mos), and appears regular in its mirror-image as well: HReC- but RoC-.

In word-final position the treatment of the laryngeals was probably a function of sentence sandhi (Chapter 4, p. 107) rules: before a word
beginning with a non-syllabic we expect lengthening (-EH%C- → -E%C-),
before a word beginning with a syllabic we expect just loss (-EH%E- →
-E%E-) after colouring. The result would be doublets of final -E/-E, or -EH/-E,
of which we find different reflexes in the dialects. Compare the Luvian
abstract suffix -ah-(id-) with the feminine abstract suffix *-ä-
or its reflex in most of the dialects (nom. sg. CGk -ā, Osc. -ū, Umbr. -u, -a, -o, Slav.
-a) from *-ah2 beside the vocatives (Gk nýmph-a, toxóta, Umbr. -a, Slav.
-o) from *-a.

Resonants
Under the term ‘resonants’ or ‘sonants’ (cover symbol R) are grouped the two
IE nasals m and n, the two liquids r and l, and the semi-vowels (glides) y (equivalent notation ï, j) and w (equivalent notation u).

Syllabicity
These phones in Proto-Indo-European could function either as non-syllabics
(‘consonantal’ R), as above, or as syllabics (‘vocalic’ R), noted µ, η, τ, δ, and
the high vowels i, u, according to context. The basic rule is that the resonant
is non-syllabic when adjacent to a vowel proper (E, see p. 46), and otherwise
becomes syllabic before a non-syllabic, iteratively from right to left. Thus in
the word for ‘dog’ /£/ /w/ /n/ before syllabic, gen. sg. -os appears as *kunós
(Gk kynós, Ved. śūnas), but before non-syllabic, instrumental plural *-bhís,
as *kwnbhís (Ved. śvábhís). Exceptions are the initial labials w and m in
clusters like *wr-, *wy, *wj-, *mr-, *ml-, *mn- (Gk mnā- ‘remember’, Skt
ā-mnā- ‘remember and transmit’, Luv. manā- [mna:] or [mna:]-, ‘see’), as
well as the zero grade of nasal infix verbs (page 57), for reasons of
paradigmatic unity: *h₃r-ne-g-ti ‘destroys’, 3 pl. *h₃r-n-g-enti (not **/z₃r-
-g-enti), Hitt. ḫarnikzi ḫarninkanzi, Ofr. orgid ‘destroys, kills’.

Reflexes
The PIE complementary distribution of non-syllabic and syllabic variants of
resonants is broken up in all IE dialects; these show divergent reflexes of R
and R. The consonantal resonants are largely preserved intact, as are the
syllabic vowels u and i (see pp. 45f.), while the syllabic liquids and nasals are
sooner or later replaced in all dialects, typically by sequences of vowel and
the consonantal liquid or nasal.

Examples of PIE consonantal liquids and nasals are such widely attested
roots as *men- ‘think’, *nem- ‘distribute, bestow’, *mer- ‘die’, *mel-it-
‘honey’. Indic and Iranian, in common with a widespread areal feature in
Asia, tend to merge the two liquids into one: Iran. r, Ved. r but dialectally l.
Indo-European distinguished /ml/ and /nl/ also in absolute final position (n. sg.
*sem ‘1’: en ‘in’, thematic acc. sg. -om : *mon-stem suffixless loc. sg. -mon,
etc.), but some of the languages have merged the two to /n/ : Anatolian, Greek,
Armenian, Celtic (except Celtiberian and Lepontic).
### Syllabic Nasals: Correspondences

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Old Irish</th>
<th>Old Church</th>
<th>Lithuanian</th>
<th>Tocharian</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>an</td>
<td>a</td>
<td>a</td>
<td>em</td>
<td>em</td>
<td>um</td>
<td>am</td>
<td>*am</td>
<td>e(u)</td>
<td>iñ/um</td>
<td>ån</td>
</tr>
<tr>
<td>n</td>
<td>an</td>
<td>a(n)₁</td>
<td>a(n)₁</td>
<td>en</td>
<td>en</td>
<td>un</td>
<td>an</td>
<td>*an</td>
<td>e(u)</td>
<td>iñ/üñ</td>
<td>ån</td>
</tr>
</tbody>
</table>

1. an before vowels.
2. The Common Celtic reflexes are given, which undergo complex changes in Irish.
3. With ‘circumflex’ (rising) intonation when accented.
4. un word-finally.

IE *pód-m* ‘foot (acc. sg.)’: in Hitt. pādan, Gk pód-a, Lat. ped-em, Arm. otn
IE *g*₇*mské*: in Ved. gāccha, Gothic Avest. jasa, Gk bás(k(e) ‘go (ipv.)’)
IE *ŋ*- ‘not-’: in Ved. á-mṛṭa-, Gk á-mbrotos, Lat. im-mortális ‘immortal’
an-canañawf ‘unknown’;
Skt an-ukta- ‘unsayable’, MIr. an-ocht ‘a metrical fault’

### Syllabic Liquids: Correspondences

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Old Irish</th>
<th>Old Church</th>
<th>Lithuanian</th>
<th>Tocharian</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>ar</td>
<td>r̥</td>
<td>r̥</td>
<td>ar</td>
<td>or</td>
<td>awr</td>
<td>ar</td>
<td>ri(ar)₁</td>
<td>ru</td>
<td>ir, ur</td>
<td>är</td>
</tr>
<tr>
<td>l</td>
<td>al</td>
<td>r̥</td>
<td>r̥</td>
<td>la</td>
<td>al</td>
<td>ul</td>
<td>ol</td>
<td>ul</td>
<td>al</td>
<td></td>
<td>al</td>
</tr>
</tbody>
</table>

1. The basic reflex is Ri; aR is limited to a few contexts (cf. Chapter 12, p. 357).
2. Myc. and Aeolic Ro, oR.

In Indo-Iranian the syllabic liquid is preserved phonetically, but it is no longer in complementary distribution with the non-syllabic r. In Greek and in Hittite there is some evidence that the syllabic liquids were similarly preserved until just before the historical period: the Homeric scansion āndrōtētā ‘manhood’ reflects *anṭātā in a poetic formula of pre-Mycenean date.

### Non-syllabic Semi-vowels: Correspondences (Word-initial)

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Old Irish</th>
<th>Old Church</th>
<th>Lithuanian</th>
<th>Tocharian</th>
</tr>
</thead>
<tbody>
<tr>
<td>j</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>z₁</td>
<td>i</td>
<td>i</td>
<td>j</td>
<td>?</td>
<td>0</td>
<td>j</td>
<td>y</td>
</tr>
<tr>
<td>w</td>
<td>w</td>
<td>v</td>
<td>v</td>
<td>(w)₀²</td>
<td>u</td>
<td>v</td>
<td>w</td>
<td>g, v</td>
<td>f</td>
<td>v</td>
<td>w</td>
</tr>
</tbody>
</table>

1. Gk h- reflects an initial cluster *Hi-.
2. w in Mycenean and many dialects ('digamma').

IE *jugóμ* ‘yoke’: Hitt. yugan, Skt yygám, Gk zygón, Lat. iugum, Goth. juk, OCS igo (< *jégo*)
IE *wih-rōs 'strong; man(ly) ': Avest. vīrō, Gk (w)īros (personal name), Lat. vir, Goth. wairs, OIr. fer, Lith. vyras, Toch. A wir 'young, vigorous'.

**Sievers' Law**

Under the term Sievers' Law are understood various manifestations in various languages, but primarily in Vedic, of a tendency to a distribution -yE-/-wE- after light sequence, -ijE-/-uwE- after heavy sequence: Ved. sūrya- 'sun', kāvya 'of the poet-seer' to be read and metrically scanned sūrīya-, kāvīya-. Even in Vedic the longer sequence is regular after heavy, but not infrequent after light sequence as well, and in some branches has been generalized: Lat. mediūs, Osc. mefiio-, OIr. mide < *medhjo- beside Ved. mādhyā- (never *mādhīya-), Gk méso- < *medhjo-. Related to Sievers' Law is Lindeman's Law, which provides for the optional realization of monosyllables of the structure CRE- as CRRE-: *djēws in Gk Zeús, Hitt. Śiuš, *diēws in Ved. diyaus.

**Vowels Proper**

The vowels of Proto-Indo-European were e, a, o plus the high vowels i and u which were in complementary distribution with the glides y and w. All five vowels occurred both long and short. These vowels were morphophonemically and distributionally not on a par. Primary short a and long ā can be reconstructed only for a few roots, albeit some of high frequency; most cases of a or ā in the dialects resulted from laryngeal colouring. Primary ĩ and ū were likewise rare; only after H-loss and compensatory lengthening of preceding vowel (including i and u), which belongs properly to the history of the individual dialects, do we get long vowels with comparable frequency to their short counterparts. The short high vowels ĩ and ū were usually zero grades (p. 52) of diphthongs ei and eu; the long vowels ē and ō were usually lengthened grades (p. 52) of a short e and o. Short e and short o were in alternation according to morphological function (p. 51); the short e was typically the basic vowel, and short o derived from it.

The resultant vowel system can be displayed as a traditional triangle subject to the distributional skewing discussed.

```
i   į   ū
 e   o   ē   ō
 a   ā
```

The system was in fact not stable; the wave of compensatory lengthening following H-loss further skewed it, and most branches of the family sooner or later lost the old quantity system and went on to develop new ones.

The disparity in frequency of ō and ā led to developments which distinguish two dialects in the IE languages of Europe. The southern languages, Celtic, Italic, Greek and Armenian, plus Tocharian, maintained the five-vowel triangle by increasing the instances of a. The northern languages, Balto-Slavic,
Germanic, and Albanian, merged a and o, and with the exception of Baltic, ā, and ē, thus creating a vowel rectangle (see Chapter 13, p. 389):

\[
\begin{array}{ccc}
\text{i} & \text{u} & \text{a} \\
\text{e} & \text{crowed} & \text{ā} \\
\end{array}
\]

The vowel ā would be differently realized according to dialect: Gmc a, ā, Slav. o, a, Baltic a, ē.

e, o: the basic e is usually the full-grade (see p. 52) of roots, like *bʰer- 'carry', *nem- 'distribute', *gʷon- 'smite': Gk pherō, Lat. ferō, Goth. bairan, OIr. berid, OCS bero, Toch. A, Toch. B pär-; Gk nēmō, Goth. niman, OIr. nemed 'privileged person'; Hitt. kuenzi [kʷentsi], Gk then-ōn 'smiting', OCS ēnoq 'I drive'. The same roots make o-grades *bʰor- *, *nem-, *gʷon-: Gk phōros 'tribute', nomōs 'pasturage', phōnos 'murder'. Without attested e-grade and possibly non-apophonic: *bʰosó- 'naked' in OEng. bær, OCS bos不断地, Arm. bok.

\[
\begin{array}{ccc}
\text{a} & \text{e} & \text{i} \\
\end{array}
\]

a: clear cases of fundamental a in common words are the roots:

*dap- in Lat. daps 'sacrificial meal', damnum 'damage entailing liability', Gk dapánē 'expenditure', Arm. tawn 'festival'

*ghans- 'goose' in Gk khēn, Lat. anser, OHG gans, OCS gos不断地, Lith. žasis;
*sal-'salt' in Skt salīla- 'sea', Gk háls 'sea', Lat. sal-, Goth. salt
*albhō- 'white' in Hitt. alpaš 'cloud', Lat. albus, Gk alphós 'leprosy'.

\[
\begin{array}{ccc}
\text{i} & \text{u} & \text{e} \\
\text{o} & \text{w} & \text{c} \\
\end{array}
\]

i, u: examples are the instrumental singular ending *-bh(-) in Ved. padbhis, Myc. Gk po-pi/popphi 'with the feet', Gaul. gobedi 'by the smiths'; *nisdo- 'nest' in Skt niṇḍa, niḍus, OIr. net, OEng., OHG nest; *putlo- 'son, boy' in Ved. putrās, Avest. putrō, Osc. puklum (acc. sg.). Underlying short u in the particles Hitt. nu, Ved. nú, Gk ny, Lat. nu-dius (tertius), Lith. nū (related to *nėwo- 'new'), but lengthened in Ved. nú, nūnām, Avest. nū, nūrōn, Palaic nu-u, Gk nyn, OCS nynē 'now'.

ē, ō: these appear typically in lengthened grade categories (see p. 52), like the nominative singular of stems in -R: Gk patēr 'father', ablauting apātōr 'fatherless', Skt pītā; IE *ukʷsō(n) (n), stem *ukʷsen- 'ox' in Ved. uksā, Avest. uxšan-, OEng. oxan, nom. pl. oxan, MWelsh ych < *uchī < *uxsā, pl. ychen < *uchem < *uxsenes, Toch. B. okso; *dōm 'house' in Hom. Gk dō, Arm. tun, OIr. dām 'retinue' < *dōm-ā abstract of *dōm-ō- 'belonging to the house'. Long ō is also the PIE product of contraction o + e. Compare the thematic (p. 66) nominative plural *-ōs < *-o-es, dative singular *-ōi, < *-o-ei: Ved. virās, virāy-a (Avest. āhūrāi), Osc. bivus, hurtūi, OIr. voc. pl. á firu < *-ūs < *-ōs, Gaul. (Gk alphabet) -oī.

ā: Rare, but compare the words for 'mother', Ved. mātār-, Dor. Gk acc. mātēra, Gaul. matīr/[mātīr], Toch. mācer, and OIr. sāl 'ocean, sea' < *sāl-ō- 'salty'.

\[
\begin{array}{ccc}
\text{a} & \text{e} & \text{i} \\
\text{ö} & \text{ù} & \text{þ} \\
\end{array}
\]
Table 2.4  Vowels proper: correspondences

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Oscan</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Old Irish</th>
<th>Old Church Slavonic</th>
<th>Lithuanian</th>
<th>Tocharian(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>e</td>
<td>e</td>
<td>e/i</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>ã^2</td>
</tr>
<tr>
<td>o</td>
<td>a</td>
<td>ā, a^3</td>
<td>ā, a^3</td>
<td>o</td>
<td>o</td>
<td>ú, o^4</td>
<td>a</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>ā</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
<td>ñ</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>ñ</td>
<td>u</td>
<td>ñ</td>
<td>ã</td>
</tr>
</tbody>
</table>

1 The reflexes are those of Toch. B.
2 With preceding palatalization.
3 ā in open syllables by Brugmann’s Law.
4 In the Latin alphabet.
### Table 2.5 Long vowels: correspondence charts

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Sanskrit</th>
<th>Avestan</th>
<th>Greek</th>
<th>Latin</th>
<th>Oscan</th>
<th>Gothic</th>
<th>Armenian</th>
<th>Old Irish</th>
<th>Old Church Slavonic</th>
<th>Lithuanian</th>
<th>Tocharian B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ē</td>
<td>e, i</td>
<td>ā</td>
<td>ā</td>
<td>ē</td>
<td>ē</td>
<td>ē</td>
<td>ē</td>
<td>ī</td>
<td>ī</td>
<td>ī</td>
<td>ī</td>
<td>ē</td>
</tr>
<tr>
<td>ō</td>
<td>a</td>
<td>ā</td>
<td>ā</td>
<td>ō</td>
<td>ō</td>
<td>uu</td>
<td>ō</td>
<td>u</td>
<td>a, ū</td>
<td>ū, ū</td>
<td>a</td>
<td>ē</td>
</tr>
<tr>
<td>ā</td>
<td>a?</td>
<td>ā</td>
<td>ā</td>
<td>ā</td>
<td>ā</td>
<td>aa</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>ā</td>
<td>a</td>
<td>a, o³</td>
</tr>
<tr>
<td>ĭ</td>
<td>i</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
<td>ĭ</td>
</tr>
<tr>
<td>ŭ</td>
<td>u</td>
<td>ŭ</td>
<td>ŭ</td>
<td>ŭ</td>
<td>ŭ</td>
<td>ŭ, ū?</td>
<td>ŭ</td>
<td>u</td>
<td>ŭ, ū?</td>
<td>ŭ, ū?</td>
<td>ŭ</td>
<td>ŭ</td>
</tr>
</tbody>
</table>

1. ė in original final syllables.
2. ūo from *oh₂, *oh₃.
3. o in some final syllables.
i ū: Likewise rare; compare *wīs-ō- in Gk ἵσ, Lat. vírus but *wīs-ō- in Skt vīsās ‘poison’. The word for ‘mouse’ (and ‘muscle’, in accord with a semantic quasi-universal, widespread in the languages of the world), Ved. mūs-, Lat. mūs, Gk μῦς, OHG, ON, OEng. mūs is a probable example, in view of Ved. muṣṇāti ‘steals’, though some have seen a laryngeal origin in the tectal of Arm. mukan.

Diphthongs
The three vowels e, o, a combined with the consonantal variants of the high vowels j, w to form six diphthongs: ej, oj, aj, ew, ow, aw. It is customary to treat the other ER cases (er, ol, en, etc.) as sequences, though in Lithuanian they are intonable in the same way as the diphthongs with semi-vowel (aī, aĩ, ān, ār, like aĩ, aũ, āi, āu), and are therefore equally legitimately diphthongs (see Chapter 15, p. 460). The diphthongs with glides are best preserved in Greek and early Italic; in most other dialects their number has been reduced by merger and monophthongization.

IE *dejwós 'god': in OLat. deiuos, Osc. deiwát, Gaul. dēvo-, Avest. daēuwa- 'demon';
IE *moj-no- 'to be exchanged', in OLat. commeine, Lat. commūnis, Osc. múṅtikû, Goth. gamains, OIr. moing 'treasure', commuin 'mutual obligation', Skt menāmenam 'in exchange', Sicilian Gk (Sophron) moīton anti moītou ‘like for like’
*kajko- 'one-eyed': in Goth. haihs, OIr. cáech, Lat. caecus 'blind', with the infrequent and therefore expressive a-vocalism frequent in words for infirmities and other semantically marked notions.

The diphthong falls together with laryngeal coloured *h₁aj < *h₂ej-, as in *h₂aj-dʰ- 'burn' in Gk aith-omai, Lat. aed-ēs 'building' < *‘hearth’, perhaps Pal. ḥa-a-ri < *ḥaj-ari 'is hot'. We have *ew in *lewk-ō-, Gk leukós 'white, bright', beside o-grade *lowk-ējeti 'lights' in Ved. rocáyati, Avest. rao- caietiti, Hitt. lukkizzi; *aw in sausō- 'dry', Gk aŭōs, Lith. saūsas, Russ. mobile sūx, n. sūxo, f. sūxa, OEng. sēar < Gmc *sauzāż.

Morphophonemics
Indo-European morphophonemics includes accent, apophony (or 'ablaut'), and root structure.

Accents
A single Indo-European word accent can be reconstructed, represented in Vedic by the udātta and in Greek by the acute, both noting high pitch: pitāram = patēra. Accent is reconstructed from other more complicated, innovated and sporadic reflexes in Vedic and Greek, from the accent and intonation systems
of Balto-Slavic, from Anatolian, and from the effects of Verner's Law in Germanic (see Chapter 13, pp. 394f.).

Aside from atonic or enclitic forms every IE word had a single accent, whose position was governed by rule of word formation and inflection, and whose presence or absence in certain grammatical categories was a function of syntactic rules. Thus in inflection some paradigms had fixed accent on the root or ending, while others had movable accent. In word formation, certain suffix morphemes were inherently accented, others were unaccented and implied root accent. In syntax the verb was unaccented or weakly accented in main clauses but accented in subordinate clauses; but when the verb in a main clause was fronted to sentence-initial position it was accented.

The inherited system underwent profound changes even in those dialects which retained it, like the restriction to the last three syllables in Greek. The remaining dialects have replaced it by newer, independent accent systems, typically linked to the word boundary: accent (stress) on the initial syllable was widespread as a sort of ‘default setting’, as in early Italic, Goidelic Celtic, and Germanic. This may represent the generalization of an accentual feature already present in the proto-language, as in the usually unstressed Vedic vocative in sentence-initial position, when it is accented on the first syllable.

**Apophony (‘Ablaut’)**

Indo-European is profoundly marked by the system of vocalic alternations, expressing morphological functions, which is termed ‘ablaut’ or apophony. The fundamental form was the vowel \( e \), which under certain conditions could appear as \( o \) while under certain other conditions the vowel could disappear entirely. We speak of the forms as exhibiting *e-grade*, *o-grade* and *zero grade*. Thus the Indo-European word for ‘knee’ appears as *\( ġ \)enu-* (Hitt. *genu*, Lat. *genu*), *\( ġ \)onu-* (Skt *jānu*, Gk *gόnu*), and *\( ġ \)nu-* (Hitt. instr. *ganut* ‘with the knees’, Ved. *jiu-bāḍh- ‘kneeling’, Gk dat. pl. * peri gnusi* ‘about the knees’).

In certain cases, termed ‘Schwebeablaut’, the full grade may appear for example not as *\( ġ \)enu-* but as *\( ġ \)neu-* (Goth. *kniu*). More limited in extent but still reconstructible for the proto-language are the ablaut vowels \( e \) and \( o \), termed ‘lengthened grade’. Compare Gk nom. sg. *patēr* ‘father’ (acc. sg. *patēra*) and *apātōr* ‘fatherless’ (acc. sg. *apātora*), or the cognate of ‘knee’ in Gk *gōma* ‘angle’ < *gōn-ih₂.

When one of the consonants is a resonant \( R \), the latter appears in zero grade as either \( R \) or \( R \), depending on the non-syllabic or syllabic nature of the following sound respectively:

\[
\begin{align*}
\text{TcR-} & \quad \text{ToR-} \\
\text{TR-T-} & \quad \text{TR-E-}
\end{align*}
\]

as in *e-grade* Gk *tēn-ōn* ‘tendon’, *o-grade* *tōn-os* ‘stretching’, and zero grade *ta-tōs*, Skt *ta-tás* < *\( t̐ \)tōs* ‘stretched’ but Skt *ta-tn-e* ‘it was stretched’.
When the sonant precedes the fundamental vowel we speak of ‘sampräsārana ablaut’:

ReT- RoT- RT-

as in e-grade wek*- in Gk (w)e̱pos, Ved. vácas- ‘word’, o-grade *wok*- in Hom. Gk (w)op- ‘voice’, zero grade *uk*-tó - in Ved. ukta- ‘spoken’.

In the fairly common situation where the root contains both a resonant R and a laryngeal H, and Schwebeablaut is in play, we find a full panoply of ablaut variants, as from IE *pelh₁- ‘fill, full’:

*pelh₁-nes- Ved. párīnas- ‘plethora’
*pelh₁-u- Goth. filu, OIr. il ‘much, many’
*polh₁-u- Gk polūs, polōi, ‘many’
*plh₁-u- Ved. puru-, Avest. paoru- ‘id.’
*pleh₁-C- Gk plēto ‘filled’, Lat. plēnus, Arm. li ‘full’
*pleh₁-isto- Gk plei̱stos, Avest. fraē̱sta- ‘most’
*plo̱h₁- Ved. perf. paprátha ‘you filled’
*plh₁-no- Ved. pûṟnā- ‘full’, Lith. pilnas, Goth. fulls < *fulnanz
*plh₁-e- Ved. reduplicated thematic āpīṟata.

Apophony in the Root

Apophony is found in the root, in suffixes, and in endings. Here we consider only the root.

Typical derivational categories showing e-grade are the singular of root athematic presents (*hxes-ti ‘is’; *gwhen-ti ‘slays’), and aorists (*e-d*ēh₁-t ‘placed’, e-stah₂-t ‘stood’), as well as thematic presents (*sēk*-or ‘follows’, *tērh₂-or ‘is able, overcomes’), whence the custom of citing roots in the e-grade, since most roots are verbal. In the noun, note the simple thematic adjectives (*néwo- ‘new’, *séno- ‘old’), thematic neuter nouns (*wērgom ‘work’, *pēdom ‘place’), and the numerous secondary suffixes (p. 64) added to e-grade roots like *-tor-, *-men, *-es-.

The o-grade is characteristic of the (late?) IE stative perfect in the singular (*wōjd-e ‘knows’, *(d)e-)*dḥōrs-e ‘dares’), the causative-iterative (*lowk-ēje-ti ‘lights’, wos-ēje-ti ‘dresses’), certain intensive reduplicated presents in the singular (Ved jaṅghanti ‘slaughters’ < *gwhen-)*gwhon-ti), numerous simple deverbal thematic nouns and adjectives (type *gwhōnos ‘slaying’, *gwhonós ‘slayer’), and numerous secondary formations.

The zero grade is found outside the singular in the apophonic paradigms of p. 60: (*h₁s-ēnti ‘they are’, *wid-mē ‘we know’, *(d*)ē-)*dḥ₁-ti ‘they place’), before accented ending in some apophonic noun paradigms, and before accented secondary nominal suffix like *-tó-, *-nó- (*uk*-tó- ‘spoken’) or verbal suffixes like *-skē- (*gwh-m-skē-ti ‘is going’).

The lengthened grade is found in the singular of some apophonic athematic
verb paradigms, alternating with e-grade outside the singular (*téks-ti *téks-
̣nti ‘fashion’), before secondary thematic adjective suffix -ó- in the derivational
process known as vrddhi (*dom/*dem- ‘household’ → *dōm-ō
‘belonging to the household’, Ofr. dām ‘retinue’; *wed-/wed- ‘water’ →
*wēd-o- ‘watery, wet’, OEng. wæt Luv. Ú.SAL¹A -anza witanza ‘in the
watery meadows’.

Morpheme Structure Rules

Under the heading of morpheme structure rules we may include the rules
for well-formed roots, and certain constraints on the phonological structure of
roots.

With a small number of exceptions, chiefly in the pronouns, the canonical
shapes for IE roots are the following (Benveniste 1935):

\[ \text{C}_1 \text{EC}_2^- \]

where \( \text{C}_1 \) and \( \text{C}_2 \) are different consonants, and \( \text{E} \) is usually the fundamental
vowel \( e/o/0 \), rarely \( a(0) \);

\[ \text{C}_1 \text{EC}_2^- \text{C}_3^- \ (\text{state I}) \]
\[ \text{C}_1 \text{C}_2^- \text{EC}_3^- \ (\text{state II}) \]

where the \( \text{C}_3 \) is different from \( \text{C}_2 \), and \( \text{C}_1 \) is equal or lower than \( \text{C}_2 \) in
sonority;

\[ \text{C}_1 \text{C}_2^- \text{EC}_3^- \text{C}_4^- \]

where \( \text{C}_4 \) is different from \( \text{C}_3 \), and equal to or lower than \( \text{C}_3 \) in sonority. The
sonority hierarchy is \( T/H < R < E \); we find root initial HT- and TH-, root-final
-TH and -HT. Well-formed roots are thus *tek-, *ter-, *ret-, but not **tet-;
*terp-, *trep-, but not **terp-, **petr-, *prekt-, *h₂weks-, but not **prekl-,
etc. In all cases \( \text{C}_1 \) may be preceded by an \( s \), with no discernible semantic
content (‘s-mobile’). Exceptionally, a root may lack an initial \( \text{C}_1 \) and begin
with a vowel (*albh-o- ‘white’; *att(a) ‘father!’). Apparent violations of the
identical consonant constraint like *ses- ‘sleep’ (Hitt. šeš-zi, Ved. sāstī),
*mēms-ō- ‘meat’ (Ved. māṁśā-, Arm. mis, Goth. mīmz) are doubtless
grammaticalized from reduplicated child-language forms.

Further root-structure constraints are that no root may begin and end with
a plain-voiced stop (**deg-, **derg-, **gʷed-), and that no root may begin
with a voiced aspirate and end with an unvoiced stop, or vice versa **bʰer-,
**bʰejk-, **tebʰ-). But the combination of initial \( s \) plus unvoiced stop with
final voiced aspirate is allowed: *stejkʰ- ‘rise’, perhaps *stebʰ- ‘crown,
wreathe’.
### Morphology

Indo-European languages, particularly in their earlier stages, share a rich and complex morphology of the synthetic type. The independence and autonomy of the word, reinforced by the accent, was such that a variety of morphological elements and processes served as the primary exponents of the syntactic functions. A large part of the success of the comparative method in this family is due to the number and the precision of the agreements among languages in the particulars of morphology. Consider the partial paradigms of the words for ‘dog’ and ‘slay’ in various dialects and as reconstructed:

<table>
<thead>
<tr>
<th>Hittite</th>
<th>Greek</th>
<th>Vedic</th>
<th>Lithuanian</th>
<th>Old Irish</th>
<th>European</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nom.</td>
<td>kuwaś</td>
<td>kūon</td>
<td>ś(u)va</td>
<td>šuō</td>
<td>cú</td>
</tr>
<tr>
<td>Acc.</td>
<td>kuwanan</td>
<td>kúna</td>
<td>svánam</td>
<td>šúnį</td>
<td>coin</td>
</tr>
<tr>
<td>Gen.</td>
<td>kúnaš</td>
<td>kúnós</td>
<td>šúnas</td>
<td>šuñis</td>
<td>con</td>
</tr>
<tr>
<td>Hittite</td>
<td>Vedic</td>
<td>Proto-Indo-European</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 sg. pres. ind.</td>
<td>kuenzi</td>
<td>hánti</td>
<td>*gʷhén-ti</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 pl. pres. ind.</td>
<td>kunanzi</td>
<td>ghnánti</td>
<td>*gʷh₁n-én-ti</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The reconstructed forms with their complex interplay of ablaut and accent serve at the same time as a shorthand for the successive morphological as well as phonological changes leading to each of the attested dialect forms.

In early IE languages the fundamental domains of morphology are inflection, derivation and composition; combining the latter two we have inflection and word formation.

- **Inflection** deals with the ‘paradigm’, the varying forms under which a given inflectible stem or lexical entry (‘word’) may appear in a sentence, as a reflex of its syntactic function.

- **Derivation** deals with the formation of inflectible stems, the formation of ‘words’ minus their inflection. We distinguish *primary* derivation and *secondary* derivation. Primary derivatives are formed at the level of the abstraction we term the *root*; secondary derivatives are inflectible stems formed from other inflectible stems which coexist in the language at the same time, that is, in the same synchronic system.

- **Composition** deals with the formation of inflectible stems from the combination of an inflectible stem with one or more other meaningful elements.

With a limited class of exceptions, including sentential particles, conjunctions and quasi-adverbial forms (preverbs, postpositions, prepositions, negations), all IE words were inflected. The structure of the inflected word was *Root* plus *Suffix* plus *Ending*.
R + S + E

(R + S) together constituted the inflectible stem. The suffix S is recursive:

(((R + S) + S) + S), etc.

The root contained the basic lexical semantic kernel, with further grammatical meaning supplied by the suffix(es); these typically determined the part of speech (‘pars orationis’) of the word. Root plus Suffix constituted the stem, which was the domain of derivation (word formation). The open set of stems were the basic lexical stock of the language. Each stem received a single ending, the domain of inflection, which specified its syntactic function in the sentence and assigned the grammatical meaning of the inflectional categories: case and number in the substantive, gender in the adjective; person, number, voice, tense/aspect, mood in the finite verb.

Consider the sentence-initial verb in the Homeric Greek phrase lissōm’ anēra toūton ‘I will entreat that man’ (Iliad 22.418). Restoration of the elided final (which implies undoing Greek sentence phonetic rules), and segmentation (which implies undoing Greek morphophonemic rules), yields the string

lit-jo-o-mai

containing:

1. a lexical morpheme lit- together with a portmanteau meaning. There is in Greek no ‘word’ lit-, an abstraction we may call a ‘root’ R. There is only (in Homer) a verb lissomai, with litēsthai, ellisāmēn; a verb of similar meaning litaneūō, a verbal adjective -llistos occurring only in composition (p. 56), and a noun occurring only in the plural, litai. The morpheme lit- in lissomai is followed by three further Greek morphemes of widely different function;

2. a derivational suffix -jo- which makes a verb stem, present tense. The suffix establishes the real lexical entry and the part of speech: a verb stem lit-jole- (R + S) ‘pray, entreat’. Only at this point can we talk about translatable meaning;

3. an inflectional suffix -ole-, the sign of the subjunctive mood, with its semantics (grammatical meaning): the narrated event filtered through the attitude of the speaker. The subjunctive stem lit-jo-o- ((R + S) + S) is part of the paradigm of the lexical verb stem lit-jo- (R + S);

4. finally, an ending -mai, a single morpheme but polyfunctional, expressing at once the grammatical categories of person (first), number (singular), voice (middle), which are semantic categories, as well as a purely formal category, ‘primary’, a distributional variant conditioned by the mood-sign subjunctive. Primary endings are associated with non-past-tense,
secondary endings with past tense. The opposition goes back to Proto-Indo-European, though Greek has partly innovated in their distribution.

Each of the morphemes, and their order, faithfully continue the IE situation, even if the -m- of the first-person singular ending -mai is a Greek innovation. The thematic present suffix *-jole- recurs in most of the family, for example IIr. -ja-. The subjunctive sign *-ole- combines with the thematic suffix to form a long vowel in IIr. -ā- just as in Gk -ōlē-. The root (-l)lit- has no known cognates outside Greek, but its root structure is perfectly canonical Indo-European, and its apparent zero grade *slit- (to a putative full grade *sleit-) is expected before the accented present forming secondary verbal suffix *-jolē-, the thematic aorist suffix *-ōlē- of litēsthai, and the nominal suffix *(< *āh₂-) of lītaī.

Composition involved the combination of two lexical items or notions into a single word: Gk trίllistos ‘thrice prayed for’, polūllistos ‘much prayed for’; IE *η-udros ‘waterless’ in Ved. anūdrās, Gk ānudros. The possessive (bahuvrihi) type, like Eng. barefoot, and the additive (dvandva), like Ved. dvādaśa ‘12’ (‘2 [and] 10’) were well represented in the proto-language; the former type was particularly frequent in personal names as well (cf. Chapter 4, p. 121). Some IE languages have extended and developed composition down to the present day; others have drastically restricted or eliminated it.

In a treatment of this size it is impossible to give a full account of the wealth of reconstructed IE morphology. Nor can we do justice here to the lively and informed controversy that surrounds some areas of PIE morphology and its ancestry. Only salient features of verbal and nominal derivation and inflexion will be given.

While Indo-European shows a clear distinction between verbal and nominal stems, certain features are common to both, like the opposition athematic : thematic stems. The minimal athematic suffix is zero (-0-); the minimal thematic suffix is the thematic vowel -ole-. A root athematic stem has the structure root + zero suffix before the ending, (R + 0) + E: nom. sg. *pōd-0-s ‘foot’, 3 sg. pres. *gwhēn-0-ti ‘slays’, more simply just *pōd-s, *gwhēn-ti, beside thematic *ēkw-o-s ‘horse’, subjunctive *gwhēn-e-t(i) ‘may slay’. More complex athematic suffixes end in consonant (*-t-, *-men-), while their thematic counterparts end in the thematic vowel (*-to-, *-mn-).

The IE verb typically expressed action (the active voice, unmarked), process (the middle voice, subject internal to the action), or state (the perfect). Compare respectively English murder, learn, know with object French. The three stood originally in a derivational rather than inflectional (‘paradigmatic’) relation; many verbs expressed only one of these functions, and where two or more were expressed it was by different stems. This situation is seen most clearly in Vedic, Homeric Greek, and Anatolian. Process and state were related, as shown by the common origin of their endings.

Primary Verbal Stems. The root athematic active formation makes
present and aorist stems. One type shows ablaut e : 0 with accent shift, as in Gk eî-mi i-men, Ved. ē-mi i-māsi ‘I, we go’, Ved. ksē-ti kṣiy-ānti ‘he, they dwell’, Myc. Gk kitijesik/tijensil, pple. mid. Gk -kti-menos; aor. ipv. 2 pl. Ved. śrotā, Avest. sroatā, 2 dual. śrūāntī/2 sg. śrūhī, Gk kīśi (with metrical or expressive lengthening). Another type has ablaut ē : e and columnar accent on the root (see Chapter 9, p. 239) Lat. ēst, ed-unt ‘he, they eat’, Ved. tāṣṭi tāksati ‘he, they fashion’.

Several types of reduplicated athematic presents are found. With i in the reduplicator and e : 0 ablaut Gk tī-thē-mītī-the-men ‘place’, di-dē-ldi-de-‘bind’, Gk pple. bi-bās (bi-bānt-) ‘striding’ = Ved. jī-gat- (3 sg. jī-gā-ti). With e in the reduplicator and o : 0 ablaut Ved. janghanti ‘slaughters’, pple. jānghanat, gen. jānghanatas.

Nasal-infix presents: the only morpheme to be infixed in Indo-European is -n-, typically forming active transitives. The element -n- is infixed between C₂ and C₃ of the root, with e : 0 apophony and mobile accent. Thus

*lejkʷ* ‘leave’ pres. *linékʷ*-tī/*linkʷ*-ēnti
*klew-* ‘hear’ *kṇēw*-tī/*kṇhw*-ēnti
*pewh₂-* ‘purify’ *punēh₂*-tī/*punh₂*-ēnti

Compare Vedic: rināk-ti riṅc-ānti, śṛṇoti śṛṇvānti, punāti punānti. The type is preserved most faithfully in Indo-Iranian; other branches have altered the original system. Reduplicated and nasal-infix presents typically formed active root aorists.


The root athematic middle aorist appears to show ablaut o : 0 and shifting accent, in the aor. ‘passives’ like ābodhi abudhran ‘awoke’. An e-grade may appear also, as an Avest. jaini ‘was slain’. Comparable ablaut relations o : e : 0 have been suggested for the ancestor of Indo-European presents showing both o and e grade, like Goth. malip, Hitt. malli ‘grinds’, Luv. malhu- ‘crushes’ but OIr. melid ‘grinds’, IE *molh₂-/melh₂-, perhaps with perfect endings.

The perfect. Until the discovery of Hittite and Tocharian the perfect was one of the most secure reconstructions in the whole IE verb: it was characterized by a special set of endings: original presental stative value, from which developed resultative, and ultimately just preterite value; a special active participle in
*-vms-/*-ws-; in most early languages reduplication with e except in the verb 'know'; root vocalism o : 0 with shifting accent. Homeric Greek shows the vocalism and value clearest: peithomai 'I am persuaded, obey', perf. peipoitha pepithmen 'I, we trust', óllumi 'I destroy' perf. pepoitha pepithmen 'I, we trust', öllumi 'I destroy' perf. pepithmen 'I, we trust', öllumi 'I destroy' perf. öllumi 'I destroy'. The ablaut and accent are both clear from the effects of Verner's Law in Germanic: OEEng. cēosan 'choose' (*gews-) pret. ic, hé cēas, (*gōws-) pl. curon 'chose' (*gūs-). The picture is now more complicated, since Tocharian shows a reflex of the reduplicated perfect participle (Toch. B obl. peparkos 'asked') but no reflex of the finite perfect, while Hittite and most other Anatolian languages have a special hi- conjugation parallel to the obviously inherited mi- conjugation, which despite many efforts resists straightforward derivation from the IE perfect as reconstructed. The case is one of the most intensely debated in IE studies today.

Primary thematic formations, present and aorist, place a large and expanding role in the dialects. Their distribution is a mark of recentness; their creation would appear to be one of the latest innovations of Common Indo-European. Their origin appears to be in the middle voice, as in p. 56, where the thematic vowel *o!e was originally a third-person singular ending. In Hittite, primary thematic presents are found only in the middle voice: 3 sg. neya(ri) 'leads' from *neh-i-o(r) like Ved. úpo naya-sva 'lead hither!' The active (the common type) 3 sg. Ved. náyati (from *-e-ti) was created later by opposition. By Hittite times this had taken place only in secondary, derived verb stems: 3 sg. -šk̂ezzi < -šk̂-ē-ti, -ižzi < -jē-ti, *-ē-je-ti.

Secondary verbal stems served to form new verbs from existing stems. The most widespread were denominatives, to make verbs from nouns and adjectives, by the suffix -jēlo-: Ved. vásnam 'sale price', vásna-yá- 'fetch a price', Gk ónos 'price', ónéomai 'I offer to buy, make a bid on', Hitt. ušne-skatta 'makes a bid on'. The thematic nominative type -e-jēlo- is most clearly attested in Indo-Iranian, Greek, Armenian (sēr. gen. siroj 'love' → sirem 'I love' < *kekre-jē-); other languages have innovated the form in part, like Latin largus → largīri (suff. *-ī-jēlo-). The suffix -jēlo- is added directly to the consonant stems: Ved. āpas- 'work' apas-yáti, Gk télos, teles- 'goal' → teleō 'achieve', Goth. riqis 'darkness' → riqiz-jan. Stems in R may show zero grade: Gk ónoma 'name' → *onomn-jēlo-, onomainō, Goth. namo, namin- 'name' → *namnjan. An old example is the zero grade of *mélit-'honey' in *mlit-jēlō- 'take out the honey from the comb', Gk blíttō.

For the denominatives to feminine abstracts in *-eh₁₂- (*-ah₂, later -ā, see p. 63) we expect -ah₂-jelō-, whence *-ājelō-, as in Lat. fuga → fugāre. In some languages the h₂ seems after colouring to have been lost before yod very early: Hitt. -aizzi/-ānzi, Gk -āō.

An old secondary causative-iterative formation with o-grade root and suffix -ējelō- is well attested throughout the family. From the root *wes- 'wear, dress' an active *wos-ēje-ti is attested in Hitt. waššižzi 'put clothes on someone', Ved. vásáyati, Gmc. *wazjan, Alb. vesh < *wasje-. From *lewk-
'light, shine' we have active *lowk-éje-ti in Hitt. lukkizzi, Ved. rocáyati, Avest. raocaiti, OLat. lúmina lúcent 'they light the lamps'.

The suffix *-skélo- forms characterized presents, typically iterative, to zero-grade roots. An old example to the root *prék- is *prék-skélo- 'ask, demand' in Ved. pṛccháti, Avest. parásaiti, Lat. poscō from *por(c)scō, OIr. arcu, Arm. harc'-anem, OHG forscon 'inquire'. Three languages attest the formation to the verb 'to be' in an existential sense: OLat. escit 'there is', Gk ἔσκε 'there was' (Alcman), ἔσκε (Odyssey, 9.508), Pal. ipv. mid. ḫśka < *h₁s-skó 'be (ipv.)'.

Two secondary suffixes are added to adjective stems: a factitive in *-ah₂- (*-eh₂-), meaning to make something what the adjective denotes, and a stative in *-eh₁- meaning to be what the adjective denotes. Compare Hittite néwāš (*néwo-) 'new' → newāh₁(ḥ)- (*newěh₂, newāh₂) 'make new', maršāš 'false' → maršāh₁- 'falsify', and → maršē- 'be false'; Lat. albus → (de-)albāre 'whiten', albēre 'be white'. In some cases this suffix is part of a so-called 'Caland-system': adj. *gʰr₂-ú- 'heavy' in Gk barús, Ved. gurité, abstract *gʰr₂-ós- in Gk bárós, static *gʰr₂-eh₁- in Aeol. Gk bóretai 'is heavy', Hom. Gk pf. pple. bebaréōs; adj. *h₁r₁udʰ-ro- 'red' in Lat. ruber, Gk eruthrós, static *h₁r₁udʰ-eh₂- 'be red' in Lat. rubēre, OIr. -rúdi, OHG rotën. A Caland-system in secondary derivation typically exhibits commutation of adjectival suffixes *-u-, *-ro-, *-ent- or substantival *-es- with *-i as first member of a compound: Ved. tji-rás, Gk arg-ós (from *arg-rós) 'swift, bright', in composition Ved. tji-, Gk argi-; note the Greek phrase kūnes argol 'swift dogs' beside the Vedic possessive compound tji-śvan-, personal name, 'having swift dogs', and Odysseus' dog Árgos.

In verb inflection, the endings of the finite IE verb marked person (1, 2, 3), number (singular, plural, dual), as well as the 'voices' (genera verbi) active, middle, stative perfect, the modal opposition indicative : imperative, and the opposition primary (marked by a particle -i emphasizing the hic et nunc) : secondary (unmarked, lacking the particle). Certain of these oppositions are neutralized outside of the singular. As noted on p 58, the thematic conjugations are apparently a late IE development out of the athematic middle, but they were fully constituted for secondary verbs before the separation of the Anatolian branch.

In athematic active paradigm, forms followed by a hyphen are completely specified. The vowel e may ablaut (o, ő, o). See Table 2.6.

In the thematic active paradigm (selections), the thematic vowel is o before R, H, e before T. See Table 2.7.

In the athematic middle paradigm (selections), the details must be tentative; this is another of the controversial areas in current Indo-European studies. See Table 2.8.

Under some circumstances (accent?) the vowel -o- may appear as -e-. The third persons in -*t-, -*nt- belong to a later chronological layer. Several dialects replaced primary -r by -i, from the active conjugation, whence
Table 2.6  Indo-European athematic active verbal endings

<table>
<thead>
<tr>
<th></th>
<th>Indo-European</th>
<th>Hittite</th>
<th>Vedic</th>
<th>Greek</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 sg.</td>
<td>*-mi</td>
<td>-mi</td>
<td>-mi</td>
<td>-mi</td>
</tr>
<tr>
<td>2 sg.</td>
<td>*-si</td>
<td>-si</td>
<td>-si</td>
<td>-(s)i</td>
</tr>
<tr>
<td>3 sg.</td>
<td>*-ti</td>
<td>-zi</td>
<td>-ti</td>
<td>-(Dor.)</td>
</tr>
<tr>
<td>1 pl.</td>
<td>*-me</td>
<td>-weni</td>
<td>-mas(i)</td>
<td>-men</td>
</tr>
</tbody>
</table>
| 2 pl.  | *
| 3 pl.  | *(e)nti       | -anzi   | -an(s)i | -ensi (Myc.) |
| 1 Dual | *-we-         | -vas    |       |       |
| 2 Dual | *
| 3 Dual | *-t(h)2-o-    | -thas   | -ton  |       |
|        | *(e)nt        | -er  | -(a)n | -n |
| 1 sg.  | *-m           | -(n)un  | -(a)m | -n |
| 2 sg.  | *
| 3 sg.  | *
| 1 pl.  | *
| 2 pl.  | *
| Secondary |               |         |       |       |
| 3 pl.  | *(e)nt        | -er, ir | -(a)n | -n |
| 1 Dual | *-we-         | -va     | -ton  |       |
| 2 Dual | *
| 3 Dual | *
| Imperative | *0, *-d\text{di} | 0, -t  | -dhi  | -thi |
| 3 sg.  | *
| 3 pl.  | *(e)ntu       | -antu   | -antu | -ntö |

*-oj/*-toj, *roj/*-ritoj, as in Indo-Iranian and Greek.

The perfect endings are clearly related to those of the middle.

From the above paradigms or something like them, the historical paradigms of the different dialects evolved, through a series of divergent innovations, analogical remodelings, and categorial reassignments. Crucial was the role of the athematic third-person singular middle ending -ole, which was evidently revalued to a suffix -ole- followed by zero ending, opening the way to the constitution of a full thematic conjugation.

The morphology of the nominal system subsumes three basic form classes: (1) nouns and adjectives; (2) demonstrative and interrogative pronouns; and (3) personal pronouns. The numerals 1–4 were adjectives, while 5 and up were indeclinable. Adjectives and demonstrative pronouns were inflected for gender, typically with a masculine-neuter stem beside which the feminine stem was a derivative: m.–n. *pih-won- ‘fat’ in Ved. p\text{iv}an-, Gk p\text{ion}- beside f. *pih-wer-ih\text{2}- in Ved. p\text{iv}ar\text{ı}, Gk p\text{ie}ira, MWelsh
Table 2.7 Indo-European thematic active verbal endings

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Vedic</th>
<th>Greek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 sg.</td>
<td>*-o-h₂</td>
<td>-emi</td>
<td>(-ā-mi), Gāthic</td>
</tr>
<tr>
<td>2 sg.</td>
<td>*-e-si</td>
<td>-eši,-iši</td>
<td>asi</td>
</tr>
<tr>
<td>3 sg.</td>
<td>*-e-ti</td>
<td>-ezzi,-izzi</td>
<td>ati</td>
</tr>
<tr>
<td>1 pl.</td>
<td>*-o-me-</td>
<td>-aweni</td>
<td>āmasi</td>
</tr>
<tr>
<td>2 pl.</td>
<td>*-e-te-</td>
<td>-atteni,-itteni</td>
<td>athas</td>
</tr>
<tr>
<td>3 pl.</td>
<td>*-o-nti</td>
<td>-anzi</td>
<td>-anti</td>
</tr>
</tbody>
</table>

| Secondary     |         |       |       |
| 1 sg.         | *-o-m | -anun | -ām | -on |
| 2 sg.         | *-e-s | -eš,-iš | -as | -es |
| 3 sg.         | *-e-t | -et,-it | -at | -et |

| Imperative    |         |       |       |
| 2 sg.         | *-e     | -i | -a | -e |
| 3 sg.         | *-e-tu  | -ittu | -atu | -etō |
| 3 pl.         | *-o-ntu | -antu | -antu | -ontō |


Accent and ablaut: many scholars now assume a complex set of interdependent ablaut alternations and fixed or mobile accents occurring with each (or many) of the suffix type stem classes (e.g. root nouns, men-stems,

Table 2.8 Indo-European athematic middle verbal endings

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Hittite</th>
<th>Vedic</th>
<th>Greek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 sg.</td>
<td>*-h₂ar (*h₂er)</td>
<td>-ha(ri)</td>
<td>-e</td>
</tr>
<tr>
<td>2 sg.</td>
<td>*-th₂ar (*th₂er)</td>
<td>-ta(ri)</td>
<td>-se</td>
</tr>
<tr>
<td>3 sg.</td>
<td><em>-or (</em>-tor)</td>
<td>-a(ri)</td>
<td>-e</td>
</tr>
<tr>
<td>3 pl.</td>
<td>*-ro(r ??)</td>
<td>-ta(ri) (Pal. -tar)</td>
<td>-te</td>
</tr>
</tbody>
</table>

| Secondary     |         |       |       |
| 1 sg.         | *-h₂(a) | -hhāḥat | -i,-a (opt.) | -mān |
| 2 sg.         | *-th₂a | -tat | -thās | -so |
| 3 sg.         | *-o     | -at | -at | |
| 3 pl.         | *-ro    | -ran | -anto | -nto |

| Imperative    |         |       |       |
| 2 sg.         | -Ø      |       |       | |
| 3 sg.         | -ow?    |       |       | |
Table 2.9 Indo-European perfect endings

<table>
<thead>
<tr>
<th>Indo-European</th>
<th>Vedic</th>
<th>Greek</th>
<th>Latin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 sg. *h₂a (*h₂e)</td>
<td>-a</td>
<td>-a</td>
<td>-(ī) (Faliscan -ai)</td>
</tr>
<tr>
<td>2 sg. *th₂a (*th₂e)</td>
<td>-tha</td>
<td>-tha</td>
<td>-(is)ū</td>
</tr>
<tr>
<td>3 sg. *e</td>
<td>-e</td>
<td>-e</td>
<td>-ē(t)</td>
</tr>
<tr>
<td>2 pl. *e</td>
<td>-a</td>
<td>(Paelign. lexē)</td>
<td></td>
</tr>
<tr>
<td>3 pl. *ēr, *ēr</td>
<td>-uh (Avest. -arās)</td>
<td>(-ēre, -erai)</td>
<td>(Venetic -ers)</td>
</tr>
</tbody>
</table>

etc.). The system of the German and Austrian school (Schindler 1975) recognizes four basic types: ‘acrostatic’ Ř-Ś-E, nom. *wód-t Šé gen. *wéd-ŋ-s ‘water’; ‘proterokinetic’ Ř-Ś-E, nom. *h₁órgh-i-t Šé gen. h₁g₂-Śt ‘testicle’; ‘hysterokinetic’ Ř-Ś-E, nom. *ph₂-tr-tr ŠÉ gen. *ph₂-tr-t ‘holokinetic’ or ‘amphikinetic’ Ř-Ś-E, nom. h₂áws-ōs (h₂éws-ōs) gen. *h₂lms-s-ēs ‘dawn’. The system has been criticized as unnatural on grounds of accent typology, and as overly rigid. The prescribed forms do not appear to belong to the same chronological layer (gen. *wed-n-s is attested nowhere, but gen. *bhra-tr-s is in Ved. bhṛtah, ON bróthtor). To the degree that for the theory the ablaut is required to be conditioned by the accent, the formations would have belonged to the remote prehistory of the proto-language; ablaut variants and accent are independent variables already in reconstructed Proto-Indo-European, cf. accented zero grade *w(kwos in Ved. vṛkas, Gk lúkos, Lith. vilkas. Competing systems have been offered, notably by the Dutch school, but comparable objections can be raised to these, and the matter is still very much sub indice.

Of obvious and immediate utility is the notion of internal derivation in the German and Austrian framework, of the type (R + S)₁ → (R + S), beside the more usual external derivation, of the type (R + S)₁ → ((R + S) + S). Compare Ved. bráhman- ‘formulation’ → brahmān- ‘formulator, brahmin’, Gk mnēma ‘remembrance’ → mnémōn ‘mindful’, *(R-ŋ) → *(R-mōn-).

A thematic nominal suffixes: the simplest is -0-, the root nouns like *pód-s, gen. *péd-s (whence *ped-ēs) ‘foot’; *dóm, gen. *dém-s ‘house’ in Gk des-pōtēs, Gāthic Avest. dōṅ patōiš ‘master of the house’. Others have a single consonant: acrostatic *nók-w-t-s ‘night’, gen. *nek-w-t-s in Hitt. nekūz mēhur ‘nighttime’; -s- in nom.-acc. neut. *men-s ‘mind’ in Gāthic Avest. maz-dā- ‘place, direct mind’, Avest. mas ... -dā-, Gāthic Avest. mōncā ... (dā-). With ablaut, -r- in nom. *g₂h₂-s-ōr ‘hand’ (OHitt. possessed kiššar=šiš, acc. g₂h₂-s-ēr-ŋ (kiššeran), loc. *g₂h₂-s-r-i (kišri=tti), kiššari=šmi, Gk kheirf). An -n-stem (a type which becomes limitlessly productive in certain dialects) appears in nom. *h₂or-ō[n] ‘eagle’, Hitt. ḫarā[š], stem ḫaran-, OHG aro, Eng. erne, Gk ór-n-[iš].

The rīn-stem, with -r- in the ‘strong’ cases (nom.-acc.) and -n- in the oblique cases, is well represented in Anatolian, including productive second-
ary formations, and residually in the other branches: an old example is the
word for ‘liver’ *h₂yēkʷ-usterity long root vowel in Gk hēpar, Lat. iēcur,
Avest. yākarə, but short in Ved. yāktṛ, and oblique stem Ved. yākān-ās. Gk
hēpat-os < *hēpə́ (t) -.

An old type of -u- stem appears in the words for ‘knee’, ‘wood’, ‘lifespan’:
nom. sg. n. *gōn-u *dōr-u h₂dōu, Gathic Avest. zānī dārū āiīu, Ved. jānu dāru
dīu, Gk gōnu dōru ou(ki) ‘not’. An apophonic stem variant *gēnw- dērw-
h₂ējw- accounts for Hitt. genu(-), Lat. genu(-), Slav. *dērv-o-, Gk ai(w)-ei; but
an alternative stem *gōn-ēw- *dēr-ēw- *h₂y-ēw- as in Goth. kniu (thematized),
Gathic Avest. yāo, gen. of āiīu, perhaps Hitt. genu-t, instr. of genu (unless zero
grade) seems very old and unlikely to be due to independent innovation. Zero
grades of these words also occur paradigmatically (Gk. gnysi), in composition
(Ved. jīu-bādh ‘kneeling’, Gk dry-tomos ‘cutting oak’), and in derivation
(*h₂ju-hén- ‘vigorous, young’ in Ved. yuvān-, Lat. iuuen-is).

A curious athematic suffix -it- marking elemental foodstuffs is found in
*mēl-it ‘honey’ (Gk mēlit-, Hitt. milit, Luv. mallit-, Gmc. mil-) and *sēp-it
‘wheat’ (sp.) (Hitt. šeppit, Gk αλφ-ιτ- with transferred epithet as root).

Ablauting secondary suffixes with two consonants (rarely three) are
common and continue to be productive in the dialects: abstracts *-men-,
*-wer/-wen-, agent *-ter-, verbal adjective *-ent-, possessive *-went-.

Feminine and abstract (collective): the derivational origin of the
feminine gender in Indo-European (Kuryłowicz 1964) is clear from the formal
relations (see p. 60); note also the Greek ‘two-ending’ adjectives, type m./f.
athēnatos, n. athēnaton). A suffix morpheme -(e)h₂ is found both in feminine
function (thematic type *senah₂-, ‘old (f.)’ in Gk hēnē, Ved. sānā, Lith. senā,
perhaps Lyc. lada ‘wife’; *swēkrūh₂- ‘husband’s mother’ in OCS svekry, Lat.
socrus, OHG swigar) and in abstract collective function (type Gk tomē
‘cutting’, neurā ‘cord of sinew’ beside neūron ‘sinew’, Luv. zidāh-[iša]
‘virility’ beside ziti- ‘man’ [earlier *zīda-, cf. PN Zidanza] and notably the
nominative-accusative neuter singular of the collective which functions as the
nominative-accusative neuter plural (Ved. yuga, Gk zugd, Lat. iuga,
Goth. juka ‘yokes’, Hitt. -a, Pal. -a/-a), still taking a singular verb in Greek, Old Avestan,
and Anatolian (Chapter 7, pp. 174f.). Whether the feminine and the collective
morpheme are ultimately the same is uncertain. More widely attested in both
feminine and collective function is *-ih₂-, with variant *yah₂-(-yeh₂-): Ved.

Thematic nominal suffixes: the simplest is *-o-, found in some very old
primary and secondary formations: masculines like *wēkʷ-os ‘wolf’,
*h₂(é)rēk-o-s ‘bear’, neuters like *jug-ō-m ‘yoke’, *wērō-m ‘work’,
adjectives like *sēn-o- ‘old’, *nēw-o- ‘new’. The line between primary and
secondary derivatives is hazy; if the first two animals have no recognizable
base, the ‘horse’ *ēkw-o-s is in all likelihood a derivative from ‘swift’ in Gk
ōkhys, Ved. āūs. Other more clearly secondary simple thematic suffixes are
the adjectives in accented -ō- marking possession of the base, and adjectives

The o-grade of the root is found in two very productive thematic types: barytone action nouns (type Gk tòmos ‘slice’) and oxytone agent nouns and adjectives (type Gk tomos ‘cutting, sharp’), both beside the Greek root tem- ‘cut’. We have zero grade in neuters of the type *jug-ôm ‘yoke’ (Hitt. iugan, Ved. yugām, Gk zygōn, Goth. juk, OCS igo), and in the second element of compounds like *ni-sd-ôs ‘nest’ (Ved. nidās, Lat. nidus, OIr. net, OEEng. nest) from *sed- ‘sit’; Gk neo- gn-ôs ‘newborn’ from *newo-gnha₁-ôs, with regular loss of laryngeal in this compound position.

A large number of thematic suffixes form secondary nouns and adjectives occurring widely in the historical languages, and commonly formed independently in each. Adjectival -jo- and -ijo- (partly continuing *-iho₂-) express relational notions. From *gʷow- ‘cow, ox’ we have Ved. gāvyā-, gāvia-, also gavyā-, (gāvyam ... śatām ‘consisting of 100 cows’, RV), Avest. gāoīa-, Gk hekatōm-boios ‘worth 100 cows’, Armenian kogi (*gʷowijo-) ‘butter, *coming from a cow’, OIr. ambe (*n-gwowijo-) ‘cowless man’. The last is semantically and sociologically identical with RV agōs (gen., ‘cowless man’) and Gk (Hesiod) andreōs abōúetō (gen. ‘cowless man’); though the formations differ slightly, we have an IE lexeme.

Verbal adjectives in accented *- tô- (and sometimes *-nô- in Indo-Iranian), with zero grade of the verbal root, are common in secondary derivation. An old example is *klu tô- ‘heard, renowned’ in Ved. śrutā-, Avest. sruta-, Gk klytós, Lat. in-clitus, OIr. ro-cloth, OHG Hlot-hari (PN, ‘whose army is renowned), Arm. lu ‘known’. Germanic *hlūdaz ‘loud’ < *klu tô- shows expressive lengthening. The suffix, not yet developed by the time of the separation of Anatolian where its function is expressed by *-ént-, marks semantically ‘the accomplishment of the notion of the object’: see Benveniste 1948.


A suffix *-zero-/-toro- marking the opposition of two notions is found in many branches, beginning with Anatolian (Hitt. nun-taras adv. gen. ‘of now’, whence ‘soon, right away’). ‘Other’ (of two) is Goth. anpar, Skt
ántaras, Lat. alter, beside 'other' (of more than two) in Skt anyás, Lat. alius, Gk állos, Goth. alís. For the syntactic distribution of the suffix, Gk deksiós – aristerós, skaiós – deksiterós, both 'right-left'), see Benveniste 1948.

The suffix -wo- is found notably in the words for 'alive; *g"ih₃-wo- in Ved. jivás, Lat. vivos, OIr. beo, Welsh byw, Goth. qius (with shortened -r-), and 'dead', *mr-wo- in OIr. marb, W. marw. The -t- of Lat. mortuos and Slavic *mirtvű probably is due to *mr-ò- 'destined to die, mortal'.

For the variants of suffix of instrument *-tro-, *-tlo-, *-dʰro-, *-dʰlo- (Lat. -trum, -c(u)lum, -brum, -bulum) see p. 39. An old example is the word for 'plough' *h₂arh₃-trom in Gk átrotron, Lat. arātrum (with analogical -ā-), OIr. arathar, Welsh aradr, Arm. arawr, Lith. árklas.

Noun inflection: the Indo-European noun was inflected for number, case, and in adjectives gender. Gender was an inherent property of substantives. The familiar three-gender system of masculine, feminine, neuter probably at an earlier date opposed just masculine (animate?) and neuter (inanimate?), with the feminine an original derivational rather than inflectional category. But already in the proto-language the three-way contrast had been attained. Ved. and Avest. gav-, Gk boüs, Lat. bós (IE*gwōw-) are both feminine 'cow, she-ox' and masculine 'bull, ox'. Number distinguished singular, plural, and dual. The singular distinguished at least eight, perhaps nine, cases; syncrétisms reduced this number in the plural and still more in the dual. The cases were nominative, vocative, accusative, genitive, dative, instrumental, locative, ablative, and probably allative. Not all the cases can be reconstructed with any confidence.

For consonant stems the clearly reconstructible endings were as in Table 2.10.

<table>
<thead>
<tr>
<th>Table 2.10 Indo-European athematic nominal endings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masculine/Feminine</strong></td>
</tr>
<tr>
<td><strong>Singular</strong></td>
</tr>
<tr>
<td>Nom.</td>
</tr>
<tr>
<td>Voc.</td>
</tr>
<tr>
<td>Acc.</td>
</tr>
<tr>
<td>Gen.</td>
</tr>
<tr>
<td>Dat.</td>
</tr>
<tr>
<td>Loc.</td>
</tr>
<tr>
<td><strong>Plural</strong></td>
</tr>
<tr>
<td>Nom.-Voc.</td>
</tr>
<tr>
<td>Acc.</td>
</tr>
<tr>
<td>Gen.</td>
</tr>
<tr>
<td>Loc.</td>
</tr>
<tr>
<td>Inst.</td>
</tr>
<tr>
<td><strong>Dual</strong></td>
</tr>
<tr>
<td>Nom.-Acc.</td>
</tr>
</tbody>
</table>
Table 2.11 Indo-European thematic nominal endings

<table>
<thead>
<tr>
<th></th>
<th>Masculine</th>
<th>Neuter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td>-os</td>
<td>-om</td>
</tr>
<tr>
<td>Voc.</td>
<td>-e</td>
<td></td>
</tr>
<tr>
<td>Acc.</td>
<td>-om</td>
<td>-om</td>
</tr>
<tr>
<td>Gen.</td>
<td>-os(y)o</td>
<td></td>
</tr>
<tr>
<td>Dat.</td>
<td>-öi &lt; -o-ei</td>
<td></td>
</tr>
<tr>
<td>Inst.</td>
<td>-ö &lt; -o-h1,?</td>
<td></td>
</tr>
<tr>
<td>Loc.</td>
<td>-oj</td>
<td></td>
</tr>
<tr>
<td>Abl. dialectal</td>
<td>-öt &lt; -o-h2at</td>
<td></td>
</tr>
<tr>
<td>Plural</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td>-öes &lt; -o-es</td>
<td>Ved. vıkäs ‘wolves’</td>
</tr>
<tr>
<td>Acc.</td>
<td>-ons</td>
<td></td>
</tr>
<tr>
<td>Gen.</td>
<td>-ömm &lt; -o-om</td>
<td></td>
</tr>
<tr>
<td>Inst.</td>
<td>-öjs</td>
<td></td>
</tr>
<tr>
<td>Loc.</td>
<td>-öjsu/i</td>
<td></td>
</tr>
<tr>
<td>Dual</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For thematic stems the reconstructible endings were as in Table 2.11.

The inflection of demonstrative and personal pronouns differed in certain critical respects from that of nouns, notably in an alternation of stem vowel -i- with the thematic vowel -ole-, as well as in certain endings.

Indo-European had a number of demonstrative pronoun stems of differing ages, some built on or coexisting with deictic particles. Thus *ki- (Lith. šis) and kó- (Luv. zaš, Hit. kāš) beside the particle *ke (OLat. hon-ce ‘hunc’, Hit. ki-nun ‘now’, Gk *ke-eno- in keînos). Others were *elono- *elowo-, and simply *elí-. In one stage of the proto-language, after the separation of Anatolian, and perhaps Italic and Celtic, the two stems *so(-) and *to- were fused into a suppletive paradigm, later IE *so *sah2 *tod (Ved. sā sā tād, Gk ho hē tó, Goth. sa so pata, Toch. B se sā te).

The relative pronoun stem *jo- (perhaps earlier *h₁jo-) is found in Vedic ya- and Avestan ya, Greek hós, Phrygian ios (ni), Celtiber. io(mui dat.). In Balto-Slavic this pronoun is suffixed to adjectives to form a definite.

The interrogative and indefinite pronoun stem *kʷo-/ *kʷi- (*kʷit- in adverbs) is found in all branches of the family; in Anatolian, Italic, Germanic and Balto-Slavic it forms the relative pronoun as well, perhaps via the indefinite function: IE yós kʷis, *yód kʷid ‘who-/whatsoever’, Gk hóstis hótti, Ved. yás cit, yác cit. Another interrogative stem *mo- is found residually in Anatolian and Tocharian.

Pronoun inflection differs from the nouns in the final dental for nominative-accusative singular neuter, the optional absence of -s in the
nominative singular masculine, and in characteristic special forms with inserted -sm- (f. -sj-) in the oblique cases. A partial paradigm of the interrogative indefinite pronoun is:

Nom. *kwis *kwid and *kwod: Lat. quis quid, quod, OLat. quo-i ‘qui’;
Avest. ciš, cit, kō, kas-, kat; OCS kš-to čh-to.
Gen. *kwol(os)j: Avest: kahibi kahibī, Gk teo, OCS česo.
Dat. *kwelosmōi
Loc. *kwešmi: Avest. kahnāi cahnāi, Celtiber. iomui, somui. OCS komu čemu, Umbr. esmei; Avest. kahmi cahmi, Umbr. esme, S. Pic. esme-n, esmi-n.

The cardinal numbers one to four were inflected adjectives, and five to ten were uninflected. (one): *sem- (Gk heis); *oj-no-, *oj-wo-, *oj-ko- (Lat. unus, OIr. den, Goth. ains; Avest. aewaa-; Skt ēka-); (two): *d(u)wo-; (three): *trej-es tri-h₂; (four): *kwetwor-es. Three and four had special feminine forms attested only in Indo-Iranian and Celtic, compounded with an old word for ‘woman’: Ved. tītras, catāras, Avest. tīsrō catagrō, OIr. teoir, catheoir, Welsh teir, pedeir. The dissimilation *tri-sr- to *ti-sr- is already Proto-Indo-European. (Five): pēnkʷe. (six) (k)sweks (seven) *sepm (eight) *h₃oktō(w) (nine) h₉něvyŋ. (ten): *dékm. One hundred is a derivative *(d)km-töm ‘that which makes 10 (decads)’. The ordinal numerals were suffixed by adjectival *-o- or -to-: for example, *tri-tó-, *sepm-ó-.

The personal pronouns show extensive and irregular allomorphy between the nominative (always stressed, syntactically emphatic) and oblique cases (with both tonic and enclitic forms). Their reconstruction poses particular problems, which cannot be addressed in a treatment of this scope. I give only the first and second singular forms in representative languages:

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Accusative (tonic)</th>
<th>(encl.)</th>
<th>Nominative</th>
<th>Accusative (tonic)</th>
<th>(encl.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greek</td>
<td>egō(n)</td>
<td>emé</td>
<td>sū, tūnē</td>
<td>sé</td>
<td>se</td>
</tr>
<tr>
<td>(Hom.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vedic</td>
<td>ahám</td>
<td>mám</td>
<td>mā</td>
<td>t(u)vām</td>
<td>t(u)vā</td>
</tr>
<tr>
<td>Hittite</td>
<td>uk</td>
<td>ammuk</td>
<td>-mu</td>
<td>zik</td>
<td>tuk</td>
</tr>
<tr>
<td>Gothic</td>
<td>ik</td>
<td>mik</td>
<td>ū</td>
<td>ū</td>
<td>ū</td>
</tr>
</tbody>
</table>

A reflexive stem *se(-), *swe(-) marked reference to the subject or the topic of the sentence; it originally referred to all three persons, but was in most branches restricted to third-person function. Non-reflexive pronominal third-person reference was normally by tonic or enclitic demonstrative pronoun.

Syntax
A syntactic description of Indo-European by the traditional method of describing the use of the different parts of speech, for example the different
cases of the noun or the moods of the verb, has been done very well in the classical manuals of Delbrück and Wackernagel, to which global reference is made. Here we confine our observations to certain syntactic rules for the simple sentence in Indo-European, relating to the order of constituents of the sentence.

**Phrase structure**: early Indo-European languages and presumably the protolanguage distinguished verbal sentences where the predicate is a verb (phrase), type NP + VP, Hitt. zik=wa URBARRA-aš kištat ‘you have become a wolf’, and nominal sentences, where the predicate is a noun phrase, type NP + NP, Ved. vṛko hṛ sās ‘for he is a wolf’, Gk. kretisōn gar basileus ‘for the king is stronger’, or other non-verbal constituent, like the adverb in Gk metà dè glaukōpis Athēnē ‘and grey-eyed Athena (was) with (them)’, Hitt. mšippa-LÚ-iš=wa=kan UL anda ‘Sippazitis (is) not in (it)’. It is not clear that the latter simply show gapped copula, since they may contrast with an overt copula sentence either stylistically or semantically: Hitt. LÚ.ULÚ.LU=ku GUD=ku UDU=ku ēši ‘whether it is man or ox or sheep’.

Indo-European was basically an SOV language, and verb-final was the ‘default’ position in most of the earlier languages and the earlier stages of the later languages; but the operation of a variety of movement rules served to complicate and obscure the word order picture in many of these same languages. The basic rule for the verb phrase was VP —> NP + V. The noun phrase was looser: we find NP —> Adj. + N or gen. + N in Hittite, where the reverse order may signal a nominal sentence (ištappuli=šet=a šulliš ‘but its stopper (is) of lead’), but we find regularly either order in Indo-Iranian and the classical languages.

To the prepositional phrases of the latter correspond in Hittite postpositional phrases (nēpišaš kattan ‘under heaven’), whose antiquity is to some extent confirmed by the accent of Greek ‘prepositions’ by ‘anastrophe’ (ommātōn ἀπο ‘from the eyes’) agreeing with that of cognates (Ved. ἀπα). When the object of the postposition was a pronoun, Old Hittite formed a possessive syntagma, with the object expressed by enclitic possessive adjective (katti=šši ‘with him’). This indicates a nominal origin for the pre-/postpositions.

The commonest movement rules in Indo-European are (1) movements of the type ‘move-WH’, typically interrogatives and relatives, to the ‘complementizer’ site at the head of the sentence to the left, and (2) ‘topicalization’ movement, to a site to the left of the complementizer site (Hale 1987). (See Figures 2.2 and 2.3.)

Ved. kāsyā brāhmāṇi jujusur yūvānāḥ
‘whose formulas do the youths enjoy?’

Ved. yó no dvēṣti, ādharah sās padiṣṭa
‘who hates us, he shall fall low’
NORMAL re only a single constituent or sub constituent appears to the left of the
WH-word, but it can be complex:

Ved. áher yátáram kám apásya indra  
‘What avenger of the serpent did you see, o Indra?’
Ved. ādevena mānasā yó riṣanyāti
‘who does harm with ungodly intent’.

Movement to the topicalization site or ‘fronting’ was a form of emphasis or focusing. The normally atonic finite verb itself was commonly so emphasized (Ved. āhann āhim ‘he/you slew the serpent’), and received its accent in this position. Sentence-initial was thus the marked position for the verb, beside unmarked sentence-final. When the verb was semantically closely linked or ‘compounded’ with an adverb (‘preverb’), the unmarked position of the (accented) preverb was before the final verb: ... P(...)V # #, optionally separated from it by a single constituent. The preverb could in such structures itself be fronted by movement to the topicalization site, yielding the marked order # # P ... V # #: Hitt. āppa=wa=mu parna tarna ‘let me go back home’. Fronted verb in topic was characteristic for certain discourse situations like text-initial cataphora and imperatives; compare the widespread tale beginning Skt āśūd rāja ‘there was a king’, Gk ἐσκε τις ... wanaśsōn, Lith. būvo karalius, OIr. boí ri, Russ. žyl byl korol’.

Extraposition of constituents to the right of the verb is another common feature. In Greek Οὔτιν ἐγὼ πῦματον ἐδομαί μετὰ ὁιὸς ἠτάρωσιν ‘Noman I will eat last among his comrades’, the prepositional phrase is extraposed to the right of the verb. The object has been moved to topic; the possessive reflexive pronoun ὁιὸς ‘his own’ refers back to the topic Οὔτιν rather than to the grammatical subject ἐγὼ, as normally.

Enclitics tend in IE (and many non-IE) languages to occupy what is loosely termed ‘second position’ in the sentence: the phenomenon known as Wackernagel’s Law after its famous codifier (1892). Recent work (Hale 1987) has shown that at least three distinct classes of Wackernagel’s Law clitics must be distinguished for Vedic, and that they end up in ‘second position’ for independent reasons. Thus in the phrase utá vā yó no marc- āyād ānāgasah ‘or also who would harm innocent us’, pronominal clitic nas (no) occupies second position before topicalization, and disjunctive clitic vā (and ca) second position defined after topicalization. A third type of clitics (smā and cit) is enclitic to the constituent they modify, and if that constituent is topicalized, it takes the clitic along to the head of the sentence: āśmānām cid yē bibhidūr vācobbhīḥ ‘who split even rock with words’. Such different syntactic history for the position of the different Wackernagel’s Law clitics will account, for example, for the consistent precedence of Gk de or te (and their Old Irish cognates -d- and -ch-, Watkins 1963) over pronominal clitics.

Relative Clause Formation: the testimony of Hittite, Vedic and Greek is unequivocal in pointing to IE correlative relative clauses, usually in the order Relative Clause + Main (Matrix) Clause, with relative pronoun fronted to the complementizer site by WH-movement. Either or both clauses may begin
with the sentence introductory particles e.g. Hitt. *nu*, which precede both the fronted WH-word, and any topicalized constituent before it, and thus do not ‘count’ for fronting. The same holds for Vedic introductory particles (adverbs) e.g. *atha* and the pronoun *sā*, and Greek introductory particles (adverbs), e.g. *éntha*.

In Hittite such sentences are semantically indefinite relatives: *kuiš paprizzi nu apaš-pat 3 GIN KÜ.BABBAR pāi* ‘Whoever commits a nuisance pays 3 shekels of silver’. In the case of semantically definite relatives some constituent must be topicalized, that is moved into the site preceding the WH-word in the complementizer site. The finite verb is so moved in Hittite *paprizzi kuiš 3 GIN KÜ.BABBAR pāi* ‘He who commits a nuisance pays 3 shekels of silver’.

There follow thematically similar correlative relative sentences in Hittite, Vedic and Greek, to point up the similarity: they may function as syntactic equations.

Hittite

*nu tarhzi kuiš nu apaš KIR₄.TAB.ANŠE epzi*

ptc wins WH ptc he bridle takes

‘And he who wins, (he) takes the bridle’

**1.1**K₄S₄E tarhzi kuiš 1 MANA KÜ.BABBAR pianzi (prob. tarhzi kuiš ḫuyatallaš)

‘the runner who wins, (to him) 1 mina of silver they give’,

Vedic

*sa yo na ujjesyati sa pratamah somasya pāsyati*

‘(ptc) he who will win, he first will drink of the soma’

*sa yo na ujjesyati tasya idam bhavisyati*

‘(ptc) he who will win, his this will be’

Greek

*hos nyn orkheston . . . atalotata paizei to tode k[ . . .)

‘he now of the dancers who most sportively plays, his (is) this k.’

*hoppóters dė ke nikēsēi . . . gynaika te oīkad’ agēsthō*

‘whichever should win . . . let him take the woman home’

See Watkins 1976.

Early IE languages are notable for the supple syntax of the participle, which may ‘transform’ finite verbs into noun phrases. With the last Greek relative clause above compare *tōi dė ke nikēsanti phītē keklēsēi ákoītis* ‘to the one having-won (= who should win) you will be called dear wife’. Note the presence of the enclitic particle *ke* (Hitt. *-kan*) in both. The performative statement, speech act of pledge and self-engagement expressed by the syntactic minor rule of the demonstrative in Vedic *ayām te asmi* ‘I here(by)
am yours', is transformed to a participle phrase in the complex má mám imám táva sántam . . . ní gärit ‘Let him not swallow . . . me, being here(by) yours’. The participle reinforces a performative statement: ‘I hereby pledge myself to you, Atri. Don’t let Svarbhānu destroy me.’ Just so the finite verb of existence in another speech act, that of the confessional formula in Hitt. ešziy=at iyawan=at ‘It is. We did it’ is transformed to the participle in the personal confession ašän=at iyanun=at (lit.) ‘It (is) being. I did it’. In the syntactic use of Hitt. ašant- here we can glimpse the background of Latin sōns, sōnitis ‘guilty’, the old present participle of esse ‘be’.

Note
For invaluable criticism, comments, and suggestions I am deeply indebted to the editors, and to Bernard Comrie, Benjamin Fortson, Andrew Garrett, Mark Hale, Eric Hamp, Ioannis Ikonomou, Stephanie Jamison, Jay Jasanoff, Joshua Katz, Craig Melchert, Steve Peter, Charles Reiss and Bert Vaux.

References
Schindler, Jochem (1975) ‘Zum Ablaut der neutralen s-Stämme des indogerman-