2. LETTERS AND SOUNDS

2.1 THE ALPHABETS OF MODERN INDO-EUROPEAN

2.1.1. Indo-European does not have an old writing system to be revived with. In the regions where PIE speakers dwelled four thousand years ago, caves and stones probably still keep some ancient pictographic writings, composed of logograms (graphemes) that represent a morpheme or a whole word, as did Egyptian hieroglyphic logographs.

2.1.2. The Indo-European dialects have adopted different alphabets during the last millennia, and all of them should be usable today – although the main alphabet for today's European Union is clearly the Latin one. This is a summary table of Proto-Indo-European phonemes and their regular corresponding letters in MIE alphabets: Greek, Latin, Cyrillic, Perso-Arabic and (alphasyllabary) Devanāgarī.

A. VOWELS AND VOCALIC ALLOPHONES

Phoneme	Greek	Latin	Persian	Armenian	Cyrillic	Devan.
[a]	Α α	A a		Ա ա	A a	अ
[e]	Ε ε	E e		Ե ե	E e	ए
[o]	O o	O o		Ո ո	O o	ओ
[a:]	Ᾱ ᾱ	Ā ā	١	Ա ա	Ā ā	आ
[e:]	Ē ē̆	Ē ē		Է է	Ē ē	ऎ
[o:]	Ω ω	Ō ō		Ո ո	Ō ō	औ

[i]	Ι ι	I i		Ի ի	И и	इ
[i:]	Ῑ ῑ	Ī ī	ى	Ի ի	Й й	ई
[u]	Υ υ	U u		Ւ ւ	У у	उ
[u:]	Ῠ ῡ	Ū ū	و	Ւ ւ	Ӯ ӯ	ऊ

[r̥]	Ρ ρ	R r	ر	Ռ ռ	Р р	ऋ (ॠ)
[l̥]	Λ λ	L l	ل	Լ լ	Л л	ऌ(ॡ)
[m̥]	Μ μ	M m	م	Մ մ	М м	म
[n̥]	N ν	N n	ن	Ն ն	Н н	ण

NOTE. The underdot diacritic might be used to mark the sonorants, as R̥ r̥, L̥ l̥, N̥ n̥, M̥ m̥; usually, however, sonorants appear between consonants, so it is not necessary to mark them, *v.i.*

B. CONSONANTS AND CONSONANTAL SOUNDS

Phoneme	Greek	Latin	Persian	Armenian	Cyrillic	Devan.
[p]	Π π	P p	پ	Պ պ	П п	प
[b]	Β β	B b	ب	Բ բ	Б б	ब
[bʰ]	Βη βη	Bh bh	بﻪ	Բh բh	Бх бх	भ
[t]	Τ τ	T t	ت	Տ տ	Т т	त
[tʰ]	Θ θ	Th th	تﻪ	Թ թ	Тх тх	थ
[d]	Δ δ	D d	د	Դ դ	Д д	द
[dʰ]	Δη δη	Dh dh	ذ	Դh դh	Дх дх	ध
[k]	Κ κ	K k	ک	Կ կ	К к	क
[kʰ]	Χ χ	Kh kh	کﻪ	Ք ք	Кх кх	ख
[g]	Γ γ	G g	گ	Գ q	Г г	ग
[gʰ]	Γη γη	Gh gh	گﻪ	Գh qh	Гх гх	घ
[kʷ]	Ϙ ϙ	Q q	ق	Ք ք	Къ къ	क
[gʷ]	Γγ γγ	C c	غ	Ղ ղ	Гъ гъ	ग
[gʷʰ]	Γγη γγη	Ch ch	غﻪ	Ղh ղh	Гъх гъх	घ
[h]	Η η	H h	ه	Հ h	Х х	ह

[j]	Ι ι (J j)	J j	ژ/ی	Յ յ, Ի ի	Й й / J j	य
[w]	Ϝ ϝ	W w	و	Ւ ւ	У у (W w)	व
[r]	Ρ ρ	R r	ر	Ռ ռ	Р р	र
[l]	Λ λ	L l	ل	Լ լ	Л л	ल
[m]	Μ μ	M m	م	Մ մ	М м	म
[n]	Ν ν	N n	ن	Ն ն	Н н	न
[s]	Σ σ ς	S s	س	Ս ս	С с	स

2.1.2. The Latin Alphabet used for Modern Indo-European is similar to the English, which is in turn borrowed from the Late Latin *abecedarium*. We also consider some digraphs part of the alphabet, as they represent original Proto-Indo-European sounds, in contrast to those digraphs used mainly for transcriptions of loan words.

NOTE. The Latin alphabet was borrowed in very early times from the Greek alphabet and did not at first contain the letter G. The letters Y and Z were introduced still later, about 50 BC.

The names of the consonants in Indo-European are as follows - **B**, **be** (pronounced *bay*); **Bh**, **bhe** (*b^{h}ay*); **C**, **ce** (*g^{w}ay*); **Ch**, **che** (*g^{wh}ay*); **D**, **de** (*day*); **Dh**, **dhe** (*d^{h}ay*); **F**, **ef**; **G**, **ge** (*gay*); **Gh**, **ghe** (*g^{h}ay*); **H**, **ha**; **K**, **ka**; **L**, **el**; **M**, **em**; **N**, **en**; **P**, **pe**; **Q**, **qu**; **R**, **er**; **S**, **es**; **T**, **te**; **V**, **ve**; **W**, **wa**; **X**, **eks**; **Z**, **zet**.

2.1.3. The Latin character **C** originally meant [g], a value always retained in the abbreviations *C.* (for *Gaius*) and *Cn.* (for *Gnaeus*). That was probably due to Etruscan influence, which copied it from Greek Γ, *Gamma*, just as later Cyrillic Г, *Ge*.

NOTE 1. In early Latin C came also to be used for [k], and K disappeared except before in a few words, as Kal. (Kalendae), Karthago. Thus there was no distinction in writing between the sounds [g] and [k]. This defect was later remedied by forming (from C, the original [g]-letter) a new character G. Y and Z were introduced from the Greek about 50 B.C., and occur mainly in loan words in Modern Indo-European.

NOTE 2. In Modern Indo-European, **C** is used (taking its oldest value) to represent the Indo-European labiovelar [g^w] in PIE words, while keeping its different European values − [k], [ts], [s], [θ], [ʃ], etc. – when writing proper names in the different modern IE languages.

2.1.4. The Latin [w] semivowel developed into Romance [v]; therefore V no longer adequately represented [w] and the Latin alphabet had to develop an alternative letter. Modern Indo-European uses **V** mainly for loan words, representing [v], while **W** is left for the consonantal sound [w].

NOTE. V originally denoted the vowel sound [u] (Eng. oo), and F stood for the sound of consonant [w] (from Gk. ϝ, called *digamma*). When F acquired the value of our [f], V came to be used for consonant [w] as well as for the vowel [u].

2.1.5. The letter **I** stood for the vowel [i], and was also used in Latin (as in Modern Greek) for its consonant sound [j]. **J** was originally developed as a swash character to end some Roman numerals in place of I; both I and J represented [i], [iː], and [j]. In MIE, **J** represents the semivowel [j]. In the Latin script, **Y** is used to represent the vowel [y] in foreign words.

NOTE. That [j] value is retained in English J only in foreign words, as *Hallelujah* or *Jehovah*. Because Romance languages developed new sounds (from former [j] and [ɡ]) that came to be represented as I and J, English J (from French J), as well as Spanish, Portuguese or Italian J have

sound values quite different from [j]. Romanisation of the sound [j] from different writing systems (like Devanagari) as Y – which originally represented in Latin script the Greek vowel [y] –, due to its modern value in English, French or Spanish, has spread a common representation of [j] as Y in Indo-European studies, while J is used to represent other sounds.

2.1.6. The consonant cluster [ks] was in Ancient Greece written as X (Chi) in Western Greek, Ξ (Xi) in Eastern Greek dialects. In the end, **X** was standardized as [kʰ] ([x] in modern Greek), while Ξ represented [ks]. In the Latin script, the X stands for [ks], as in English or Latin, whereas in the Cyrillic alphabet it stands for [h] (and aspiration), as well as for [x] in foreign words.

NOTE. The Etruscans took over X from Old Western Greek, therefore it stood for [ks] in Etruscan and then in Latin, and also in most languages which today use an alphabet derived from the Roman, including English. Cyrillic X was taken with its standard Greek value [x], but is also used as [h] in those languages that need it; as, Macedonian, and Bulgarian and Serbian dialects.

2.1.7. As in Ancient and Classic Greek, in the Greek alphabet **X** stands for [kʰ], **Φ** for [pʰ], and **Θ** for [tʰ].

NOTE. Because of its use in Modern Greek, they also represent (mainly foreign) [x], [f] and [θ].

2.1.8. **Ē** represents [ɛː] in the Greek alphabet, because **H** was originally used in most Greek dialects to represent the sound [h], and it is therefore used with this value in IE writings, as well as to mark aspirated phonemes.

NOTE. For more on the problem of historical Eta and its representation in the modern Greek alphabet, see <http://www.tlg.uci.edu/~opoudjis/unicode/unicode_aitch.html>.

2.2. CLASSIFICATION OF SOUNDS

2.2.1. The Vowels are short [a], [e], [i], [o], [u], written **a**, **e**, **i**, **o**, **u**, and long [aː], [eː], [iː], [oː], [uː], written **ā**, **ē**, **ī**, **ō**, **ū**, respectively. The other sounds are Consonants.

The Modern Indo-European Diphthongs proper are [ei̯], [oi̯], [ai̯], written **ei**, **oi**, **ai**, and [eu̯], [ou̯], [au̯], written **eu**, **ou, au**. In these diphthongs both vowel sounds are heard, one following the other in the same syllable.

NOTE. For the so-called *long diphthongs* [eːi̯], [oːi̯], [aːi̯], written **ēi**, **ōi**, **āi**, and [eːu̯], [oːu̯], [aːu̯], written **ēu**, **ōu**, **āu**, which remained only in Indo-Iranian, Greek and partly in Baltic languages, Schulze (1885) interpreted a regular correspondence of the type **āi**/**ā**/**ī**, which came respectively from the full grade of the long diphthong, the full grade before consonant (where the second

element was lost), and the zero-grade (a contraction of *schwa* with the semivowel). Martinet (1953) proposed that laryngeals were behind those long diphthongs.

In any case, in the languages in which they are retained, *long diphthongs* have not a longer duration than normal diphthongs; phonologically they are equivalent, as Vedic and Greek metric shows. After Adrados (1995), "The difference, therefore, is not on the duration of the group, but on the relative duration of their components; in other words, for example **ei** and **ēi** have the same phonological duration (they are long, as opposed to a brief vowel), but in **ei** both elements have approximately the same duration, whereas in **ēi** the duration of **i** is perceptibly shorter than **e**. Because of that, the name 'long first element diphthongs' is more appropriate to refer to these phonemes". Cf. Allen (1976) for an analysis of these diphthongs.

Strictly speaking, phoneticians do not consider the so-called rising diphthongs, [je], [jo], [ja], [jeː], [joː], [jaː], nor [we], [wo], [wa], [weː], [woː], [waː], as diphthongs proper, but rather sequences of glide and vowel.

NOTE. Whilst most Indo-Europeanists differentiate between sequences of approximant and vowel (rising diphthongs) from true falling diphthongs in their transcriptions, i.e. writing [je] (from [i]+[e]) but [ei] or [ei] (from [e]+[i]), some use a different approach, considering all of them combinations of vowel plus glide or glide plus vowel, i.e. writing [je] and [ej], or [i̯e] and [ei̯].

Therefore, there are no real triphthongs. The formations usually called triphthongs are [jei̯], [joi̯], [jai̯], [jeu̯], [jou̯], [jau̯], as well as [wei̯], [woi̯], [wai̯], [weu̯], [wou̯], [wau̯]; and none can be named strictly triphthong, as there is a consonantal sound [j] or [w] followed by a diphthong. The rest of possible formations are made up of a diphthong and a vowel.

2.2.2. Consonants are either voiced (sonant) or voiceless (surd). Voiced consonants are pronounced with vocal cords vibration, as opposed to voiceless consonants, where the vocal cords are relaxed.

a. The voiced consonants are [b], [d], [g], [gʷ], [l], [r] and [ɾ], [m], [n], [z], [j], [w].

b. The voiceless consonants are [p], [t], [k], [kʷ], [s].

c. The digraphs **bh**, **dh**, **gh** and **ch** represent the Indo-European voiced aspirates proper, i.e. [bʰ], [dʰ], [gʰ], [gʷʰ], whereas **ph**, **th**, and **kh** represent voiceless aspirates [pʰ], [tʰ], [kʰ], mostly confined to words of Greek origin, as well as foreign [ɸ], [θ] and [x], respectively.

d. The consonants [r], [l], [m], [n], and the semivowels [j] and [w], can function both as consonants and vowels, i.e. they can serve as syllabic border or center.

NOTE. There is a clear difference between the vocalic allophones of the semivowels and those of the sonants, though: the first, [i] and [u], are very stable as syllabic center, while [r̥], [l̥], [m̥], [n̥], aren't, as they cannot be pronounced more opened. Hence the big differences in their evolution, depending on the individual dialects.

2.2.3. The Mutes are classified as follows:

	voiceless	*voiced*	*aspirated*
labials	p	b	bʰ
dentals	t	d	dʰ
velars	k	g	gʰ
labiovelars	kʷ	gʷ	gʷʰ

Labialized velars or Labiovelars [kʷ] (written **q**), [gʷ] (written **c**), [gʷʰ] (written **ch**), are pronounced like [k], [g], [gʰ] respectively, but with rounded lips.

NOTE 1. German Neogrammarians reconstructed a fourth series of phonemes, the voiceless aspirates *pʰ, *tʰ, *kʰ, to explain some irregularities in the outputs of the voiceless row. Most Indo-Europeanists reject this fourth independent row of phonemes, and findings of Indo-Iranian, Armenian and Greek have been explained as 1) expressive in origin, 2) contact of a voiceless with a laryngeal phoneme, and 3) effect of a prior **s**. For support of the fourth row, cf. Szemerényi (1985).

NOTE 2. The modern mainstream Proto-Indo-European reconstruction, that accepts only these two rows of velars as the most logical PIE phonetic system, has been confronted with the question of the actual existence of the groups [kw], [gw], and [gʰw], different from (and similar or identical in their dialectal outputs to) labialized [kʷ], [gʷ], and [gʷʰ]. A distinction between both is often found, though, whether an independent row of palatalized velars is accepted or not; as, **kwōn**, *dog*, **ekwos**, *horse*, **ghwer**-, *wild*, **kweidos**, *white*, **kwet**-, *cook* (cf. O.Ind. *kwathati*), **tekw**-, *run*, etc. which might be found reconstructed as *__qōn__, *__eqos__, *__cher__-, etc. For a defence of such unified forms, see e.g. Jussi Halla-aho at <http://ethesis.helsinki.fi/julkaisut/hum/slavi/vk/halla-aho/problems.pdf>.

2.2.4. The so-called Liquids are **l**, which represents the alveolar lateral approximant [l], an **r**, pronounced in PIE and in most modern IE languages (at least occasionally) as alveolar trill [r], today often allophonic with an alveolar tap [ɾ], particularly in unstressed positions. These sounds are voiced.

NOTE. About Indo-European **r** and **l**, cf. Ban'czerowski (1968).

For foreign words, the group **rh** represents an [h] sound coming just after the [ɾ], mainly in words of Greek origin. Other groups include **rr**, the alveolar trill [r], and its aspirated counterpart **rrh**. The palatal lateral approximant [ʎ] shall be represented as **lj**.

2.2.5. The Nasals are labial [m], written **m**, and dental [n], written **n**. These are voiced. The velar nasal [ŋ] – as *ng* in English *sing* – could have existed in IE as allophone of [n] before velars.

NOTE. Erhart (1970) reconstructs three nasals, N, M1 and M2, this one a fricative seminasal with which he explains the results of alternating m and w in some suffixes and roots; as, -***ment***-/-***went***-, ***men***-/***wen***-, etc. It was left unexplained, though, under which conditions did it change.

The palatal nasal [ɲ] of foreign words (similar to the [n] sound in English *onion* or *canyon*) is represented by the pair **nj**.

2.2.6. The Fricatives are voiceless [s] and voiced [z], **z** being usually the output of **s** before voiced consonants.

NOTE. [z] was already heard in Late Proto-Indo-European, as a different pronunciation (allophone) of [s] before voiced consonants, as can be clearly seen in PIE **nizdos** (for **nisdos**), *nest*, which comes from PIE roots **ni**-, *down*, and zero-grade -***sd***- of **sed**, *sit*. Because of that it is preferred to write **s** for [z] in MIE.

It is also possible to write voiceless and voiced pairs from foreign words: labiodentals, **f** and **v**; dorsal voiceless **h** and [x], written **kh**; and postalveolar [ʃ] as **sh** and [ʒ] as **zh**. Possible groups include **ks**, **ts**, **dz**, **tsh** (for [tʃ]), **dzh** (for [dʒ]), etc.

2.2.7. The Semivowels are usually written **j**, and **w**. These are voiced.

NOTE. Some authors make a distinction between consonantal [j], [w], and vocalic [i], [u]. Actually, however, both appear as TIT and EYE (where T = consonant, E = vowel, I = i,u, Y = j,w), and never as TYT or EIE. Against it, see Schmitt-Brandt (1967) and Szemerényi (1985) and Mayrhofer (1986).

2.2.8. Gemination appears in phonemes whose duration is long enough to be perceived – their implosion and explosion, both audible – as distributed in two syllables. They existed in PIE: in stops, as **appās**, **attās** (and **tātā**), *dad*, **pappājō**, *eat*, or **kakkājō**, *shit*; in nasals, as **anna**-, **ammā** (and **mammā**), *mother, mum*; in liquids, as **bōullā**, *buble*; and in the sibilant, as **kussō**, *kiss*.

NOTE. They appear mostly in words of expressive origin, children vocabulary, onomatopoeia, etc., which makes it more likely that PIE inherited gemination as an expressive resource, different from its central phonological system; a resource that was retained for a long time by most IE languages as a recurrent possibility.

2.2.10. A synoptic table of the Proto-Indo-European phonetic system:

	Labials	Coronals	Palatovelars	Velars	Labiovelars	Gutturals
Voiceless	**p**	**t**	*kʲ*	**k**	**kʷ**	
Aspirated	*pʰ*	**tʰ**		**kʰ**		
Voiced	**b**	**d**	*gʲ*	**g**	**gʷ**	
Aspirated	**bʰ**	**dʰ**	*gʲʰ*	**gʰ**	**gʷʰ**	
Nasals	**m**	**n**				
Fricatives		**s , z**				**h**, *H*
Liquids		**r , l**				
Approximant	**w**		**j**			

NOTE 1. The existence of a distinctive row of PIE 'satemizable' velars, the so-called palatovelars, has been the subject of much debate over the last century of IE studies. Today a majority of modern scholars support only two types of velars in Late PIE – generally Velars and Labiovelars, although other solutions have been proposed, see Appendix II.2.

Palatovelars could be found in PII, though, and are to be represented with Ḳ ḳ, Ǵ ǵ, Ǵh ǵh.

The support of German Neogrammarians to the 'palatals' in Proto-Indo-European, as well as its acceptance in Brugmann's *Grundriß* and Pokorny's *Wörterbuch*, extended the distinction to many (mainly etymological) works, which didn't deal with the phonological reconstruction problem directly. As Adrados (2005) puts it, about the *standard* [=Brugmannian] theories nowadays, "Indo-Europeanists keep working on a unitary and flat PIE, that of Brugmann's reconstruction. A reconstruction prior to the decypherment of Hittite and the study of Anatolian! This is but other proof of the terrible conservadurism that has seized the scientific discipline that is or must be Indo-European linguistics: it moves forward in the study of individual languages, but the general theory is paralised. It is sad when our students go to Germany and come back brainwashed".

NOTE 2. The cover symbol *H*, traditionally *ə, stands for the uncertain Late PIE ouput of the (for Middle PIE) reconstructed laryngeal phonemes h_1, h_2, h_3, which had evolved differently already by the time when Late PIE and Proto-Anatolian were independent languages. There is no

consensus as to what these phonemes were like, or how many of them (if any) survived into Late PIE, but it is widely accepted that PIH *h₂ was probably uvular or pharyngeal, and that *h₃ was labialized. Commonly cited possibilities are ʔ, ʕ, ʕʷ and x, χ~ħ, xʷ. See Appendix II.3.

2.3. SOUNDS OF THE LETTERS

2.3.1 The following pronunciation scheme is substantially that used by the common Europe's Indo-European speakers in ca. 2500 BC, when the laryngeal phonemes had already disappeared, having coloured following vowels, and lengthened preceding ones.

NOTE. MIE cannot permit dialectal phonetic differences, whether vocalic or consonantal – like Grimm's Law effects in PGmc. consonants, already seen –, because a homogeneous pronunciation system is especially needed when targeting a comprehensible common language. Some differences exist in sister dialects Hellenic, Aryan and Anatolian, though.

2.3.2. Vowels:

ā as in *father*	**a** as in *idea*
ē as in *they*	**e** as in *met*
ī as in *meet*	**i** as in *chip*
ō as in *note*	**o** as in *pot*
ū as in *rude*	**u** as in *put*

NOTE 1. Following the mainstream laryngeals' theory, Proto-Indo-Hittite knew only two vowels, *e and *o, while the other commonly reconstructed vowels were earlier combinations with laryngeals. Thus, short vowels PIE *a* < *h₂e; *e* < *(h₁)e; *o* < *h₃e, *(h₁)o; long vowels *ā* < *eh₂; *ē* < *eh₁; *ō* < *eh₃, *oh. The output of *h₂o in Late PIE was either *a* or *o*, after the different schools. Short and long vowels *ĭ* and *ŭ* were just variants of the semivowels PIH *j and *w.

NOTE 2. The sonants may have been lengthened too (usually from older sequences of sonant + laryngeal, or because of compensatory lengthenings), especially in the conjugation of verbs, giving thus [r̥ː], [l̥ː], [m̥ː], [n̥ː], written as r̄, l̄, m̄, n̄. The semivowels can also have a prolonged pronunciation, giving allophones **ij** and **uw**. For more details on this see § 2.7.2.

2.3.3. Falling Diphthongs and equivalents in English:

ĕi as in *vein*	**ĕu** *e* (*met*) + *u* (*put*)
ŏi as in *oil*	**ŏu** as *ow* in *know*
ăi as in *Cairo*	**ău** as *ou* in *out*

There are several ways to generate breathy-voiced sounds:

1. To hold the vocal cords apart, so that they are lax as they are for [h], but to increase the volume of airflow so that they vibrate loosely.

2. To bring the vocal cords closer together along their entire length than in voiceless [h], but not as close as in modally voiced sounds such as vowels. This results in an airflow intermediate between [h] and vowels, and is the case with English intervocalic [h].

3. To constrict the glottis, but separate the arytenoid cartilages that control one end. This results in the vocal cords being drawn together for voicing in the back, but separated to allow the passage of large volumes of air in the front. This is the situation with Hindustani.

2.3.4. Consonants:

1. **b**, **d**, **h**, **l**, **m**, **n**, are pronounced as in English. **n** might also be pronounced as guttural ŋ[] when it is followed by another guttural, as in Eng. *sing* or *bank*.

3. **p**, **k**, **t** are plain as in Romance, Slavic or Greek languages, not aspirated as in English; **t** is never pronounced as *sh*, as in English *oration* or *creation*.

4. **g** always as in *get*. It had two dialectal pronunciations, the common simple velar and the 'eastern' (later generalized in PII) palatovelar. Compare the initial consonants in *garlic* and *gear*, whispering the two words, and it will be observed that before *e* and *i* the *g* is sounded farther forward in the mouth (more 'palatal') than before *a* or *o*. That is what we represent as **ǵ**, similar to **ḱ**, pronounced as *k* in *key*, compared to *c* in *cold*.

5. **c** is pronounced similar to [g] but with rounded lips. Compare the initial consonant in *good* with *get* to feel the different articulation. The voiceless **q** is similar to [k] but pronounced with rounded lips; as *c* in *cool*, compared to *c* in *car*.

6. **bh**, **dh**, **gh**, **ch** are uncertain in sound, but the recommended pronunciation is that of the Hindustānī's "voiced aspirated stops" *bh, dh, gh*, as they are examples of living voiced aspirates in an Indo-European language (see note to the left).

7. The voiceless aspirated **ph**, **kh**, **th**, frequently of Hellenic origin, are pronounced very nearly like English word-initial *p, k, t*, as in *pen, ten, Ken*. Their sound is also described as equivalent to *p+h, t+h,*

k+h, i.e. to the corresponding mutes with a following breath, as in *loop-hole*, *hot-house*, *block-house*.

8. **j** as the sound of *y* in *yes*, never the common English [dʒ], as *j* in *join*; **w** as *w* in *will*.

9. Indo-European **r** was probably slightly trilled with the tip of the tongue (still common today in many IE languages), as in Scottish English *curd*. Another pronunciation is common today among modern IE languages, and was possibly heard in PIE, the alveolar tap [ɾ], pronounced like the intervocalic *t* or *d* in American or Australian English, as in *better*.

10. **s** is voiceless as in *sin*, but there are situations in which it is voiced, depending on the surrounding phonemes. Like the aforementioned [r], modern speakers will probably pronounce [s] in slightly different ways, but this should not usually lead to misunderstandings, as there are no proper IE roots with original [z] or [ʃ], even though the former appeared in some phonetic environments, v.s.

11. Doubled letters, like **ll**, **mm**, **tt**, etc., should be so pronounced that both members of the combination are distinctly articulated.

12. Regarding foreign sounds:

- **kh** might represent [x], whether strong, with '*ach-laut*', such as *kh* in Russian *Khrushenko,* or *ch Chanukah*, or soft, with '*ich-laut*', such as *ch* in German *Kirche* or *Lichtenstein*. Also, **th** might be pronounced as English *th* in *thing*, and **dh** as *th* in *this*.
- **z**, **v**, **f**, **sh**, are pronounced as in English.
- **zh** is pronounced as *s* in English *leisure*.
- **tsh** corresponds to English *ch* in *chain*, and **tzh** to *j* in *jump*.

2.4. SYLLABLES

2.4.1. In many modern languages, there are as many syllables in a word as there are separate vowels and diphthongs. This is not exactly so in Modern Indo-European. It follows, indeed, this rule too:

swe-**sōr**, *sister*, **skrei**-**bhō**, *write*, **ne**-**wā**, *new*, **ju**-**góm**, *yoke*.

NOTE. The semivowels are always written **j** and **w**. So in **trejes**, *three*, **newos**, *new*, **dn̥ghwās** ['dngh-waːs], *languages*, etc.

2.4.2. Indo-European has also consonant-only syllables. It is possible to hear similar sound sequences in English *cattle* or *bottom*, in German *Haben*, in Czech *hlt*, Serbian *srpski*, etc. In this kind of syllables, it is the vocalic sonant [r̥], [l̥], [m̥], or [n̥] – constrained allophones of [r], [l], [m], [n] –, the one which functions as syllabic centre, instead of a vowel proper:

kr̥-di, *heart*, **wl̥-qos**, *wolf*, **de**-**km̥**, *ten*, **nō**-**mn̥**, *name*.

NOTE 1. Words derived from these groups, represented TRT (where T = consonant, R = sonant), are unstable and tend to add auxiliary vowels before or after the sonants, i.e. T°RT or TR°T. Because of that, their evolutions differ greatly in modern IE languages. For example, **dn̥ghwā**, *language*, evolved as [ˈdᵊn-gʰwa:] into PGmc. *tung(w)ō*, and later English *tongue* or German *Zunge*, while in archaic Latin it was pronounced *dingwa*, and then the initial *d* became *l* in Classic Latin, written lingua, which is in turn the origin of Modern English words "linguistic" and "language". For **wl̥qos** (cf. Ved. *vr̥kas* < PII wr̥kas), it evolved either as [ˈwᵊl-kwos], later into PGmc. **wulxʷaz* (cf. O.H.G. *wolf*) or BSl. **wilkas* (cf. O.C.S. *vъlkъ*) or as [ˈwlᵊ-kwos], which gave Common Greek **wlukwos* (cf. Gk. *lykos*), Ita. **wlupos* (cf. Lat. *lupus*).

NOTE 2. Apart from the common scheme TRT, another, less stable scheme has been proposed for a common PIE, a certain TRE (where E = vowel); as, PIE **gʷ°nā*, for MIE **cenā**, *woman*, or **k°rwos*, for **kerwos**, *deer*, etc. – conventionally, the symbol ° under the sonant is placed before it in these schemes. Nevertheless, it is commonly accepted that Late PIE dialects did in fact add an auxiliary vowel to this sequence at early times, probably before the first dialectal split: as early Indo-Iranian and Balto-Slavic dialects show, vocalization of TRE had already happened when TRT hadn't still been vocalized, i.e. T°RE > TERE. Also, many dialects show a common vocalization in [a] for the sonant in some TERE groups, while showing different outputs (even non-vocalization) for TRT. Therefore, even if this theory might make some irregularities fit into a common Late (or Middle) PIE sound, it is not applicable to those early PIE dialectal words, whose vocalization might be inferred using the comparative grammar. Some TRE groups persisted in early IE dialects, though, often from older sequences that included laryngeals, and they are kept in MIE.

2.4.3. In the division of words into syllables, these rules apply:

1. A single consonant is joined to the following vowel or diphthong; as **ne**-**wos**, **me**-**dhjos**, etc.

2. Combinations of two or more consonants (other than the vocalic ones) are regularly separated, and the first consonant of the combination is joined to the preceding vowel; as **ok-tōu**, *eight*, **pen-qe**, *five*, etc. but **a-gros**, *field*, **sqa-los**, *squalus*.

3. In compounds, the parts are usually separated; as **a-pó-sta-tis**, *distance*, from **apo** + **statis**; or **am-bhí-qo-los**, *servant*, from **ambhí** + **qolos.**

2.4.4. The semivowels [j], [w] are more stable than sonants when they are syllable centres, i.e. [i] or [u]. However, when they are pronounced *lento*, they give the allophones (or *allosyllables*) **ij**, **uw**. Examples of alternating forms in PIE include **médhijos** (cf. Lat. *medius*), and **medhjos** (cf. O.Ind. *mádhjas* or Gk. μέσσος); **dwōu**, *two* (cf. Goth. *twai*, Gk. δω-,), and **duwōu** (cf. O.Ind. *duva*, Gk. δύω < *δύϝω, Lat. *duo*).

2.5. QUANTITY

2.5.1. Syllables are distinguished according to the length of time required for their pronunciation. Two degrees of Quantity are recognized, *long* and *short*.

NOTE. In syllables, quantity is measured from the beginning of the vowel or diphthong to the end of the syllable. Such distinctions of long and short are not arbitrary and artificial, but are purely natural, a long syllable requiring more time for its pronunciation than a short one.

2.5.3. A syllable is long usually,

   a. if it contains a long vowel; as, **mā-tér**, *mother*, **kē-lā-jō**, *hide*,

   b. if it contains a diphthong; as, **lai-wós**, *left*, **oi-nos**, *one*,

   c. if it contains any two non-syllabic consonants (except a mute followed by *l* or *r*); as, **pneu-sō**, *breathe strongly*, **tmā-mi**, *cut*.

2.5.4. A syllable is short usually,

   a. if it contains a short vowel followed by a vowel or by a single consonant; as, **pel-nis**, *skin*, or **e-í-mi**, *go,*

   b. if it contains a vocalic sonant; as, **qr̥-mis**, *worm*, **cm̥-tis**, *march.*

2.6. ACCENT

2.6.1. There are stressed as well as unstressed words. The last could indicate words that are always enclitic, i.e., they are always bound to the accent of the preceding word, as -**qe**, *and*, -**ṛ**, *for;* while another can be proclitics, like prepositions.

2.6.2. The oldest PIE was a stress language in which syllable strength was chiefly a matter of pitch differences and, presumably, of intensity (loudness).

NOTE. Following Gąsiorowski, "[i]n this respect it was similar to Spanish or Polish, but not to English with its emphatic 'expiratory' stress (...) It thus stood close to the borderline between stress systems and pitch accent systems. Indeed, some linguists have attributed pitch accent contrasts to PIE on the strength of accentual correspondences between Balto-Slavic and Greek. However, scholars such as Jerzy Kuryłowicz and – more recently – Paul Kiparsky have convincingly argued that such contrasts arose independently in the branches in question. The best evidence for the original location of stress in PIE comes from Vedic (Classical Sanskrit developed its own stress system, similar to that of Latin). The location of pitch accent in Classical Greek (especially in Greek noun paradigms) also reflects the PIE stress pattern. There are, to be sure, some specifically Greek constraints on the distribution of pitch accents, but in the environments where such restrictions do not apply, Greek usually agrees with Vedic. In the Germanic languages the original location of stress is sometimes reconstructible thanks to the phonetic 'fingerprints' of Verner's Law. Germanic spectacularly bears out the testimony of Vedic and Classical Greek. Finally, the evolution of pitch-accent systems in Balto-Slavic makes most sense if we adopt the stress system reconstructed on the basis of Vedic, Greek and Germanic as its starting-point".

2.6.4. The Stress is free, but that does not mean anarchy. On the contrary, it means that each non-clitic word has an accent and only one accent, and one has to know – usually by way of practice – where it goes. Its location depended on the inflectional type to which a given word belonged.

NOTE. Indo-European stress is (at least partly) unpredictable. Rather, it is lexical: it comes as part of the word and must be memorized, although orthography can make stress unambiguous for a reader, and some stress patterns are ruled out. Otherwise homophonous words may differ only by the position of the stress, and it is thus possible to use stress as a grammatical device.

2.6.5. Adjectives are often stressed on the ending, especially if they are derivatives; as, **ghḷtnós**, *golden*, from **ghḷtom**, *gold*, **ṇgnōtós**, *unknown*, from **gnōskō**, *know*. Nevertheless, nouns and adjective might be stressed on any syllable.

NOTE. There are some accent rules to be followed in the declension of nouns and in the conjugation of verbs, which will be later studied.

2.7. VOWEL CHANGE

2.7.1. Vowel Change was common in Proto-Indo-European. In many words the vowel varies because of old alternating forms that gave different derivatives.

NOTE. With the creation of zero-grade stems, vocalization appears, as the original radical vowels disappear and new ones are added. That happens, for example, in root **bhṛ**- [bhr̥], *carry*, (cognate with English *bear*), which can be reconstructed from IE languages as **bher**-, **bhor**- or **bhṛ**-. The same can be said of the semivowels [j] and [w] when they are syllable edges, being syllable centres [i] and [u] in zero-grades.

So for example in o-grade **domos**, *house*, which gives **dómūnos**, *lord*, as Lat. *dominus*, Skr. *dá ́mūnaṣ* but full grade root **dem**-, which gives **demspóts**, *master, lord,* later *despot*, as Gk. δεσπότης (*despótēs*), Skr. *dampati*, Av. *dəng patōiš*, (with *fem.* **demspotnjā**).

NOTE. The forms attested in Indo-Iranian (and maybe Greek) come from i-stem **potis**, probably derived from the original Late PIE form **dems**-**póts**, cf. **ghósti**-**pots**, *guest*, as Lat. *hospēs*, *hospitis*, O.Russ. *gospodь<*-ostьpot*-; compare, for an original PIE ending -**t** in compounds, Lat. *sacerdōs* < MIE **sákrodhots**, O.Ind. *devastút*-, "*who praises the gods*", etc. The compound is formed with **pot**-, *lord, husband*, and **pot**-**njā**, *mistress, lady*.

2.7.2. Different vocalizations appeared in IE dialects in some phonetic environments, especially between two occlusives in zero-grade, impossible to pronounce without adding a vowel; as e.g. **skp**-, which evolved as Lat. *scabo* or Got. *skaban*.

NOTE. Although the dialectal solutions to such consonantal groups aren't unitary, we can find some general PIE timbres. As **a**, **i** with a following dental (especially in Gk. and BSl.) or **u**, also considered general, but probably influenced by the context, possibly when in contact with a labial, guttural or labiovelar.

2.7.3. Sometimes different reconstructions might account for some vowel differences; **a** for **o**, as **lawō* for **lowō**, *wash*; a vocalic sonant for **a** or **e** plus sonant, as **Sṃos* for **Samos**, *summer*, or **kṛwos* for **kerwos**, *deer*, etc.

NOTE. Different reconstructions might be equally valid, depending on the criteria employed. Sometimes different PIE language stages have to be taken into account; as, for root **neqt**-, *night*, a

common PIH full-grade *_neqts_ is reconstructible, according to Hitt. _nekut_; however, Late PIE dialects show that an o-grade noun was later generalized; cf. O.Gk. _nuks_, _nuktós_, O.Lat. _nox_, _noctis_, for an old PIE consonant stem *_noqts_. The newer i-stem **noqtis** was the general Late PIE (and later also PII, EIE) form, cf. O.Ind. _nakti_, Gmc. _naxti_, Sla. _notjь_, Bal. _nakti_.

The phonological reconstruction of Late PIE includes generally the _Schwa Indogermanicum_, uncertain in sound, which usually stands for an older laryngeal *_h₂_. In North-West IE, PIE reconstructed *_ə_ usually appears as **a**; as, **statis**, _standing post_, from zero-grade *_sth₂_- of root **stā**- (<_steh₂_-) _stay_; or **patḗr**, from older *_ph₂tér_-.

NOTE. Other examples are a-stems in *-_ī/-jə_, from older *-_ih₂_, and neuter plural in *-_ə_<*-_h₂_.

2.8. CONSONANT CHANGE

2.8.1. Regarding Consonant Change, different reconstructions might appear, too; as, for **ghortos**, _garden, enclosure_, later _town_ (cf. Gmc. _gardan_, Lat. _hortus_, Gk. _khortos_, Phry. -_gordum_, O.Ir. _gort_, Lith. _gardas_, O.C.S. _gradu_, Alb. _garth_, etc.), some would reconstruct an alternative ***_ghordhos_**, so that both forms (in -**t**- and -**dh**-) fit perfectly into the schemes of dialectal phonological laws.

2.8.2. The so called s-Mobile (_mobile_ pronounced as in Latin, it is a neuter adjective) refers to the phenomenon of alternating word pairs, with and without **s** before initial consonants, in stems with similar or identical meaning. This "_moveable_" prefix **_s_**- is always followed by another consonant. Typical combinations are with voiceless stops **(s)p**-, **(s)t**-, **(s)k**-, with liquids and nasals, **(s)l**-, **(s)m**-, **(s)n**-; and rarely **(s)w**-.

NOTE. Examples include **(s)ten**-, compare O.Ind. _stánati_, Gk. _sténō_, O.Eng. _stenan_, Lith. _stenù_, O.Sla. _stenjo_, and without **s**- in O.Ind. _tányati_, Gk. Eol. _ténnei_, Lat. _tonare_, O.H.G. _donar_, Cel. _Tanaros_ (name of a river). For **(s)pek**-, cf. O.Ind. _spáśati_, Av. _spašta_, Gk. _skopós_ (<**_spokós_**), Lat. _spektus_, O.H.G. _spehon_, without **s**- in O.Ind. _páśyati_, Alb. _pashë_. For **(s)ker**-, cf. O.Ind. _ava_-, _apa-skara_-, Gk. _skéraphos_, O.Ir. _scar(a)im_, O.N. _skera_, Lith. _skiriù_, Illyr. _Scardus_, Alb. _hurdhë_ (<***_skṛd_**-), without **s**- in O.Ind. _kṛnáti_, Av. _kərəntaiti_, Gk. _keíro_, Arm. _kcorem_, Alb. _kjëth_, Lat. _caro_, O.Ir. _cert_, O.N. _horund_, Lith. _kkarnà_, O.Sla. _korŭcŭ_, Hitt. _kartai_-, and so on.

Such pairs with and without **s** are found even within the same dialect, as Gk. _(s)tégos_, "roof", _(s)mikrós_, "little", O.Ind. _(s)tṛ_, "star",_ and so on._

NOTE. Some scholars believe it was a prefix in PIE (which would have had a causative value), while others maintain that it is probably caused by assimilations of similar stems – some of them

beginning with an **s**-, and some of them without it. It is possible, however, that the original stem actually had an initial **s**, and that it was lost by analogy in some situations, because of phonetic changes, probably due to some word compounds where the last -**s** of the first word assimilated to the first **s**- of the second one. That helps to explain why both stems (with and without s) are recorded in some languages, and why no regular evolution pattern may be ascertained: so for example in **wĺqons spekjont**, *they saw wolves*, becoming **wĺqons 'pekjont**. See Adrados (1995).

2.8.3. Before a voiced or aspirated voiced consonant, **s** was articulated as voiced, by way of assimilation; as, **nisdos** ['niz-dos], *nest*, **misdhom** ['miz-dʰom], *meed, salary*, or **osdos** ['oz-dos], *branch*. When **s** forms a group with sonants there is usually assimilation, but such a trend was sometimes reversed by adding a consonant; as Lat. *cerebrum (<Ita. kereθrom)*, from **kersrom** ['kerz-rom], *brain*.

NOTE. Related to the later assimilation of [s] into [z] between vowels, they became very unstable in some IE dialects, showing sometimes *rhotacism*; as, **snusós**, *daughter-in-law*, cf. Lat. *nurus*, O.H.G. *snur*; or **genos**, *race, stock, kind*, cf. Lat. *genus, generis (<*geneses)*.

2.8.4. Similarly, the manner of articulation of an occlusive usually depends on its environment. Thus, voiced stops turn voiceless in final position; as, **pods**, *foot*, gives voiceless O.Ind. *pāt*, **qid** gives O.Ind. *cit*, **agtós** gives voiceless Gk. ακτος (*aktos*) or Lat. *actus*. The same happens with voiced aspirates, as in **legh**-, *lie* (cognate to Eng. *log*), giving Gk. λεκτρον (*lektron*), Lat. *lectus*, O.H.G. *Lehter*. Voiceless occlusives become voiced before voiced consonants; as, zero-grade **ped**- in Gk. επιβδα (*epi-bd-a*).

2.8.5. A sequence of two dentals, such as -**tt**-, -**dt**-, -**tdh**-, -**ddh**-, etc. was eliminated in all Indo-European dialects, but the process of this suppression differed among branches; Vedic Sanskritshowing little change, some others an intermediate -*sT*-, and others -*ss*- or -*s*-. Compounds were not affected by this trend; as, **kréd**-**dhēmi**, *believe*.

NOTE. This trend began probably in Late PIE, and thus all IE speakers knew such evolutions, which we sum up into a common intermediate stage -**st**-, -**sdh**-, etc., which was followed in some early IE dialects, and probably known to the rest of them. See the section *Conventions Used in this Book* for more on this question. For phonetic changes in Aryan dialects, see Appendix II.

Examples in MIE are e.g. forms derived from PIE root **weid**-, *know, see*, which gave verb **widējō**, cf. Lat. *vidēre*, Goth. *witan*, O.C.S. *videǐi*, Lith. *pavydéti*; p.p. **wistós**, *seen*, from **wid-tó**-, (cf. O.Ind. *vitta*-, but Av. *vista*-, O.Pruss. *waist*, O.Sla. *věstъ*, or

Gmc. *wīssaz*, Lat. *vīsus*, O.Gk. ϝιστος, O.Ir. *rofess*, etc.); noun **wistis**, *sight, vision*, from **wid-ti-**, cf. Goth *wizzi*, Lat. *vīsiō*; Greek **wistōr**, *wise, learned man*, from **wid-tor**, cf. Gk. ἵστωρ<*ϝίστωρ (*wístōr*), PGk **wistorjā**, *history*, from Gk. ἱστορία (*ʰistoría*); Imp. **wisdhi!** *know!*, from **wid-dhí**, cf. O.Ind. *viddhí*, O.Gk. ϝίσθι, O.Lith. *veizdi*, and so on.

2.9. PECULIARITIES OF ORTHOGRAPHY

2.9.1. Indo-European words may show a variable orthography, although a unified one should be strongly encouraged.

2.9.1. Vowel Changes that influence the way MIE is written include the alternating PIE forms that gave different frozen derivatives.

A vowel change that should not affect MIE orthography is what many reconstruct as PIE [ə] or *schwa*, generally evolved as North-West IE **a**; as, PIH **ph₂tér-* → PIE **pətér-* → EIE **patér-**, *father*; PIH *bʰh₂tis → PIE *bʰətis → EIE **bʰatis**, *appearance*; PIH *anh₂mos → PIE *anəmos → EIE **ánamos**, *breath*, and so on.

NOTE. This Late PIE reconstructed *schwa* (see §2.2.1) is important for the different vocalism of EIE, PII and PGk; cf. MIE **patér-** with Aryan *pitár-*, or MIE **ánamos** with Hellenic *ánemos*.

2.9.2. Consonant Changes that should not affect MIE orthography, already seen, include voiced sibilants, as **nisdos** ['niz-dos], **kersrom** ['kerz-rom]; and voiceless occlusives, as **pods** [pots], **agtós** [ak-'tos], **leghtrom** ['lek-trom], -**pd**- [bd].

NOTE. Although the accuracy of some allophones in PIE is certain, for practical reasons the phonetically correct notation is therefore avoided in favour of the phonemically correct notation.

Changes that usually affect how MIE is written include commonly reconstructed variants, as **egh-**, **ek-**, *outside, out, from*; and doubious cognates, as **necr-**, *dark*, and **neqt-**, *night*, maybe from a common PIH suffixed **negʷ-*, *to dawn*.

2.9.3. About semivowels, as a general exception, they are not written when the semivowel is the last sound of the first word in a compound; e.g., for **triathlom** (from **tri-**, *three*, and Gk. *athlon*, "*contest*"), *triathlon,* we won't write ***trjathlom**; **sindhueurōpājóm**, and not ***sindhweurōpājóm**; etc.

NOTE. In Modern Indo-European, compounds may be written with and without hyphen, as in the different modern Indo-European languages. Nevertheless, the older, not hyphenated version is preferred for formal writings; as, **sindhueurōpājóm**, and not ***sindhu-eurōpājóm**, compare

Eng. *Indo-European*, Ger. *Indoeuropäisch*, Fr. *Indo-européen*, It., Sp. *indoeuropeo*, Gal.-Pt. *Indo-européu*, Cat. *indoeuropeu*, Du. *Indo-Europees*, Pol. *indoeuropejski*, Lit. *indoeuropiečių*, Ir. *Ind-Eorpach*, Russ. *индоевропейский*, Gk. *ινδοευρωπαϊκή*, Ira. ﻩﺩﻭﺍﺭﻭﭘﻴﻠﻰ , Hin. हिन्द-यूरोपीय, etc.

2.9.4. The vocallic allophones [r̥], [l̥], [m̥], [n̥] may be written, as in Latin transliterations of Sanskrit texts, as *r̥, l̥, m̥,* and *n̥,* to help the reader clearly identify the sonants; therefore, alternative writings **n̥mr̥tós**, *inmortal*, **km̥tóm**, *hundred*, **wodr̥**, *water*, etc. are also possible.

2.9.5. An Apostrophe is used to mark the ommited letter of a contraction in word-final position, usually in elisions at the end of imperative verbs, especially in spoken language; as **cemj**' for **cemje**, *come here*; or **takej**' for **takēje**, *shut up.*

2.9.6. An Acute Accent is written over the vowel or semivowel in the stressed syllable, except when stress is on the second to last syllable (or *paenultima*) and in monosyllabic words. Accented long vowels and sonants are represented with special characters. The weak vowel of a possible diphthong is also accented; so in **eími** [e-'i-mi], *I go*, and not ***eimi***, pronounced ['ei-mi] if left unaccented.

2.9.7. The forms with the copulative -**qe**, *and*, and disjunctive -**wĕ**, *or*, are usually written by adding it to the preceding word, as in Latin -*que*, but with a hyphen.

2.9.8. The capital letters are used at the beginning of the following kinds of words:

a. the names of days, months, seasons and public holidays; as, **Jānwārjos**, *January*, **Samos**, *summer*, **Newos Atnos**, *New Year*, etc.

b. the names of people and places, including stars and planets; as, **Sāwel**, *Sun*, **Aleksanốr**, *Alexander*, **Deiwos**, *God*, **Sindhu**, *Indus (river)*, **Teutiskolondhom**, *Germany* (cf. O.H.G. *Diutisk-lant<*þeudiska-landam*), etc.

NOTE. In old IE languages demonyms were not written in capital letters; as, **Eurṓpā**, **eurōpājós**; **Angljā** or **Angljolondhom**, *England* (cf. O.E. *Engla-land*, "*land of the Angles*"), but **angljós**, *English*; **Hispānja**, *Spain*, but **hispānós**, *Spanish*; **teutiskós**, *German*; and so on.

c. people's titles, as **Prōbhastṓr**, *Professor*, **Kelomnelis**, *Colonel*, **Rēgtṓr**, *rector*,

d. **Skeuros**, *North*, **Déksinā**, *South*, **Áusteros**, *East*, **Éperom**, *West*, and their derivatives.

NOTE. Germanic **Nertros**, *lower*, later *North*, from **ner**-, *lower, bottom*, and **Suntos**, *"of the Sun"*, later *South*, possibly from alternative root **Sun**- of **Sāwel**, *sun*; Gmc. *West* comes probably from the same root as **wespros**, *evening*.

e. in official or well-established place names; as **Plátejā**, *the Square*, etc.

2.10. KINDRED FORMS

Compare the following Europe's Indo-European words and their evolution in Germanic and Latin, with their common derivatives in Modern English.

EIE	*PGmc.*	*O.Eng.*	*Latin*	*English (Lat.)*
patếr, *father*	faðer	fæder	pater	father (*paternal*)
septm̥, *seven*	sibun	seofon	septem	seven (*September*)
trebhō, *dwell*	þurp-	þorp	trabs/trabēs	thorp (*trabecula*)
globjō, *hold, clench*	klupjō	clyppe	globus	clip (*globe*)
bhrātēr, *brother*	brōþēr	brōþor	frāter	brother (*fraternal*)
bherō, *carry*	berō	bere	ferō	bear (*infer*)
wertō, *turn*	werþō	weorþe	uertō	worth (*versus*)
trejes, *three*	þrejez	þrēo	trēs	three (*trinity*)
dekm̥, *ten*	texan	ten,tien	decem	ten (*decimal*)
edmi, *eat*	et-	ete	edō	eat (*edible*)
dhēmi, *do, make*	dōmi	dōm	faciō (<*dha-k*-)	do (*factor*)
dhersō, *be adroit*	dersō	dearr	festus (<*dhers-t*-)	dare (*manifest*)
leuk-, *light*	leux-	lēoh-	lūc-	light (*lucid*)
kr̥d-, *heart*	xert-	heort-	cord-	heart (*core*)
augō, *increase*	aukō	eacie	augeō	eke (*augment*)
gn-, *know*	kunnō	cunne	(g)nōtus	can (*notice*)
ghostis, *guest*	gastiz	gæst, giest	hostis	guest (*hostile*)
bhergh-, *mountain*	burg-	beorg	fortis (<*forctus*)?	barrow (*force*)
leiq-, *leave*	leixʷ-	læne	līqu-	lend (*relic*)
qi-/**qo**-, *what, who*	hʷi- / hʷo-	hwi-	qui- / quo-	why/what (*quote*)
cemjō, *come*	kʷemjō	-cwem-	ueniō	come (*venue*)
cīwós, *alive*	kʷi(k)waz	cwic	uīuus	quick (*vivacity*)
leghús, *light*	lextaz	līht, lēoht	leuis	light (*levity*)
chormos, *warm*	warmaz	wearm	formus	warm (*furnace*)

3. WORDS AND THEIR FORMS

3.1. THE PARTS OF SPEECH

3.1.1. Words are divided into eight Parts of Speech: Nouns, Adjectives (including Participles), Pronouns, Verbs, Adverbs, Prepositions, Conjunctions, and Interjections.

3.1.2. A Noun is the name of a person, place, thing or idea; as, **Klewopatrā**, *Cleopatra*, **dānus**, *river*, **dhworis**, *door*, **wr̥dhom**, *word*.

Names of particular persons and places are called Proper Nouns; other nouns are called Common.

NOTE. An Abstract Noun is the name of a quality or idea. A Collective Noun is the name of a group or a class.

3.1.3. An Adjective is a word that attributes a quality; as, **patrjóm**, *parental*, **leukós**, *bright*, **kartús**, *hard*, **gr̥ndhís**, *grown*.

NOTE 1. A Participle is a word that attributes quality like an adjective, but, being derived from a verb, retains in some degree the power of the verb to assert.

NOTE 2. Etymologically there is no difference between a noun and an adjective, both being formed alike. So, too, all names originally attribute quality, and any common name can still be so used. Thus, **Regeinā Elísabet II**, *Queen Elizabeth II*, (cf. Gk. *Ελισ(σ)αβετ*, from Hebrew *Elisheva*, "*God is an oath*"), distinguishes this *Elizabeth* from other *Elizabeths*, by the attribute expressed in the name **Regeinā**, *Queen*.

3.1.4. A Pronoun is a word used to distinguish a person, place, thing or idea without either naming or describing it: as, **egṓ**, *I*, **tewós**, *thine*, **wejes**, *we*.

Nouns and pronouns are often called Substantives.

3.1.5. A Verb is a word capable of asserting something: as, **bherō**, *I carry, bear*; **bhāti**, *it shines*.

NOTE. In English the verb is usually the only word that asserts anything, and a verb is therefore supposed to be necessary to complete an assertion. Strictly, however, any adjective or noun may, by attributing a quality or giving a name, make a complete assertion; as, **wīrós bhilis** (**esti**), *the man is good (decent)*, unlike **bhilis wīrós**, *the good man*; or **autom dwenos** (**esti**), *the car is*

good (efficient), unlike **dwenos autom**, *the good car*. In the infancy of language there could have been no other means of asserting, as the verb is comparatively of late development.

3.1.6. An Adverb is a word used to express the time, place, or manner of an assertion or attribute: as, **per**, *in front*, **epi**, *near*, **antí**, *opposite*.

NOTE. These same functions are often performed in Indo-European by cases of nouns, pronouns and adjectives, and by phrases or sentences.

3.1.7. A Preposition is a word which shows the relation between a noun or pronoun and some other word or words in the same sentence; as, e.g., **ad**, *at, to*, **dē**, *from upwards*, **kom**, *with*, **ek**(**sí**), *outside*, **upo**, *under*, and so on.

3.1.8. A Conjunction is a word which connects words, or groups of words, without affecting their grammatical relations: as, -**qe**, *and*; -**wĕ**, *or*, -**ma**, *but*, -**r**, *for*.

3.1.9. Interjections are mere exclamations and are not strictly to be classed as parts of speech; as, **alā!** *hello!*; **ō!** *O* (vocative); **wai!** *alas* (grief); **ha ha!** (laughing sound), or older **kha kha!** as in **khákhatnos**, *laugh*; **ha!** (surprise); etc.

NOTE. Interjections sometimes express an emotion which affects a person or thing mentioned, and so have a grammatical connection like other words.

3.2. INFLECTION

3.2.1. Indo-European is an inflected language. Inflection is a change made in the form of a word to show its grammatical relations.

NOTE. Some modern Indo-European languages, like most Germanic and Romance dialects, have lost partly or completely their earliest attested inflection systems – due to different simplification trends –, in nominal declension as well as in verbal conjugation.

3.2.2. Inflectional changes sometimes take place in the body of a word, or at the beginning, but oftener in its termination:

bhabhā, *the* or *a bean*, **snichwós**, *of the snow*, **reidhō**, *I ride*, **idhi!** *go!*

3.2.3. Terminations of inflection had possibly originally independent meanings which are now obscured. They probably corresponded nearly to the use of prepositions, auxiliaries and personal pronouns in English.

Thus, in **ghórdejos**, *of the barley* (Gen.), the termination is equivalent to "*of the*"; in **deikō**, *I show* (Indicative), and **dikóm**, *I showed* (Aorist), the change of vowel grade and accent signifies a change in the aspect.

3.2.4. Inflectional changes in the body of a verb usually denote relations of tense or mood, and often correspond to the use of auxiliary verbs in English:

(**tu**) **déikesi**, *(thou) show*; **doike**, *he showed*; (**gí**)**gnṓsketi**, *he knows*, **gégona**, *I knew* (see Verbal Inflection for Reduplication and its meaning).

3.2.5. The inflection of Nouns, Adjectives, Pronouns and Participles to denote gender, number and case is called Declension, and these parts of speech are said to be *declined*.

The inflection of Verbs to denote voice, mood, tense, number and person is called Conjugation, and the verb is said to be *conjugated*.

NOTE. Adjectives are often said to have inflections of comparison. These are, however, properly stem-formations made by derivations.

3.2.6. Adverbs, Prepositions, Conjunctions and Interjections are not inflected, and together form the group of the so-called Particles.

3.3. ROOT, STEM AND BASE

3.3.1. The body of a word, to which the terminations are attached, is called the Stem. The Stem contains the idea of the word without relations; but, except in the first part of compounds (e.g. **somo**-**patōr**, "*of the same father*", *sibling*, **mṇ**-**dōmi**, *commit*), it cannot ordinarily be used without some termination to express them.

NOTE. According to Mallory & Adams (2006): "To the root might be added a variety of suffixes to create a stem and then finally the case endings depending on number and perhaps gender. In some cases, the so-called root-nouns, there are no suffixes before the case ending. Using R for 'root', S for 'stem-creating suffix', and E for 'case-number-ending', we might establish the formula for an inflected word in Proto-Indo-European as R-(S)-E".

Thus the stem **pater**- denotes *father*; **patḗr**, Nominative, means *a father* or *the father*, as the Subject or Agent of an action; **patér** (or **pater**) is the Vocative, as in *O father!*; **patérṃ** is the means *to a father* or *to the father*, as the Direct Object; **patrós** is the Genitive and indicates *of a father* or *of the father*, and so on.

NOTE. In inflected languages like Indo-European, words are built up from Roots, which at a very early time were possibly used alone to express ideas. Roots are then modified into Stems, which, by inflection, become fully formed words. The process by which roots are modified, in the various forms of derivatives and compounds, is called stem-building. The whole of this process is originally one of composition, by which significant endings are added one after another to forms capable of pronunciation and conveying a meaning.

3.3.2. A Root is the simplest form attainable by analysis of a word into its component parts. Such a form contains the main idea of the word in a very general sense, and is common also to other words either in the same language or in kindred languages; cf. for **stā**-, *stand*, reduplicated present **sí**-**stā**-**mi**, *I stand*, noun **stā**-**men**-, *place for standing*, zero-grade p.p. **sta**-**tós**, *placed, standing*, or noun **sta**-**tis**, *erection, standing*.

For example, the root of verb **spekjō**, *look*, is **spek**-, which does not necessarily mean *to look*, or *I look*, or *looking*, but merely expresses vaguely the idea of *looking*, and possibly cannot be used as a part of speech without terminations.

3.3.3. The Stem may be the same as the root; as, **dō**-, *give*, **dakru**, *tear*; but it is more frequently formed from the root.

1. By changing or lengthening its vowel; as, from athematic root verb **dā**-, *divide*, common derivative **dai**-**mai**, *divide up, distribute*.

2. By the addition of a simple suffix; as, from root **dā**-, *divide*, derivative **dā**-**mos**, *people, people's division*, cf. Dor. Gk. δημος, O.Ir. *dām*, Hitt. *da-ma-a-iš*.

NOTE. Some suffixes probably conveyed an earlier underlying meaning, e.g. the suffix -**trom** tends to indicate an instrument, as **arā**-**trom**, *plough*, from a verb **arājō**, *plough*, while kinship names tend to have the suffix -**er** or -**ter**, cf. **swes**-**ōr**, *sister*, **bhrā**-**tēr**, *brother*.

3. By two or more of these methods; from the same root, suffixed derivative **dai**-**tis**, *time, period*, cf. Gmc. *tīÞ*, Arm *ti*, as well as Gk. δαιτύς, O.Ind. *dātu*-.

4. By derivation and composition, following the laws of development peculiar to the language, which we will see in the corresponding chapters.

3.3.4. The Base is that part of a word which is unchanged in inflection: as, **chorm**- in **chormos**, *warm*, **eus**- in **eusō**, *burn*; **cou**- in **cōus**, *cow*,etc.

The masculine functions as the negative term in the opposition, i.e. when the gender is not defined, the masculine is used. This is a grammatical utility, one that is only relevant for concordance, and which has to do with the evolution of the language and its inflection.

*The earliest PIE had probably no distinction of gender; when the inanimate appeared, it was marked by a different inflection, and the animates remained as the negative term in the opposition. After that, probably at the same time as the thematic declension (in -**e/o**) appeared, the feminine was differentiated from the remaining animates, with marks like the different stem vowel (usually -**a**) or vowel length (as -**ī**, -**ū**). Therefore, the feminine is the positive term of the opposition within the animates, because when we use it we reduce the spectrum of the animates to the feminine, while the masculine still serves as the negative (non-differentiated) term for both, the general and the animates, when used in this sense, i.e. when not differentiating the masculine from the other genders.*

a. The Base and the Stem are often identical, as in many consonant stems of nouns (as **eghṛ**, *boundary*). If, however, the stem ends in a vowel, the latter does not appear in the base, but is variously combined with the inflectional termination. Thus the stem of **nócodos**, *naked*, is **nocod**-; that of **ceri**, *mountain*, is **cer**-.

3.3.5. Inflectional terminations are modified differently by combination with the final vowel or consonant of the Stem, and the various forms of Declension and Conjugation are so developed.

3.4. GENDER

3.4.1. The Genders distinguished in Modern Indo-European are three: Masculine, Feminine (both are referred to as Animate) and Neuter or Inanimate.

3.4.2. The gender of Indo-European nouns is either *natural* or *grammatical*.

a. Natural Gender is distinction as to the sex of the object denoted: **bhrātēr** (m.), *brother*; **cenā** (f.), *woman, wife*.

b. Grammatical Gender is a formal distinction as to sex where no actual sex exists in the object. It is shown in the form of the adjective joined with the noun: as **swādús noqtis** (f.), *a pleasant night*; **mṛghús kanmṇ** (m.), *brief song*. The gender of the adjective is simply a gender of concordance: it indicates to which noun of a concrete gender the adjective refers to.

NOTE 2. Names of classes or collections of persons may be of any gender. For example, **wolgos** (m.), *(common) people*, or **teutā**, *people (of a nationality)*.

3.4.3. The neuter or inanimate gender differs from the other two in inflection, not in the theme vowel. The gender of the animates, on the contrary, is usually marked by the theme vowel, and sometimes by declension, vocalism and accent.

3.4.4. The neuter does not refer to the lack of sex, but to the lack of liveliness or life. Sometimes, however, animates can be designated as inanimates and *vice versa*.

While the distinction between masculine and feminine is usually straightforward, sometimes the attribution of sex is arbitrary; thus, different words for parts of the body are found feminine, as **nāsis**, *nose*, **kanmā**, *leg*; masculine, as **kolsos**, *neck*, **armos**, *arm*, *upper arm*; and neuter, as **kaput**, *head*, or **genu**, *knee*.

3.4.5. The animate nouns can have:

a. An oppositive gender, marked:

    I. by the lexicon, as in **patḗr**/**mātḗr**, *father*/*mother*, **bhrātēr**/**swesōr**, *brother*/*sister*, **sūnús**/**dhugtēr**, *son*/*daughter*;

    II. by the stem ending, as in general **ekwos**/**ekwā**, *horse*/*mare*, or the rare **wlqos**/**wlqīs**, *wolf*/*she-wolf*, **djēus**/**djewja**, *sky(-god)*/*sky-goddess*;

    III. by both at the same time, as in **swekros**/**swekrús**, *father-in-law-mother-in-law*, **wīrós**/**cenā**, *man-woman*, **regs**/**regeinā**, *king-queen*.

b. An autonomous gender, that does not oppose itself to others, as in **nāus** (f.), *ship*, **pods** (m.), *foot*, **egnis** (m.), *fire*, **owis** (f.), *sheep*, **jewos** (n.) or **legs** (f.), *law*.

c. A common gender, in nouns that are masculine or feminine depending on the context; as, **cōus**, *cow* or *bull*, **deuks**, *leader*, **ghostis**, *foreigner*.

d. An epicene gender, which, although being masculine or feminine, designates both sexes; as, **médodiks**, *doctor*, **nawāgós**, *sailor*, **nemots**, *enemy*, **setis**, *visitor*.

3.4.6. The gender of a noun can thus be marked by the stem vowel (or sometimes by inflection), or has to be learnt: it is a feature of a word like any other. In its context, concordance is a new gender mark; a masculine noun has a masculine adjective, and a

feminine noun a feminine adjective. However, not all adjectives differentiate between masculine and feminine, a lot of them (those in -***i**-**s***, -***u**-**s***, -***ē**s***, -***ō**n***, and some thematic in -***os***) are masculine and feminine: only the context, i.e. the noun with which they agree, helps to disambiguate them. This happens also in nouns with a common gender.

3.4.7. Most endings do not indicate gender, as in **patḗr** and **mātḗr**. Only by knowing the roots in many cases, or from the context in others, is it possible to determine it. Some of the suffixes determine, though, totally or partially if they are masculine or feminine. These are the following:

1. -***os*** marks masculine when it is opposed to a feminine in -***ā*** or -***ī**/-**ja***, as in **ekwos**/**ekwā**, **deiwos**/**deiwā**, *god*/*goddess*, etc. This happens also in adjectives in the same situation, as in **newos**/**newā**, **bheronts**/**bherontja**. In isolated nouns, -***os*** is generally masculine, but some traces of the old indistinctness of gender still remained in Late PIE, as in the names of trees (among others). In adjectives, when the ending -***os*** is not opposed to feminine, concordance decides.

2. -***ā*** marks the feminine in oppositions of nouns and adjectives. It is usually also feminine in isolated nouns, in the first declension. But there are also some masculines in -***ā***; as, **sloughā**, *servant*, cf. O.Sla. *slŭga*, Lith. *slauga* "*service*", O.Ir. *sluag*, "*army unit*", etc.

3. -***ī**/-**ja*** (<*-*ih₂*), although feminine in Late PIE, shows remains of its old Abstract-Collective value, as neuter plural. It appears in nouns, adjectives and pronouns.

4. Roots ending in long vowels -***ī*** and -***ū*** are always feminines.

3.5. GENERAL RULES OF GENDER

3.5.1. Names of Male beings, and of Rivers, Winds, Months, and Mountains are masculine:

patḗr, *father*, **Góralos**, *Charles*, **Reinos**, *the Rhine*, **Áusteros**, *south wind*, **Magjos**, *May*, **Urales**, *the Urals*.

NOTE. For **Uraliskā́ Cerja**, *Ural Mountains*, cf. Russ. Ура́льские го́ры (*Uralskiye gory*).

a. A few names of Rivers ending in -***ā*** (as **Wolgā**), and many Greek names ending in -*ē(s)*, which usually corresponds to IE -***ā***, are feminine; others are variable or uncertain, generally retaining their oldest attested IE gender in MIE.

b. Some names of Mountains are feminines or neuter: as, **Alpes** (f. pl.), *the Alps.*

3.5.2. Names of Female beings, of many Cities, Countries, Plants, Trees, Gems, Animals, and abstract Qualities, are feminine:

mātḗr, *mother*, **Djówiljā**, *Julia*, **Frankjā**, *France*, **Rōmā**, *Rome*, **bharwos**, *pinetree*, **saniprijós**, *sapphire* (Gk. *sáppheiros*, ult. from Skr. *sani-priyaḥ*, lit. "*sacred to Saturn*"), **aqamarinā́**, *aquamarine*, **wḗrā**, *true.*

a. Some names of Towns and Countries are masculine: as, **Kṛsnomontis**, *Montenegro*; or neuter, as, **Jugtóm Regnom**, *United Kingdom*, **Swjoregnom**, *Kingdom of Sweden*, **Finnlondhom**, *Finland.*

b A few names of Plants and Gems follow the gender of their termination; as, **kṃtaurjom** (n.), *centaury*, **úpolos** (m.), *opal.*

NOTE. The gender of most of the above may also be recognized by the terminations, according to the rules given under the different declensions.

3.5.3. Indeclinable nouns, infinitives, terms or phrases used as nouns, and words quoted merely for their form, are neuter: **sestum** (<***sed-tu***-) *to sit, be sitting*, "**wétānom smeughtum**", *"smoking prohibited";* **gummi**, *gum.*

NOTE. Eng. *gum* comes from O.Fr. *gomme*, from L.Lat. *gumma*, from Lat. *gummi*, from Gk. *kommi*, from Coptic *kemai*, hence MIE loan **gummi**, PGk. *kommi.*

3.5.4. Nouns, Pronouns, Adjectives and Participles are declined in MIE in two Numbers, singular and plural – Late PIE had also possibly a dialectal dual – and up to eight cases, Nominative, Vocative, Accusative, Genitive and Oblique - which is found subdivided into combinations of Dative, Locative, Instrumental and Ablative.

NOTE 1. European dialects show around six cases, but most of the oldest attested ones (PII, PGk, Ita.) and Balto-Slavic show remains of up to eight original cases. Even though the situation has evolved differently due to migrations and linguistic contacts, linguists generally agree that the most securely reconstructed are the nominative, vocative, accusative, and genitive of the singular and plural. Traditional theories maintain that the original common PIE situation is a complex system of eight noun cases. On the contrary, a five-case system is for other scholars the oldest situation (of Middle PIE, as Anatolian dialects apparently show), later changed by Late PIE by way of merging or splitting the five original cases. An eight-case system would have been, then, an innovation of individual dialects, just as the phonetic satemization. It is thus a general opinion

that already in Late PIE both trends (split and convergence of Obliques) coexisted. In this MIE Grammar we follow the general, oldest trend, i.e. an eight-case inflection system, due to the findings in North-West IE.

NOTE 2. In the number we use singular and plural, and not dual, not only because of its doubtful existence in IE II and the objections to its reconstruction for Late PIE, but because it is also more practical in terms of modern Indo-European languages.

I. The Nominative is the case of the Subject of a sentence.

II. The Vocative is the case of Direct Address.

III. The Accusative is the case of the Direct Object of a verb. It is used also with many prepositions.

IV. The Genitive may generally be translated by the English Possessive, or by the Objective with the preposition *of*.

V. The Obliques might be found as:

a. The Dative, the case of the Indirect Object. It may usually be translated into English by the Objective with the preposition *to* or *for*.

b. The Locative, the place *where*.

c. The Instrumental, the thing *with*.

d. The Ablative, usually the Objective with *from*, *by*, *with*, *in* or *at*. It is often found with prepositions.

NOTE. The oblique cases appear in the English pronoun set; these pronouns are often called *objective pronouns*; as in *she loves me* (accusative), *give it to me* (dative) or *that dirt wasn't wiped with me* (instrumental), where me is not inflected differently in any of these uses; it is used for all grammatical relationships except the genitive case of possession and a non-disjunctive nominative case as the subject.

3.6. VOWEL GRADE

3.6.1. The vowel grade or Ablaut is normally the alternation between full, zero or lengthened grade vocalism. Proto-Indo-Eropean had a regular ablaut sequence that contrasted the five usual vowel sounds called *Thematic*, i.e. **e/ē/o/ō/Ø**. This means that in different forms of the same word, or in different but related words, the basic vowel, a

short **e**, could be replaced by a long **ē**, a short **o** or a long **ō**, or it could be omitted (transcribed as Ø).

NOTE. The term Ablaut comes from Ger. *Abstufung der Laute*, 'vowel alternation'. In Romance languages, the term Apophony is preferred.

3.6.2. When a syllable had a short **e**, it is said to be in the "*e grade*"; when it had no vowel, it is said to be in the "*zero grade*", when in **o**, in "*o grade*", and they can also be "*lengthened*". The e-grade is sometimes called "*full grade*".

A classic example of the five grades of *ablaut* in a single root is provided by the following different case forms of EIE **patḗr**, *father*, and **ṇpatōr**, *fatherless*.

Ablaut grade	EIE	Greek	(translit.)	Case
e-grade or full grade	**pa-ter-ṃ**	πα-τέρ-α	*pa-tér-a*	Accusative
lengthened e-grade	**pa-tḗr**	πα-τήρ	*pa-tḗr*	Nominative
zero-grade	**pa-tr-ós**	πα-τρ-ός	*pa-tr-ós*	Genitive
o-grade	**ṇ-pá-tor-ṃ**	ἀ-πά-τορ-α	*a-pá-tor-a*	Accusative
lengthened o-grade	**ṇ-pa-tōr**	ἀ-πά-τωρ	*a-pá-tōr*	Nominative

NOTE. Another example of the common Ablaut is t-stem **nepot**-, *grandson*, which gives lengthened grade Nominative, **nep-ōts**, full-grade Genitive **nép-ot-os**, and zero-grade feminine **nep-t-is**, *grand-daughter*. The study of declensions and practice with vocabulary should help the reader learn such special genitives.

3.6.3. Synoptic table of common examples of different vowel grades:

Vowel Grade	Full (F)	Zero (Ø)	Lengthened (L)
e/o - Ø - ē/ō	***dom***-	***dm***-	***dōm***-
je/jo - i - jē/jō	***djeu***-	***diw***-	***djēu***-
we/wo - u - wē/wō	***kwon***-	***kun***-	***kwōn***
ei/oi/ai - u/i - ēi/ōi/āi	***bheid***-	***bhid***-	***bhēid***-
eu/ou/au - u/i - ēu/ōu/āu	***bheud***-	***bhud***-	***bhēud***-
au/ai - u/i - āu/āi	***pau***-	***pu***-	***pāu***-
ā/ē/ō - a - ā/ē/ō	***stā***-	***sta***-	***stā***-
ēi/ōi - ū/ī - ēi/ōi	***pōi***	***pī***	***pōi***

3.6.4. There are also some other possible vowel grade changes, as *a-grade*, *i-grade* and *u-grade*, which usually come from old root endings, rather than from systematized phonetic changes.

NOTE. It seems that the alternation full-grade/zero-grade in PIH was dependent on the accent. Compare **klewos**/**klutós**, **eími**/**imés**, **patérm̥**/**patrós**, etc., where the unstressed morpheme loses its vowel. This happens only in the oldest formations, though, as Late PIE had probably lost this morphological pattern, freezing such older alternations and creating a new (more stable) vocabulary without changes in vowel grade.

3.7. WORD FORMATION

3.7.1. Word Formation refers to the creation of new words from older ones. Indo-European scholars show an especial interest in Derivational Affixes (most commonly Suffixes), i.e. morphemes that are attached to a base morpheme, such as a Root or a Stem, to form a new word. The main affixes are as follows.

3.7.2. Athematic suffixes:

a. The most simple is the zero-ending, i.e. full-grade root nouns like **dem-s** (Gk. *des-*), *house*, in consonant, as **neq-t-s** (Hitt. *nekuz*), *night*, or **men-s** (Av. *maz-*), *mind*, in -**r**, as **ghes-ōr** (Hitt. *kiššar*), *hand*, with apophony, Ac. **ghes-er-m̥** (Hitt. *kišširan*), Loc. **ghes-r-i** (Hitt. *kišri*, Gk. *kheiri*), with ending -**n**, as **or-ōn** (Hitt. *ḫara*[š], stem *ḫaran-*, from PIH *h₃or-o-,* cf. O.H.G. *aro*, Eng. *erne*, Gk. *or-n-*[*is*]), *eagle*. Common examples include **regs**, as Lat. *rex*, Cel. *ri*, Gmc. *rīh*, Skr. *rāt*, **cōus**, as Lat. *bou*, Cel. *bó*, Gmc. *ko*, Skr. *gáu/go*, **mūs**, Lat. *mūs*, Gk. μῦς, Gmc. *mūs*, Sla. *mys*, Skr. *mū*, etc.

b. Also, the stem **r**/**n**, with -**r**- in 'strong' cases (Nom-Acc.) and -**n**- in the Obliques, is well represented in Anatolian; see Variable Nouns in the next chapter for more on these so-called *heteroclites*.

c. An old stem in -**u**- appears e.g. in the words **gon-u**, *knee*, **dor-u**, *wood*, and **oj-u**, *lifetime*, cf. Av. *zānū*, *dārū*, *āiiū*, Skr. *jā́nu*, *dā́ru*, *ā́yu*, Gk. *góny*, *dóry*, *ou*(*kî*), "*no*", etc. Apophonic variants are found as full-grade **genu-**, **deru-**, **eju-**, cf. Hitt. *genu-*, Lat. *genu-*, Sla. *dérw-o*, Gk. *ai(w)-eí*, etc., and as zero-grade **gn-eu**, **dr-eu**, **j-eu**-, as in Goth. *kniu*, Av. *yaoš*, Hitt. *ganu-t*, etc. Such zero-grades are found within Declension, in Composition (cf. Skr. *jñu-bādh-*, "*kneeled*", Gk. *dru-tómos*, "*timber-cutter*"), and in Derivation, as e.g. **ju-wen-**, *vigorous, young* (cf. Skr. *yuván-*, Lat. *iuuen-is*).

d. A suffix -*it*-, which refers to edible substances, as **mel**-**it**, *honey* (cf. Gk. *mélit*-, Hitt. *milit*, Luw. *mallit*, Gmc. *mil*-), **sep**-**it**, *wheat* (cf. Hitt. *šeppit*, Gk. *álphit*), etc.

3.7.3. Feminine and Abstract (Collectives):

a. A general PIH suffix *-*(e)h₂* is found in Feminine, as in **senā**, *old* (<*seneh₂*, cf. Gk. *hénē*, Skr. *śanā*-, Lith. *senà*), **swekrús**, *mother-in-law* (<*swekrúh₂*-, cf. O.Sla. *svekrŭ*, Lat. *socrus*, O.H.G. *swigar*), in Abstract Collectives, as in Gk. **tomā́**, cut, or **neurā́**, *rope made from sinew* (cf. **neurom**, Eng. *neuron*), etc., and in the Nom.-Acc. Neuter singular of the collective that functions as Nom.-Acc. Plural (cf. Skr. *yugā́*, Gk. *zygá*, Lat. *iuga*, Goth. *juka*, "*jokes*", Hitt. -*a*, Pal. -*a/-ā*, etc.).

b. The Feminine and Abstract Collective PIH *-*ih₂* gives two alternative phonetic outputs in Late PIE, represented *-*jə* and (contracted) *-*ī*, written and pronounced as MIE -**ja** (not to be confused with general -**jā**, from PIH *-*jeh₂*) and -**ī**; cf. Skr. *devī́* (Gen. *dévyās*), "*goddess*", etc.

NOTE. According to Mallory & Adams (2006), "[t]he fact that Proto-Indo-European also forms collectives in *-*h₂*- (e.g. the Hittite collective *alpa*, '*group of clouds*' from a singular *alpeš*, '*cloud*') has suggested that this was its original use and that it later developed the specifically feminine meaning". Abstract nouns are usually built in PIE with suffix -**tāt**; as, **kommoinitā́ts**, *community*, **solwotāts**, *totality*, **áiwotāts**, *eternity*, etc.

3.7.4. Thematic Suffixes, the most abundant affixes found in PIE Nominal and Adjectival derivation:

a. A simple -**o**-, which appears in some primary and secondary old formations, as **wĺqo**-**s**, *wolf*, **ŗtko**-**s**, *bear*, neuters **jug**-**ó**-**m**, *joke*, **werg**-**o**-**m**, *work*, adjectives **sen**-**o**-, *old*, **new**-**o**-, *new*, etc.

NOTE. The Distinction into primary and secondary is not straightforward, unless there is an older root attested; compare e.g. PIE **ekw**-**os**, *horse*, which has been deemed a derivation from PIH *h₁ek*-, "*quick*", the root behind adjective **ōkús**.

Accented -**ó**- is deemed a secondary suffix which marks the possession of the base, as well as adjectives in -**ó**- with lengthened grade root, cf. PIE **cjā**, *bow's string*, as Skr. *jyá*, but **cjos**, *bow* (< "*that has a bow's string*"), as Gk. *biós*, or *swekrós (>**swekros**), *father-in-law*, from **swekrús**, *mother-in-law*, *deiwós (>**deiwos**), from **djēus**, etc.

b. About the Root Grade, o-grade roots are found in two thematic types, barytone Action Nouns (cf. Gk. **tomos**, "*slice*"), and oxytones Agent Nouns and Adjectives (cf.

Gk. **tomós**, "*who cuts, acute*"), both from PIE **tem**-, *cut*; zero-grade in neuters **jug-óm**, *joke*, from **jeug**-, *join*, and in second elements of compounds like **ni-sd-os**, *nest*, from **sed**, *sit*, or **newo-gn-ós**, "*newborn*", as Gk. *neognós*.

c. Adjectival suffixes -**jo**- and -**ijo**- have a relational sense, as in **cow-jós**, "*of a cow/ox*", from **cow**-, *cow, ox*, as in Av. *gaoya*-, Skr. *gavyá* or *gávya*, Gk. *hekatóm-boios*, "*that costs a hundred cows*", Arm. kogi (<**cow**-**ijo**-), "*derived from the cow*", O.Ir. *ambuæ* (<**n̥**-**cow**-**ijo**-, as in Skr. *ágos*, Gk. *aboúteō*), "*man without cows*", or e.g. **patr-jós**, *paternal*, **ped-jós**, "*of the foot*", etc. As a nominal suffix, cf. Lat. *ingenium*, *officium*, O.Ir. *cride, setig*, Skr. *vairya, saujanya*, Sla. *stoletie, dolia*, etc.

d. Verbal adjectives in -**tó**- (Ind.-Ira. -**nó**-), with zero-grade verbal root, are common in secondary derivation, as in **klu-tós**, *heard, famous*, from **kleu**-, *hear*, cf. Skr. *śrutá-*, Av. *sruta-*, Gk. *klytós*, Lat. *in-clitus*, M.Ir. *rocloth*, O.H.G. *Hlot-*, Arm. *lu*, etc. They were incorporated to the Verbal inflection as participles and gerunds. For nouns in -**to**-, -**no**-, -**ti(j)-o**-, -**ni(j)-o**-, -**tu(w)-o**-, -**nu(w)-o**-, etc. cf. Skr. *svápn(i)ya, prāvīnya*, Lat. *somnium, dominium*, O.Ir. *blíad(a)in*, Sla. *sunie, cozarenie*, etc.

e. Common thematic suffixes include -**nó**-, -**ro**-, -**mo**-, and diminutives in -**ko**-, -**lo**-, -**isko**-, etc. which may also be participial, ordinal or adjectival (from nouns) lengthenings. They are usually preceded by a vowel, as in -**e/onó**-, -**e/oro**-, and so on. Compare for example from **cher**-, *warm*, adjective **chor-mos**, *warm*, cf. Gmc. *warmaz*, Lat. *formus*, Skr. *gharmá*, Av. *garəma-*, Gk. *thermós*, Toc. A. *särme*, Phryg. *Germiai*, Arm. *jerm*, Alb. *zjarm*. -**bhó**- gives names of animals, as e.g. Gk. *éribhos*, *kid*.

f. A secondary suffix -**tero**-/-**toro**- marks the opposition of two notions, and is found in Anatolian (cf. Hitt. *nun-taras*, Adv. gen. "*from now*"), **ál-teros**, "*the other (of two)*" (cf. Goth. *anþar*, Skr. *ántaras*, Lat. *alter*, etc.) opposed to a simple "*other*", **aljos** (cf. Skr. *anyás*, Lat. *alius*, Gk. *állos*, Goth. *aljis*). This suffix is also found in some syntactic formations, as Gk. *deksiós – aris-terós, skaiós – deksi-terós*, both meaning "*right-left*" (Benveniste 1948).

g. The suffix -**wó**- is particularly found in words for "*alive*", as **cī-wó**- (cf. Skr. *jīvás*, Lat. *uīuos*, O.Ir. *béo*, Welsh *buw*, Goth. *qius*) and "*dead*", as **mr̥-wó**- (cf. O.Ir. *marb*, Welsh *marw*, and also Lat. *mortuos*, Sla. *mr̥tvŭ*, where the -**t**- was possibly inserted influenced by **mr̥-tó**-, "*mortal*").

h. There are some instrumental suffixes, as -***tro***-, -***tlo***-, -***klo***-, -***dhro***-, -***dhlo***-, as Lat. -*trum*, -*c(u)lum*, -*brum*, -*bulum*, etc.; e.g. **arā**-**trom**, *plough*, cf. Gk. *árotron*, Lat. *aratrum*, O.Ir. *arathar*, Welsh *aradr*, Arm. *arawr*, Lith. *árklas*, etc.; also, Gk. *báthron*, O.Ind. *bharítram*, Goth. *fōdr*, etc.

i. Other common suffixes (also participial) are -***men***-, -***mon***-, -***mn***-, with secondary -***mn***-***to***-, -***men***-***o***-, -***men***-***t***- (and -***wen***-, *v.s.*), etc., cf. Lat. *augmentum*, or Goth. *hliumant*, equivalent to O.Ind. *s͙rómatam*, both meaning "*reputation*", from **kleu**-, *hear*, and so on.

NOTE. Detailed information on Proto-Indo-European word morphology with dialectal examples is in Appendix III.3, and online at <http://dnghu.org/indoeuropean_noun_morphology.pdf>.

3.8. COMPOUND WORDS

3.8.1. *Nominal Compositum* or nominal composition is the process of putting two or more words together to form another word. The new word, called a Compound Word, is either a Noun or an Adjective, and it does not necessarily have the same meaning as its parts.

3.8.2. The second term of a Compound Word may be

a) a Noun (Gk. *akró-polis*, "*high city, citadel*")

b) an Adjective (Gk. *theo-eíkelos*, "*similar to the gods*") or

c) a Noun adapted to the adjectival inflection (Gk. *arguró-tozos*, "*silver arc*")

NOTE. Sometimes a suffix is added (cf. Gk. *en-neá-boios*, "*of nine cows*"), and the Compound Noun may have a different gender than the second term (cf. Lat. *triuium*, "*cross roads*", from *trēs* and *uia*).

3.8.3. The first term is a Pure Stem, without distinction of word class, gender or number. It may be an Adverb, a Numeral (Gk. *trí-llistos, "supplicated three times", polú-llistos, "very supplicated"*) or a Pronoun (cf. O.Ind. *tat-puruṣa*, "*that man*"), as well as a Nominal-Verbal stem with Nominal (Gk. *andra-phónos*, "*who kills a man*"), Adjetival (Gk. *akró-polis*), or Verbal function (Gk. *arkhé-kakos*, "*who begins the evil*"), and also an Adjective proper (Gk. *polú-tropos*, "*of many resources*").

3.8.4. Usually, the first term has zero-grade, cf. O.Ind. *ṇr-hán*, Gk. *polú-tropos*, Lat. *aui-*(*caps*), etc. Common exceptions are stems in -***e/os***, as Gk. *sakés-palos*, "*who shakes the shield*" (Gk. *sákos*, "*shield*"), and some suffixes which are substituted by a

lengthening in -**i**, cf. Gk. *kudi-áneira*, "*who glorifies men*" (Gk. *kudrós*), Av. *bərəzi-čaxra-*, "*of high wheels*" (Av. *bərəzant-*).

In Thematic stems, however, the thematic -**e∕o** appears always, as an **o** if Noun or Adjective (Gk. *akró-polis*), as an **e** if Verb (Gk. *arkhé-kakos*).

3.8.5. The first term usually defines the second, the contrary is rare; the main Compound types are:

A. Formed by Verbs, cf. O.Ind. *ņr-hán*, Gk. *andra-phónos* (Gk. *andro-* is newer) Lat. *auceps*, O.Sla. *medv-ĕdĭ*, "*honey-eater*", *bear*, and also with the second term defining the first, as Gk. *arkhé-kakos*.

B. Nominal Determiners (first term defines the second), with first term Noun (cf. Gk. *mētro-pátōr*, "*mother's father*", Goth. *þiudan-gardi*, "*kingdom*"), Adective (cf. Gk. *akró-polis*, O.Sla. *dobro-godŭ*, "*good time*", O.Ir. *find-airgit*, "*white plant*", Lat. *angi-portus*, "*narrow pass*"), or Numeral (cf. Lat. *tri-uium*, from *uia*, Gk. *ámaza*, "*chariot frame*", from *ázōn*).

C. Adjectival Determiners (*tatpuruṣa-* for Indian grammarians), with first term Noun (cf. Gk. *theo-eíkelos*, Goth. *gasti-gods* "*good for the guests*"), Adverb (cf. O.Ind. *ájñātas*, Gk. *ágnotos*, "*unknown*", *phroudos*, "*who is on its way*", from *pró* and *odós*).

D. Possessive Compounds (*bahu-vrihi-*, "*which has a lot of rice*", for Indian grammarians), as in Eng. *barefoot*, "*(who goes) with bare feet*", with the first term Noun (cf. Gk. *arguró-tozos*, O.Sla. *črŭno-vladŭ*, "*of black hair*"), Adjective (cf. Lat. *magn-animus*, "*of great spirit*"), Adverb (cf. O.Ind. *durmanā́s*, GK. *dus-menḗs*, "*wicked*").

The accent could also distinguish Determiners from Possessives, as in O.Ind. *rāja-putrás*, "*a king's son*", from O.Ind. *rajá-putras*, "*who has a son as king, king's father*".

3.9. NAMES OF PERSONS

The use of two-word compounds for personal names was common in PIE; as, **Suklewos**, *of good fame*, with cognates found in poetic diction, cf. Gk. *Eukleḗs*, and Skr. *Suśráva-*, or Illyr. *Vescleves-*.

NOTE. The use of two-word compound words for personal names is common in IE languages. They are found in in Ger. *Alf-red*, "*elf-counsel*", O.H.G. *Hlude-rīch*, "*rich in glory*", O.Eng. *God-gifu*, "*gift of God*" (Eng. *Godiva*), Gaul. *Orgeto-rix*, "*king who harms*", Gaul. *Dumno-rix*, "*king of the world*", Gaul. *Epo-pennus*, "*horse's head*", O.Ir. *Cin-néide* (Eng. *Kennedy*) "*ugly head*", O.Ind.

Asva-ghosa, "*tamer of horses*", O.Ind. *Asvá-medhas,* "*who has done the horse sacrifice*", O.Pers. *Xša-yāršā* (*Gk. Xérxēs*) "*ruler of heroes*", O.Pers. *Arta-xšacā*, "*whose reign is through truth/law*", Gk. *Sō-krátēs,* "*good ruler*", Gk. *Mene-ptólemos*, "*who faces war*", Gk. *Hipp-archus*, "*horse master*", Gk. *Cleo-patra, Pátro-klos*, "*from famous lineage*", Gk. *Arkhé-laos*, "*who governs the people*", O.Sla. *Bogu-milŭ*, "*loved by god*", Sla. *Vladi-mir*, "*peaceful ruler*", from *volodi-mirom*, "*possess the world*"; etc. Individual names may further be modified through the use of suffixes to form hypocorisms.

Other area in which it is suspected the retention of ancient Proto-Indo-European personal names is the use of animal names or numerals, composed of one stem; as **Wĺqos**, *wolf*, cf. O.Ir. *Olc*, O.Eng. *Wulf*, Gk. *Lukos*, Skr. *Vŕka*; or, **Qétwr̥tos**, *fourth*, cf. Lat. *Quārta*, Lith. *Keturai*, Russ. *Četvertoj*, Gk. *Tetartíōn*.

NOTE. The word for 'name' and possible Indo-European names can be found in Beekes (1987), Markey (1981), Pinault (1982), Schmitt (1973), and Watkins (1970).

Further, the syntactical indication of the father's name also dates from Proto-Indo-European, whether by adding the name of the father in the genitive, in the sense of 'son of X', or by adding a possessive adjective that is derived from the name of the father.

NOTE 1. An example of the former is *Hadubrand Heribrandes suno*; an example of the later is Myc. *a-re-ku-tu-ru-wo e-te-wo-ke-re-we-i-jo*, i.e. *Alektruwōn Etewoklewehijos*, "Alektruwōn, son of Etewoklewēs", or Russ. *Nikolaj Sergejevich*. Patronymics ending in -*ios* (later -*ius*) led to what is called the nomine gentile in Rome, cf. *Gaius Iulius Caesar* with *Gaius = praenomen <* individual name, *Iulius = nomen gentile <* patronymic and *Caesar = cognomen*.

When considering the giving of names to individuals, one departs generally from the basis of the free men.

NOTE. Whereas the man is addressed using the individual name, a simple 'oh woman' suffices in the case of woman. "The woman is treated more as a typus, the man as an individual". Wackernagel (1969) makes clear that the same forms of address were adopted for interactions with the gods. To say that the Indo-Europeans were not very different from the Romans and Greeks would not likely be too far from the mark. In Rome, women generally carried only the nomen gentile, cf. *Cornelia*, *Julia*, etc. In the case of the Greeks, most names of women are simply feminine forms of masculine names of individuals, e.g. Myc. *a-re-ka-sa-da-ra*, i.e *Aleksandrā* (corresponding to *Aleks-anōr*, "who fights off men"), Hom. *Andromákhē*, from *Andrómakhos*, "who fights with men", etc. (Meier-Brügger 2003).