A GRAMMAR OF

MODERN

INDO-EUROPEAN

Second Edition

Language and Culture

Writing System and Phonology

Morphology

Syntax

Texts and Dictionary

Etymology

DNGHU

Carlos Quiles

with   Fernando López-Menchero

Table of Contents

PREFACE

This second edition of _A Grammar of Modern Indo-European_ is a renewed effort to systematize the reconstructed phonology and morphology of Europe's Indo-European.

Modern Indo-European is common to most Europeans, and not only to some of them, as Latin, Germanic, or Slavic. Unlike Lingua Ignota, Solresol, Volapük, Esperanto, Quenya, Klingon, Lojban and the thousand invented languages which are imagined by individuals daily, PIE dialects are natural, i.e. they evolved from an older language – Proto-Indo-European, of which we have extensive knowledge –, and were spoken by prehistoric communities at some time roughly between 2500 and 2000 BC, having themselves evolved into different dialects already by 2000 BC.

Proto-Indo-European and its dialects have been reconstructed in the past two centuries (more or less successfully) by hundreds of linguists, having obtained a rough phonological, morphological, and syntactical system, equivalent to what Jews had of Old Hebrew before reconstructing a system for its modern use in Israel. Instead of some inscriptions and oral transmitted tales for the language to be revived, we have a complete reconstructed grammatical system, as well as hundreds of living languages to be used as examples to revive a common Modern Indo-European.

Some known philologists, university professors, experts in Classical Languages, still consider the Proto-Indo-European language reconstruction an "_invention_"; also, Spanish Indo-Europeanist Bernabé, a brilliant Spanish IE professor, has left its work on IE studies to dedicate himself to "_something more serious_". Francisco Villar, professor of Greek and Latin at the University of Salamanca, deems a complete reconstruction of PIE "_impossible_"; his opinion is not rare, since he supports the glottalic theory and the Armenian Homeland hypothesis (against the view of the majority), and supports the use of Latin instead of English within the EU. The work of Elst, Talageri and others defending the 'Indigenous Indo-Aryan' viewpoint by N. Kazanas, and their support of an unreconstructible and hypothetical PIE nearest to Vedic Sanskrit opens still more the gap between the mainstream reconstruction and minority views supported by political or personal opinions. Also, among convinced Indo-Europeanists, there seems to be no possible consensus between the different 'schools' as to whether Common PIE

distinguished between ŏ and ă (as Gk., Lat. or Cel.) or if those vowels were all initial ă, as in the other attested dialects (Villar), or if the Preterites were only one tense (as Latin *praeteritum*) with different formations, or if there were actually an Aorist and a Perfect.

Furthermore, José Antonio Pascual, a member of the Royal Spanish Academy (RAE), considers that "*it is not necessary to be a great sociologist to know that 500 million people won't agree to adopt Modern Indo-European in the EU*" (Spa. journal *El Mundo*, 8th April 2007). Of course not, as they won't agree on any possible question – not even on using English, which we use in fact –, and still the national and EU's Institutions keep working, adopting decisions by majorities, not awaiting consensus for any question. And it was probably not necessary to be a great sociologist a hundred years ago to see e.g. that the revival of Hebrew under a modern language system was a utopia (an "*impossible*", "*unserious*" "*invention*" then), and that Esperanto, the 'easy' and 'neutral' IAL, was going to succeed by their first so-called 'World Congress' in 1905. Such learned opinions are only that, opinions, just as if Hebrew and Semitic experts had been questioned a hundred years ago about a possible revival of Biblical Hebrew in a hypothetic new Land of Israel.

Whether MIE's success is more or less probable and why is not really important for our current work, but hypotheses dealt with by sociology, anthropology, political science, economics, psychology, etc. or usually just by personal opinions with no strict rational and reasonable basis. It remains unclear whether the project will be accepted by the different existing social movements, such as Pan-Latinism, Pan-Americanism, Pan-Sanskritism, Pan-Arabism, Pan-Iranism, Pan-Slavism, Pan-Hispanism, *Francophonie*, Anglospherism, Atlanticism, and the hundred different pan-nationalist ideas, as well as the different groups supporting anti-globalization, anti-capitalism, anti-communism, etc.

What we do know now is that the idea of reviving Europe's Indo-European as a modern language for Europe and international organizations is rational, that it is not something new, that it doesn't mean a revolution – as the use of Spanglish, Syndarin or Interlingua – nor an involution – as regionalism, nationalism, or the come back to French, German or Latin predominance –, but merely one of the many different ways in which the European Union linguistic policy could evolve, and maybe one way to unite different peoples from different cultures, languages and religions (from the Americas to East Asia) for the sake of stable means of communication. Just that tiny possibility is enough for us

to "lose" some years trying to give our best making the main Proto-Indo-European dialects as usable and as known as possible.

PREFACE TO THE FIRST EDITION

According to Dutch sociologist Abram de Swaan, every language in the world fits into one of four categories according to the ways it enters into (what he calls) the *global language system.*

- *Central*: About a hundred languages in the world belong here, widely used and comprising about 95% of humankind.
- *Supercentral*: Each of these serves to connect speakers of central languages. There are only twelve *supercentral* languages, and they are Arabic, Chinese, English, French, German, Hindi, Japanese, Malay, Portuguese, Russian, Spanish and Swahili.
- *Hypercentral*: The lone *hypercentral* language at present is English. It not only connects central languages (which is why it is on the previous level) but serves to connect supercentral languages as well. Both Spanish and Russian are supercentral languages used by speakers of many languages, but when a Spaniard and a Russian want to communicate, they will usually do it in English.
- *Peripheral*: All the thousands of other languages on the globe occupy a peripheral position because they are hardly or not at all used to connect any other languages. In other words, they are mostly not perceived as useful in a multilingual situation and therefore not worth anyone's effort to learn.

De Swaan points out that the admission of new member states to the European Union brings with it the addition of more languages, making the polyglot identity of the EU ever more unwieldy and expensive. On the other hand, it is clearly politically impossible to settle on a single language for all the EU's institutions. It has proved easier for the EU to agree on a common currency than a common language.

Of the EU's current languages, at least 14 are what we might call a 'robust' language, whose speakers are hardly likely to surrender its rights. Five of them (English, French, German, Portuguese and Spanish) are supercentral languages that are already widely used in international communication, and the rest are all central.

In the ongoing activity of the EU's institutions, there are inevitably shortcuts taken - English, French and German are widely used as 'working languages' for informal discussions. But at the formal level all the EU's official languages (i.e. the language of each member state) are declared equal.

Using all these languages is very expensive and highly inefficient. There are now 23 official languages: Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish Gaelic, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish and Swedish, and three *semiofficial* (?): Catalan, Basque and Galician. This means that all official documents must be translated into all the members' recognized languages, and representatives of each member state have a right to expect a speech in their language to be interpreted. And each member state has the right to hear ongoing proceedings interpreted into its own language.

Since each of the twenty one languages needs to be interpreted/translated into all the rest of the twenty, 23 x 22 (minus one, because a language doesn't need to be translated into itself) comes to a total of 506 combinations (not taking on accound the '*semiofficial*' languages). So interpreters/translators have to be found for ALL combinations.

In the old Common Market days the costs of using the official languages Dutch, English, French, and German could be borne, and interpreters and translators could be readily found. But as each new member is admitted, the costs and practical difficulties are rapidly becoming intolerably burdensome.

The crucial point here is that each time a new language is added, the total number of combinations isn't additive but multiplies: 506 + one language is not 507 but 552, i.e. 24 x 23, since every language has to be translated/interpreted into all the others (except itself).

It is not hard to see that the celebration of linguistic diversity in the EU only lightly disguises the logistical nightmare that is developing. The EU is now preparing for more languages to come: Icelandic and Norwegian might be added in the future, with the incorporation of these two countries to the EU, as well as Albanian, Macedonian, Serbian, Bosnian and Croatian (the three formerly known as Serbo-Croatian, but further differentiated after the Yugoslavian wars) if they are admitted to the EU as expected; and many other regional languages, following the example of Irish Gaelic, and the three *semi-official* Spanish languages: Alsatian, Breton, Corsican, Welsh, Luxemburgish and Sami

are likely candidates to follow, as well as Scottish Gaelic, Occitan, Low Saxon, Venetian, Piedmontese, Ligurian, Emilian, Sardinian, Neapolitan, Sicilian, Asturian, Aragonese, Frisian, Kashubian, Romany, Rusin, and many others, depending on the political pressure their speakers and cultural communities can put on EU institutions. It will probably not be long before Turkish, and with it Kurdish (and possibly Armenian, Aramaic and Georgian), or maybe Ukrainian, Russian and Belarusian, are other official languages, not to talk about the *eternal candidates*' languages, Norwegian (in at least two of its language systems, *Bokmål* and *Nynorsk*), Icelandic, Romansh, Monegasque (Monaco) and Emilian-Romagnolo (San Marino), and this could bring the number of EU languages over 40. The number of possible combinations are at best above 1000, which doesn't seem within the reach of any organization, no matter how well-meaning.

Many EU administrators feel that to a great extent this diversity can be canceled out by ever-increasing reliance on the computer translation that is already in heavy use. It is certainly true that if we couldn't count on computers to do a lot of the translation 'heavy lifting', even the most idealistic administrator would never even dream of saddling an organization with an enterprise that would quickly absorb a major part of its finances and energy. But no machine has yet been invented *or probably ever will be* that is able to produce a translation without, at the very least, a final editing by a human translator or interpreter.

The rapidly increasing profusion of languages in the EU is quickly becoming intolerably clumsy and prohibitively expensive. And this doesn't even count the additional expense caused by printing in the Greek alphabet and soon in the Cyrillic (Bulgarian and Serbian). Everyone agrees that all languages must have their 'place in the sun' and their diversity celebrated. But common sense suggests that the EU is going to be forced to settle on a very small number of working languages, perhaps only one, and the linguistic future of the EU has become the subject of intense debate.

Only in public numbers, the EU official translation/interpretation costs amount to more than 1230 M€, and it comes to more than 13% of today's administrative expenditure of the EU institutions. There are also indirect costs of linguistic programmes aimed at promoting the learning of three or more languages since the *Year of Languages* (2001), which also means hundreds of millions of Euros, which haven't been counted in

the EU's budget as linguistic expenditure, but are usually included in budget sections such as Cohesion or Citizenship. It is hard to imagine the huge amount of money (real or potential) lost by EU citizens and companies each day because of communication problems, not only because they *can't* speak a third party's language, but because they *won't* speak it, even if they can.

Preserving the strict equality is the EU's lifeblood, and it is a very disturbing thought that the strongest candidate for a *one-language EU* is the one with an established dominance in the world, English, which is actually only spoken by a minority within Europe. Latin and Artificial languages (as Esperanto, Ido or Interlingua) have been proposed as alternatives, but neither the first, because it is only related to romance languages, nor the second, because they are (too) artificial (invented by one person or a small group at best), solve the linguistic theoretical problems, not to talk about the practical ones.

The Indo-European language that we present in this work, on the contrary, faces not only the addressed theoretical problems - mainly related to cultural heritage and sociopolitical proud - but brings also a practical solution for the European Union, without which there can be no real integration. European nations are not prepared to give up some of their powers to a greater political entity, unless they don't have to give up some fundamental rights. Among them, the linguistic ones have proven harder to deal with than it initially expected, as they are raise very strong national or regional feelings.

Indo-European is already the grandmother of the majority of Europeans. The first language of more than 97% of EU citizens is Indo-European, and the rest can generally speak at least one of them as second language. Adopting Indo-European as the main official language for the EU will not mean giving up linguistic rights, but enhancing them, as every other official language will have then the same status under their common ancestor; it won't mean losing the own culture for the sake of unity, but recovering it altogether for the same purpose; and, above all, it will not mean choosing a *lingua franca* to communicate with foreigners within an international organization, but accepting a National Language to communicate with other nationals within the same country.

NOTE. The above information is mainly copied (literally, adjusted or modified) from two of Mr. William Z. Shetter *Language Miniatures*, which he published in his (now dead) website.

WHAT'S NEW IN THIS EDITION

This is _A Grammar of Modern Indo-European_, _Second Edition_, with Modern Indo-European Language Grammar in _Version 4_, still adjusting some important linguistic questions, and lots of minor mistakes, thanks to the contributions of experts and readers.

NOTE. A version number (N) is given to full revisions of the grammar, and each minor correction published must be given a different number to be later identified, usually ranging from N.01 to N.99. This book includes a full correction of version 3, following Pre-Version 4, which means the correction was finished, and it its therefore 4.xx.

"_Europe's Indo-European_" version 4 continues "_Modern Indo-European_" version 3 (first printed edition, since June 2007), and this in turn version 2, which began in March 2007, changing most features of the old "_Europaio_"/"_Sindhueuropaiom_" concept of version 1 (_Europaio: A Brief Grammar of the European Language_, 2005-2006).

1. Apart from the unified "_Modern Indo-European_", based on Europe's Indo-European (also _residual or_ North-Western Indo-European, or _Proto-European_), this grammar makes reference to other coeval PIE early dialects, especially Proto-Greek, Proto-Indo-Iranian (or _Proto-Aryan_) and Common Anatolian.

2. One of the main changes of this version is the adoption of a writing system with a clear phonological distinction between **i**, **u** and their semivocalic allophones **j**, **w**. The artificial distinction of **i/j** and **u/w** in PIE roots and derivatives, hold in versions 1-3, was untenable in the long term, as it was a labile decision, open to future changes. With the traditional written differentiation we get a greater degree of stability.

3. Emphasis is on the old Latin-only alphabet, but attention is paid to Greek and Cyrillic writing systems. Stubs of possible Armenian, Arabo-Persian and Devanagari (Abugida) systems are also included. The objective is not to define them completely (as with the Latin alphabet), but merely to show other possible writing systems for Modern Indo-European, Modern Aryan, and Modern Hellenic languages.

4. The traditional distinction in writings of the controversial _palatovelar_ phonemes has been extensively discussed and rejected. Whether _satemization_ appeared already as a dialectal phonological trend in Late PIE, or were just similar individual dialectal innovations restricted to some phonetic environments (**k**- before some sounds, as with

Latin *c-* before *-e* and *-i*), is not important. Reasons for not including the *palatovelars* in MIE writing system are 1) that, although possible, their existence is not sufficiently proven (see Appendix II.2); 2) that their writing because of tradition or even 'etymology' is not justified, as this would mean a projective writing (i.e., like writing Lat. *casa*, but Lat. *ĉentum*, because the k-sound before *-e* and *-i* evolves differently in Romance).

5. The historically alternating *Oblique* cases *Dative, Locative, Instrumental* and *Ablative*, are shown on a declension-by-declension (and even pronoun-by-pronoun) basis, as Late PIE shows in some declensions a simpler reconstructible paradigm (for some more archaic, for others an innovation) while others show almost the same Late PIE pattern of four differentiated oblique case-endings. The 8 cases traditionally reconstructed are used – and its differentiation recommended – in MIE.

5. The so-called Augment in *é-*, attested almost only in Greek, Indo-Iranian and Armenian, is sometimes left due to tradition of Indo-European studies, although recent research has shown that it was neither obligatory, nor general in the earliest PIE dialects. It is believed today that it was just a prefix that had a great success in the southern dialects, just like *per-* (<PIE **per**-) in Latin, or *ga-* (<PIE **ko**-) in Germanic.

6. The syntactical framework of Late PIE has been dealt with extensively by some authors, but, as the material hasn't still been summed up and corrected within *mainstream* Indo-European linguistics – Indo-Europeanists usually prefer the phonological or morphological reconstruction –, we use literal paragraphs from possibly the most thorough work available on PIE syntax, Winfred P. Lehmann's <u>Proto-Indo-European Syntax</u> (1974), adding comments and corrections made since its publication by other scholars

7. The whole section on Morphosyntax is taken from Michael Meier-Brügger's <u>Indo-European Linguistics</u> (2003).

8. Appendices I and III were written by Fernando López-Menchero and published 2007-2009. The rest of this book has been written thanks to his countless corrections and additions in those years.

ACKNOWLEDGEMENTS

CONVENTIONS USED IN THIS BOOK

1. *Modern Indo-European* (MIE), **Eurōpājóm** or *European* are used only to refer to the *European language*, i.e. to the modern language system based on the reconstructed *North-West* or *Europe's* Indo-European (EIE), also *Old European* proto-language.

2. The roots of the reconstructed PIE language are basic morphemes carrying a lexical meaning. By addition of suffixes, they form stems, and by addition of desinences, these form grammatically inflected words (nouns or verbs).

NOTE. PIE reconstructed roots are subject to *ablaut*, and except for a very few cases, such ultimate roots are fully characterized by its constituent consonants, while the vowel may alternate. PIE roots as a rule have a single syllabic core, and by ablaut may either be monosyllabic or unsyllabic. PIE roots may be of the following form (where K is a voiceless stop, G an unaspirated and G^h an aspirated stop, R a semivowel (r̥, l̥, m̥, n̥, **w**, **j**) and H a laryngeal (or *s*). According to Meillet, impossible PIE combinations are voiceless/aspirated (as in **teub^h* or **b^{h}eut*), as well as voiced/voiceless (as in **ged* or **deg*). The following table depicts the general opinion:

stops	-	*K-*	*G-*	*G^h-*
-	[HR]e[RH]	K[R]e[RH]	G[R]e[RH]	G^h[R]e[RH]
-K	[HR]e[RH]K	-	G[R]e[RH]K	G^h[R]e[RH]K
-G	[HR]e[RH]G	K[R]e[RH]G	-	G^h[R]e[RH]G
-G^h	[HR]e[RH]G^h	K[R]e[RH]G^h	G[R]e[RH]G^h	G^h[R]e[RH]G^h*

*This combination appears e.g. in **b^{h}eud^h**-, *awake*, and **b^{h}eid^h**-, *obey, believe*.

A root has at least one consonant, for some at least two (e.g. PIH **h₁ek-* vs. EIE **ek**-, "*quick*", which is the root for MIE adj. **ōkús**). Depending on the interpretation of laryngeals, some roots seem to have an inherent vowel **a** or **o**; as, EIE **ar**- (vs. PIH **h₂ar-*), *fit*, EIE **ong^w**- (vs. PIH **h₃eng^w-*) "*anoint*", EIE **ak**- (vs. PIH **h₂ek-*) "*keen*".

By "root extension", a basic CeC (with C being any consonant) pattern may be extended to CeC-C, and an s-mobile may extend it to s-CeC.

The total number of consonant, sonant and laryngeal elements that appear in an ordinary syllable are three – i.e., as the triliteral Semitic pattern. Those which have less than three are called 'Concave' verbs (cf. PIH **Hes-*, **Hei-*, **g^{w}em-*); those extended are called 'Convex' verbs (cf. Lat. *plangō*, *spargō*, *frangō*, etc., which, apart from the extension in -*g*, contain a laryngeal); for more on this, *vide infra* on MIE Conjugations.

3. In this book PIE roots are usually written with laryngeals. Therefore, we only assume certainty in the non-laryngeal nature of early PIE dialects, from ca. 2500 BC on. Whether Late PIE lost them all sooner (ca. 3500 BC?) or later (ca. 2500 BC?), etymological roots which include laryngeals will often be labelled as PIH, or just as (general) PIE, while specific Late PIE vocabulary will be shown with an undetermined laryngeal output *ə.

NOTE. Common PIE (or PIH) roots are reconstructed by most modern Indo-Europeanists with laryngeals; so e.g. different vowel outputs of early PIE dialects (like North-West IE or Proto-Greek) are explained through the phonological effects of old aspirated phonemes on adjacent vowels. See Appendix II.3 for more on this question.

Some linguists still follow the old non-laryngeal PIE concept (see Walde-Pokorny's *lexica*), while many only conceive a PIE with laryngeals. However, it is not logical to assume that, whereas in Proto-Anatolian laryngeals were lost or evolved, in Late PIE they were the same (*h_1, *h_2, *h_3, or any other scheme) as in their common ancestor, Middle PIE. Therefore, some scholars have adapted the Late PIE reconstruction to a *partially laryngeal* or *non-laryngeal language* (see Adrados, Nikolayev, etc.), coeval with the *partially laryngeal* PAn, thus supposing a similar laryngeal loss in both Middle PIE dialects, usually implying a quicker loss in Late PIE, due to the conservation of laryngeals in Anatolian, and their complete disappearance in Late PIE dialects. Some still reconstruct for Late PIE an uncertain laryngeal (or vowel) *H or *ə, in some phonetic environments, otherwise difficult to explain, prior to its full loss in early PIE dialects.

4. Proto-Indo-European vowel apophony or Ablaut is indeed normal in MIE, but dialectal Ablaut must be corrected when loan-translated. Examples of these Lat. *confessus* (cf. Lat. *fassus sum*), from EIE **bʰā**-; Lat. *facilis/difficilis*, from EIE **dʰē**-; Lat. *saliō/insiliō/insultō,* etc.

NOTE. Such Ablaut is linked to languages with musical accent, as Latin. In Italic, the tone was always on the first syllable; Latin reorganized this system, and after Roman grammarians' "*paenultima* rule", Classic Latin accent felt on the second to last syllable if long, on the third to last syllable, or *antepaenultima*, if short (hence Lat. *pudícus* but *módicus*), thus triggering off different inner vocalic timbres or *Ablauts*. Other Italic dialects, as Oscan or Umbrian, didn't suffered such apophony; cf. Osc. *anterstataí* , Lat. *interstitae*; Umb. *antakres*, Lat. *integris*; Umb. *procanurent*, Lat. *procinuerint*, etc. Germanic also knew such tone variations. More on Latin phonotactic development at <http://www.cunyphonologyforum.net/SYLLPAPERS/Senhandoutnew.pdf>.

5. In Romance languages, *Theme* is used instead of *Stem*. Therefore, *Theme Vowel* and *Thematic* refer to the Stem endings, usually to the **e/o** endings. In the Indo-European

languages, *Thematic* roots are those roots that have a "*theme vowel*"; a vowel sound that is always present between the root of the word and the attached inflections. *Athematic* roots lack a theme vowel, and attach their inflections directly to the root itself.

NOTE. The distinction between thematic and athematic roots is especially apparent in the Greek verb; they fall into two classes that are marked by quite different personal endings. Thematic verbs are also called *-ω* (*-ô*) verbs in Greek; athematic verbs are *-μι* (*-mi*) verbs, after the first person singular present tense ending that each of them uses. The entire conjugation seems to differ quite markedly between the two sets of verbs, but the differences are really the result of the thematic vowel reacting with the verb endings. In Greek, athematic verbs are a closed class of inherited forms from the parent IE language. Marked contrasts between thematic and athematic forms also appear in Lithuanian, Sanskrit, and Old Church Slavonic. In Latin, almost all verbs are thematic; a handful of surviving athematic forms exist, but they are considered irregular verbs.

The thematic and athematic distinction also applies to nouns; many of the old IE languages distinguish between "vowel stems" and "consonant stems" in the declension of nouns. In Latin, the first, second, fourth, and fifth declensions are vowel stems characterized by *a*, *o*, *u* and *e*, respectively; the third declension contains both consonant stems and *i* stems, whose declensions came to closely resemble one another in Latin. Greek, Sanskrit, and other older IE languages also distinguish between vowel and consonant stems, as did Old English.

6. PIE *$d+t$, *$t+t$, *$dh+t$ → MIE **st**; PIE *$d+d$, *$t+d$, *$dh+d$ → MIE **sd**; PIE *$d+dh$, *$t+dh$, *$dh+dh$ → MIE **sdh**; because of the common intermediate phases found in Proto-Greek, cf. Gk. *st*, *sth* (as *pistis*, *oisqa*), and Europe's Indo-European, cf. Lat. *est*, "*come*", and O.H.G. examples. For an earlier stage of this phonetic output, compare O.Ind. *sehí*<**sazdhi*, 'sit!', and not **satthi* (cf. O.Ind. *dehí*, Av. *dazdi*).

NOTE. It has been proposed an old PIE TT→TsT (where T = dental stop), i.e. that the cluster of two dental stops had a dental fricative **s** inserted between them. It is based on some findings in Hittite, where cluster *tst* is spelled as *z* (pronounced as *ts*), as in PIH *h_1ed-ti, "*he eats*" → *h_1etsti → Hitt. *ezzi*. Confirmation from early intermediate and common (Late PIE) *-st- are found e.g. in O.Ind. *mastis*, "*measure*", from *med-tis, or Av. *-hasta-*, from *sed-tós. This evolution was probably overshadowed by other Aryan developments, see Appendix II.

7. The Feminine Late PIE *-*jə/-ī*, old Abstract Collective PIH *-*ih₂*, gives EIE *-**ja**/-ī*. While both were still interchangeable in the common North-West IE (as the different dialectal outputs show), we prefer to use the ending *-**ja*** for feminines, and *-**ī*** for neuters; as, **smīghslī**, *thousand* (neuter), but **trja**, *three* (fem.).

The following abbreviations apply in this book:

IE	: Indo-European
IE II	: Middle PIE or PIH
PIH	: Proto-Indo-Hittite
IE III	: Late PIE
PIE	: Proto-Indo-European
EIE	: Europe's Indo-European
MIE	: Modern Indo-European

PAn	: Proto-Anatolian
CA	: Common Anatolian
Hitt.	: Hittite
Luw.	: Luwian
Lyc.	: Lycian
Pal.	: Palaic
Lyd.	: Lydian

PII	: Proto-Indo-Iranian
Ind.	: *Proto-Indo-Aryan*
O.Ind.	: Old Indian
Skr.	: Sanskrit
Hind.	: Hindustani
Hi.	: Hindi
Ur.	: Urdu
Ira.	: *Proto-Iranian*
Av.	: Avestan
O.Pers.	: Old Persian
Pers.	: Persian
Kur.	: Kurdish
Oss.	: Ossetian
Kam.	: Kamviri

PGk	: Proto-Greek
Gk.	: (Ancient) Greek
Phryg.	: Phrygian
Thr.	: Thracian
Dac.	: Dacian
Ven.	: Venetic
Lus.	: Lusitanian
A.Mac.	: Ancient Macedonian
Illy.	: Illyrian
Alb.	: Albanian

Ita.	: Proto-Italic
Osc.	: Oscan
Umb.	: Umbrian
Lat.	: Latin
O.Lat.	: Archaic Latin
V.Lat.	: Vulgar Latin
L.Lat.	: Late Latin
Med.Lat.	: Mediaeval Latin
Mod.Lat.	: Modern Latin
O.Fr.	: Old French
Prov	: Provenzal
Gl.-Pt.	: Galician-Portuguese
Gal.	: Galician
Pt.	: Portuguese
Cat.	: Catalan
Fr.	: French
It.	: Italian
Spa.	: Spanish
Rom.	: Romanian

PGmc.	: Pre-Proto-Germanic
Gmc.	: Proto-Germanic
Goth.	: Gothic
Frank.	: Frankish
Sca.	: *Scandinavian (N. Gmc.)*
O.N.	: Old Norse
O.Ice.	: Old Icelandic
O.S.	: Old Swedish
Nor.	: Norwegian
Swe.	: Swedish
Da.	: Danish
Ice.	: Icelandic
Fae.	: Faeroese
W.Gmc.	: *West Germanic*
O.E.	: Old English (*W.Saxon, Mercian*)
O.Fris.	: Old Frisian
O.H.G.	: Old High German
M.L.G.	: Middle Low German
M.H.G.	: Middle High German
M.Du.	: Middle Dutch
Eng	: English
Ger.	: German
L.Ger.	: Low German
Fris.	: Frisian
Du.	: Dutch
Yidd.	: Yiddish

Cel.	: Proto-Celtic
Gaul.	: Gaulish
O.Ir.	: Old Irish
Sco.	: Scottish Gaelic
Ir.	: Irish Gaelic
Bret.	: Breton
Cor.	: Cornish
O.Welsh	: Old Welsh

BSl.	: Balto-Slavic
Bal.	: *Proto-Baltic*
O.Lith.	: Old Lithuanian
O.Pruss.	: Old Prussian
Lith.	: Lithuanian
Ltv.	: Latvian
Sla.	: *Proto-Slavic*
O.C.S.	: Old Church Slavonic
O.Russ.	: Old Russian
O.Pol.	: Old Polish
Russ.	: Russian
Pol.	: Polish
Cz.	: Czech
Slo.	: Slovenian
Slk.	: Slovak
Ukr.	: Ukrainian
Bel.	: Belarusian
Bul.	: Bulgarian
Sr.-Cr.	: Serbo-Croatian

1. INTRODUCTION

1.1. THE INDO-EUROPEAN LANGUAGE FAMILY

1.1.1. The Indo-European languages are a family of several hundred modern languages and dialects, including most of the major languages of Europe, as well as many in Asia. Contemporary languages in this family include English, German, French, Spanish, Portuguese,
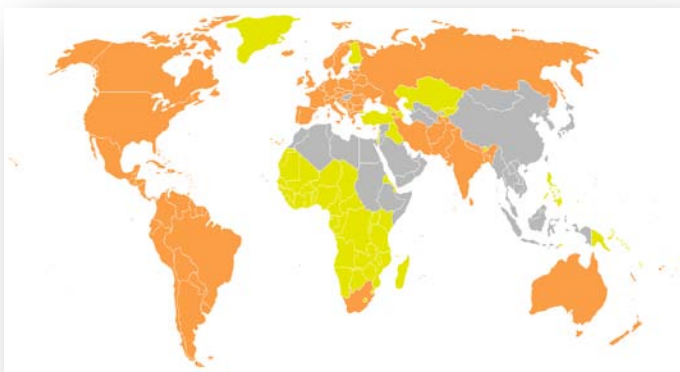
In dark, countries with a majority of Indo-European speakers; in light color, countries with Indo-European-speaking minorities.

Hindustani (i.e., Hindi and Urdu among other modern dialects), Persian and Russian. It is the largest family of languages in the world today, being spoken by approximately half the world's population as mother tongue. Furthermore, the majority of the other half speaks at least one of them as second language.

1.1.2. Romans didn't perceive similarities between Latin and Celtic dialects, but they found obvious correspondences with Greek. After Grammarian Sextus Pompeius Festus:

Suppum antiqui dicebant, quem nunc supinum dicimus ex Graeco, videlicet pro adspiratione ponentes <s> litteram, ut idem ὔλας dicunt, et nos silvas; item ἕξ sex, et ἑπτά septem.

Such findings are not striking, though, as Rome was believed to have been originally funded by Trojan hero Aeneas and, consequently, Latin was derived from Old Greek.

1.1.3. Florentine merchant Filippo Sassetti travelled to the Indian subcontinent, and was among the first European observers to study the ancient Indian language, Sanskrit. Writing in 1585, he noted some word similarities between Sanskrit and Italian, e.g. *deva/dio*, "*God*", *sarpa/serpe*, "*snake*", *sapta/sette*, "*seven*", *ashta/otto*, "*eight*", *nava/nove*, "*nine*". This observation is today credited to have foreshadowed the later discovery of the Indo-European language family.

1.1.4. The first proposal of the possibility of a common origin for some of these languages came from Dutch linguist and scholar Marcus Zuerius van Boxhorn in 1647. He discovered the similarities among Indo-European languages, and supposed the existence of a primitive common language which he called "*Scythian*". He included in his hypothesis Dutch, Greek, Latin, Persian, and German, adding later Slavic, Celtic and Baltic languages. He excluded languages such as Hebrew from his hypothesis. However, the suggestions of van Boxhorn did not become widely known and did not stimulate further research.

1.1.5. On 1686, German linguist Andreas Jäger published *De Lingua Vetustissima Europae*, where he identified an remote language, possibly spreading from the Caucasus, from which Latin, Greek, Slavic, 'Scythian' (i.e. Persian) and Celtic (or 'Celto-Germanic') were derived, namely *Scytho-Celtic*.

1.1.6. The hypothesis re-appeared in 1786 when Sir William Jones first lectured on similarities between four of the oldest languages known in his time: Latin, Greek, Sanskrit and Persian:

"The Sanskrit 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 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 Gothic and the Celtic, though blended with a very different idiom, had the same origin with the Sanskrit; and the old Persian might be added to the same family"

1.1.7. Danish Scholar Rasmus Rask was the first to point out the connection between Old Norwegian and Gothic on the one hand, and Lithuanian, Slavonic, Greek and Latin on the other. Systematic comparison of these and other old languages conducted by the young German linguist Franz Bopp supported the theory, and his *Comparative Grammar*, appearing between 1833 and 1852, counts as the starting-point of Indo-European studies as an academic discipline.

NOTE. The term Indo-European itself now current in English literature, was coined in 1813 by the British scholar Sir Thomas Young, although at that time there was no consensus as to the naming of the recently discovered language family. Among the names suggested were *indo-*

germanique (C. Malte-Brun, 1810), *Indoeuropean* (Th. Young, 1813), *japetisk* (Rasmus C. Rask, 1815), *indisch-teutsch* (F. Schmitthenner, 1826), sanskritisch (Wilhelm von Humboldt, 1827), indokeltisch (A. F. Pott, 1840), arioeuropeo (G. I. Ascoli, 1854), Aryan (F. M. Müller, 1861), aryaque (H. Chavée, 1867), etc.

In English, Indo-German was used by J. C. Prichard in 1826 although he preferred Indo-European. In French, use of indo-européen was established by A. Pictet (1836). In German literature, Indo-Europäisch was used by Franz Bopp since 1835, while the term Indo-Germanisch had already been introduced by Julius von Klapproth in 1823, intending to include the northernmost and the southernmost of the family's branches, as it were as an abbreviation of the full listing of involved languages that had been common in earlier literature, opening the doors to ensuing fruitless discussions whether it should not be *Indo-Celtic*, or even *Tocharo-Celtic*.

1.1.8. There are certain common linguistic ancestors of modern IE languages, and some of them are well-attested dead languages (or language systems), such as Latin for modern Romance languages – French, Spanish, Portuguese, Italian, Romanian or Catalan –, Sanskrit for some modern Indo-Aryan languages, or Greek for Modern Greek.

Furthermore, there are some still older IE languages, from which these old formal languages were derived and later systematized. They are, following the above examples, *Archaic* or *Old Latin*, *Archaic* or *Vedic Sanskrit* and *Archaic* or *Old Greek*, attested in older compositions or inscriptions, or inferred through the study of oral traditions and even foreign texts, like the Indo-Aryan superstrate of the Mitanni.

And there are also some old related dialects, which help us reconstruct proto-languages, such as Osco-Umbrian for an older *Proto-Italic* (and with Proto-Celtic, Proto-Germanic and Proto-Balto-Slavic for Europe's Indo-European), Indo-Aryan dialects for Proto-Indo-Aryan (and with Proto-Iranian for *Proto-Indo-Iranian*) or Mycenaean for an older *Proto-Greek*.

NOTE. Mallory and Adams (2006) argue, about (Late) Proto-Indo-European: "How real are our reconstructions? This question has divided linguists on philosophical grounds.

o There are those who argue that we are not really engaged in 'reconstructing' a past language but rather creating abstract formulas that describe the systematic relationship between sounds in the daughter languages.

o Others argue that our reconstructions are vague approximations of the proto-language; they can never be exact because the proto-language itself should have had different dialects (yet we reconstruct only single proto-forms) and our reconstructions are not set to any specific time.

o Finally, there are those who have expressed some statistical confidence in the method of reconstruction. Robert Hall, for example, claimed that when examining a test control case, reconstructing proto-Romance from the Romance languages (and obviously knowing beforehand what its ancestor, Latin, looked like), he could reconstruct the phonology at 95% confidence, and the grammar at 80%. Obviously, with the much greater time depth of Proto-Indo-European, we might well wonder how much our confidence is likely to decrease.

Most historical linguists today would probably argue that [laryngeal PIE] reconstruction results in approximations. A time traveller, armed with this book and seeking to make him- or herself understood would probably engender frequent moments of puzzlement, not a little laughter, but occasional instances of lucidity".

1.2. TRADITIONAL VIEWS

1.2.1. In the beginnings of the Indo-European or Indo-Germanic studies using the comparative grammar, the Indo-European proto-language was reconstructed as a unitary language. For Rask, Bopp and other Indo-European scholars, it was a search for *the* Indo-European. Such a language was supposedly spoken in a certain region between Europe and Asia and at one point in time – between ten thousand and four thousand years ago, depending on the individual theories –, and it spread thereafter and evolved into different languages which in turn had different dialects.

1.2.2. The *Stammbaumtheorie* or Genealogical Tree Theory states that languages split up in other languages, each of them in turn split up in others, and so on, like the branches of a tree. For example, a well known old theory about Indo-European is that, from *the* PIE language, two main groups of

Modern tree diagram of the IE languages by Eric Hamp (1990), Mallory & Adams (2007).
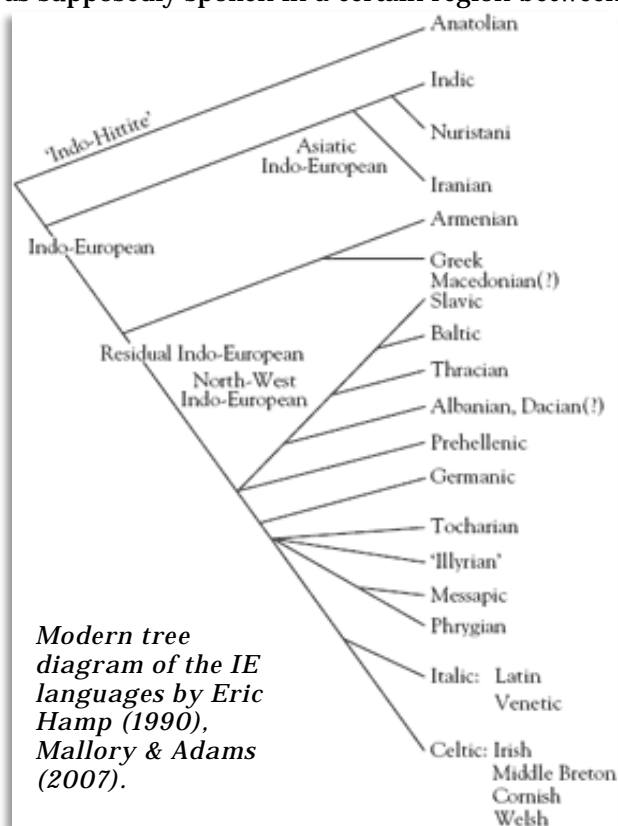
dialects known as *Centum* and *Satem* separated – so called because of their pronunciation of PIE *$kmtóm$, "*hundred*", in Latin and Avestan. From these groups others split up, as *Centum* Proto-Germanic, Proto-Italic or Proto-Celtic, and *Satem* Proto-Balto-Slavic, Proto-Indo-Iranian.

NOTE. The *Centum* and *Satem* isogloss is one of the oldest known phonological differences of IE languages, and is still used by many to classify them in two main dialectal groups – postulating the existence of *proto-Centum* and a *proto-Satem* –, disregarding their relevant morphological and syntactical differences. The isogloss is based on a simple vocabulary comparison; as, from PIE *$kmtóm$ (possibly earlier *$dkmtóm$, from *$dekm$, "*ten*"), *Satem*: O.Ind. *śatám*, Av. *satəm*, Lith. *šimtas*, O.C.S. *sto*, or *Centum*: Gk. ἑκατόν, Lat. *centum*, Goth. *hund*, O.Ir. *cet*, etc.

It remains the most used model for understanding the Indo-European language reconstruction, since it was proposed by A. Schleicher (*Compendium*, 1866). The problem with its simplicity is that "the branching of the different groups is portrayed as a series of clean breaks with no connection between branches after they have split, as if each dialectal group marched away from the rest. Such sharp splits are possible, but assuming that all splits within Proto-Indo-European were like this is not very plausible, and any linguist surveying the current Indo-European languages would note dialectal variations running through some but not all areas, often linking adjacent groups who may belong to different languages" (Mallory & Adams, 2006).

1.2.3. The *Wellentheorie* or Waves Theory, of J. Schmidt, states that one language is created from another by the spread of innovations, the way water waves spread when a stone hits the water surface. The lines that define the extension of the innovations are called isoglosses. The convergence of different isoglosses over a common territory signals the existence of a new language or dialect. Where isoglosses from different languages coincide, transition zones are formed.
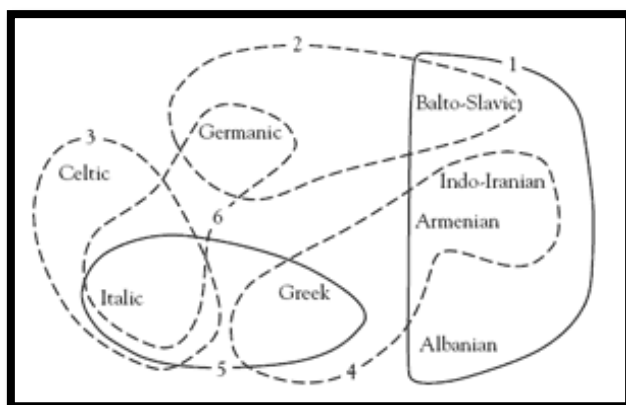
"Wave model" of some of the interrelationships of the IE languages, Mallory & Adams (2007).

NOTE. These old models for our understanding of language reconstructions are based on the hypothesis that there was one common and *static* Proto-Indo-European language, and that all features of modern Indo-European languages can be explained in such unitary schemes, by classifying them either as innovations or as archaisms of one old, rigid proto-language. After Mallory and Adams (*The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, 2006), "their criteria of inclusion, why we are looking at any particular one, and not another one, are no more solid than those that define family trees. The key element here is what linguistic features actually help determine for us whether two languages are more related or less related to one another. A decision in this area can be extraordinary difficult because we must be able to distinguish between features that may have been present throughout the entire Indo-European world (*Indoeuropeia* has been employed to describe this concept) and have dropped out in some but not others against those features that are innovations in only some of the different groups. The historical linguist is principally looking for shared innovations, i.e. are there traces of corresponding developments between two or more language groups that would indicate that they shared a common line of development different from other language groups? Only by finding shared innovations can one feel confident that the grouping of individual Indo-European linguistic groups into larger units or branches of the tree is real".

1.2.4. Because of the difficulties found in the modelling of Proto-Indo-European branches and daughter languages into the traditional, unitary 'Diverging Tree' framework, i.e. a uniform Proto-Indo-European language with its branches, a new model called '*Converging Association of Languages*' was proposed, in which languages that are in contact (not necessarily related to each other) exchange linguistic elements and rules, thus developing and acquiring from each other. Most linguists have rejected it as an implausible explanation of the irregularities found in the old, static concept of PIE.

NOTE. Among the prominent advocates is N.S. Trubetzkoy (*Urheimat*, 1939): "The term 'language family' does not presuppose the common descent of a quantity of languages from a single original language. We consider a 'language family' a group of languages, in which a considerable quantity of lexical and morphological elements exhibit regular equivalences (...) it is not necessary for one to suppose common descent, since such regularity may also originate through borrowings between neighboring unrelated languages (...) It is just as conceivable that the ancestors of the Indo-European language branches were originally different from each other, but though constant contact, mutual influence, and borrowings, approached each other, without however ever becoming identical to one another" (Meier-Brügger, 2003).

Agreeing with Neumann (1996), Meier-Brügger (2003) states that "that the various Indo-European languages have developed from a prior unified language is certain. Questionable is, however, the concrete 'how' of this process of differentiation", and that this "thesis of a 'converging association of languages' may immediately be dismissed, given that all Indo-European languages are based upon the same Proto-Indo-European flexion morphology. As H. Rix makes clear, it is precisely this morphological congruence that speaks against the language association model, and for the diverging tree model", even if the traditional language tree models were unable to explain the newest findings.

1.3. THE THEORY OF THE THREE STAGES

1.3.1. Even some of the first Indo-Europeanists had noted in their works the possibility of older origins for the reconstructed (Late) Proto-Indo-European, although they didn't dare to describe those possible older stages of the language.

NOTE. The development of this theory of three linguistic stages can be traced back to the very origins of Indo-European studies, firstly as a diffused idea of a non-static PIE language, and later widely accepted as a dynamic dialectal evolution, already in the 20th century, after the deciperment of the Anatolian scripts. Most
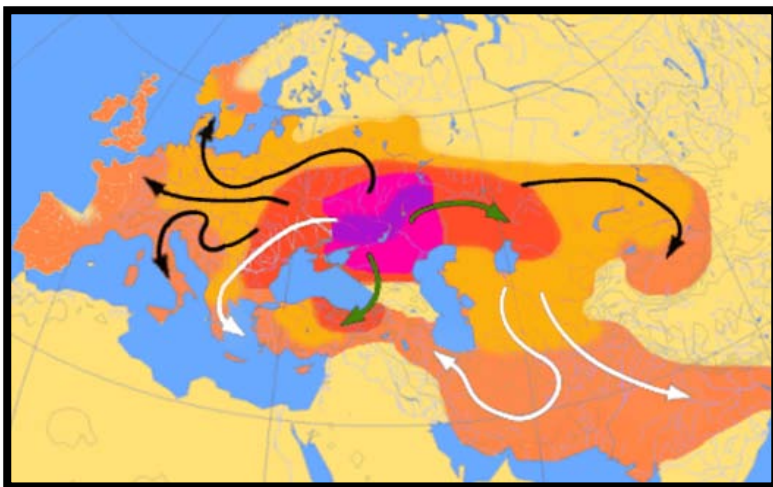
Expansion of Proto-Indo-European and its dialects ca. 4000BC-500AD, according to the Kurgan hypothesis.

linguists accept that Proto-Indo-European must be the product of a long historical development. On this question, H. Rix (Modussystem, 1986) asserts "Whereby comparative reconstruction is based upon a group of similar forms in a number of languages, internal reconstruction takes its point of departure from irregularities or inhomogeneities of the system of a single language (...) The fundamental supposition of language-internal reconstruction is that such an irregularity or inhomogeneity in the grammar of a language is the result of a diachronic process, in which an older pattern,, or homogeneity is eclipsed, but not fully suppressed". According to Meier-Brügger (2003), "Rix works back from Late Proto-Indo-European Phase B (reconstructible Proto-Indo-European) using deducible information about an Early Proto-Indo-European Phase A, and gathers

in his work related evidence on the Proto-Indo-European verbal system". On that question, see also G.E. Dunkel *Early, Middle, Late Indo-European: Doing it My Way* (1997), Adrados, Bernabé, Mendoza, *Manual de Lingüística Indoeuropea* (1998); etc.

1.3.2. Today, a widespread Three-Stage Theory divide PIE (internal) language evolution into three main historic layers or stages, including a description of branches and languages as clean breaks from a common source (e.g. PAn from IE II) or from intermediate *dialect continua* (e.g. Germanic and Balto-Slavic from EIE), or classifying similarities into continued linguistic contact (e.g. Italic and Celtic, or Balto-Slavic and Indo-Iranian):

1) Early PIE (also IE I, for some *Pre-Proto-Indo-European*) is the hypothetical ancestor of Middle PIE, and probably the oldest stage of the language that comparative grammar could help reconstruct using internal reconstruction. There is, however, no common position as to how it was like or when and where it was spoken.

2) The second stage corresponds to a time before the separation of Proto-Anatolian from the common linguistic community where it should have coexisted (as a Pre-Anatolian dialect) with Pre-IE III. That stage of the language is called Middle PIE or IE II, also *Indo-Hittite*, identified with early Kurgan cultures in the Kurgan Hypothesis.

NOTE. On the place of Anatolian among IE languages, the question is whether it separated first as a language branch from PIE, and to what extent was it thus spared developments common to the remaining (Late) Proto-Indo-European language group. There is growing consensus in favor of its early split from Indo-European (Heading, among others, 'Indo-Hittite'); see N. Oettinger '*Indo-Hittite' – Hypothesen und Wortbildung* (1986), A. Lehrman *Indo-Hittite Revisited* (1996), H. Craig Melchert *The Dialectal Position of Anatolian within IE in IE Subgrouping* (1998), etc.

For Kortlandt (1989): "Since the beginnings of the Yamnaya, Globular Amphora, Corded Ware, and Afanasievo cultures can all be dated between 3600 and 3000 BC, I am inclined to date Proto-Indo-European to the middle of the fourth millennium, and to recognize Proto-Indo-Hittite as a language which may have been spoken a millennium earlier".

3) The common immediate ancestor of most of the earliest known IE proto-languages – more or less the same *static PIE* searched for since the start of Indo-European studies, before Hittite was deciphered – is usually called Late PIE, also *IE III*, often simply *Proto-Indo-European*, often dated some time ca. 3500-3000 BC using linguistic or archaeological models, or both.

NOTE. According to Mallory & Adams (2006), "Generally, we find some form of triangulation based on the earliest attested Indo-European languages, i.e. Hittite, Mycenaean Greek, and Indo-Aryan, each of these positioned somewhere between c. 2000 and 1500 BC. Given the kind of changes linguists know to have occurred in the attested histories of Greek or Indo-Aryan, etc., the linguist compares the difference wrought by such changes with the degree of difference between the earliest attested Hittite, Mycenaean Greek,
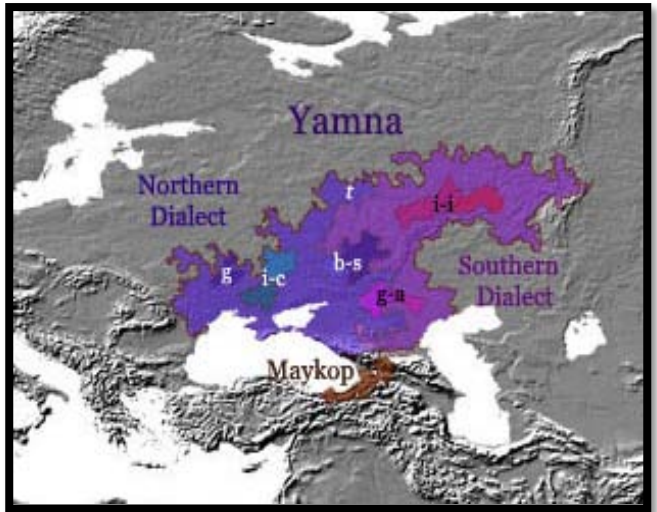
Yamna culture ca. 3000 BC, roughly the time when Late PIE and Proto-Anatolian were spoken, according to the Kurgan framework.

and Sanskrit and reconstructed Proto-Indo-European. The order of magnitude for these estimates (or guesstimates) tends to be something on the order of 1,500-2,000 years. In other words, employing some form of gut intuition (based on experience which is often grounded on the known separation of the Romance or Germanic languages), linguists tend to put Proto-Indo-European sometime around 3000 BC plus or minus a millennium (...) the earliest we are going to be able to set Proto-Indo-European is about the fifth millennium BC if we want it to reflect the archaeological reality of Eurasia. We have already seen that individual Indo-European groups are attested by c. 2000 BC. One might then place a notional date of c. 4500-2500 BC on Proto-Indo-European. The linguist will note that the presumed dates for the existence of Proto-Indo-European arrived at by this method are congruent with those established by linguists' 'informed estimation'. The two dating techniques, linguistic and archeological, are at least independent and congruent with one another".

Likewise, Meier-Brügger (2003) thinks that "No precise statement concerning the exact time period of the Proto-Indo-European linguistic community is possible. One may only state that the ancient Indo-European languages that we know, which date from the 2nd millennium BC, already exhibit characteristics of their respective linguistic groups in their earliest occurrences, thus allowing one to presume the existence of a separate and long pre-history (...) The period of 5000-3000 BC is suggested as a possible timeframe of a [common] Proto-Indo-European language".

However, on the early historic and prehistoric finds, B. Hänsel (1994) states that "Linguistic development may be described in steps that, although logically comprehensible, are not precisely

analyzable without a timescale. The archaeologist pursues certain areas of cultural development, the logic of which (if one exists) remains a mystery to him, or is only accessible in a few aspects of its complex causality". On the other hand, he is provided with concrete ideas with regard to time, as vague as these may be, and works with a concept of culture that the Indo-European linguist cannot attain. For the archaeologist, culture is understood in the sense of a sociological definition (...) The archaeological concept of culture is composed of so many components, that by its very nature its contours must remain blurred. But languages are quite different. Of course there are connections; no one can imagine cultural connections without any possibility of verbal communication. But it is too much to ask that archaeologists equate their concept of culture, which is open and incorporates references on various levels, to the single dimension of linguistic community. Archaeology and linguistics are so fundamentally different that, while points of agreement may be expected, parallels and congruency may not. The advantage of linguistic research is its ability to precisely distinguish between individual languages and the regularity of developments. The strength of archaeology is its precision in developing timelines. What one can do, the other cannot. They could complement each other beautifully, if only there were enough commonality.

1.3.3. Another division has to be made, so that the dialectal evolution and this revival project is properly understood. Late PIE had at least two main *inner* dialectal branches, the *Southern* (or IE IIIa) and the *Northern* (or IE IIIb) ones. As far as we know, while speakers of *Southern* or *Graeco-Aryan* dialects (like Pre-Greek and Pre-Indo-Iranian Indo-European) spread in different directions with the first Late PIE migrations ca. 3000-2500 BC, speakers of Northern dialects remained still in close contact for centuries in Europe, but for those (like Pre-Tocharian IE speakers) who migrated to Asia.

NOTE. "Within the group of IE languages, some individual languages are more closely associated with one another owing to morphological or lexical similarities. The cause for this, as a rule, is a prehistoric geographic proximity (perhaps even constituting single linguistic community) or a common preliminary linguistic phase, a middle mother-language phase, which would however then be posterior to the period of the [common PIE] mother language" (Meier-Brügger, 2003).

About Tocharian, Adrados (1998) says "even if archaic in some respects (its *Centum* character, subjunctive, etc.) it has shares common features with Balto-Slavic, among other languages: they must be old isoglosses, shared before it separated and migrated to the East. It is, therefore, an IE III B [=Northern] language. It shows great innovations, too, something normal in a language that evolved isolated".

On the Southern (*Graeco-Aryan* or *Indo-Greek*) PIE dialect, see Gamkrelidze, Ivanov (1995) *Indo-European and the Indo-Europeans*, and Clackson (1994) *The Linguistic Relationship Between Armenian and Greek,* Adrados (1998), etc. Even in Mallory & Adams (2006): "Many have argued that Greek, Armenian, and Indo-Iranian share a number of innovations that suggest that there should have been some form of linguistic *continuum* between their predecessors"; "An Irish-Indic cognate looks a damn sight stronger than a Greek-Iranian and linguists have long noted that there are a whole series of words that seem to be confined largely to Greek and Indo-Iranian".
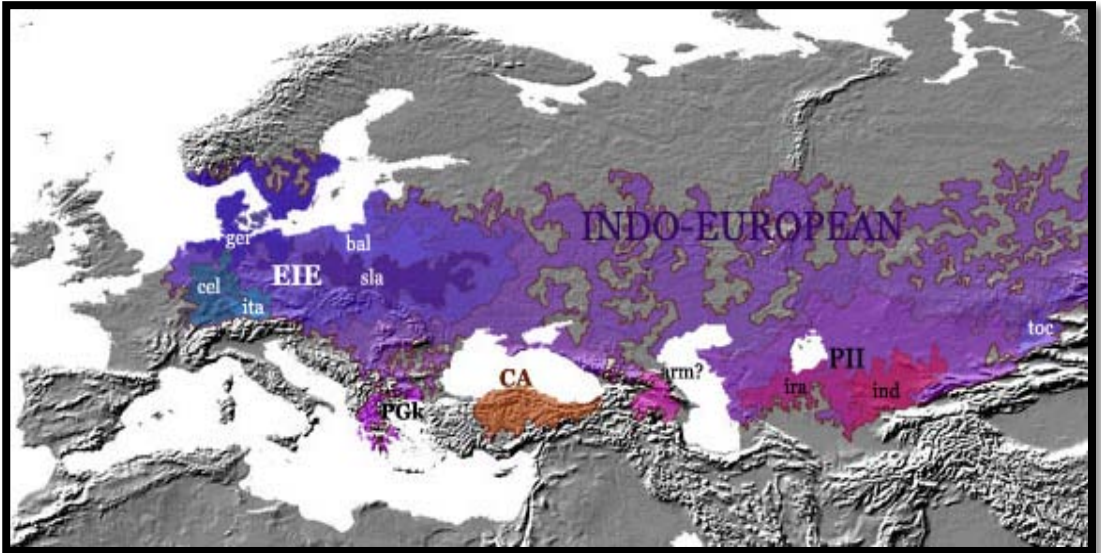
1.3.4. The so-called *North-West*, *Residual*, or *Europe's* Indo-European language, also *Old European* or *Proto-European*, linguistic ancestor of Celtic, Italic, Germanic, Balto-Slavic and other (fragmentary or unattested) European dialects, is believed to have formed the last common IE dialect *continuum* in Northern Europe during the centuries on either side of 2500 BC. It was therefore a language coeval with Common Anatolian, Proto-Indo-Iranian and Proto-Greek, able to fill in the time gap between the dates estimated for North-West IE proto-languages and the Central and Eastern ones.

NOTE 1. On the so-called "*Nort-West Indo-European*" language or *dialect continuum*, see N. Oettinger *Grundsätzliche überlegungen zum Nordwest-Indogermanischen* (1997), and *Zum nordwestindogermanischen Lexikon* (1999); M. E. Huld *Indo-Europeanization of Northern Europe* (1996); Adrados, Bernabé, Mendoza, *Manual de Lingüística Indoeuropea* (1998); etc.

Furthermore, regarding modern IE studies, even without accepting a common, independent North-West IE language, "There are so many of these words that are confined within these five language groups (Celtic, Italic, Germanic, Baltic, and Slavic) that most linguists would regard

cognates found exclusively between any two or among all of these groups as specifically North-West Indo-European and not demonstrably Proto-Indo-European. To accept a series of cognates as reflections of a PIE word requires that the evidence come from further afield than a series of contiguous language groups in Europe", Mallory & Adams (2006).

   NOTE 2. About the 'Dark Ages', the time gap between the reconstructed North-West Indo-European proto-languages (dated ca. 1500 BC) and the other IE proto-languages reconstructed (dated ca. 2500 BC), Mallory & Adams (2006): "If one reviews discussion of the dates by which the various Indo-European groups first emerged, we find an interesting and somewhat disturbing phenomenon. By c. 2000 BC we have traces of Anatolian, and hence linguists are willing to place the emergence of Proto-Anatolian to c. 2500 BC or considerably earlier. We have already differentiated Indo-Aryan in the Mitanni treaty by c.1500 BC so undifferentiated Proto-Indo-Iranian must be earlier, and dates on the order of 2500-2000 BC are often suggested. Mycenaean Greek, the language of the Linear B tablets, is known by c.1300 BC if not somewhat earlier and is different enough from its Bronze Age contemporaries (Indo-Iranian or Anatolian) and from reconstructed PIE to predispose a linguist to place a date of c. 2000 BC or earlier for Proto-Greek itself. So where we have written documentation from the Bronze Age, we tend to assign the proto-languages to an earlier period of the Bronze Age, i.e. earlier than at least 2000 BC if not 2500 BC. When we turn to western and northern Europe, however, both our attestation of the different groups and the estimates of their proto-languages tend to be shallower. The Germanic languages, for example, are all derived from Proto-Germanic. Now the earliest runic inscriptions are so close to reconstructed Proto-Germanic that there is a tendency to date the Germanic proto-language to about 500 BC. Similarly, if we examine the earliest Celtic inscriptional evidence, be it Continental or even the much more recent Irish ogam stones, these inscriptions are not that far removed from the reconstructed Proto-Celtic and again we tend to have dates suggested on the order of 1000 BC. The Slavic languages only began differentiating from one another during the historical period, and Proto-Slavic is generally set to about the beginning of the Christian era while Proto-Baltic and Proto-Balto-Slavic (assuming its existence) are probably envisaged as a second millennium BC phenomenon. In short, where the Indo-European groups are more recently attested, we tend to find that they are also regarded as having differentiated at a more recent time, i.e. between c.1500 and 500 BC. One explanation for the relatively short time depths of the attested northern and western Indo-European groups is that these groups are the only survivors of a long process of linguistic assimilation that has occurred as small demographic and linguistic groups moved, interacted, and merged. We can see precisely such a process in action in the historic period as Latin assimilated and replaced all the other Italic languages, Umbrian, Oscan, etc., and then went on to assimilate and replace much of the Celtic languages. Also within the historic period Slavic

assimilated and replaced such other Indo-European languages as Thracian, and Koine Greek replaced nearly all other varieties of Greek. If we had only contemporary data to work with, we would have to conclude that both Proto-Italic (now equivalent to Proto-Romance) and Proto-Greek flourished around the beginning of the Christian era. These 'extinction events' in the history of Italic and Greek had the effect of 'resetting' the time depth of the proto-language. This process must have been repeated time and again in the prehistoric period".

Similarly, Kortlandt (1989) detected what "seems to be a general tendency to date proto-languages farther back in time than is warranted by the linguistic evidence. When we reconstruct Proto-Romance, we arrive at a linguistic stage which is approximately two centuries later than the language of Caesar and Cicero (cf. Agard 1984: 47-60 for the phonological differences). When we start from the extralinguistic evidence and identify the origins of Romance with the beginnings of Rome, we arrive at the eighth century BC, which is almost a millennium too early. The point is that we must identify the formation of Romance with the imperfect learning of Latin by a large number of people during the expansion of the Roman empire".

IE dialects ca. 500 BC.
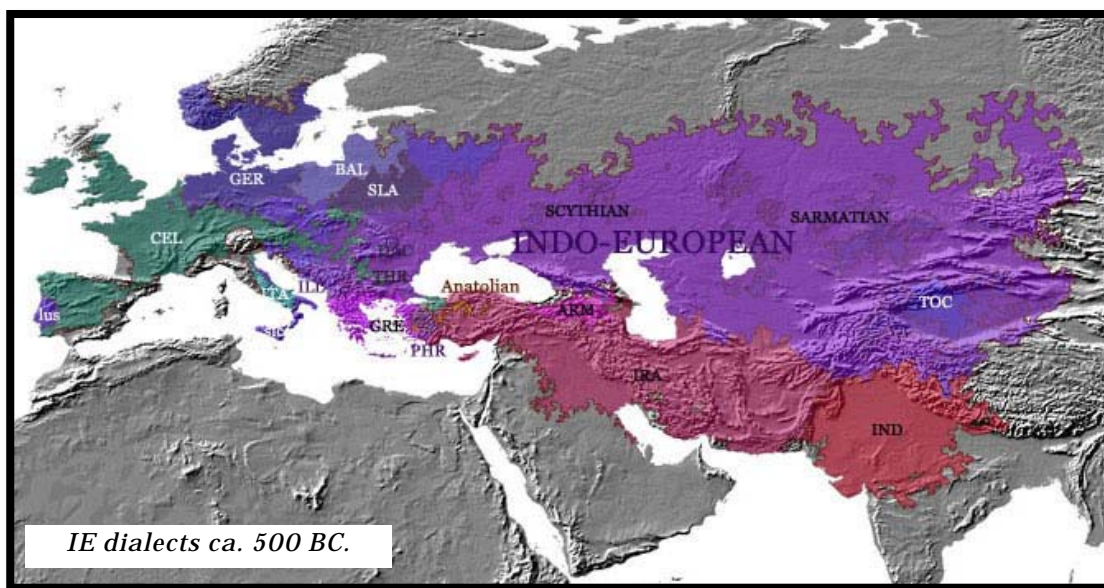
1.3.4. Apart from the <u>shared phonology and vocabulary</u>, Europe's Indo-European shows other common features, as a trend to reduce the noun inflection system, shared innovations in the verbal system, the -***r*** endings of the middle or middle-passive voice, etc. The southern dialects, which spread in different directions and evolved without forming a *continuum*, show therefore a differentiated phonology and even vocabulary, but common developments like the augment in ***é***-, etc.
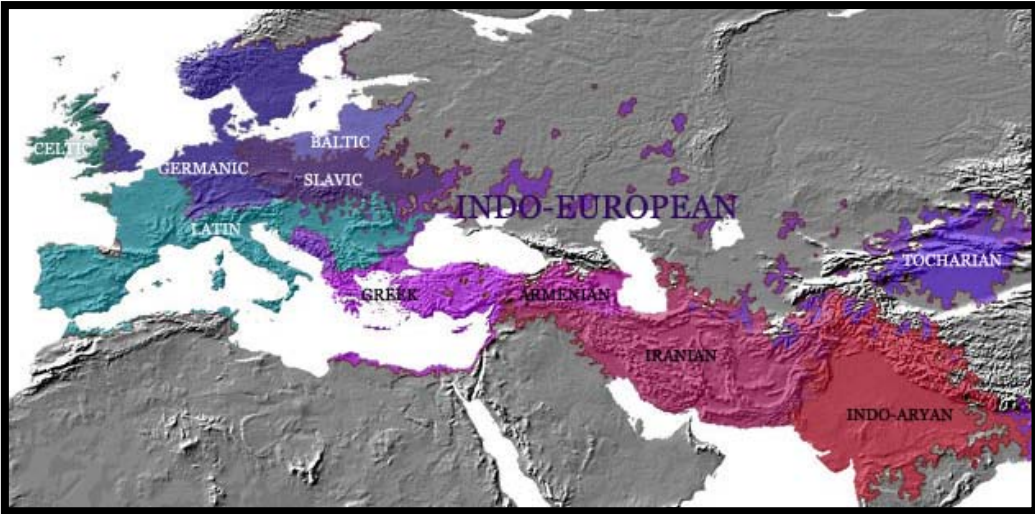
NOTE. Some European dialects like Balto-Slavic and, to some extent, Italic (or Latin, if differentiated from Osco-Umbrian), either because of general PIE innovative or archaic trends that only they maintained, or because of their original situation within the prehistoric dialectal territories in relation with the origin of innovations – or just because they remained in contact with Southern Indo-European dialects after the first PIE split (e.g. through the Scythian or general Iranian expansions) – show features usually identified with Indo-Iranian, as an 8-case noun declension and phonetic satemization, while having morphological features clearly common to Germanic and Celtic dialects, like their verbal system. *Images show IE languages ca. 500 AD (top) and 1500 AD (under these lines).*

Indo-European Language Association <http://dnghu.org/>
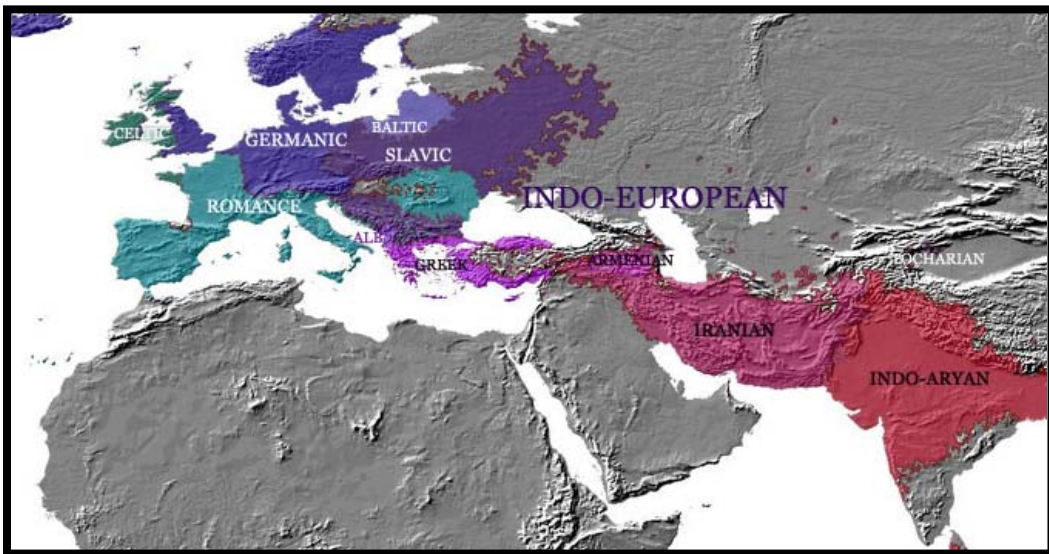
1.4. THE PROTO-INDO-EUROPEAN *URHEIMAT* OR 'HOMELAND'

1.4.1. The search for the *Urheimat* or 'Homeland' of the prehistoric Proto-Indo-Europeans has developed as an archaeological quest along with the linguistic research looking for the reconstruction of the proto-language.

NOTE. According to A. Scherer's *Die Urheimat der Indogermanen* (1968), summing up the views of various authors from the years 1892-1963, still followed by mainstream Indo-European studies today, "Based upon the localization of later languages such as Greek, Anatolian, and Indo-Iranian, a swathe of land in southern Russia north of the Black Sea is often proposed as the native area of the speakers of Proto-Indo-European" (Meier-Brügger, 2003).

1.4.2. The *Kurgan hypothesis* was introduced by Marija Gimbutas in 1956 in order to combine archaeology with linguistics in locating the origins of the Proto-Indo-Europeans. She named the set of cultures in question "Kurgan" after their distinctive burial mounds and traced their diffusion into Eastern and Northern Europe.

Photo of a Kurgan (Archaeology Magazine).

1.4.3. According to her hypothesis, PIE speakers were probably a nomadic tribe of the Pontic-Caspian steppe that expanded in successive stages of the Kurgan culture and three successive "waves" of expansion during the 3rd millennium BC:

- Kurgan I, Dnieper/Volga region, earlier half of the 4th millennium BC. Apparently evolving from cultures of the Volga basin, subgroups include the Samara and Seroglazovo cultures.

- Kurgan II–III, latter half of the 4th millennium BC. Includes the Sredny Stog culture and the Maykop culture of the northern Caucasus. Stone circles, early two-wheeled chariots, anthropomorphic stone *stelae* of deities.

- Kurgan IV or Pit Grave culture, first half of the 3rd millennium BC, encompassing the entire steppe region from the Ural to Romania.

o Wave 1, predating Kurgan I, expansion from the lower Volga to the Dnieper, leading to coexistence of Kurgan I and the Cucuteni culture. Repercussions of the migrations extend as far as the Balkans and along the Danube to the Vinča and Lengyel cultures in Hungary.

o Wave 2, mid 4th millennium BC, originating in the Maykop culture and resulting in advances of "*kurganized*" hybrid cultures into northern Europe around 3000 BC – Globular Amphora culture, Baden culture, and ultimately Corded Ware culture.

o Wave 3, 3000-2800 BC, expansion of the Pit Grave culture beyond the steppes; appearance of characteristic pit graves as far as the areas of modern Romania, Bulgaria and eastern Hungary.
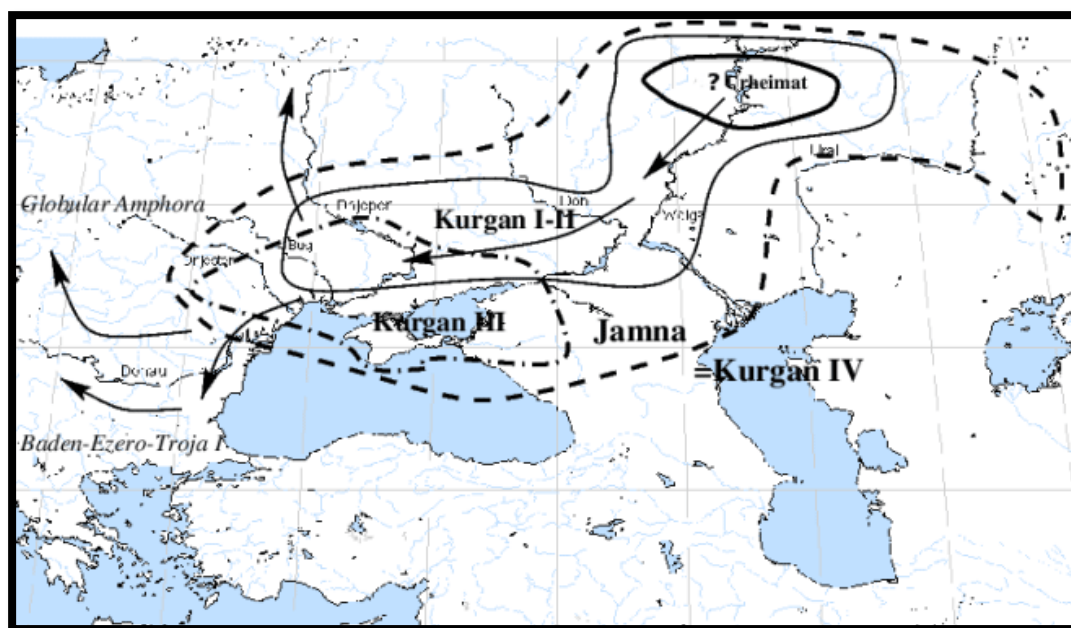
Hypothetical Homeland or Urheimat of the first PIE speakers, from 4500 BC onwards. The Yamna (Pit Grave) culture lasted from ca. 3600 till 2200 BC. In this time the first wagons appeared. People were buried with their legs flexed, a position which remained typical for the Indo-Europeans for a long time. The burials were covered with a mound, a kurgan. During this period, from 3600 till 3000 IE II split up into Pre-IE III and Pre-Proto-Anatolian. From ca.3000 B.C on, Late PIE dialects began to differentiate and spread by 2500 westward (Europe's Indo-European), southward (Proto-Greek) and eastward (Proto-Aryan, Pre-Proto-Tocharian).

1. Introduction

NOTE. On the Kurgan hypothesis, Mallory & Adams (2006) say that "[t]he opposite method to a retrospective approach is a prospective approach where one starts with a given archaeological phenomenon and tracks its expansion. This approach is largely driven by a theory connected with the mechanism by which the Indo-European languages must have expanded. Here the trajectory need not be the type of family tree that an archaeologist might draw up but rather some other major social phenomenon that can move between cultures. For example, in both the nineteenth century and then again in the later twentieth century, it was proposed that Indo-European expansions were associated with the spread of agriculture. The underlying assumption here is that only the expansion of a new more productive economy and attendant population expansion can explain the widespread expansion of a language family the size of the Indo-European. This theory is most closely associated with a model that derives the Indo-Europeans from Anatolia about the seventh millennium BC from whence they spread into south-eastern Europe and then across Europe in a Neolithic 'wave of advance'.

A later alternative mechanism is the spread of more pastoral societies who exploited the horse (and later the chariot) and carried a new language across Europe and Asia from the fourth millennium bc onwards. The underlying assumption here is that the vector of Indo-European language spread depended on a new, more aggressive social organization coupled with a more mobile economy and superior transportation technology. As this theory sets the homeland in the steppelands north of the Black and Caspian seas among different cultures that employed barrows for their burials (Russian *kurgan*), it is generally termed the Kurgan theory.

Although the difference between the Wave of Advance and Kurgan theories is quite marked, they both share the same explanation for the expansion of the Indo-Iranians in Asia (and there are no fundamental differences in either of their difficulties in explaining the Tocharians), i.e. the expansion of mobile pastoralists eastwards and then southwards into Iran and India. Moreover, there is recognition by supporters of the Neolithic theory that the 'wave of advance' did not reach the peripheries of Europe (central and western Mediterranean, Atlantic and northern Europe) but that these regions adopted agriculture from their neighbours rather than being replaced by them.
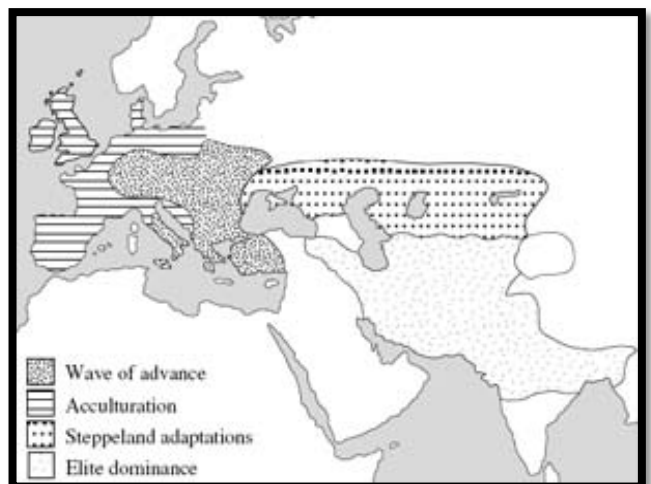
I. LINGUISTIC RECONSTRUCTION AND ARCHAEOLOGICAL TIMELINE

The *Manual de Lingüística Indoeuropea* (Adrados, Bernabé, Mendoza, 1998) makes a summary of main linguistic facts, supported by archaeological finds: "Remember the recent date of the 'cristalization' of European languages. 'Old European' [=North-West Indo-European], from which they derive, is an already evolved language, with opposition masculine/feminine, and must be located in time ca. 2000 BC or before. Also, one must take into account the following data: the existence of Tocharian, related to IE IIIb [=Northern PIE], but far away to the East, in the Chinese Turkestan; the presence of IE IIIb [=Graeco-Aryan] languages to the South of the Carpathian Mountains, no doubt already in the 3rd millennium (the ancestors of Thracian, Iranian, Greek speakers); differentiation of Hittite and Luwian, within the Anatolian group, already ca. 2000 BC, in the documents of Kültepe, what means that Common Anatolian must be much older.

NOTE. Without taking on account archaeological theories, linguistic data reveals that:

a) IE IIIb, located in Europe and in the Chinese Turkestan, must come from an intermediate zone, with expansion into both directions.

b) IE IIIa, which occupied the space between Greece and the Northwest of India, communicating both Paeninsulas through the languages of the Balkans, Ukraine and Northern Caucasus, the Turkestan and Iran, must also come from some intermediate location. Being a different linguistic group, it cannot come from Europe or the Russian Steppe, where Ural-Altaic languages existed.

c) Both groups have been in contact secondarily, taking on account the different 'recent' isoglosses in the contact zone.

d) The more archaic Anatolian must have been isolated from the more evolved IE; and that in some region with easy communication with Anatolia.

(...) Only the Steppe North of the Caucasus, the Volga river and beyond can combine all possibilities mentioned: there are pathways that go down into Anatolia and Iran through the Caucasus, through the East of the Caspian Sea, the Gorgan plains, and they can migrate from there to the Chinese Turkestan, or to Europe, where two ways exist: to the North and to the South of the Carpathian mountains.

NOTE. For Kortlandt (1989), too, "Starting from the linguistic evidence (...) The best candidate for the original IE homeland is the territory of the Sredny Stog culture in the eastern Ukraine".

1. Introduction
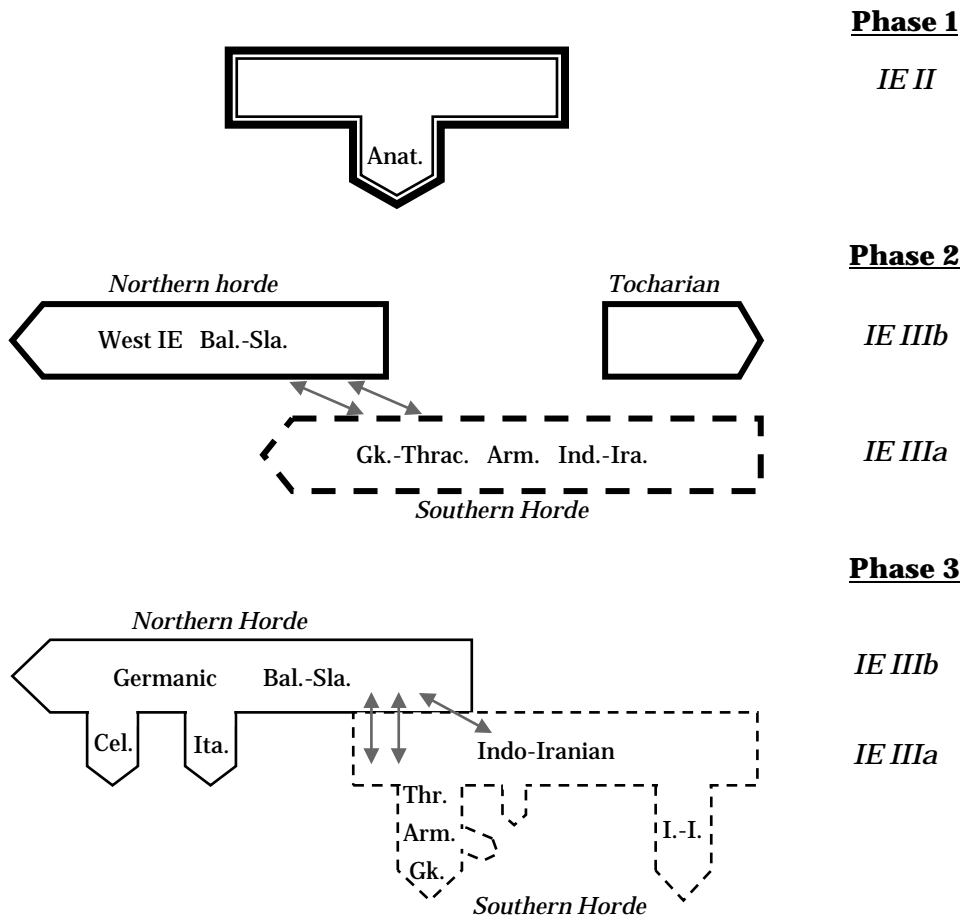
Diagram of the expansion and relationships of IE languages, Adrados (1979).

These linguistic data, presented in a diagram, are supported by strong archaeological arguments: they have been defended by Gimbutas 1985 against Gamkrelidze-Ivanov (...) This diagram proposes three phases. In the first one, IE II [=Middle PIE] became isolated, and from it Anatolian emerged, being first relegated to the North of the Caucasus, and then crossing into the South: there must Common Anatolian be located. Note that there is no significant temporal difference with the other groups; it happens also that the first IE wave into Europe was older. It is somewhere to the North of the people that later went to Anatolia that happened the great revolution that developed IE III [=Late PIE], the 'common language'.

The following phases refer to that common language. The first is that one that saw both IE III B [=Northern IE] (to the North) and IE III A [=Southern IE] (to the South), the former being fragmented in two groups, one that headed West and one that migrated to the East. That is a proof that somewhere in the European Russia a common language III B emerged; to the South, in Ukraine or in the Turkestan, IE III A.

The second phase continues the movements of both branches, that launched waves to the South, but that were in contact in some moments, arising isoglosses that unite certain languages of the IE IIIa group (first Greek, later Iranian, etc.) with those of the rearguard of IE IIIb (especially Baltic and Slavic, also Italic and Germanic)".

II. ARCHEOGENETICS AND INDO-EUROPEAN MIGRATIONS

Cavalli-Sforza and Alberto Piazza argue that Renfrew (*v.i.*) and Gimbutas reinforce rather than contradict each other, stating that "genetically speaking, peoples of the Kurgan steppe descended at least in part from people of the Middle Eastern

Distribution of haplotypes R1b (light color) for Eurasiatic Paleolithic and R1a (dark color) for Yamna expansion; black represents other haplogroups.

Neolithic who immigrated there from Turkey".

NOTE. The genetic record cannot yield any direct information as to the language spoken by these groups. The current interpretation of genetic data suggests a strong genetic continuity in Europe; specifically, studies of mtDNA by Bryan Sykes show that about 80% of the genetic stock of Europeans originated in the Paleolithic.

Spencer Wells suggests that the origin, distribution and age of the R1a1 haplotype points to an ancient migration, possibly corresponding to the spread by the Kurgan people in their expansion across the Eurasian steppe around 3000 BC, stating that "there is nothing to contradict this model, although the genetic patterns do not provide clear support either".

NOTE. R1a1 is most prevalent in Poland, Russia, and Ukraine, and is also observed in Pakistan, India and central Asia. R1a1 is largely confined east of the Vistula gene barrier and drops considerably to the west. The spread of Y-chromosome DNA haplogroup R1a1 has been associated with the spread of the Indo-European languages too. The mutations that characterize haplogroup R1a occurred ~10,000 years bp.

Haplogroup R1a1, whose lineage is thought to have originated in the Eurasian Steppes north of the Black and Caspian Seas, is therefore associated with the Kurgan culture, as well as with the postglacial Ahrensburg culture which has been suggested to have spread the gene originally.
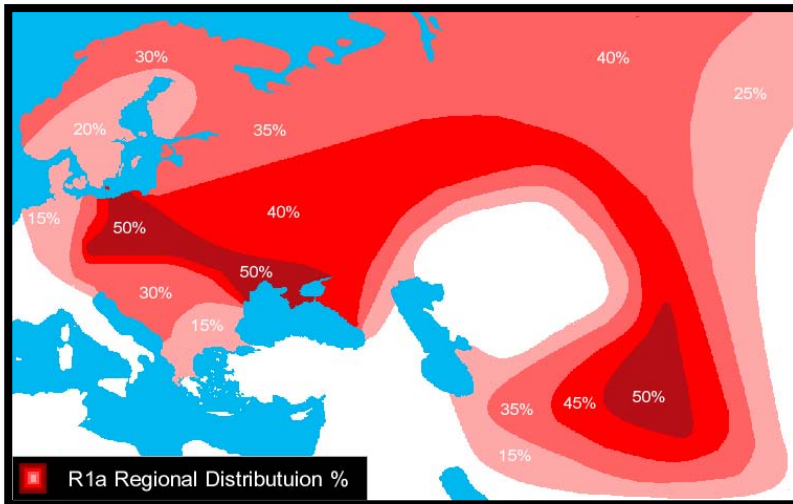
The present-day population of R1b haplotype, with extremely high peaks in Western Europe and measured up to the eastern confines of Central Asia, are believed to be the descendants of a refugium in the Iberian peninsula (Portugal and Spain) at the Last Glacial Maximum, where the haplogroup may have achieved genetic homogeneity. As conditions eased with the Allerød Oscillation in about 12000 BC, descendants of this group migrated and eventually recolonised all of Western Europe, leading to the dominant position of R1b in variant degrees from Iberia to Scandinavia, so evident in haplogroup maps.

NOTE. High concentrations of Mesolithic or late Paleolithic YDNA haplogroups of types R1b (typically well above 35%) and I (up to 25%), are thought to derive ultimately of the robust Eurasiatic Cro Magnoid homo sapiens of the Aurignacian culture, and the subsequent gracile leptodolichomorphous people of the Gravettian culture that entered Europe from the Middle East 20,000 to 25,000 years ago, respectively.

43
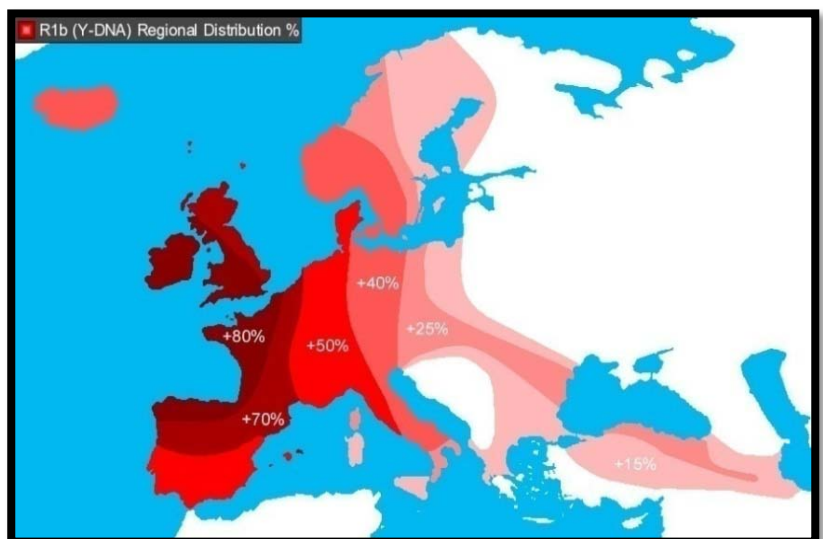
III. THE KURGAN HYPOTHESIS AND THE THREE-STAGE THEORY

ARCHAEOLOGY (Kurgan Hypothesis)	LINGUISTICS (Three-Stage Theory)
ca. 4500-4000 BC. Sredny Stog, Dnieper-Donets and Sarama cultures, domestication of the horse.	*Early PIE spoken, probably somewhere in the Pontic-Caspian Steppe.*
ca. 4000-3500 BC. The Yamna culture, the kurgan builders, emerges in the steppe, and the Maykop culture in northern Caucasus.	*Middle PIE. Pre-IE III and Pre-Proto-Anatolian dialects evolve in different communities but presumably still in contact*
ca. 3500-3000 BC. Yamna culture at its peak: stone idols, two-wheeled proto-chariots, animal husbandry, permanent settlements and hillforts, subsisting on agriculture and fishing, along rivers. Contact of the Yamna culture with late Neolithic Europe cultures results in kurganized Globular Amphora and Baden cultures. Maykop culture shows earliest evidence of the beginning Bronze Age; bronze weapons and artifacts introduced.	*Proto-Anatolian becomes isolated south of the Caucasus, and has no more contacts with the linguistic innovations of the common Late PIE language.*
ca. 3000-2500 BC. The Yamna culture extends over the entire Pontic steppe. The Corded Ware culture extends from the Rhine to the Volga, corresponding to the latest phase of IE unity. Different cultures disintegrate, still in loose contact, enabling the spread of technology.	*Late PIE evolves into dialects, at least a Southern and a Northern one. Dialectal communities remain still in contact, enabling the spread of phonetic and morphological innovations, and loan words. PAn, spoken in Asia Minor, evolves into Common Anatolian.*
ca. 2500-2000 BC. The Bronze Age reaches Central Europe with the Beaker culture of Northern Indo-Europeans. Indo-Iranians settle north of the Caspian in the Sintashta-Petrovka and later the Andronovo culture.	*The breakup of the southern IE dialects is complete. Proto-Greek spoken in the Balkans; Proto-Indo-Iranian in Central Asia; North-West Indo-European in Northern Europe; Common Anatolian dialects in Anatolia.*
ca. 2000-1500 BC. The chariot is invented, leading to the split and rapid spread of Iranians and other peoples from the Andronovo culture and the Bactria-Margiana Complex over much of Central Asia, Northern India, Iran and Eastern Anatolia. Greek Darg Ages and flourishing of the Hittite Empire. Pre-Celtic Unetice culture.	*Indo-Iranian splits up in two main dialects, Indo-Aryan and Iranian. European proto-dialects like Germanic, Celtic, Italic, and Balto-Slavic differentiate from each other. Anatolian languages like Hittite and Luwian are written down; Indo-Iranian attested through Mitanni; a Greek dialect, Mycenaean, is already spoken.*
ca. 1500-1000 BC. The Nordic Bronze Age sees the rise of the Germanic Urnfield and the Celtic Hallstatt cultures in Central Europe, introducing the Iron Age. Italic peoples move to the Italian Peninsula. Rigveda is composed. The Hittite Kingdoms and the Mycenaean civilization decline.	*Germanic, Celtic, Italic, Baltic and Slavic are already different proto-languages, developing in turn different dialects. Iranian and other related southern dialects expand through military conquest, and Indo-Aryan spreads in the form of its sacred language, Sanskrit.*
ca. 1000-500 BC. Northern Europe enters the Pre-Roman Iron Age. Early Indo-European Kingdoms and Empires in Eurasia. In Europe, Classical Antiquity begins with the flourishing of the Greek peoples. Foundation of Rome.	*Celtic dialects spread over Western Europe, German dialects to the south of Jutland. Italic languages in the Italian Peninsula. Greek and Old Italic alphabets appear. Late Anatolian dialects. Cimmerian, Scythian and Sarmatian in Asia, Palaeo-Balkan languages in the Balkans.*

1.5. OTHER LINGUISTIC AND ARCHAEOLOGICAL THEORIES

1.5.1. A common development of new hypotheses has been to revise the Three-Stage assumption. It is actually not something new, but the come back to more traditional views, reinterpreting the new findings of the Hittite scripts, trying to insert Anatolian into the old, static PIE concept.

1.5.2. The most known new alternative theory concerning PIE is the Glottalic theory. It assumes that Proto-Indo-European was pronounced more or less like Armenian, i.e. instead of PIE *p, *b, *b^h, the pronunciation would have been *p', *p, *b, and the same with the other two voiceless-voiced-voiced aspirated series of consonants. The IE *Urheimat* would have been then located in the surroundings of Anatolia, especially near Lake Urmia, in northern Iran, hence the archaism of Anatolian dialects and the glottalics found in Armenian.

NOTE. Those linguistic and archaeological findings are supported by Th. Gamkredlize-V. Ivanov (1990: "*The early history of Indo-European languages*", *Scientific American*, where early Indo-European vocabulary deemed "of southern regions" is examined, and similarities with Semitic and Kartvelian languages are also brought to light. This theory has been criticized by Meid (1989)

1.5.3. Alternative theories include:

I. The *European Homeland thesis* maintains that the common origin of the IE languages lies in Europe. These theses are more or less driven by Archeological. A. Häusler (1981, 1986, 1992) continues to defend the hypothesis that places Indo-European origins in Europe, stating that all the known differentiation emerged in the *continuum* from the Rhin to the Urals.

NOTE. It has been traditionally located in 1) *Lithuania* and the surrounding areas, by R.G. Latham (1851) and Th. Poesche (1878: *Die Arier. Ein Beitrag zur historischen Anthropologie*, Jena); 2) *Scandinavia*, by K.Penka (1883: *Origines ariacae*, Viena); 3) *Central Europe*, by G. Kossinna (1902: "*Die Indogermanische Frage archäologisch beantwortet*", *Zeitschrift für Ethnologie*, 34, pp. 161-222), P.Giles (1922: *The Aryans*, New York), and by linguist/archaeologist G. Childe (1926: *The Aryans. A Study of Indo-European Origins*, London).

a. The *Old European* or *Alteuropäisch Theory* compares some old European vocabulary (especially river names), which would be older than the spread of Late PIE dialects through Northern Europe. It points out the possibility of an older, pre-IE III

spread of IE, either of IE II or I or maybe some other Pre-IE dialect. It is usually related to the PCT and Renfrew's NDT.

b. The *Paleolithic Continuity Theory* posits that the advent of IE languages should be linked to the arrival of Homo sapiens in Europe and Asia from Africa in the Upper Paleolithic. The PCT proposes a continued presence of Pre-IE and non-IE peoples and languages in Europe from Paleolithic times and allowing for minor invasions and infiltrations of local scope, mainly during the last three millennia.

NOTE. There are some research papers concerning the PCT available at <http://www.continuitas.com/>. Also, the PCT could in turn be connected with Frederik Kortlandt's Indo-Uralic and Altaic studies <http://kortlandt.nl/publications/> – although they could also be inserted in Gimbutas' early framework.

On the temporal relationship question, Mallory & Adams (2006): "How early a solution is admitted depends on individual decisions regarding the temporally most diagnostic vocabulary. That the vocabulary is clearly one reflecting at least a Neolithic economy and technology, i.e. domesticated plants and animals, ceramics, means that it cannot be set anywhere on this planet prior to c. 8000 BC. Although there are still those who propose solutions dating back to the Palaeolithic, these cannot be reconciled with the cultural vocabulary of the Indo-European languages. The later vocabulary of Proto-Indo- European hinges on such items as wheeled vehicles, the plough, wool, which are attested in Proto-Indo-European, including Anatolian. It is unlikely then that words for these items entered the Proto-Indo-European lexicon prior to about 4000 BC. This is not necessarily a date for the expansion of Indo-European since the area of Proto-Indo-European speech could have already been in motion by then and new items with their words might still have passed through the continuum undetected, i.e. treated as inheritances rather than borrowings. All that can be concluded is that if one wishes to propose a homeland earlier than about 4000 bc, the harder it is to explain these items of vocabulary".

c. The PCT is, in turn, related to the theories of a *Neolithic revolution* causing the peacefully spreading of an older pre-IE language into Europe from Asia Minor from around 7000 BC, with the advance of farming. It proposes that the dispersal (discontinuity) of Proto-Indo-Europeans originated in Neolithic Anatolia.

NOTE. Reacting to criticism, Renfrew by 1999 revised his proposal to the effect of taking a pronounced Indo-Hittite position. Renfrew's revised views place only Pre-Proto-Indo-European in 7th millennium BC Anatolia, proposing as the homeland of Proto-Indo-European proper the Balkans around 5000 BC, explicitly identified as the "Old European culture" proposed by Gimbutas.

1. Introduction

As of 2005, Colin Renfrew seems to support the PCT designs and the usefulness of the Paleolithic assumptions. He co-authored a paper concluding: *Our finding lends weight to a proposed Paleolithic ancestry for modern Europeans* The above quotation coming as results of archaeogenetic research on mtDNA where 150 x greater N1a frequency was found. The first European farmers are descended from a European population who were present in Europe since the Paleolithic and not coming as a wave of Neolithic migration as proposed in Renfrew's NDT.

Talking about these new (old) theories, Adrados (1998) makes an interesting remark about the relevance that is – wrongly – given to each new personal archaeological 'revolutionary' theory: "[The hypothesis of Colin Renfrew (1987)] is based on ideas about the diffusion of agriculture from Asia to Europe in [the 5th millennium Neolithic Asia Minor], diffusion that would be united to that of Indo-Europeans; it doesn't pay attention at all to linguistic data. The [hypothesis of Gamkrelidze-Ivanov (1980, etc.)], which places the Homeland in the contact zone between Caucasian and Semitic peoples, south of the Caucasus, is based on real or supposed lexical loans; it disregards morphological data altogether, too. Criticism of these ideas – to which people have paid too much attention – are found, among others, in Meid (1989), Villar (1991), etc.".

II. Another hypothesis, contrary to the European ones, also mainly driven today by a nationalistic view, traces back the origin of PIE to Vedic Sanskrit, postulating that it is very *pure*, and that the origin of common Proto-Indo-European can thus be traced back to the Indus Valley Civilization of ca. 3000 BC.

NOTE. Such Pan-Sanskritism was common among early Indo-Europeanists, as Schlegel, Young, A. Pictet (1877: *Les origines indoeuropéens*, Paris) or Schmidt (who preferred Babylonia), but are now mainly supported by those who consider Sanskrit almost equal to Late Proto-Indo-European. For more on this, see S. Misra (1992: *The Aryan Problem: A Linguistic Approach*, Delhi), Elst's *Update on the Aryan Invasion Debate* (1999), followed up by S.G. Talageri's *The Rigveda: A Historical Analysis* (2000), both part of "*Indigenous Indo-Aryan*" viewpoint by N. Kazanas, the "*Out of India*" theory, with a framework dating back to the times of the Indus Valley Civilization.

III. The Black Sea deluge theory dates the origin of the expansion of IE dialects in the genesis of the Sea of Azov, ca. 5600 BC, which would in turn be related to the deluge myth, which would have remained in oral tails until its description in the biblical story of Noah's Ark, the Hindu Puranic story of Manu, through Deucalion in Greek mythology or Utnapishtim in the Epic of Gilgamesh. This date is generally considered as rather early for the PIE spread under frameworks which include the *Urheimat* near the Black Sea.

NOTE. W.Ryan and W.Pitman published evidence that a massive flood through the Bosporus occurred about 5600 BC, when the rising Mediterranean spilled over a rocky sill at the Bosporus. The event flooded 155,000 km² of land and significantly expanded the Black Sea shoreline to the north and west. This has been connected with the fact that some Early Modern scholars based on Genesis 10:5 had assumed that the 'Japhetite' languages (instead of the 'Semitic' ones) are rather the direct descendants of the Adamic language, having separated before the confusion of tongues, by which also Hebrew was affected. That was claimed by *Blessed* Anne Catherine Emmerich (18th c.), who stated in her private revelations that the purest descendants of the Adamic language were the main Proto-Indo-European dialects, *v.i.*

1.6. RELATIONSHIP TO OTHER LANGUAGES

1.6.1. Many higher-level relationships between PIE and other language families have been proposed. But these speculative connections are highly controversial. Perhaps the most widely accepted proposal is of an Indo-Uralic family, encompassing PIE and Proto-Uralic, a language from which Hunarian, Finnish, Estonian, Saami and a number of other languages belong. The evidence usually cited in favor of this is the proximity of the proposed *Urheimaten* for both of them, the typological similarity between the two languages, and a number of apparent shared morphemes.

NOTE. Other proposals, further back in time (and correspondingly less accepted), model PIE as a branch of Indo-Uralic with a Caucasian substratum; link PIE and Uralic with Altaic and certain other families in Asia, such as Korean, Japanese, Chukotko-Kamchatkan and Eskimo-Aleut (representative proposals are Nostratic and Joseph Greenberg's Eurasiatic); etc.

1.6.2. Indo-Uralic or Uralo-Indo-European is therefore a hypothetical language family consisting of Indo-European and Uralic (i.e. Finno-Ugric and Samoyedic). Most linguists still consider this theory speculative and its evidence insufficient to conclusively prove genetic affiliation.

NOTE. Dutch linguist Frederik Kortlandt supports a model of Indo-Uralic in which its speakers lived north of the Caspian Sea, and Proto-Indo-Europeans began as a group that branched off westward from there to come into geographic proximity with the Northwest Caucasian languages, absorbing a Northwest Caucasian lexical blending before moving farther westward to a region north of the Black Sea where their language settled into canonical Proto-Indo-European.

1. Introduction

The problem with lexical evidence is to weed out words due to borrowing, because Uralic languages have been in contact with Indo-European languages for millennia, and consequently borrowed many words from them.

1.6.3. The most common arguments in favour of a relationship between Early PIE and Uralic are based on seemingly common elements of morphology, such as:

Meaning	Early PIE	Proto-Uralic
"I, me"	*me, "me"* (Acc.), *mene, "my"* (Gen.)	*mun, *mina, "I"
"you" (sg)	*tu (Nom.), *twe (Acc.), *tewe "your"* (Gen.)	**tun, *tina*
1st P. singular	*-m	*-m
1st P. plural	*-me	*-me
2nd P. singular	*-s (active), *-tHa (perfect)	*-t
2nd P. plural	*-te	*-te
Demonstrative	*so, "this, he/she"* (animate nom)	*ša (3rd person singular)
Interr. pron. (An.)	*kʷi-, "who?, what?"; *kʷo-, "who?, what?"*	*ken, "who?", *ku-, "who?"
Relative pronoun	*jo-	*-ja (nomen agentis)
Accusative	*-m	*-m
Ablative/partitive	*-od	*-ta
Nom./Acc. plural	*-es (Nom. pl.), *-m̥-s (Acc. pl.)	*-t
Oblique plural	*-i (pronominal pl., cf. we-i- *"we"*, to-i- *"those"*)	*-i
Dual	*-H₁	*-k
Stative	*-s- (aorist); *-es-, *-t (stative substantive)	*-ta
Negative particle	*nei, *ne	*ei- [negative verb] , *ne
"to give"	*deh₃-	*toHe-
"to wet", "water"	*wed-, "to wet", *wodr̥-, "water"*	*weti, "water"
"water"	*mesg-, "dip under water, dive"*	*muśke-, "wash"
"to assign", "name"	*nem-, "to assign, to allot", *h₁nomn̥-, "name"*	*nimi, "name"
"metal"	*h₂weseh₂-, "gold"*	*waśke, "some metal"
"trade"	*mei-, "exchange"*	*miHe-, "give, sell"
"fish"	*(s)kʷalo-, "large fish"*	**kala, "fish"*
"sister-in-law"	*galou-, "husband's sister"*	**kälɜ, "sister-in-law"*
"much"	*polu-, "much"*	**paljɜ, "thick, much"*