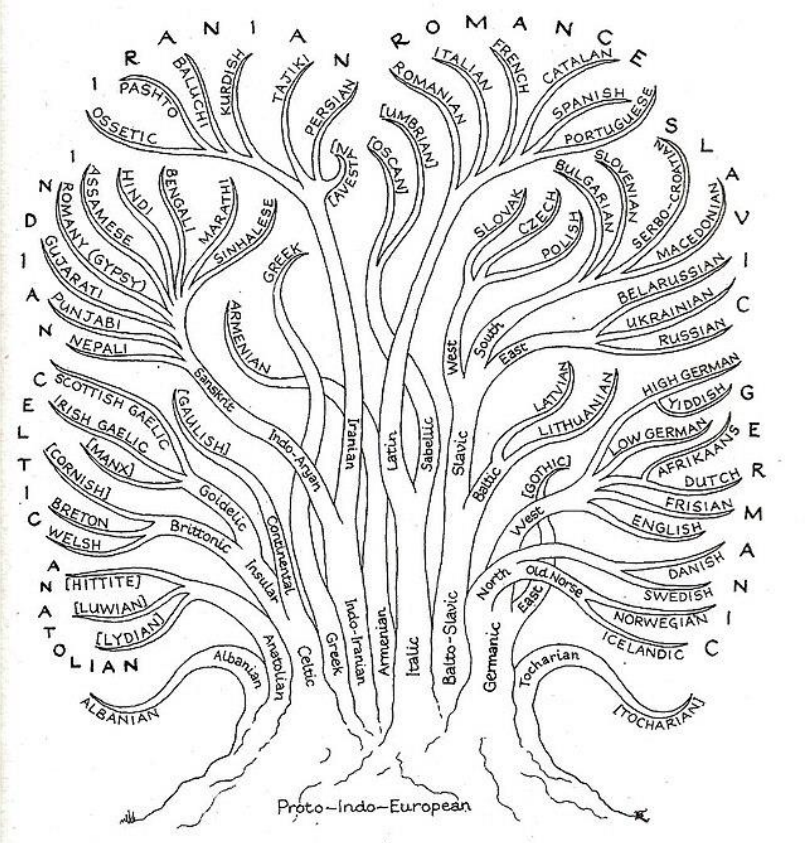
Do we speak dialects of Proto-Indo-European?

The reconstruction of ancient languages through the example of Proto-Indo-European



Family tree of the Indo-European languages

School year: 2023-2024

Students: Noémie Turquin and Viljami Viinamäki

Directrice de mémoire: Eloïse Hartheiser

Table of contents

1. Introduction

For this school year of 2023-2024, we, Noémie Turquin & Viljami Viinamäki, chose the topic of linguistic reconstruction as the main focus of our *mémoire collectif* in 3ème, assisted & guided by our *directrice de mémoire*, Mrs. Eloïse Hartheiser. More specifically, as evidenced by the title of our *mémoire*, *Reconstruction of ancient languages as explained through Proto-Indo-European*, we will be examining linguistic reconstruction & related concepts using examples of its application to the Indo-European language family.

We will thus first explain general linguistic theory necessary to understand reconstruction in particular, before explaining the general principles, tools & precise methodology that historical linguists make use of in order to reconstruct unattested languages.

Furthermore, the exact question we decided on is: "Are we all still just speaking different dialects of Proto-Indo-European?", with *we* referring to speakers of Indo-European languages, such as English. It is thus aimed at questioning the distinction between separate languages & highlighting the evolution & differences between Proto-Indo-European (often abbreviated to PIE) & its daughter languages, an answer being given in the conclusion of our mémoire (Chapter 9).

We chose this *mémoire* as we both share an interest in literature & languages generally & thought it might be interesting to take a deep dive into such an enigmatic discipline that is frequently overlooked & misunderstood.

2. Linguistic baseline

In this first chapter, we will introduce the reader to some elementary concepts in linguistics, specifically those we deem irreplaceable to understanding the content of this *mémoire*.

Firstly, let us start off with main focus of this very work: linguistic reconstruction. Briefly explained, it consists in the efforts of linguists to reconstruct (i.e. recreate) an unattested language, recognizable by the prefix *proto-* for which we have neither oral (living speakers) nor written (manuscripts, written records) evidence, by utilizing evidence preserved in a "corrupted" form in related languages. The following disciplines & concepts are linked to this process:

Phonology, the study of systems of sounds, at the core of which lies the so-called phoneme, defined as a sound whose distinction from other sounds is crucial in differentiating words. For such phonemes, we can usually make out a minimal pair, i.e. 2 words which differ only in that one phoneme, take for example the words *sink* & *think*, which may be rendered as /sɪŋk/ & /θɪŋk/ in phonetic transcription. They differ only in their first components, /s/ & /θ/, which must thus be maintained as separate sounds in order to uphold the distinction between these 2

words. It is important though to consider that as simple phonemes, they function more as symbols to differentiate from one another & thus don't yet have precise pronunciations. Their actual pronunciations manifest as so-called allophones, variations of phonemes depending on the environment, which do not differentiate between words on their own: In Spanish, for instance, the phoneme /b/ has 2 main allophones, [b] (a bilabial stop) & [β] (a bilabial fricative). The former occurs word-initially, after a nasal consonant & after a pause, while the latter occurs mostly intervocalically. They never overlap in terms of position & can thus be said to only occur in complementary distribution, i.e. one occurs when the other cannot. The domain of linguistics that deals with allophones & the specific of pronunciation is called phonetics.

An equally important field of study in linguistics is that of morphology, where the main focus lies on morphemes, the smallest units of meaning in a language. These do not necessarily always correspond to independent words, see the word *unbreakable*, which contains a total of 3 morphemes when looked at through surface analysis: *un-* (expressing inability), *break* (to fracture, crack, separate), & *-able* (denoting the ability to do something). These cannot further be broken down into smaller meaningful bits, thus constituting morphemes as per our previous definitions.

Syntax, the configuration & coordination of terms in a sentence, as well as grammar, the general internal functioning of a language to accurately communicate information, are also often discussed in historical linguists.

3. Language changes

3.1 Introduction

This chapter "Language changes" is about changes inside a language, about how a language evolves trough time. Those changes occur often, and it is possible to notice them in our daily language use. You surely have notice that you speak an even slightly different language than your grandparents or even parents, though you still understand each other. You are using new words that they do not use, or they use some old-fashioned vocabulary no one uses in your generation. Or you pronounce differently some words than them. Those are language changes and are part of the evolution of languages. Languages are not fixed, they evolve over time and space and, yes, changes occur. Linguist have discovered that it is possible to categorize them and so understand better how they occur and why. This is what this chapter is about: what kind of language change exists? How do they happen and why?

3.2 Sound changes

We shall start with the sound changes. We will first see the different types of sound changes with examples to understand them better. Linguistic where able to categorize them into those different types of changes:

Change A > B: a phoneme[1] (A) transforms into a new phoneme (B) which did not exist previously in a language. For example, in PIE[1] **k'* became in Sanskrit *ś*, a sound that did not exist previously in that language. So, in PIE **k'mtóm* "hundred" became in Sanskrit *śatám*.

Merger A, B > B: a phoneme (A) merges with another phoneme (B), who already exists. This means that a sound will melt with another sound and disappear, like in Proto Germanic where the phoneme *o* completely merges with *a*. In PIE we see **g^{h}ostis* "stranger, guest" which became *gasts* in Gothic.

Loss A>Ø: A phoneme (A) is lost (Ø means "zero" or "nothing). A *p* disappeared in Celtic for example as in the word father where the PIE was **ph$_2$tēr* and became in old Irish *athir.*

Splits A>A, B or B, C and so forth: a phoneme (A) diverges into two different phonemes. It can be the same phoneme(A) and another (B), or two different ones (B, C).

Emergence Ø > A: A phoneme (A) appears where there was none before. As in Greek, where the genitive of "man" *anēr* is *andrós*. There the sound *d* came into being between the *n* and the *r*.

What is interesting in the topic of sound change is that sound changes occur in a ***regular way.*** If a certain change happens in a language, it will not only occur to a word but to every word where the affected phoneme is found.

In Greek, a *o* always corresponds to an *à* in Sanskrit as in the word "cloud." Greek *néphos*, Sanskrit *nàbhas.*

They are so regular that linguist speak of "sound law" and even "Ausnahmlosigkeit der Lautgesetzte," which means that the sound law does not know any exception. This is of course a hyperbole as there are exception, but even does are government by rules and therefore are not really exception. We shall now explain those different exceptions:

As we have seen a *o* in Greek always corresponds to an *à* in Sanskrit, but in the word for "knee" we see in Greek *gónu* and Sanskrit *jānu.* An *o* in Greek suddenly corresponds to a long *ā* in Sanskrit. This can be easily explained as it

[1] PIE stands for Proto-Indo-European

has been remarked that an *o* in Greek at the end of a syllable corresponds to a long *ā* in Sanskrit.

Another explanation for exceptions in sound changes is the borrowing of words from another language or dialect. Indeed, as they are borrowed, they do not get affected the same way.

The Latin word *bōs* "cattle" was borrowed from a dialect of Latin. If it came from Latin, it would have been **gōs*, but since it is borrowed it was not affected by this sound change.

Then there is the situation of sporadic sound changes, sound changes that happen irregularly and are exceptional. There are several types of sporadic sound changes:

Dissimilation: similar consonants or vowels in a word become less similar, as for example /r/ and /l/. Those two sounds can be similar and therefore change as in the Latin *peregrīnus* (foreigner) who evolved into *pelegrīnus.*

Metathesis: is the changing of position. This means that a sound will change of place in a word. For example, in Dutch *wesp* was *weps* "wasp." As we can see both are similar, only the *s* changed of position.

Haplology: simplified it is saying a syllable only once instead of twice. This means that if similar syllables are next to each other, one might disappear. For example, in German "female wizard" used to be *Zauberin* but now is *Zauberin*. There we see that *re* disappeared making the pronunciation easier.

Onomatopoeias: Onomatopoeias are word who comes from sound. They are the most common exception. For example, the sound that goat makes is in Greek *bē*. Theoretically it should have become **vi* according to sound law but of course it stayed *bē* as the goat does "beeeee" and not "viiii."

Another important notion to understand sound changes is the fact that those ***changes are limited in space and time.*** Those boundaries are called isogloss. This means a specific sound change happens at one place at a time, which can explain exceptions.

In Greek, an *s* was lost between two vowels, as we can see in *méné-os* "the will of power" who comes from PIE **ménés-os*. Though there is the word *tísis* "revenge" where there is a *s* between two vowels. How is this possible? Simple explication, the *s* appeared **after** the sound law had effect. So chronologically

seen this happened: at first the sound changes which makes *s* between vowels disappear took place **and afterwards** another sound law took place and *t* transformed into *s* before *i*. So, the word *tísis* was before **títis* and then became *tísis.* This can explain the exception of *tísis* having an *s* between two vowels.

A crucial point is that every change is particular to a language. To also explain the exceptions, we need to know that sound changes depend on the sound context. This means that specific changes only happen after a specific sound. We speak then of conditioned change. Here are some examples:

As we have seen in the example above, in Greek *t* transformed into *s* before an *i* but never after an *s*. So, we find *phé-si "*he says," but *és-ti* "he is." In *és-ti* the *t* stayed, even if it is after an *s* because of the *i*.

Some changes can also only happen at the end of a word ("auslaut") as in Sanskrit where the last syllable does not have more than one consonant. So **á-bhār-s-t* "he carried" became *ábhār.*

Short and long vowels can also be a condition for the syllable and the position of the sound elements. In West Germanic the *i* and the *u* in the last syllable of two syllables-long word disappeared after a long vowel but not after a short vowel: In Old high Germanic *gastiz* "guest, stranger" became *gast*, but **uiniz* "friend" became *wini*, as *a* is a long vowel, the *i* disappeared, but *i* is a short vowel and therefore the *i* stayed. In the third syllable *i* and *u* always disappeared.

If there are conditioned changes, there are also unconditioned changes who does not depend on the sound context. As in Proto Germanic where each PIE *o becomes an *a*: PIE **g^{h}ostis* "stranger, guest" became in Gothic *gasts*.

So now we may ask why does change happens. It is difficult to give precise explanation or reason for a sound change. An inherent cause (causes inside a language) could be to make the pronunciation easier; another cause could be the phonological stress[2]. Some linguists believe that children never learn a language to 100%. It has been noted that sound changes spread quicker among children. There are also external factors (outside of a language), the influence of external dialects or language can bring new vocabulary or new phonemes. For example, the sound *g* does not exist in Dutch except in the words coming from other language like *garcon, garage,* who added the phoneme *g* to the Dutch language. Another plausible reason for sound change is when a group of people must learn a new language, but they keep their own articulation habits, in situation of migration for example. In conclusion it is difficult to find a reason for sound changes.

3.3 Analogical changes

Analogical changes are changes modeled on the example of other words or forms. They happen based on a perceived similarity. Those changes are completely different from sound changes. Indeed, sometimes analogy undoes sound changes as we shall see later. What is important to understand with analogy is that it does not affect the root of a word, but analogy affects the suffix and/or the ending. Therefore, for analogy to be able to work it needs to be a model. Though it is important to remember that models are theoretical constructs and not simple matter of fact.

A nominative in -us is constructed on base of an accusative in -um, trough the u-stem (it has a u-suffix) who have nominative in -us and an accusative in -um.

This will become clearer with the next example. Two distinct types of analogy exist: proportional analogy and non-proportional analogy. We shall begin with the proportional analogy. Proportional analogy often works on the model of **a: b = c: x,** where x is the new form of a word. Let us take the example of the word for "foot" in Gothic.

accusative	nominative		accusative	nominative
sūnu-m :	*sūnu-s*	=	*fotu-m* :	x

X = *fotus*

There we clearly see that *fotus* was interpreted to have a *-us* ending (even if it originally maybe did not), because of *fotum* who was thought to be based on a u-stem like *sūnum*. In Sanskrit something similar happened:

The word "night" got the nominative *náktam* based on the accusative *náktam*, based on the model of the neuter who had *-am* in both nominative and accusative. The word became so neutral though it was originally feminine.

This can be expressed with *a:b =c:x* :

accusative	nominative		accusative	nominative
neuter-am :	*neuter-am*	=	*nákt-am* :	x

X = *náktam*

We shall now explore non-proportional analogy. In non-proportional analogy changes are based on other forms whereby non-proportions are seen.

In Greek, the verb "to follow" is conjugated following *hépomai* "I follow/obey," *hépēi* "you follow" *hépetai* "he/she/it follows" and so forth. According to regular sound law it should be **hetēi, *hetetai,* but the *p* sound spread over the paradigm due to *hépomai*, this is called leveling. This change occurred based on *hépomai* "I follow."

With the example above one notices that sound law and analogy seem to work against each other. This is not exact. It is better to say that analogy has the power to restore where a sound law had effect. For example, in Greek the aorist *é-lū-s-a* has a restored -s-, though the -s- disappeared between two vowels in Greek. The explanation for why there is an -s is that analogy restored the -s. Let us go back to non-proportional analogy. Non-proportional analogy can change the usual ending of cases, also influences suffixes, by elongating them.

In Sanskrit the ending of the instrumental plural is *-bhis*, in Greek it is *-phi,* as in *náu-bhis* and *naũ-phi* "with the ships." It is very probable that the *-s* was added in Sanskrit. This would have occurred based on the model of the nominative plural and accusative plural where the endings are -as, similarly as in the dative plural where it is *-bhyas*. Therefore, *-s* can be seen as a characteristic of the plural.

In Latin, the suffix *-nus* is used for the formation of adjectives as in *domi-nus* "master (of the house)." The same suffix also follows roots ending in *-ā* like *silvānus* "of the woods." From such words, derived the suffix *-ānus* who is used in *urb-ānus* "of the city" and *hum-ānus* "human."

Now that we have seen proportional and non-proportional analogy, we will now look at the causes. Indeed, analogy needs a reason to occur. The most common one is the simplification of the system because less different endings leads to a reduction of types which make a system simpler. Also, words that occur regularly are less likely to change through analogy. In Greek, we see *oĩda* "I know" and *ídmen* "we know." There is the old interchange *oi/i* is preserved, while elsewhere this ablaut disappeared, as the verb "to know" is a common one.

3.4 Morphological changes

Morphological changes are the changes in how words are identified. Morphological changes are due to unknown causes, one of them could be sound changes, when a syllable disappears at the end of a word for example. This can lead to syntactic changes as we will see in the next session. In those parts we will mostly see how it was in PIE and how it changed into different languages. We will now see different examples of morphological changes.

Disappearance of morphological categories: PIE had eight cases: nominative, accusative, genitive, dative, ablative (to express the origin), locative (for the place), instrumental case (to express that something happens) and the vocative (respect to the addressee). In Greek, the instrumental and locative have merged with the dative as we can see in the sentence *oíkad'iōn sùn nēusí te s~eis kaì soĩs hetároisi Murmidónessin ánasse*. "Go home with (*sùn*) your ships and your men and be king of the Myrmidons." There we see that the dative has taken over the function of the instrumental case and that a preposition is used. The dative (*Murmidónessin*) is also used in the sense of "at, in the midst of" where PIE would have used the locative.

Change in the case system: In PIE, each case had his own ending who was completely different in singular and in plural as in the dative here the singular ending was *-ei* and the plural was *-mus*. In Tocharian we see a reorganization of the case system. Tocharian has an obliquus case, and the ending are the same in plural and in singular. The endings were added based on the obliquus singular and plural as in the declination of kässi "teacher," in a so called "case plus ending" system.

---

	singular	plural
Nominative	*kässi*	*kässiñ*
Genitive	*kässiyāp*	*kässiśśi*

Obliquus	*kässim*	*kässis*
Instrumental *-yo*	*kässin-yo*	*kässis-yo*
Perlative *-ā*	*kässin-ā*	*kässis-ā*
Comitative *-assäl*	*kässin-aśśäl*	*kässis-aśśal*
Allative *-ac*	*kässin-ac*	*kässis-ac*
Ablative *-äs*	*kässin-äs*	*kässis-äs*
Locative *-am*	*kässin-am*	*kässis-am*

Changes in verbal tenses: PIE had imperfect, perfect, and aorist[3]. In Greek, the three still exist while in Latin the aorist has merged with the perfect.

The passive: Greek has a passive aorist with -thē-, though PIE had no passive. This form must come from forms with -ē which where intransitives but also used with a passive meaning for example e-míg-ē, who later became the passive e-míkh-thē, meant "it was in a mixed state" and later became to mean "it was mixed."

Change in gender: PIE had three genders: masculine, feminine and neuter. In French, the neuter disappeared, in German the three where preserved, in Hittite the masculine and feminine have merged (*genus commune*) and in English there are no longer any gender.

3.5 Syntactic changes

Morphological changes (see section above) can lead to syntactical change due to the disappearance of morphological categories. In fact, syntactical changes are changes in how the sentences are constructed, therefore how relationships between words are expressed. We shall see the types of changes there are.

Articles: As surprising as it can sound PIE did not have any articles, as Latin and Russian do. The articles developed later as in Greek for example.

> Classical Greek did not have articles, but modern Greek does now have *énas* that developed from *hén-* "one."

Prepositions: A lot of language have developed prepositions when morphological categories disappeared (see section above). They developed where a single form was not sufficient anymore.

> This can be seen in Greek. In *tóks' ōmoisin ékhōn* "with the bow over the shoulder" the dative is sufficient, but in *amphì d'ár ōmoisin báleto ksíphos arguróēlon* "around his shoulder he threw (the carrying strap of) his sword with

silver buttons" the preposition *amphì* "around" has been added, since it was needed.

Construction: In PIE when the subject of the verb was a plural neuter, the verb was conjugated as if it was singular. This changed and when the subject of the verb is a plural neuter, the verb is conjugated in plural.

> In Homer (Old Greek), one can see the old and new construction side by side in the sentence *kaì dē doũra sésēpe neõn kaì spárta léluntai* "and see the beams (the woodwork) of the ships (is) rotted and the ropes are hanging slack." There *doũra* and *spárta* are neutral plural, *sésēpe* is perfect singular, but *léluntai* is plural.

Word order: The word order in IE languages also changed. The oldest IE languages were OV-languages, Object-Verb languages. This means that the object was placed before the verb, as in Hittite. (PCL= particle) (LOC.PL. =Locative plural)

LÚ^{MEŠ} ^{GIŠ}BANŠUR-kan 2^{NINDA}mitgaimiš^{4} appanzi n-as-kan appa suppaias ^{GIŠ}BANŠUR$^{HI.A}$-as tyanzi. Man^{pl} ^{wood}table-PCL 2^{bread}mitgaimi take and-them-PCL back pure ^{wood}table^{s}-LOC.PL. put. "The table people take the 2 mitgaimi-loaves and put them again upon the pure table."

There we see that the object 2^{NINDA}*mitgaimiš* is placed before the verb *appanzi*. We can also see that the adjective *suppaias* comes before the noun ^{GIŠ}BANŠUR$^{HI.A}$*-as*. Most of the IE language became later verb-object language, with the reverse order and where the adjectives follow the noun. For example, in Old Irish:

> *Im-díched in cú Laigiu huili*
> Defended the dog Leinster all
> "The dog defended all of Leinster."

There we see that the object *Laigiu* is placed after the verb *Im-díched* and that the adjective *huili* is placed after the object, wich is a clear change compared to the Hittite language.

Subordinate clauses: PIE probably didn't use subordinate clauses, therefore all the syntax developed after the split up of the PIE language.

The causes for syntactic changes are unsure, one of the most probable hypotheses is (again) the influence of foreign languages. Indeed, people who learn a new language tend to keep the syntax of their mother language. Another cause may be borrowing a different language.

This chapter about language change is important to us, then to reconstruct a language one needs to know about how languages evolve and change and in which way they do. As we have seen, one can classify those changes into various categories and sub-categories.

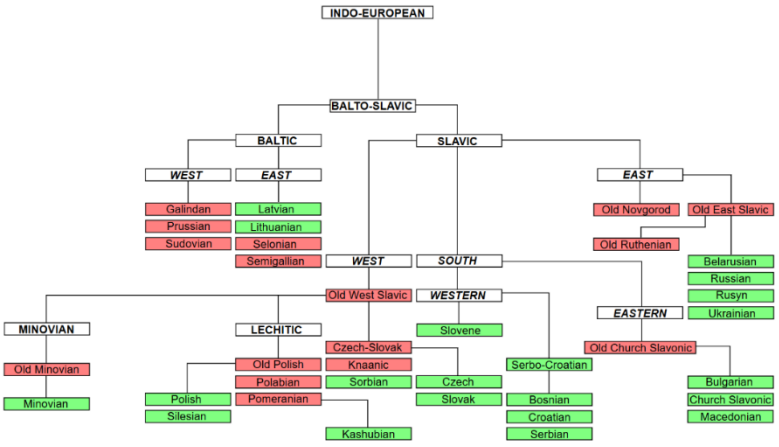
4. Language families / genealogical relationships

Another fundamental concept in reconstruction is that of the relationships between languages, especially the so-called “genetic model of language relations”, where languages & their origin are considered in terms analogous to genetics in a biological context.

According to this model, as language *A* evolves & changes over time (as explained in Chapter 3), eventually, over centuries or even decades depending on the overarching circumstances, it starts to differ from its initial form to such an extent that a speaker of the earlier variety (*A*) could no longer effectively communicated with a speaker of the contemporary form (*B*), at which point they would be considered mutually unintelligible, thus constituting 2 different languages in mainstream linguistics.

But, as language *B* evolved & originates from language *A*, the two are considered related, with the daughter language (*B*) having evolved from the ancestor language (A). The parent language can be either attested or unattested: historical records document the evolution of Latin into the Romance languages, while Proto-Germanic, the ancestor of the Germanic branch of Indo-European, is only sparsely documented in a couple Runic inscriptions.

This model can generally be applied to most languages, safe for a handful of exceptions including constructed languages like Damin in Australia, and therefore most languages can be grouped together into language families, having all descended from one common ancestor language (in our case Indo-European) & diverged into branches (e.g. Germanic), that further split of into subbranches (e.g. Ingvaeonic) & finally individual languages. Thus, similarly to a family tree, language families can be depicted schematically as shown in the following graph showcasing the Balto-Slavic branch of Indo-European (living languages shown in green, dead ones in red):



Now, when it comes to how interlinguistic genealogical relationships are determined there are usually a list of criteria to consider, with the deciding factors being clear similarities (cognates) in core vocabulary, even when accounting for borrowing, onomatopoeia & so forth, as well as shared innovations & features in grammar & phonology especially. In this context, the term *cognate* is used to describe terms in different languages which share a common origin (i.e. descend from the same ethnonym), while reflex refers to the modern rendition of an inherited phoneme, as it appears in daughter languages.

Of course, some family relationships may be excessively obscured due to long histories of interlinguistic contact, as is the case for the Oceanic language family & the neighboring Papuan languages, for example. A too far-reaching time-depth, that is, a family connection that is so ancient that it is practically undetectable, also presents problems in the reconstruction of super families, also called phylums. These connections, exemplified by Nostratic (linking Indo-European, Afroasiatic, Uralic & more), Austric (linking Austronesian, Austroasiatic & more), are regarded as relatively weak proposals, and thus contrast heavily with the well-established families they wish to regroup, remaining fringe theories.

5. Indo-European people

5.1 Introduction

If there is a proto-Indo-European language there must be people who spoke this language. Indeed, those people existed, archaeologist found traces of the Indo-European people (8000-4000 B.C.E.). In this chapter we shall give a summary of the Indo-European society, religion, and homeland. As this work is about linguistic and not history, most of the information given are taken from the reconstructed language. Indeed, vocabulary can reveal a good deal of information about a folk. If one has a word for "seashell" for example, it therefore means that the people who spoke that language knew what seashells are and had seen some. And so, the simple vocabulary can help understand how a folk lived and the organisation of their society.

5.2 The Indo-European Society

Family and habitation

The Indo-Europeans did not have any cities, they lived in villages (**uiḱ-*) where families, tribes, clans lived together, in houses (**dōm*) who were surely made of wood since this root is found in Gothic *timrjan* "to carpenter." As the word door (**dʰuer-*) is mostly found in plural, one can deduce that there were double doors. For the families one can find words for father *(*ph₂tḗr*), mother (**méh₂t(ē)r*), brother (**bʰréh₂t(ē)r*), sister (**suésōr*) and so one. What is noticeable is that one can find words designating the family of the husband, such as **suéḱuros* "the husband's father" for example, but non for the wife's family. Therefore, it was concluded that the wife left her family to live with her husband's family.

Agriculture and animals

Indo-European people were settled (see section above) and knew agriculture. They planted and cultivated different kind of cereals. As they migrated and expanded, they lived in very varied types of climates and used to cultivate many varied species of plants and cereals, therefore the words who were reconstructed are believed to mean "cereals" in a more general term than one specific plant. For example, the word **yewos* who developed in Sanskrit as *yáva* which means "céréales" Other words we can find, like **pūros* whose future meaning depends on the daughter-language ("wheat" in Greek, "spelt" in old Slavic), could have been borrowed to some older languages of tribes that the Indo-European conquered. The proto-Indo-European did not just cultivate plants, animals were more important to show once patrimony, hence we know more words for animals then for the farmed plants. We know that they had cows **g^{w}ous* and taurus **tauros*, as well as sheep **péku,* who were the most common farm animals and used to estimate one's richness. The Indo-European volks also had horses **ékwo*-which is a reason they could expand and migrate so much compared to other volks. *kwōn- (dog) were also known and popular since the word is found in a lot of languages and did not change so much (i.e. Hund in German). We also know that proto-Indo-European knew the wheel and had wheeled vehicles. It is also probable that they had small ships or boats that they would use for fishing.

Hierarchy

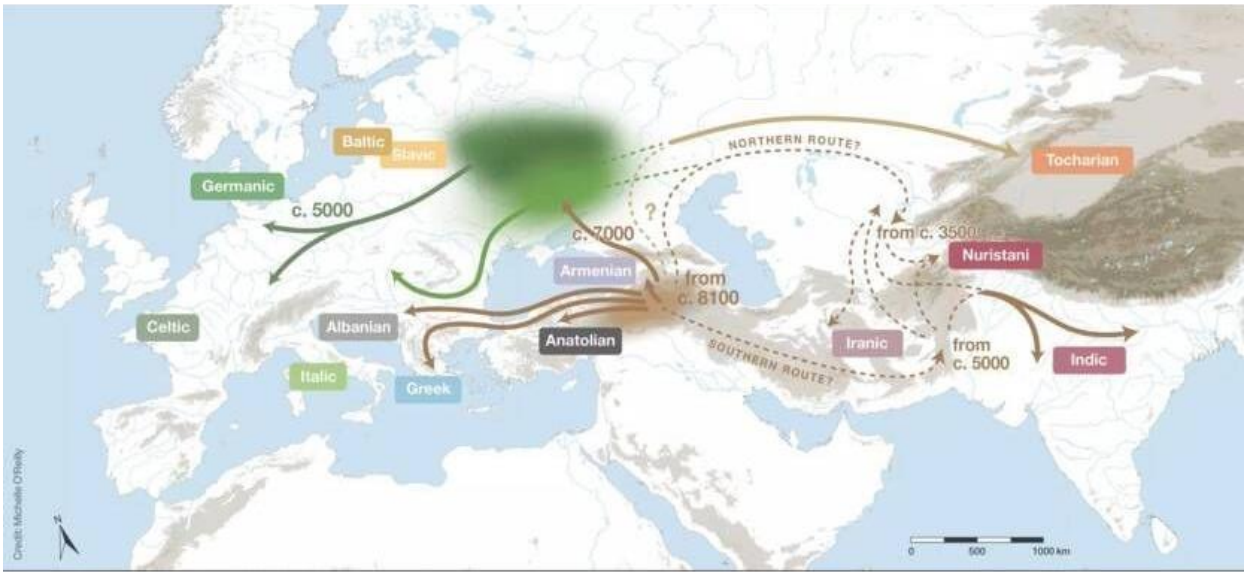
The Indo-European society was a patriarchy and hierarchical society divided in clans or tribes. Their society was based on wealth and poverty with priest, warriors, and commoners. It is possible though that the commoners were also fighters. One can find a word for "king" (**h$_3$rēǵ-s*) but it was more sort of chief of the tribe than a king as we understand today. We also know that Indo-Europeans probably had slaves though no words for it have been found yet. If there were slaves there also were "free mans" (**uiHró*). The theory of a hierarchical society is also supported by graves that were found with some larger than others which indicates a higher status in the society. We can also deduce that it was a patriarchal society since the word for "husband" (**potis*) also means "master (of the house)." What is also interesting is the word for stranger **g^{h}osti-* could have mean "guest" but also "enemy."

Religion

It is known that the Indo-Europeans worshiped a god called **Dyēus* or **Dyēus ph$_2$tḗr* which means so much as *sky-God* or *sky-God father,* this word comes from PIE **d(e)i-*, "to shine, be bright." The ancient Greek word *Ζεύς (Zeus)* descends from this word. It was also reconstructed that his consort was *D^{h}ég^{h}om, the earth mother, his children were *H$_2$wésōs (a dawn-goddess), *Seh$_2$ul (a sun goddess) and *Meh$_2$not (a moon god). Other gods are possible to have been worshipped but only in a few cultures and could therefore not be reconstructed. Some myths, like the creation of the world, have also been reconstructed, but this is not in our focus and should therefore not be further mentioned. The priests were important, and it is possible that the high-priest and the king were one and only person. For the rituals, evidence have been found for animal sacrifice, horses seemed to hold an important place, to gain the favor of the gods or for oaths.

5.3 Origine (Urheimat) and migration

The exact geographical origine of the Indo-European is unsure. The two main hypotheses are the "Steppe hypothesis" and the "Anatolian hypothesis." The steppe hypothesis considers that the origine of the PIE lies in the Pontic-Caspian steppes and emerged around 4500 B.C.E[1]., the Indo-European volks then migrated fast trough the breeding of horses. The Anatolian theory suggest that the PIE speakers are originating from Anatolia and were settled farmer around 8500 B.C.E. and then migrated around 6500 - 5500 B.C.E. Recently scholars of the Max Planck institute have proposed a third hypothesis which combines those two hypotheses, published in July 2023 in the journal "Science"[2.] In the study, they confirm the appearance of the PIE around 8000 B.C.E. which supports the Anatolian hypothesis. Though the scholars also found some evidence for the Steppes theory since they found out that the Indo-European people already migrated around 7000 B.C.E. So according to this study, the language and their people would have emerged around 8000 B.C.E. in Anatolian, and then migrated and it is probable that a part went north in the steppes. On the map below, coming from this study, we can see the result of this study more clearly: In brown one can see the Anatolian homeland and the possible migration movement leaving it. In green we see the Pontic-Caspian Steppe who could have been a second homeland between 7000 and 5000 B.C.E. Hence the two theories may be seen as one.



A map of the study with the hybrid hypothesis who shows the possible migration movement - P.Heggarty / Revue Science [3]

6. Principles & assumptions in reconstructive linguistics

In the following chapter, we will establish the principles & assumptions that linguists adhere to in their attempts to reconstruct any given proto language. First let us discuss basic guiding principles:

6.1 Principles in reconstructive linguistics

Occam's razor, also known as **the *law of parsimony* (*lex parsimoniæ*)**, is a philosophical principle which is widely utilized in scientific research, and historical linguistics in particular, where it serves as one of the core foundations of the methodology of the discipline.

Generally, the principle states that, in order to explain any given phenomena, one should always prefer the option which makes the least amount of assumptions while remaining satisfactory, or, to quote Phillip Baldi in introduction to *Patterns of Change: Change of Patterns – Linguistic change & reconstruction methodology*: "(…) the solution to a problem should be the simplest one which accounts for all the known aspects of what is being solved, or for as many of them as possible."

In the context of the reconstruction of languages, it is most often applied when determining the original proto phoneme from which reflexes in daughter languages are derived (discussed in detail in Chapter 4): take for example the following correspondence set:

Language *A*	Language *B*
h	Ø

For these two phonemes, we may think of two possible versions of the proto phoneme:

1. *f > h > Ø: the proto phoneme is *f, which then, by process of lenition, changes to h, further weakening & disappearing completely in language *B* while language *A* retains the h.
2. *h > Ø: the proto phoneme is *h which weakens to Ø in language *B*, while remaining stable in language *A*.

While in theory both are equally possible, Occam's razor tells us to choose the second option, as it presents fewer hypothesized changes. It also presupposes that if two languages share a common feature, that feature is assumed to have arisen only once & then having been transmitted to both languages, & not two separate times, unless there is evidence to the contrary.

The reasoning behind the law of parsimony is grounded in the fact that by making less & less assumptions & staying as close to verifiable data as possible, while the result may not always be 100% accurate (e.g. the proto phoneme could have also been *f), it allows us to significantly

reduce the margin of error & uncertainty, while still explaining everything to a satisfactory extent, for if new proof surfaces that discredits a previous hypothesis, it can simply be amended to once again fit the data as tightly as possible.

Another important guideline in reconstructing the phonology of a proto language is the **principle of directionality**, or, in other words, the adherence to general trends when it comes to language change. For example, it is universally observed that certain types of sound changes are more common, and therefore maybe more natural, than others: the change s > h, i.e. the weakening a.k.a. lenition of a sound is much more common than the inverse h > s. Patterns of chain reactions, as seen in the following example, are also commonly observed: /t/ > /k/ > /ʔ/ , with the place of articulation shifting further & further back in the throat. If a change opposite to those described general patterns is to be hypothesized (which is surely an option worth considering), take for example h > f (fortition), satisfactory proof is recovered, for example by supplementing the analysis with philological material (manuscripts, orthographical evidence), alongside a valid reason as to why the usual trend doesn't apply.

Next, the majority principle is also employed to hypothesize a proto phoneme: If a majority of cognates share similar reflexes, one can usually assume that the proto phoneme was also similar in articulation, as that would be the simplest yet still comprehensive option according to Occam's razor. Thus, if 3 out of 4 reflexes were listed as s, while only one differed, we would be inclined to also choose *s as the reconstruction, though as a rule of thumb, this should only be applied in cases of >50% agreement.

6.2 Assumptions in reconstructive linguistics

In order to recognize the limits of reconstructive methods, linguists place a great deal of emphasis on acknowledging assumptions made in the pursuit of their goals, as this, once again, helps to put their observations in context & avoid any major oversights.

Now, as discussed previously already in Chapter 3, probably the most important assumption that is made in historical linguistics & on which all others are based in the **Neogrammarian principle**, named after its initial advocates, the Neogrammarians. In summary, this principle states that all sound changes, to some extent, are regular & predictable, i.e. that they have no true exceptions to any sound change. This means that the reflex in a daughter of a given phoneme in the parent language will always be the same, so long as it is subject to the same environment & phonological conditions. If the reflexes differ, according to this **principle of the persistent regularity of sound change**, it is thus due to some change in that conditioning environment.

A good example of this can be found in several sound changes from PIE to Proto Germanic:

1. Grimm's law: This law, first systematically established by Jacob Grimm (one of the brothers Grimm), posits that, firstly, PIE voiceless stops (*p, *t, *k, *k^w) became voiceless fricatives (*f, *þ, *x, x^w), secondly, that voiced stops (*b, *d, *g, *g^w) became unvoiced stops (*p, *t, *k, *k^w) &, thirdly, that voiced, aspirated stops (b^h, d^h, g^h, g^{wh}) became voiced, unaspirated stops (b, d, g, g^w), thus forming the following chain of sound changes:
    - *b^h -> *b -> *p -> *f
    - *d^h -> *d -> *t -> *þ
    - *g^h -> *g -> *k -> *x
    - *g^{wh} -> *g^w -> *k^w -> x^w
2. Verner's law: Discovered by Karl Verner's, this law complements the earlier Grimm's law, explaining discrepancies in edge cases where following the latter doesn't lead to the expected or attested result. Firstly, it suggests that voiceless stops in PIE become voiceless fricatives in Proto Germanic (as according to Grimm's law) **except** in a medial position immediately preceding an accented (i.e. stressed or emphasized) syllable, at which point they became voiced stops. Secondly, in a similar vein, it also says that voiceless fricatives in Proto Germanic changed to voiced fricatives in a medial position also immediately before an accented syllable.

Thus, these 2 laws show that even in cases where one (Grimm's law) doesn't explain recorded data to a complete extent, a subrule (Verner's law) is usually responsible for said divergent behavior. So it follows different rules, but nonetheless remains regular according to those. To quote Phillip Baldi yet again: "[…] sound change has no exceptions that are not governed by some recoverable rule."

One other important thing to keep in mind during reconstruction is the **inherent unreliability of the proposed result**. At best, a reconstruction represents the most likely approximation of the proto language, remaining a hypothetical, merely meant to present the most plausible theoretical model. This principle is the reason that reconstructions are always preceded by an asterisk (*): to mark them as hypothetical & thus uncertain.

Nonetheless, advancements in this field can certainly not be said to rely on complete conjecture, as some hypotheses established early were later confirmed to be near-certain, take for example de Saussure's theory of the laryngeals:

After observing several series of vowel discrepancies which were unexplainable when following the assumed sound changes of the different Indo-European languages, Ferdinand de Saussure eventually hypothesized the existence of the so-called PIE laryngeals in PIE, a set of consonants, all produced roughly around the larynx (hence the name): *h_1, *h_2 & *h_3. According to de

Saussure, these phonemes, though absent in PIE's daughter languages, influenced surrounding vowels *e & *o, & thus causing the aforementioned vowel discrepancies in its descendants, before themselves disappearing Saussure's theory remained obscured, that is, until the decipherment of Hittite Cuneiform in the 1930's and its subsequent addition to the Indo-European language family, after which it quickly rose to prominence: Hittite was found to have a genuine reflex of at least one of the laryngeals, transcribed as *ḫ*, being the only known language to do so, agreeing with de Saussure's hypothesized roots. Thus, the laryngeal theory was confirmed beyond a reasonable doubt & is now considered widely accepted.

Other limitations that our reconstructions may be subject to are our inability to account for dialectal variation in a proto language, and thus also assess its impact on the current linguistic landscape, as our reconstruction, being an approximation, simply represents a rough generalization of the features of such a language.

We may also need to keep in mind that languages do not usually split suddenly into perfectly distinguishable descendants (as seen in our reconstructions) but go through transitionary stages over untold periods of time, as discussed in Chapter 4.

7. Internal reconstruction and the comparative method

7.1 Introduction

As the title of our work says, we are working on the reconstruction of antique languages. This chapter is about the two main reconstruction methods that linguistic scholars use: the internal reconstruction and the comparative method (also called external reconstruction). This chapter will not go deep into the details, since there is no exact systematic way that says how it is supposed to be done. Every language is different from each other and so is all linguistic problem scholars meet, same for how to solve them. Still there are those two main tools that you shall now meet.

7.2 Internal reconstruction

As one can see in "internal reconstruction" we can find "internal." Indeed, this reconstruction method is focused on the reconstruction of an older state (of development) of a chosen language, called pre-language. Therefore, the data used for the reconstruction is coming only from this language and not from a sister-language or dialects, though those can be used to support the outcoming hypothesis or theory. The internal reconstruction helps to find the change that happened and find some sound law. One can say that the goal is to "undo" the previous change. This is done by comparing phonemes within a language (word, root, sound…). With internal reconstruction we can understand how a language evolved (i.e., sound laws).

> Example: in Greek, the genitive of *patḗr* (eng. father) is *patrós* which can be deconstruct in *pat-ēr* and *pat-rós*, similarly we find that the genitive of *anḗr* (eng. man) is *andrós* if we deconstruct the same way we get *an-ḗr* and *an-d-rós*.

Therefore, it can be concluded that *-rós* is typical for genitive. Though this let us with a mysterious *-d-*. Thus, it can be concluded that a *-d-* came phonetically in being between *-nr- (-nr- > -ndr-).*

7.3 Comparative method (external reconstruction)

The goal of the comparative method is to reconstruct a proto language by comparing two or more related or considered as related languages. This is the main difference between the internal reconstruction and the comparative method, one uses data from one language, while the other compares language together to reconstruct their common ancestor. Internal and external reconstruction are both used together in historical linguistics. Internal reconstruction can help by reconstructing a previous state of the language to compare it better, since the more data and knowledge one has the better the comparative methods can be applied. In applying the comparative method, it is important to note that not everything who seems to correspond or to be related is in fact related or corresponding. The comparative method is done by selecting a set of cognates[2], better basic vocabulary like close family member (since they are less likely to be borrowed or undergone big changes).

Example: comparison of romance languages descending from Latin

	Italian	Spanish	Portuguese	French	Gloss
1)	*corpo*	*cuerpo*	*corpo*	*corps*	body
2)	*crudo*	*crudo*	*cru*	*cru*	raw
3)	*catena*	*cadena*	*cadeia*	*chaîne*	chain
4)	*cacciare*	*cazar*	*caçar*	*chasser*	to hunt

There we see that the correspondence is k:k and k:ʃ. We also see that ʃ only appears in French before an *a* and when the other languages also have *a*. Thus, it can be assumed that the original phoneme in the mother language was *k. The original Latin words are *corpus, crudus, catena* and *captiare* all with the *k phoneme.

It is also important to note that languages that are chosen for the comparative methods can have different degrees of relatedness (see chapter about the Indo-European language family). Also, the result, so the reconstructed change pattern and system, must be probable.

8. Reconstruction process of phonology

In this penultimate chapter, we will finally showcase specific examples of reconstruction, and the exact procedure used to arrive at those conclusions.

[2] Related words (morpheme) from tow sister languages who descend from one common ancestor.

Let it be said that in this chapter, we will concern ourselves almost exclusively with the reconstruction of, firstly, phonology, &, to the extent that it is possible, morphology. This is due to the comparative method being applicable to these areas to a greater degree than for example grammar & syntax, with especially the latter being considered essentially uncharted territory for reconstructionist. As Phillip Baldi puts it, "[…] there is no regularity principle or Neogrammarian hypothesis for either morphology or syntax; the method keeps coming back to phonology.", which of course severely limits the efficacy of research in any discipline not directly linked to the latter.

8.1 Choosing languages to analyze

Before beginning the reconstruction process, we need to of course first establish our goal, i.e. the proto language we wish to reconstruct, for which we, in turn, must select a language family that is its supposed product.

In our case, we choose to reconstruct PIE, as the Proto-Indo-European languages present several advantages concerning that endeavor, as the enormous amount of research & documentation of even their oldest members is readily available, & the oldest preserved attestations date to even 2000 BC, demonstrating a remarkable time-depth. Furthermore, it allows for a large collection of comparative data due to its relative size & diversity, which still doesn't take away from the fact that in many key areas, its major branches are in relative agreement.

Thus, we will be working with languages such as Ancient Greek (An. Gr.), Latin (Lat.), Sanskrit (Skr.) & Gothic (Goth.) if possible, as each represents a separate branch of Indo-European while also being relatively ancient & thus likely closer to the initial PIE from which they all descend.

8.2 Assembly of datasets (cognates)

To start off, it is necessary to assemble the data we will be working with, that is to say, we need to create cognate lists, comparing terms which are theorized to have descended from a common form in order to recognize their similarities.

Said cognates lists ordinarily focus on essential basic vocabulary items which have a high resistance to lexical replacement and are thus likely to be preserved across languages. These usually include close kinship terms (mother, father, brother etc.), lower numerals (numbers up to 10), general body parts, in some cases pronouns & so forth. The cognates often share relatively evident phonological similarities (the key factor here), & any semantic differences may be explained by semantic drift.

Here a basic list of cognates:

	Meaning	Ancient Greek	Latin	Sanskrit	Gothic
1)	*"mouse"*	mûs	mūs	mū́ṣ-	mūs

2)	*"mother"*	mā́tēr	māter	mātár-	mōðir[3]
3)	*"father"*	patēr	pater	pitar	faðar
4)	*"nine"*	~~en~~né(w)a	novem	náva	niun
5)	*"dead"*	á~~m~~brotos	mortuus	mṛtá	maúrþr
6)	*"dog"*	kúōn	canis	śván-	hunds
7)	*"race, kind"*	genos	genus	jánas	kuni
8)	*"race, kind" (gen.)*	géneos	generis	jánasaḥ	kunjis
9)	*"old"*	hénos	senex	sána-	sineigs
10)	*"old" (gen.)*	hénos	senex	sánaḥ	sineigs
11)	*"(I) am"*	eimí	sum	ásmi	im
12)	conjugations of "to be"	éō	erō	ástu	ist
13)	*"(I) vomit"*	eméō	vomō	vámiti	vāma[4]
14)	*"to stand"*	hí-stāmi	stare	sthā-	standan

At this stage, to accurately reveal the cognates, an extensive grasp of philology, the study of ancient (attested) languages & documents, is required: For example, Ancient Greek *enné(w)a*, *ámbrotos* & *hí-stāmi* all contain prefixes, *en-*, *á-* & *hí-* which may be removed to isolate the cognate. Furthermore, the *-b-* in *ámbrotos* was added later by epenthesis as to avoid the *-mr-* cluster, which is disallowed in Ancient Greek. As a result, the relevant portions of these terms can be reduced to *né(w)a* & *mrotos*.

8.3 Establishing (sound) correspondence sets

Next up, we have to note down the phoneme correspondences in each set of terms. During this procedure, we should stay as precise as possible & also note down the exact environment in order to represent it most accurately & avoid making unnecessary generalizations.

	Environment	Ancient Greek	Latin	Sanskrit	Gothic
1)	word in.	m-	m-	m-	m-
2)	word in.	m-	m-	m-	m-
3)	word in.	p-	p	p-	f-
4)	word in.	n-	n-	n-	n-
5)	word in.	m-	m-	m-	m-
6)	word in.	k-	c-	ś-	h-
7)	word med.	-n-	-n-	-n-	-n-
8)	word fin.	-s	-s	-ḥ	-s
9)	word in.	h-	s-	s-	s-
10)	word fin.	-s	-s	-ḥ	-s
11)	word med. / fin.	-m-	-m	-m-	-m
12)	word med.	-Ø-	-r-	-s-	-s-

[3] Old Icelandic term, as Gothic cognate is unattested
[4] Old Icelandic term, as Gothic cognate is unattested

13)	word med.	-m-	-m-	-m-	-m-
14)	word in.	s-	s-	s-	s-

So, in grouping these correspondence sets together we would get:

	Environment	Ancient Greek	Latin	Sanskrit	Gothic
A)	In. & med.	m	m	m	m
B)	In. & med.	n	n	n	n
C)	fin.	s	s	ḥ	s
D)	In.	h	s	s	s
E)	med.	Ø	r	s	s
F)	in. (before consonants)	s	s	s	s
G)	in.	p	p	p	f
H)	in.	k	c	ś	h

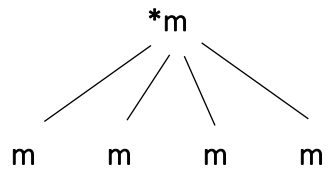
8.4 Reconstruction of proto phonology & further

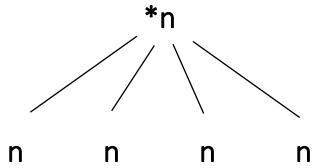
Now that we have effectively analyzed the comparative or horizontal dimension of our cognates, it is time to go for the vertical, otherwise called the historical dimension. It is here that we will try to reconstruct the underlying proto phonemes that these sound correspondences arise from & are reflexes of.

Let us start with the nasal phonemes of sets A) & B), as those present the simplest solutions:

	Environment	Ancient Greek	Latin	Sanskrit	Gothic
A)	in. & med.	m	m	m	m
B)	in. & med.	n	n	n	n

We choose the simplest possible reconstruction, which, with this particularly neat correspondence set, leads us to reconstruct the proto phonemes as *m & *n, for A) & B) respectively, directly inherited & preserved by our sample languages:



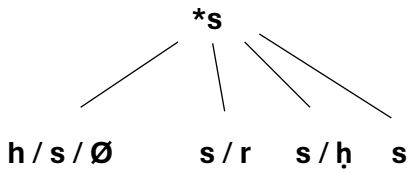


Next up, let's examine the correspondence sets of C), D), E) & F):

	Environment	Ancient Greek	Latin	Sanskrit	Gothic
C)	fin.	s	s	ḥ	s
D)	in.	h	s	s	s

E)	med.	Ø	r	s	s
F)	in. (before consonants)	s	s	s	s

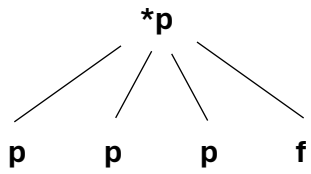
Sets C) & D) both leave room for the application of the majority principle, with three out of four candidates sharing a *s*, as well as the principle of directionality, with h/ḥ > s being much more common than the inverse s > h/ḥ. While set F) shows a uniform correspondence of s, set E) presents 2 outliers, which can nonetheless be conveniently explained: Latin r is the product of intervocalic rhotacism (s > r / V_V), while intervocalic s in Greek can be hypothesized to have weakened to h first, before completely disappearing eventually. We are able to reconstruct all of the above sets as *s, which remained in all the languages before other consonants, turning into h word-initially in Ancient Greek & Sanskrit ḥ in certain word-final positions. In a medial position between vowels, it changed to r in Latin & disappeared in Ancient Greek and remaining the same in all other languages in every case covered unless specified otherwise:

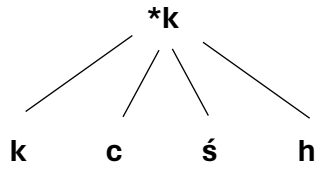


Lastly, let us analyze sets G) & H):

	Environment	Ancient Greek	Latin	Sanskrit	Gothic
G)	in.	p	p	p	f
H)	in.	k	c (k)	ś	h

Yet again the principle of directionality & that of majority inform our efforts: set G) shows three out of four in agreement, while the Gothic reflex f represents a weakening of a stop consonant to a fricative, cross-linguistically extremely common. Similarly, while Ancient Greek k & Latin c are in agreement, Sanskrit ś & Gothic h may also be explained the same way as f in set G). Thus, the proto phonemes for set G) & H) are likely to be *p & *k, respectively:





When enough of these original phonemes have been reconstructed, one may first double check all outcomes (see whether the proposed sound changes apply to words not already analyzed etc.), before potentially moving on to reconstruct the morphology of our proto language, or even further beyond, using our observations about phonology.

9. Conclusion

Here we are, at the end of this work. We have a long journey behind us. Together, we explored the topic of language reconstruction, using the example of Indo-European and making it accessible to people unfamiliar with linguistics.

It was the first time for both of us to work on such a project with someone else. I would say we did it quite well. For me who never investigated the research field of linguistics, except for etymology, this was quite a challenge, an interesting one though. I learned a lot and Viljami helped me understand some nuances better.

The only real problem we encountered was communication. It was a bit difficult to tell clearly who was doing what and so on. Thus, a big thanks to Mrs. Éloïse Hartheiser, our director of work, who helped us solve this issue and accompanied us through the whole creation process of this work. Her experience, comments and advice were a major help.

Thanks also to Viljami who is a trustworthy coworker: he did his part and did it well. He gave me (N. Turquin) advice to enhance my work's quality and precision.

The goal of this work was to answer the question "Do we just speak proto-Indo-European dialects?" We tried to find this answer by looking at the roots of languages, how proto-languages are reconstructed, specifically proto-Indo-European. It may not have been the fastest way, but we found an answer....

The different languages from the Indo-European family are not "just" dialects.

They are descendant from one mother-language indeed, but they changed so much that a French-speaker will not understand Hittite, even though they are both Indo-European languages.

They each have a unique, long, and complex history behind them. Every language has its own history, changes, grammar, uses, vocabulary. They can have similar points and common origin, but they are not shades of one color. They are different colors.

They evolved, are evolving and will evolve.

And that.... That is the beauty of languages.

Noémie Turquin and Viljami Viinamäki

10. Sources

10.1 Books & literature

- « Les Indo-européens : histoire, langues, mythes », by Bernard Sergent, Éditions Payot & Rivages, 1995, ISBN 2228889563
- “Comparative Indo-European: an Introduction”, by Robert S.P. Beekes, John Benjamins publishing company Amsterdam/Philadelphia, 1955, ISBN 9027221502
- “Patterns of change, change of patterns: linguistic change and reconstruction methodology”, by Phillip Baldi, Mouton de Gruyter Berlin/New York, 1991 ISBN 3110134055
- “Historical linguistics: A cognitive grammar introduction”, Margaret E. Winters, Benjamins Publishing Company Amsterdam/Philadelphia, 2020, ISBN 9789027205506

10.2 Web & internet

- https://upload.wikimedia.org/wikipedia/commons/6/6d/Family_tree_of_the_indo-european_languages_%28English%29.jpg
- https://en.wikipedia.org/wiki/Haplology#:~:text=As%20a%20general%20rule%2C%20haplology,%E2%86%92%20gently%3B%20ablely%20%E2%86%92%20ably.
- https://en.wiktionary.org/wiki/wesp
- https://www.ipachart.com/
- https://en.wikipedia.org/wiki/International_Phonetic_Alphabet
- https://phoible.org/
- https://en.wikipedia.org/wiki/Language_change
- https://en.wikipedia.org/wiki/Analogical_change
- https://en.wikipedia.org/wiki/Help:IPA
- https://en.wikipedia.org/wiki/Vowel_diagram
- https://indo-european.info/
- https://en.wikipedia.org/wiki/Proto-Indo-European_language#In_popular_culture
- https://en.wikipedia.org/wiki/Category:Splits_and_mergers_in_English_phonology#:~:text=A%20split%20in%20phonology%20is,with%20vowels%2C%20although%20not%20exclusively.
- https://en.wikipedia.org/wiki/Stress_(linguistics)#
- https://en.wiktionary.org/wiki/phonemic_split
- https://www.assyrianlanguages.org/hittite/en_lexique_hittite.htm#m
- https://upload.wikimedia.org/wikipedia/commons/b/b1/Indo-European_migrations.jpg
- https://en.wikipedia.org/wiki/File:Indo-European_expansions.jpg
- https://en.wikipedia.org/wiki/Proto-Indo-European_homeland

- https://en.wikipedia.org/wiki/Proto-Indo-European_society#Social_structure
- https://en.wikipedia.org/wiki/Proto-Indo-European_mythology
- https://en.wikipedia.org/wiki/*Dy%C4%93us
- https://en.wikipedia.org/wiki/Proto-Indo-Europeans#Notes
- https://hlw.id.ucsb.edu/ch12/PDF/Ch12GuideHistoricalReconstruct.pdf
- https://aryamccarthy.github.io/campbell1998historical/
- https://www.studysmarter.co.uk/explanations/english/the-history-of-english-language/comparative-method/
- https://hlw.id.ucsb.edu/ch12/PDF/Ch12GuideHistoricalReconstruct.pdf
- https://hlw.id.ucsb.edu/ch12/PDF/Ch12WritingLingHistoricalReconstruct.pdf
- https://alic.sites.unlv.edu/chapter-15-2-linguistic-reconstruction/