Indo-Slavic lexical isoglosses and the prehistoric dispersal of Indo-Iranian

Palmér, A.I.

2. Theory and methodology

2.1. Introduction

As the discussion in 1.3 above has shown, the internal structure of the Indo-European language family has been described using the tree model or the wave model, which are often seen as complementary. In the following sections, the theoretical principles and methodological practices associated with both models will be discussed and evaluated. Special emphasis will be given to lexical evidence in subgrouping methodology. It will be argued that both the tree and wave models have their place, but occupy different stages in the workflow of subgrouping research. Finally, linguistic palaeontology and its role in debates on the homelands of prehistoric linguistic communities will be discussed.

2.2. Phylogenetic subgrouping

The idea to represent the internal structure of the Indo-European language family in the form of a tree diagram goes back to Schleicher (1853; 1861).[16] As indicated by the title of his 1853 paper "Die erste Spaltungen des indogermanischen Urvolkes", Schleicher envisioned the splits between the branches in migrationist terms, i.e., the splits were caused by physical separation of speech communities. He operated with binary splits and indicated the longevity of branches by their relative length in the diagram. For example, "graecoitalokeltisch" represents a shorter period of unity than "arisch" (i.e., Indo-Iranian), despite the fact that these nodes are on the same level in the family tree (cf. Figure 2).

[16] However, already Rask (1818: 84) used a tree-like diagram of the Celtic languages, and there are even older examples, e.g., by the 17th century scholars Georg Stiernhielm and Georg Hickes (Sutrop 2012).

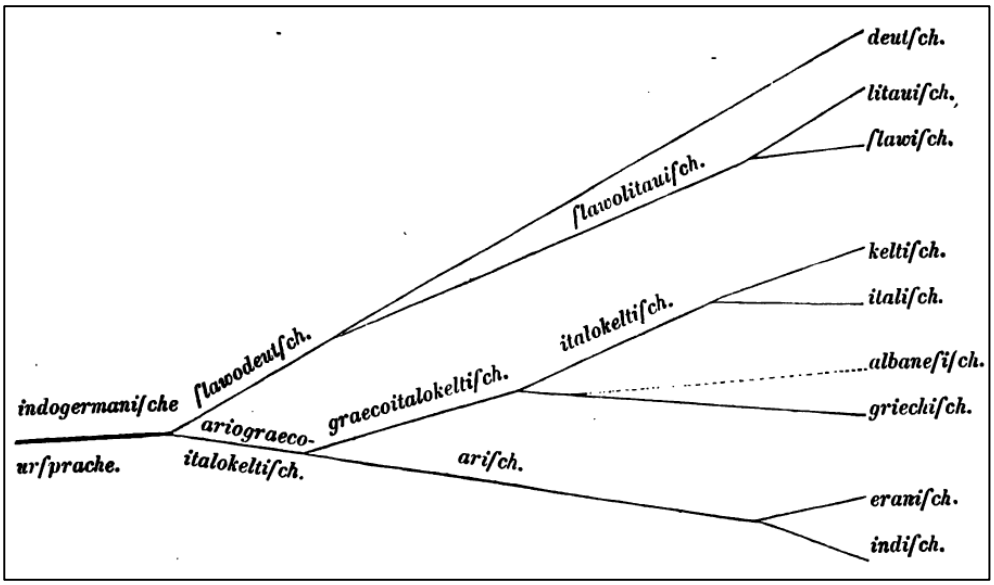

Figure 2. Schleicher's Indo-European family tree (1861).

It has often been remarked that the tree model is too abstract or simplistic to reflect the actual process of diversification of a language family as a historical reality (Hoenigswald 1966; Schlerath 1981; Clackson 2022: 26). Sudden migrations, causing clear-cut splits of speech communities, as Schleicher describes them, are rare. Already in the mid-19th century, scholars like Pictet (1859–1863: I, p. 48) and Schuchardt (1866: 103) argued that linguistic divergence is gradual. This realization was certainly an important motivation behind Schmidt's (1872) wave model (Geisler & List 2013). However, rather than reflecting the complex nature of actual linguistic divergence, the tree model may be seen as a post-hoc representation of the relationship between the branches. In other words, a split in the tree model does not, in most cases, correspond to a sudden split of a linguistic community, but rather to the observable result of a gradual diversification process. As we will see, the methodological strength of the tree model, in contrast to the wave model, lies in its rigidity (Schlerath 1981): it forces us to make explicit hypotheses regarding the phylogenetic relationship between the branches, which can be tested (and falsified) against the data.

2.2.1. The shared innovation principle

A core principle in subgrouping methodology, or phylogenetics, is that subgroups, or clades, should be based on shared innovations, not shared retentions or archaisms (Fick 1873; Leskien 1876; Delbrück 1880; Brugmann 1884; Dyen 1953: 581; Porzig 1954: 55; Greenberg 1957: 49; Hoenigswald 1966; Ringe 2017: 62).[17] An innovation is any linguistic

[17] Although Brugmann is usually credited for formalizing the principles of subgrouping methodology, already Fick (1873: 164) clearly states that subgroups must be based on shared innovations (here in an argument in favour of a European subgroup): "Um den Schluss auf eine ehemalige Spracheinheit aller Europäer wirklich zu begründen,

feature (phonological, inflectional, derivational, lexical, syntactical) of a language that was not present in its immediate ancestor. When two related languages have undergone the same innovation, it either reflects independent developments[18] or a shared development in a common ancestor. By compiling shared innovations of related languages, a family tree can be constructed, each branch or subgroup reflecting a set of shared innovations.

In the terminology of Ringe, Warnow & Taylor (2002), which is borrowed from cladistics, linguistic features may be described as *characters* with variable *states*. A character is an abstraction of phonology, morphology, lexicon or syntax, i.e., a phoneme, morpheme, lexeme, or syntactic structure. A state is the representation of a character in a particular language. For example, the Proto-Indo-European phoneme */p/ is a character that is reconstructed to explain a set of regular sound correspondences in the Indo-European languages. The character */p/ is reflected in the attested languages by different states, such as Latin *p*, Greek *p*, Germanic *f*. Based on the typological principle that *p* > *f* is a more likely sound change than the opposite, we hypothesize that Proto-Indo-European had the state **p*, which may be termed the *ancestral state*.[19] In this example, Germanic shows an innovative state, whereas Latin and Greek continue the ancestral state.

The identification of innovations presupposes that the ancestral state of a character can be determined. In practice, this is often extremely difficult. As discussed in section 1.3 above, the ancestral state of PIE **ḱ* has been reconstructed as a palatal/fronted velar or as a plain velar. Depending on the preferred reconstruction, the status of the attested states in the satem languages (e.g., Skt. *ś*, Av. *s* < PIIr. **ć*) as innovations or retentions changes.

---

müssen die Differenzen zwischen europäischer und arischer Sprache also derart sein, dass die europäische Eigenthümlichkeit eine Abweichung von der Ursprache enthält, dann zwingt uns dieselbe allerdings ein einheitliches sprachlich verbundenes Volk als Urheber dieser Umwandlung des früheren Bestandes anzunehmen, und wie uns die Differenz das Faktum der Scheidung verbürgt, so bezeugt *die gleichmässige Durchführung der sprachlichen Neuerung sprachlichen Zusammenhang unter den diese Neuerung durchführenden Individuen* [emphasis added]." It is possible that the shared innovation principle goes back to even earlier scholarship.

Leskien (1876: vii) argues: "Die Kriterien einer engeren Gemeinschaft können nur in positiven Uebereinstimmungen der betreffenden Sprachen, die zugleich Abweichungen von den übrigen sind, gefunden werden." Later (p. xxii), in a discussion on Schmidt's (1872) wave model, Leskien states that "Man bemerke, das es sich um lauter Verluste einst gemeinsamer indogermanischer Bildungen handelt. Sie beweisen für die nähere oder fernere Beziehung der betreffenden Sprachen nichts." Further on (p. xxiv), regarding an alleged Indo-Iranian-Balto-Slavic derivational correspondence, he argues that "Die Uebereinstimmung … beschränkt sich also … darauf, dass … in beiden Sprachgruppen eine gleichartige Weiterbildung mit Suffix *-ti-* vorgenommen ist, ein Umstand, dem ich bei der Häufigkeit des Suffixes in beiden keine besondere Bedeutung beilegen kann." With this, Leskien stresses that subgroups must be based on shared innovations, not shared archaisms or independent innovations.

Delbrück (1880: 135) contends that "nicht jede Gleichheit zwischen zwei Sprachen als argument für eine Urgemeinschaft betrachtet weden kann. […] [E]s bleiben streng genommen nur gemeinsam vollzogene Neuerungen als beweiskräftig übrig".

A few years later, Brugmann (1884: 231), in a dedicated methodological paper, concludes that "wirkliche Beweisgründe für die engere Zusammengehörigkeit zweier oder mehrerer Sprachen können nur solche Übereinstimmungen sein, welche Abweichungen von den übrigen Sprachen desselben Stammes sowie zugleich von der allgemeinen Grundsprache sind, also gemeinsam vollzogene Neuerungen".

[18] Independent developments refer to innovations that are independent in the phylogenetic sense. As such, the term encompasses parallel innovations as well as areal developments, such as borrowing.

[19] In the case of phonemes, the character and ancestral state are often identical, but refer to different aspects of the reconstruction; the character is a representation of the correspondence set responsible for postulating the phoneme, whereas the ancestral state refers to its phonological representation in the protolanguage. For morphological characters, the difference is more obvious; e.g., the 3sg. middle ending is a character with attested states such as Skt. *-te*, *-e*, Gr. -ται, -τοι, Lat. *-tur*. In this case, the reconstruction of the ancestral state is much more debated.

Even if a state that is shared by two languages is decidedly innovative, it is often difficult to determine whether the innovation is shared or independent. Continuing with the centum/satem example, provided that the ancestral state is a three-way system *$\acute{K}$, *K, *K^{w}, the merger of *K with either *$\acute{K}$ or *K^{w} is an evident innovation. However, if Luwian preserves the ancestral three-way system, Hittite must have merged *$\acute{K}$ and *K independently of the other centum languages.[20] This calls into question whether the centumization and satemization of the other branches could not also have happened independently (cf. Ringe 2017: 64).

2.2.2. Typology of shared innovations

Different types of characters are generally given different weight for subgrouping purposes (Porzig 1954: 59; Clackson 2022: 25). Innovations in inflectional morphology have been seen as the most significant,[21] since such morphemes are not easily borrowed, and often alter the morphosyntactic structure of the language (Greenberg 1957: 52; Clackson 1994: 25–26; Klingenschmitt 1994: 236). Yet, Greenberg (1957: 46) cautions that related languages may develop independently in the same direction, since they share the same starting point. Innovations in derivational morphology are also given considerable weight, as new morphemes are unlikely to develop independently. However, they are less resistant to borrowing (Thomason 2001: 70–71). Phonological innovations, i.e., sound changes, are generally ranked lower than morphological innovations, since they are often typologically common and therefore may affect languages independently (Greenberg 1957: 50; Clackson 1994: 20). However, a chain of shared changes that feed each other seems more significant (Greenberg 1957: 51). An advantage of a particular type of phonological innovation, namely phonological mergers, is that their directionality is provable (Hoenigswald 1966). Less trivial sound changes are given more weight, but judging which sound changes are trivial often becomes subjective. Syntactic innovations are often disregarded for subgrouping purposes,[22] since the risk of chance resemblance is high, and since syntactic structures of the protolanguage are more difficult to access with the comparative method than phonological and morphological features (Clackson 2007: 157ff; 2022: 23; Gildea, Luján & Barðdal 2020). For some syntactic characters, e.g., basic word order, the number of possible states is so limited that the risk of independent innovations is high (cf. Ringe & Eska 2013: 262).

2.2.3. Lexical characters as evidence for subgrouping

Lexical innovations have been given relatively little importance for inferring Indo-European phylogeny (Leskien 1876: xxiii; Delbrück 1880: 135; Hoenigswald 1966: 8). According to Meillet (1908: 126), this is because there are no two branches that do not share at least a few unique lexical correspondences. Porzig (1954: 59) argues that lexical

[20] Here and elsewhere in this work, the validity of the primary branches of Indo-European is taken for granted.
[21] Cf. Schleicher (1858b: 12): "…ich lasse nach einem bei mir feststehenden grundsatze nur den grammatischen ban als masstab der verwandtschaft zweier sprachen eines und desselben sprachstammes gelten und betrachte andere übereinstimmungen nur als willkommene zuthat…"
[22] Notable exceptions include Longobardi & Guardiano (2009) and Longobardi et al. (2013).

correspondences only show that two branches were in contact, i.e., developed from contiguous dialect groups, but not that they form a subgroup in the strict sense. In Olander's (2022) volume on Indo-European phylogeny, lexical evidence is either seen as complementary to phonology and morphology, or disregarded, with the exception of Olsen & Thorsø (2022: 211–12), who take lexical innovations as the main evidence for a Graeco-Armenian subgroup (see also Martirosyan 2013, who interprets this as a dialectal grouping, however).

The main problem associated with lexical evidence in subgrouping methodology is that it is unclear if and how lexical characters can be defined in a way that makes them relevant for subgrouping purposes. There are essentially two options:

1) *Character = etymon*: If the lexeme itself is the character, e.g., **h₂erh₃-tro-* 'plough', all branches attesting the lexeme share the same state. However, as Ringe, Warnow & Taylor (2002: 71) argue, the branches that do not attest **h₂erh₃-tro-* must count as having different states, since they could have lost the lexeme independently. The consequence is that such a character is compatible with any tree structure, i.e., it can always be back-projected to the root of the tree, and is thus uninformative for subgrouping (cf. Kortlandt 2016). This problem was recognized by Porzig (1954: 58–59), who argued that isoglosses must be based on two positive states rather than presence and absence, which requires the use of semantic concepts rather than reconstructed etyma as characters. Peyrot (2022) describes the problem in terms of lack of *identifiability*. This term refers to the methodological criterion that "the linguistic element adduced as a shared innovation in the lower node should be clearly identifiable in the higher as well as in the lower node" (Peyrot 2022: 90). In other words, if a feature such as a lexeme is absent from an attested language, it is not identifiable, because we cannot determine whether it was lost in that language or never existed. By implication, the lexeme is not identifiable in the common ancestor; it may or may not have existed there. The result is that the ancestral state cannot be determined.

It is important to realize that identifiability is not only a concern regarding lexical characters. As exemplified by Peyrot (2022: 91), the comparative and superlative suffixes of, e.g., Greek and Indo-Iranian, are unattested in Anatolian and Tocharian. While one may be inclined to analyse them as innovations of the non-Tocharian Core Indo-European branches, it cannot in principle be excluded that these morphemes were lost in Anatolian and Tocharian. Therefore, in the strictest application of the identifiability criterion, the comparative and superlative suffixes are uninformative for the phylogeny. Many morphological characters in Ringe, Warnow, and Taylor (2002: 117–18), such as the augment and thematic optative, suffer from the same problem.

2) *Character = semantic concept*: If, as Porzig argued, semantic concepts are used as characters (thus, e.g., Ringe, Warnow & Taylor 2002; Kassian et al. 2021), e.g., PLOUGH, each branch that attests the same formation for a particular semantic concept is assigned the same state. As long as every branch has a word for PLOUGH, the identifiability criterion is met.[23] However, this approach does not resolve the problem of determining the ancestral state, i.e., which state should be reconstructed to Proto-Indo-European for that particular

[23] In reality, it is not necessarily the case that each branch attests a word for a given semantic concept.

semantic concept. Furthermore, it restricts the object of study to basic vocabulary, since this is the only part of the lexicon where all languages are expected to attest one basic lexeme for each semantic concept, based on linguistic typology.[24] For non-basic vocabulary, such as PLOUGH, or more abstract concepts (e.g., BEAUTY, DOWNWARDS), languages show great variation in how semantic concepts are mapped. Therefore, using such concepts as characters is not justifiable.

However, even if exclusively basic vocabulary is considered, it is far from straightforward to accurately determine which lexeme occupied a certain basic vocabulary slot in modern languages, let alone in ancient languages or in the protolanguages of each branch (cf. the different approaches in Dyen, Kruskal & Black 1992; Kassian et al. 2021; Heggarty et al. 2023). As Peyrot puts it, "several etyma may have similar, overlapping or even identical meanings, and it is therefore difficult to prove that a certain meaning came to be expressed with a different etymon" (2022: 91).

Additionally, from a theoretical perspective, it is questionable whether the replacement of the lexical form mapped onto a particular semantic concept is in itself a significant process, comparable to a sound change or replacement of an inflectional ending, especially if the form itself is not a unique formation. Consider, for example, the character SWIM in Kassian et al. (2021: S86). Indo-Iranian and Balto-Slavic are argued to share the innovative state **pleu*- 'to swim', as opposed to, e.g., Gr. νέω 'to swim'. Yet, Greek also attests περιπλέω 'to bypass while sailing or swimming'. This implies that the meaning 'to swim' of **pleu*-, which elsewhere means 'to float, flow' (cf. LIV: 487), may be archaic. In Kassian et al.'s methodology, however, it is the fact that Indo-Iranian and Balto-Slavic use the same lexical material as the basic word for the same semantic concept that counts as a shared state, irrespective of whether a cognate is attested elsewhere. Even more questionable is the treatment of the basic vocabulary item FIRE in Kassian et al. (2021: S41–42). Although Indo-Iranian, Balto-Slavic, and Latin all attest reflexes of **h₁ngʷni*- 'fire', only the former two are argued to share the same state. Latin *ignis* 'fire' is coded as a separate state, since the other Italic languages attest reflexes of **peh₂ur*/*n*- 'fire', which is argued to be the ancestral state, as it is found in Anatolian (e.g., Hitt. *paḫḫur*-/*paḫḫuen*-). Therefore, the replacement of **peh₂ur*/*n*- by **h₁ngʷni*- in Latin is argued to be a post-Proto-Italic development, not shared with Indo-Iranian and Balto-Slavic. However, this argumentation fails to take into account that the alleged replacement of the basic word for fire could be an independent process in Indo-Iranian and Balto-Slavic as well. Thus, this methodology clearly represents a departure from the principles of the traditional comparative method.

In any case, for lexical comparison of words outside of the basic vocabulary, we are forced to retain the etymon, or lexeme itself, as the defining unit of the lexical character (*character = etymon*), as opposed to the semantic concept (*character = semantic concept*). As we have seen, this methodology carries with it two problems. First, the possibility of loss in one or several branches prevents identification of the lexical character in those

[24] The notion of "basic vocabulary" is by no means unproblematic or objective, but cf. Tadmor, Haspelmath & Taylor (2010) for an empirically motivated basic vocabulary list.

branches. By extension, the ancestral state of lexical characters (which is limited to presence vs. absence) cannot be determined.

To address the difficulty of determining their ancestral state, lexical characters must be divided into subtypes, each with different limitations.

1) For root isoglosses, i.e., when branches share a unique root but no verbal or nominal derivatives, their status as archaisms or innovations is in general not possible to determine. The default assumption is that they are archaisms. An exception is if it can be demonstrated that a root is derived from another root by a suffix or *root extension*.

2) In the case of nominal and verbal derivatives, an important factor is whether it is reasonable to believe that the root was synchronically productive in the subgroup from which the branches that attest the formation are hypothesized to be descended. If the root is isolated, i.e., not found in any other formations, it may indicate, though not prove, that the shared derivative is an archaism. If the root has a solid Indo-European etymology, i.e., is found in other branches, it becomes more likely, although by no means proven, that the shared derivative is an innovation. Another important factor is the productivity of the derivational morphology in question. Shared formations with derivational morphemes that are highly productive may indicate independent innovations. Conversely, rare or obscure derivations may indicate archaisms.

3) Semantic isoglosses, i.e., when branches share a specific meaning of a lexeme attested in other branches, are powerful in the sense that they more easily fulfil the identifiability criterion. However, the directionality of the semantic change, and thus the ancestral state, is not always possible to determine.

4) Finally, shared lexemes may reflect borrowings from other (non-Indo-European) languages. Such loanwords are most plausibly identified by irregular correspondences in other Indo-European languages by or violations of Proto-Indo-European phonotactics (Meillet 1908–1909; Cuny 1910; Jakob 2023a; Thorsø 2023; Wigman 2023).

The list above represents some general considerations, but in practice, possible shared lexical innovations must be analysed case-by-case. As will be shown in Chapter 3, a number of plausible innovations may be found among the Indo-Slavic lexical isoglosses, despite the numerous problems and caveats presented here.

To address the problem of identifiability of lexical characters (as well as characters more broadly), it is important to realize that a single character is rarely informative for the phylogeny of the whole family. For example, suppose a language family with five branches *ABCDE*. For a character *xy*, state *x* is attested in branches *A* and *B*, and can be shown to be a shared innovation as opposed to state *y*, attested in branches *C* and *D*. Branch *E* attests neither state *x* or *y*. Based on this situation, state *x* suggests that *A* and *B* derive from a subgroup to the exclusion of *C* and *D*. However, several scenarios are possible for branch *E*: 1) it could theoretically have taken part in the innovation of *x*, but subsequently lost it, in which case *E* would belong to the same subgroup as *A* and *B*, 2) it may have lost the ancestral state *y*, or 3) *E* had neither state *x* or *y*, in which case *y* is rather a shared innovation of *ABCD*, after which *AB* replaced *y* with *x*. Ideally, by combining the evidence of a large number of characters, each informing on different subsets of the branches of the language family, a true phylogeny may be inferred.

The above considerations notwithstanding, lexical evidence also has its advantages over morphology and phonology. New lexemes develop continuously through derivational processes and language contact, without altering the system of the language as a whole. Since the lexicon consists of a large number of discrete units, even a short-lived subgroup would be expected to show lexical innovations. Therefore, lexical isoglosses have the potential of disentangling phylogenetic relationships that cannot otherwise be defined by morphological or phonological isoglosses, either because there were none or because they were obscured through later developments. Furthermore, lexical evidence indirectly encompasses derivational morphology and might reveal innovative patterns of derivation. Additionally, the lexicon can reveal material conditions of the speakers, such as familiarity with flora and fauna or technological innovations, which with the methodology of linguistic palaeontology can be compared with the archaeological record to trace subgroups in time and space (cf. 2.5).

2.2.4. Quantity of shared innovations in subgrouping

Next to the shared innovation principle, discussions on subgrouping methodology generally assert that subgroups should be posited only when the number of shared innovations is high enough (Brugmann 1884: 253; Dyen 1953: 581). This is to avoid reliance on a small number of innovations that may in fact have been independent (Porzig 1954: 55). With respect to lexical innovations, a large number is argued to be especially important, since they are not given much weight individually (Clackson 1994: 25).

However, the required number of shared innovations is difficult to define objectively. As Porzig (1954: 55) notes, Brugmann & Delbrück (1897: 20–21) present "only" seven shared innovations as evidence for Balto-Slavic.[25] Clackson (1994: 199–200) concluded that the five lexical innovations (among 25 lexical isoglosses) he found in Greek and Armenian (with no supporting phonological or morphological innovations) were too few to support a Graeco-Armenian subgroup. Holst (2009: 53–54), on the other hand, argues that Clackson's conclusion is too conservative, and that the aggregate of evidence (including less convincing cases of shared innovations) rather tips the balance in favour of Graeco-Armenian.[26] These conflicting interpretations are mostly a consequence of different approaches to subgrouping: in Clackson's view, the paucity of lexical innovations is not only problematic due to the ever-present possibility that they were after all independent innovations, but also because such innovations would not have altered the structure of the ancestor of Greek and Armenian sufficiently to justify calling it a subgroup. Rather, Clackson attributes such shared innovations to dialectal developments within Proto-Indo-European. Holst (2009: 52) describes such statements as "Wischi-Waschi-Klassifizierungen", and calls for more explicit conclusions in terms of subgrouping. This

[25] Although, in their defense, the list is called "Einige Kennzeichen des baltisch-slavischen Zweigs", which implies that additional shared innovations may have been excluded.

[26] However, Holst (2009: 65) ultimately operates with a Balkan Indo-European subgroup, where Greek is closer to Albanian, Macedonian, and Phrygian than to Armenian. This is not insignificant for the interpretation of Clackson's material, since one of his five Graeco-Armenian lexical innovations (**wes-nu-* 'to clothe) is shared to the exclusion of Albanian, cf. Alb. *vesh* 'to clothe' ~ Skt. *vāsáyati* 'id.'.

debate illustrates the difficulty associated with the quantitative assessment of shared innovations.

A related question is whether a large number of lexical isoglosses is significant even if not all of them can be shown to be innovations. This is essentially what Holst (2009: 54) alludes to regarding Clackson's (1994) compilation of Graeco-Armenian isoglosses: "Wenn aber große Datenmengen in eine Richtung weisen, dann ist dies aussagekräftig." Holst's statement rests on the assumption that similarity (= a high number of lexical isoglosses) can serve as a proxy for subgroupiness (= a high number of shared innovations), which need not be the case (Holm 2003). Moreover, the problem is again how large that number must be to be significant in this regard, i.e., not a result of chance. Ideally, all Indo-European lexemes should be classified according to which branches attest them, to determine whether certain branches share disproportionate numbers of lexical isoglosses. Even with such a dataset, however, the Indo-European languages are not equal in terms of time depth, wealth and type of attestation. Again, due to lexical replacement, languages that are abundantly attested (modern and/or with rich ancient literature) have a higher chance of preserving more lexical material. Some languages tend to be more conservative, whereas others are known to have undergone heavy lexical replacement. This makes it difficult to assess to what extent the number of isoglosses shared by various branch pairs reflects actual relatedness.

2.2.5. The Indo-Iranian bias

As discussed above, several methodological constraints apply to the distinction of lexical innovations from archaisms. On top of these, however, progress has been hindered, especially with regards to Indo-Iranian, by the practice of back-projecting all Indo-Iranian lexemes with cognates in other branches to Proto-Indo-European (e.g., Scherer 1952: 6–7), without considering the possibility that not all such cases need necessarily be archaisms. I call this the Indo-Iranian bias.

For example, Mallory & Adams (1997; also Mallory 2013; 2019) reconstruct any lexeme to Proto-Indo-European that has cognates in at least one European and one Asian branch (which here means Indo-Iranian and Tocharian).[27] Thus, the principle has a clear methodological purpose, namely to account for the probability that most Proto-Indo-European words are not retained in all the branches. To avoid back-projecting words attested in geographically contiguous branches, for which the possibility of post-Proto-Indo-European developments increases, the geographical distance between Indo-Iranian (+ Tocharian) and the rest of the non-Anatolian branches is invoked to make the reconstruction more plausible (cf. also Gaitzsch & Tischler 2017). While this is not entirely unreasonable per se, it is not surprising that, with this methodology, Indo-Iranian will have preserved the largest number of Proto-Indo-European lexemes, since a cognate in any of the seven "European" branches leads to back-projection. Yet, while a larger geographic spread might make it more likely that a word goes back to Proto-Indo-European, it does not prove it, since branches that are now far apart may have been closer in prehistory, or even part of

[27] This is one of two possible criteria for Proto-Indo-European reconstruction in Mallory and Adams (1997): the other is attestation in Anatolian and any other branch, since they accept the Indo-Anatolian hypothesis.

a subgroup after the split of (Core) Proto-Indo-European. In fact, this methodology introduces a bias in the phylogeny, since it implicitly assumes a primary split between Indo-Iranian and the European branches. The result is that any potential evidence for a subgroup consisting of Indo-Iranian and a European branch, such as Balto-Slavic, is rejected out of hand.

2.3. Dialectal subgrouping

As mentioned above (p. 4, fn. 5), Pictet (1859–1863: I, p. 48ff) argued that linguistic divergence is gradual. He used an abstract geographical model (cf. Figure 3) to describe the dispersal of the Indo-European branches from a common origin. The model implies a period of dialectal differentiation where the branches gradually emerged from what was once a unified Proto-Indo-European language.

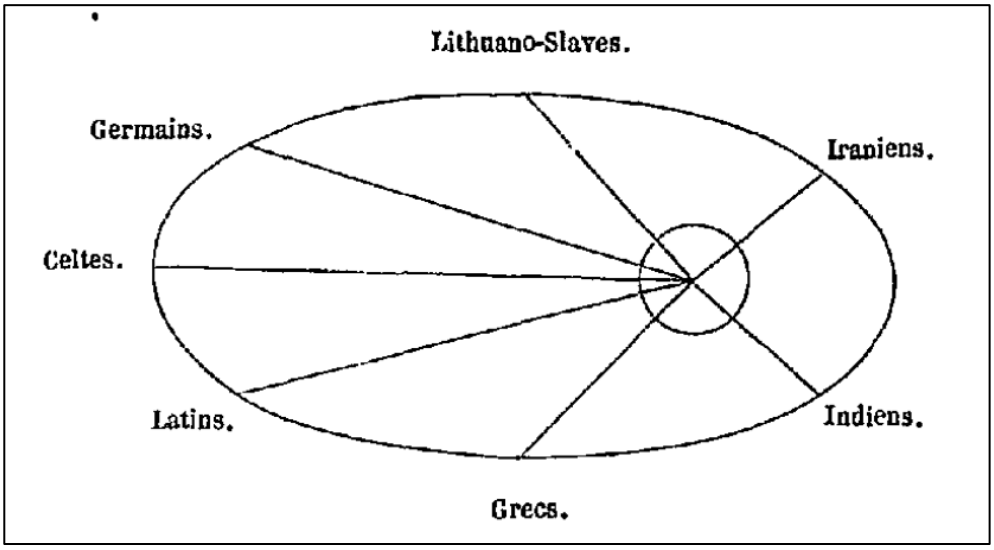

Figure 3. Pictet's divergence model (1859).

Schmidt (1872) formulated the wave model as a direct reaction to Schleicher's (1861) tree model.[28] Based on a rich dataset of overlapping lexical isoglosses shared by various combinations of Indo-European branches, Schmidt argued that it was not possible to view the diversification of the family in terms of splits from a common source. Instead, the branches reflect the remnants of a prehistoric dialect continuum, which was broken up when centres of innovation emerged at various points in the continuum, gradually making certain dialects more similar, and others more differentiated. This process can be conceived of as a stairway, where the steps become bigger over time.

[28] "Wollen wir nun die verwandtschaftsverhältnisse der indogermanischen sprachen in einem bilde darstellen, welches die enstehung irer verschiedenheiten veranschaulicht, so müssen wir die idee des stammbaumes gänzlich aufgeben. Ich möchte an seine stelle das bild der welle setzen, welche sich in concetrischen mit der entfernung vom mittelpunkte immer schwächer werdenden ringen ausbreitet" (Schmidt 1872: 27).

Meillet (1908) adopted the premises of Schmidt's wave model, but focused on phonological and morphological isoglosses to establish dialectal groups among the branches of Indo-European. Importantly, Meillet based these groups on shared innovations, just like subgroups in the tree model. His most important result was that the branches that share innovations are geographically contiguous, which was interpreted as evidence that the relative position of the branches reflects the relative position of prehistoric Proto-Indo-European dialect areas (Meillet 1908: 10–11, 134–35).

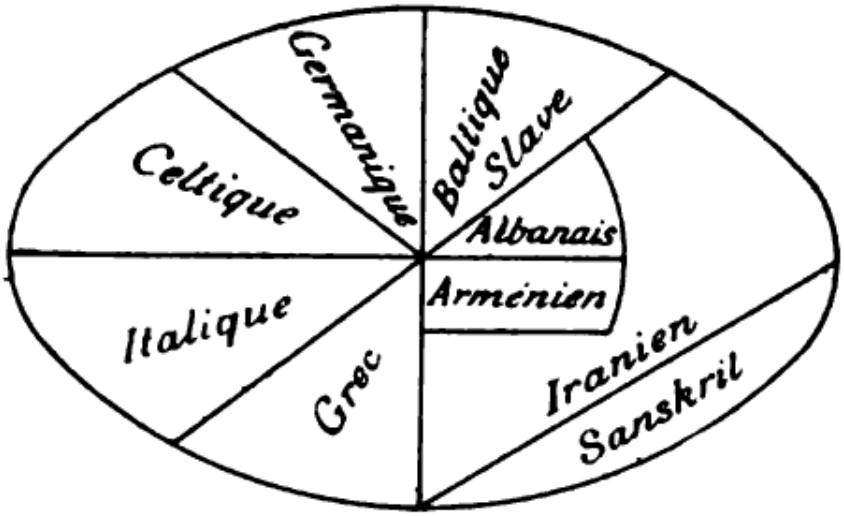

Figure 4. Meillet's Indo-European dialectal model (1908).

Bonfante (1976, reprint of 1931 original) reached a similar conclusion, supporting a fundamental east-west division corresponding roughly to the centum/satem isogloss, with Greek and Balto-Slavic occupying intermediate positions. Bonfante (1976: 116–17) further argued that apart from Celtic and Indo-Iranian, any Indo-European branch can share isoglosses with any other branch, without overlap in geographically intermediate branches. This is represented in the following model (Figure 5):

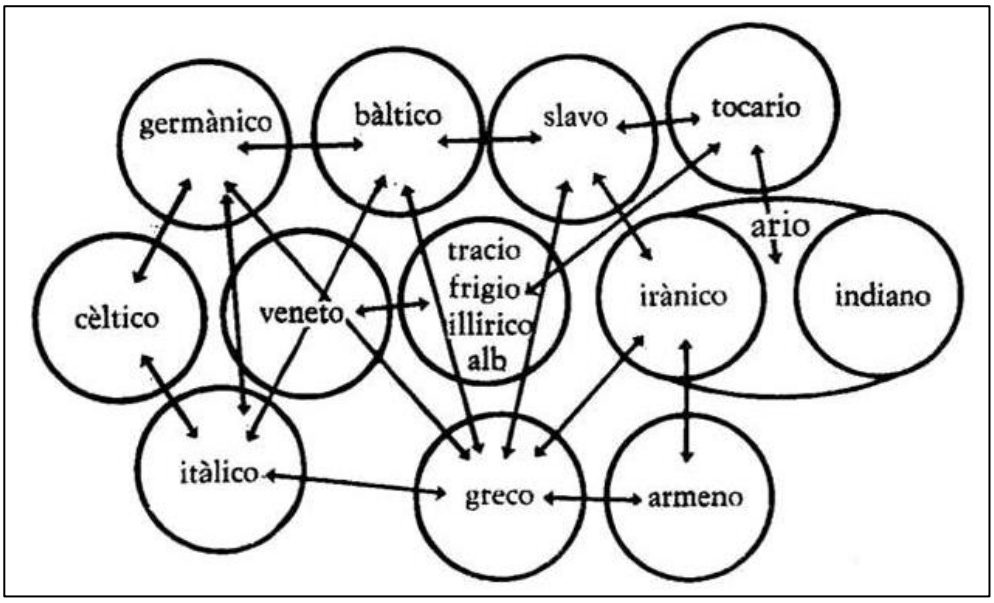

Figure 5. Bonfante's Indo-European dialectal model (1976/1931).

Other dialectal models represent the isoglosses themselves, rather than just the relative position of the branches, cf. Figure 6 (from Bloomfield 1935: 316) and Figure 7 (from Anttila 1972: 305). It should be noted that the authors of these models do not claim to present a complete picture of Indo-European dialect relationships, but they show that Meillet's methodology has had a lasting impact on Indo-European studies.

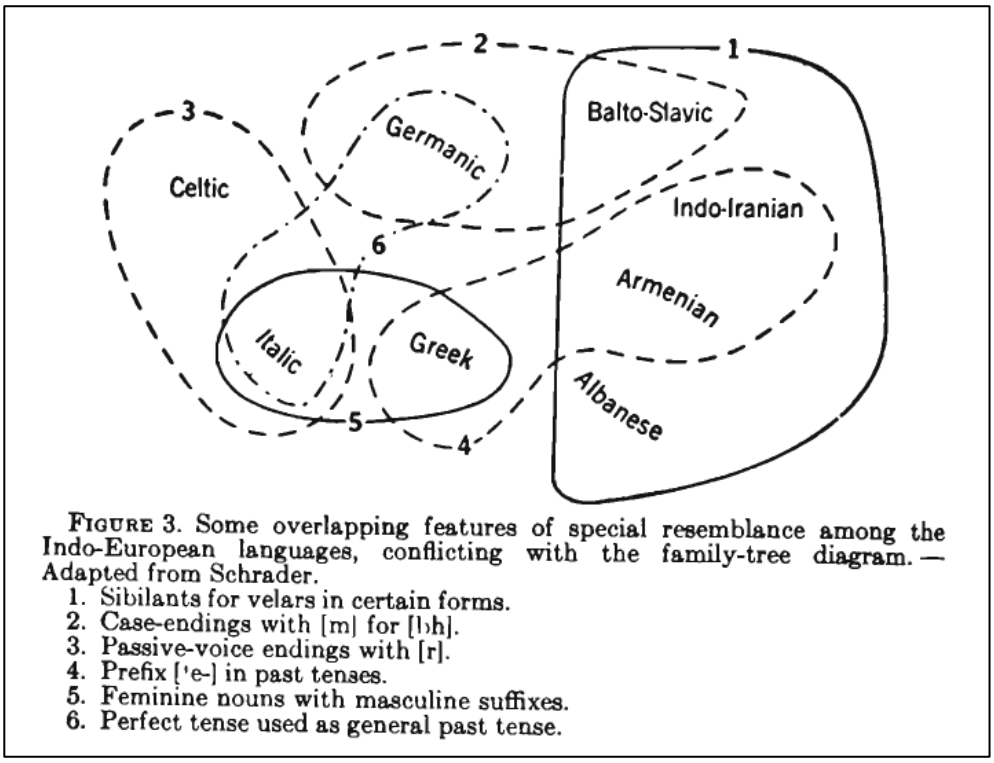

FIGURE 3. Some overlapping features of special resemblance among the Indo-European languages, conflicting with the family-tree diagram. — Adapted from Schrader.

1. Sibilants for velars in certain forms.
2. Case-endings with [m] for [bh].
3. Passive-voice endings with [r].
4. Prefix [ᵗe-] in past tenses.
5. Feminine nouns with masculine suffixes.
6. Perfect tense used as general past tense.

Figure 6. Bloomfield's Indo-European dialectal model (1935: 316).

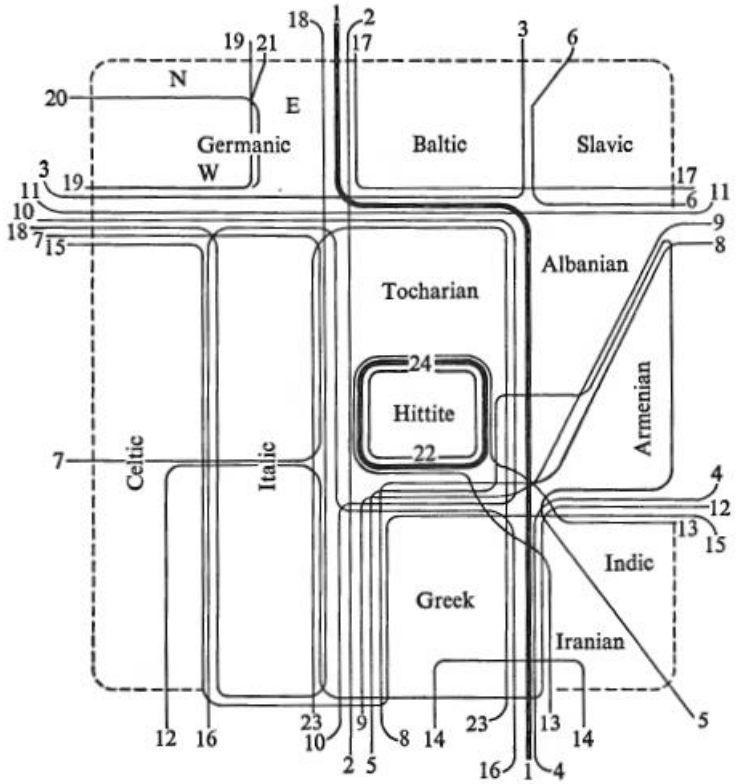

FIGURE 15-2. A dialect map of the Indo-European languages.

DEFINITIONS OF THE ISOGLOSSES

1. centum | satem [right] (§ 11.12)
2. *-ss-* | *-st-*, *-tt-* [right]
3. *aoə* | *a*, *āō* | *ō* [inside]
4. *eao* | *a* [inside]
5. *s* | *h* [inside]
6. *CVRC* | *CRVC* [inside] (§ 4.18)
7. *kʷ* | *p* [inside] (§§ 18.13, 18.16)
8. *e-* | *ø* 'past' [left, outside] (§ 19.10)
9. *-osyo* 'genitive' [right, inside]
10. *-r* | *-i* 'present' [right, outside] (§ 19.10)
11. *-m-* | *-bh-* 'case marker' [below]
12. *-to-* | *-mo-* 'ordinal' [below]
13. *-u* 'imperative' [inside] (§ 19.10)
14. *proti* | *poti* 'preposition' [inside]
15. secondary endings (without no. 10 *-i*) [below] (§ 19.10)
16. feminine nouns with masculine e..d-ings [inside]
17. *-ad* 'ablative' | 'genitive' [inside]
18. new tense system from perfect [inside] (Chapter 12)
19. umlaut [inside] (§ 4.5)
20. *-ww-*, *-jj-* | stop + *w*, *j* [outside]
21. *-ggj-* | *-ddj-* [right] (no. 20)
22. laryngeals as *h*'s [inside] (§ 12.4)
23. uncontracted reflexes of sequence **yH* [inside]
24. unit pronouns | particles + enclitic pronouns [inside] (§§ 19.8, 19.9)

Figure 7. Anttila's Indo-European dialectal model (1972: 305).

The advantage of the wave model and dialectal models above is that they allow more complex interrelations between branches to be represented, as opposed to the sharp splits of the tree model. Since dialect continua are the rule rather than exception in the historical period, it may be argued that dialectal models more accurately reflex the reality of linguistic

divergence in prehistory. This was certainly the motivation behind Meillet's (1908) study of the internal relationships of the Indo-European language family.

However, the wave model is not only motivated by typological-theoretical considerations: for many proponents, the motivation stems directly from the data. According to Bloomfield (1935: 317), "scholars, who insisted upon the family-tree diagram, faced an insoluble problem", since the discovery of more and more overlapping isoglosses did not allow the Indo-European branches to be neatly divided into subgroups. Instead, explaining such isoglosses required the assumption that a branch could undergo shared innovations with different branches independently.

Yet, the alleged overlapping innovations usually cited are not unambiguous. For example, Bloomfield (1935: 315) points out that the shared instrumental-dative endings in *-*m*- of Germanic and Balto-Slavic contrast with endings in *-*bʰ*- of the other branches (cf. Schleicher 1858b: 13; Leskien 1876: 157), which conflicts with the centum/satem isogloss that otherwise divides Germanic and Balto-Slavic. However, even if we accept Bloomfield's premise, i.e., that Germanic and Balto-Slavic share an innovative state here,[29] the ancestral state is not straightforwardly reconstructed. Hirt (1895a) argued that both endings existed in the protolanguage with different function. In this case, the generalization of either one in the branches may reflect independent innovations (Pronk 2022: 280). As for the centum/satem isogloss, the discussion in 1.3 above has shown that its value for subgrouping is limited, since the ancestral state is ambiguous, preventing us from proving beyond reasonable doubt that the satem languages have undergone a shared innovation.[30] In contrast to Meillet and Bloomfield, Anttila's (1972: 305) model (Figure 7 above) makes no attempt to distinguish archaisms from shared or independent innovations. Consequently, it does not demonstrate the need to assume overlapping innovations to account for the relationship between the branches.

A further example of an alleged overlapping innovation is Meillet's (1908: 57–61) discussion on the development of the Proto-Indo-European dental cluster **tt*. Since the reflex **ss* shared by Italic, Celtic, and Germanic bears the least resemblance to the Proto-Indo-European reconstruction, Meillet interprets this as a shared innovation, which may well be the case. However, the development to **st* in the eastern branches, which Meillet considers "moins instructive, bien qu'encore notable", since it correlates with his idea of

[29] The main argument for this is OCS dat.sg. *tebě* 'to you', showing that Balto-Slavic too retains an ending in *-*bʰ*-, which due to its marginal position may be considered a relic of an older inflectional system.

[30] All isoglosses mentioned by Bloomfield (1935: 316) turn out to be uninformative: 1) Satemization; 2) Case-endings with *-*m*- for *-*bʰ*- (cf. above); 3) Middle endings with *-*r*. These are archaic, given that they are found in Anatolian. Since *-*r* is reflected in the Indo-Iranian 3pl. but lacking from the Celtic 2pl., the ancestral system cannot be reduced to featuring either a marker *-*r* or *-*i* (cf. Beekes 2011: 268–69); 4) The augment **hₑe*-. The function of this morpheme is not identical in the branches that have it. This suggests that **hₑe*- was grammaticalized independently in the branches, in which case the branches that do not show the augment may have lost it. Furthermore, it has been argued that Anatolian preserves traces of the augment (Norbruis 2021: 209ff); 5) Feminine nouns with masculine suffixes. Rather than reflecting an innovation, such cases surely go back to the Proto-Indo-Anatolian gender system, which (like Anatolian) did not include the feminine gender; 6) Perfect tense used as general past tense. This is a typologically common change (also in Classical Sanskrit, for example) that need not be a shared innovation of Germanic and Italic, especially since these categories are morphologically divergent in the branches (e.g., Italic perfect resulting from a merger of the Proto-Indo-European perfect and aorist).

shared developments in contiguous branches, turns out to be uninformative. First, the change from **tt* to **tst* is shared with Anatolian and therefore reflects the ancestral state, cf. Hitt. 2pl.pres. *azzaštēni* 'you eat' < **$h_1$ed-t°*. In Indo-Aryan and Iranian, the outcome of **tst* (Skt. *-tt-*, Av. *-st-*) is posterior to branch-specific post-Proto-Indo-Iranian developments, i.e., Indo-Aryan loss of **s* between stops, cf. Skt. 3sg.aor. *ábhakta* < **Ha-b^hak-s-ta* from *bhaj-* 'to distribute', and Iranian loss of **t* before **s*, cf. YAv. *masiia-* m. 'fish' ~ Skt. *mátsya-* 'id.' < **matsia-*. Unlike Iranian, in Greek, the development of **tst* > στ is not identical to the development of **ts*, which is retained in some dialects (cf. Cretan αναδαζαθαι from δατέομαι 'to divide').[31] Based on these considerations, the evidence points to independent innovations in Greek and Indo-Iranian (thus also Porzig 1954: 78) and does not contradict the group defined by satemization, which Meillet (1908: 51) takes as a shared innovation.

One case where the wave model seems necessary to disentangle the internal structure of a language group is the Germanic branch (cf. Agee 2021). On the one hand, North and East Germanic share Holzmann's law, causing fortition of **-jj-* and **-ww-*. On the other hand, North and West Germanic share several innovations, such as deictic pronouns in **-si* and introduction of secondary diphthongs in strong verbs (Kroonen & Hansen 2022: 159). While the value of Holzmann's law for subgrouping has been questioned (Rasmussen 1990), it seems plausible that it spread in a Proto-Germanic dialect continuum, after which Northwest Germanic formed a subbranch defined by a large number of shared innovations (Kroonen & Hansen 2022: 160).

However, applying a wave model perspective to the ten branches of Indo-European is different from applying it to Germanic, which has only three subbranches and more shallow time depth. Thus, while the typological-theoretical motivation behind the wave model is sound, much remains uncertain regarding the extent to which it is actually necessary to apply it to the diversification of Indo-European. It seems a fair assessment that the inclination to abandon the tree model for the wave model has not been based on the strength of one or a few plausible examples of overlapping innovations (since the ones discussed above generally do not stand up to scrutiny). Rather, the observation that so many *possible* overlapping shared innovations are found in the first place has led to the suspicion that the tree model imposes artificial constraints on subgrouping (i.e., the disallowance of overlapping shared innovations) that do not reflect the reality of linguistic diversification.

The problem is that once the wave model is accepted, the distinction between shared and independent innovations becomes less important (e.g., François 2014: 177, who deliberately does not distinguish between them), since the model can easily accommodate overlapping innovations. As Ringe (2017: 65) puts it, "[a] major weakness of the dialect geography model is that it is difficult to falsify; new evidence that is at variance with the evidence already in hand can often be accommodated on an abstract dialect "map" without major revisions". Therefore, Ringe & Eska (2013: 263) argue that "…the *Stammbaum* hypothesis is always preferable *as a first hypothesis* because it is falsifiable".

[31] According to Norbruis (2023), the regular outcome of **tst* in Greek is rather -σθ-, in which case it would be different from Iranian altogether.

In this sense, the wave model is secondary to the tree model in the workflow of subgrouping research. If (or rather, when) efforts to infer clear-cut splits in the family tree fail, wave-like developments may be assumed for those shared innovations that seem to contradict the tree structure. Both methods are – or should be – based on the shared innovation principle, but only the wave model allows overlapping shared innovations.

2.4. Hybrid models

Already Leskien (1876: xii) saw the tree and wave models as complementary rather than contradictory, describing different aspects of language divergence and relatedness. For this reason, attempts have been made to devise a hybrid model that eliminates the shortcomings from which each model suffers on its own.

Meid (1975) proposed a model in which the branches of Indo-European descend not from a uniform protolanguage but rather from dialectally and chronologically diverse varieties of the protolanguage. To the Early Indo-European (*Frühindogermanisch*) layer belong features where all branches agree. Middle Indo-European (*Mittelindogermanisch*) and Late Indo-European (*Spätindogermanisch*) refer to less archaic stages with increasing dialectal differentiation. Anatolian split off before the Late Indo-European stage, but the remaining dialects are not argued to form a subgroup, because they did not innovate sufficiently together, rather forming a dialect continuum. Schlerath (1981) criticized Meid's approach, on the grounds that the comparative method by definition produces (at least in theory) a chronologically and dialectally uniform protolanguage. Therefore, Meid's model makes assumptions about Proto-Indo-European that, while not themselves implausible, do not follow from the comparative method, and thus cannot be falsified.

Gamkrelidze & Ivanov (1995: 363) present a hybrid "areal-genetic" model (cf. Figure 8), which follows Meid's division of the protolanguage into three stages, the latter two being dialectally differentiated. Later, the non-Anatolian part of the family splits into dialectal subgroups that in some cases nevertheless continue to innovate together (as in the case of Armenian, Indo-Iranian, and Balto-Slavic), long after what may be termed "dialectal Proto-Indo-European". However, in this model, it is unclear why certain innovations are treated as independent and others as shared by branches that had already split in the sense of the tree model. For example, the outcome of the dental clusters in Italic, Celtic, and Germanic (**ss*) is treated as independent innovations that have no implications for subgrouping, since the Italic development is argued to be posterior to Lachmann's Law (Gamkrelidze & Ivanov 1995: 356). Conversely, satemization is taken as an areal shared innovation at chronological level 6 (postdating the split of Aryan-Greek-Armenian and Balto-Slavic-Germanic), even though it is argued to be posterior to the vocalization of syllabic resonants at level 5 (Gamkrelidze & Ivanov 1995: 364). By treating relative chronology so inconsistently as a tool for distinguishing between shared and independent innovations, the model becomes difficult to falsify.

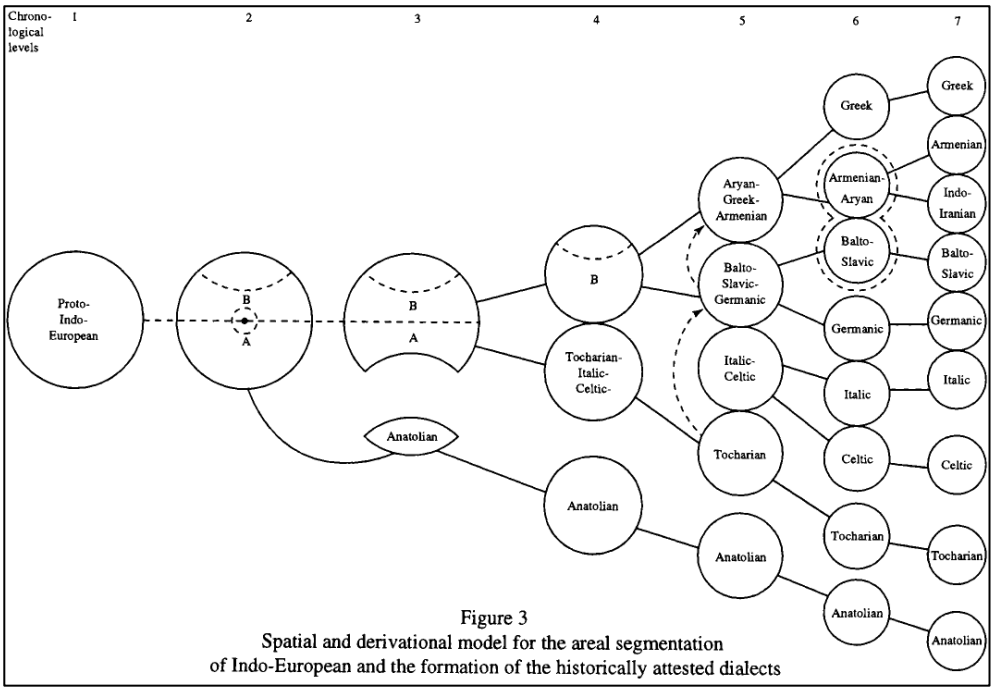

Figure 8. Gamkrelidze & Ivanov's Indo-European areal-genetic model (1995: 363).

When branches share only a small number of innovations, Dyen (1953: 581) argues that this reflects shared dialectal features of the protolanguage, whereas a high number of shared innovations warrants the postulation of a subgroup. Similarly, Clackson (1994: 17) states that "[t]he difference between dialect and sub-group is therefore one of time and degree". This approach is similar to Leskien's (1876) and Meid's (1975) models and is hybrid in the sense that it recognizes both wave-like, dialectal developments and tree-like splits as important factors in language diversification. However, if the difference between wave-like dialect groups and tree-like subgroups were only one of degree, I see no reason why both types could not be represented in the tree model. Schleicher (1861) intended the length of a branch in his tree to reflect its longevity and by implication the number of shared innovations that the subgroup is based on. Thus, rather than a difference in degree, the fundamental difference between the models lies, as we have seen, in whether or not they accommodate overlapping shared innovations.

Ross (1997) developed a hybrid model that distinguishes "language fission" (= tree-like splits), cf. Figure 9, from "linkage breaking" (= breaking of a dialect continuum). The latter is preceded by "lectal differentiation", i.e., a period when overlapping innovations develop in a dialect continuum, cf. Figure 10. Both processes, termed "speech community events", are based on shared innovations, but only linkage breaking involves overlapping innovations. Linkage breaking is of course very similar to Schmidt's (1872) original formulation of the wave model: gradually increasing distance between dialects in a continuum eventually leads to sharp language boundaries. However, an important

difference – besides the fact that tree-like splits, or fissures, are also included in Ross' model – is that the emergence and eventual differentiation of dialect continua, or linkages, are recursive, i.e., these processes may happen more than once in the history of a language family. As Ringe & Eska (2013: 262) point out, linkage breakings may be caused by the disappearance of intermediate dialects. Such dialect pruning may occur when neighbouring dialects are "pulled" apart toward different centres of innovation, as in Schmidt's (1872) wave model, or simply because the speakers of intermediate dialects undergo a language shift or die out.

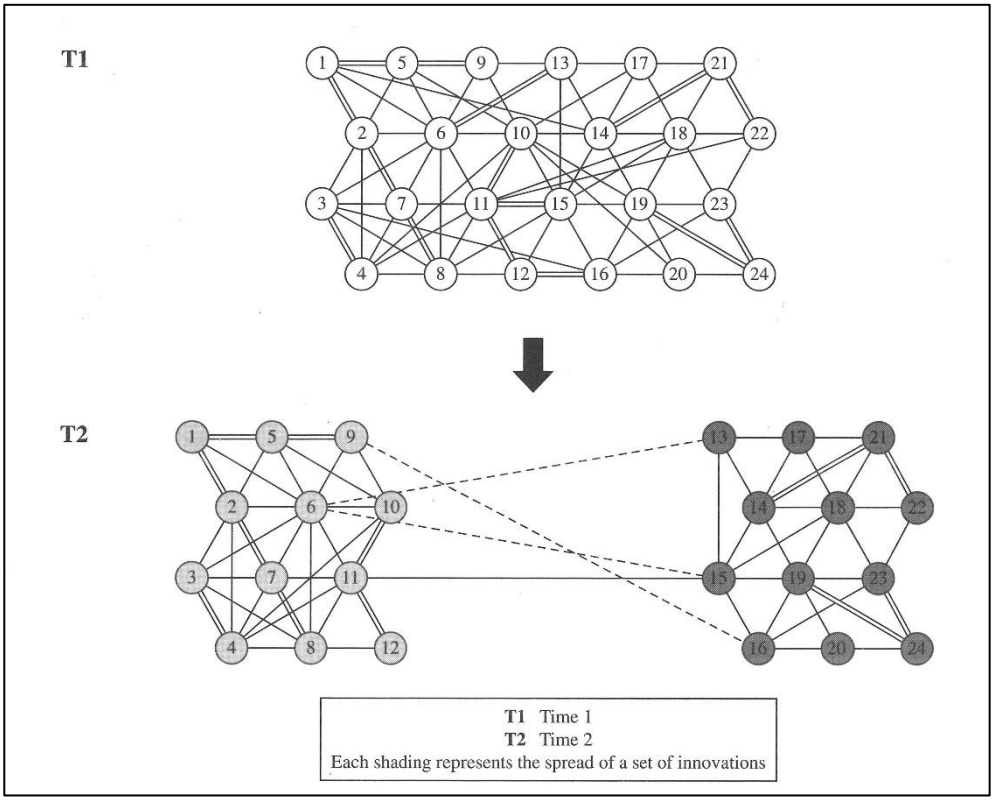

Figure 9. Language fissure according to Ross (1997).

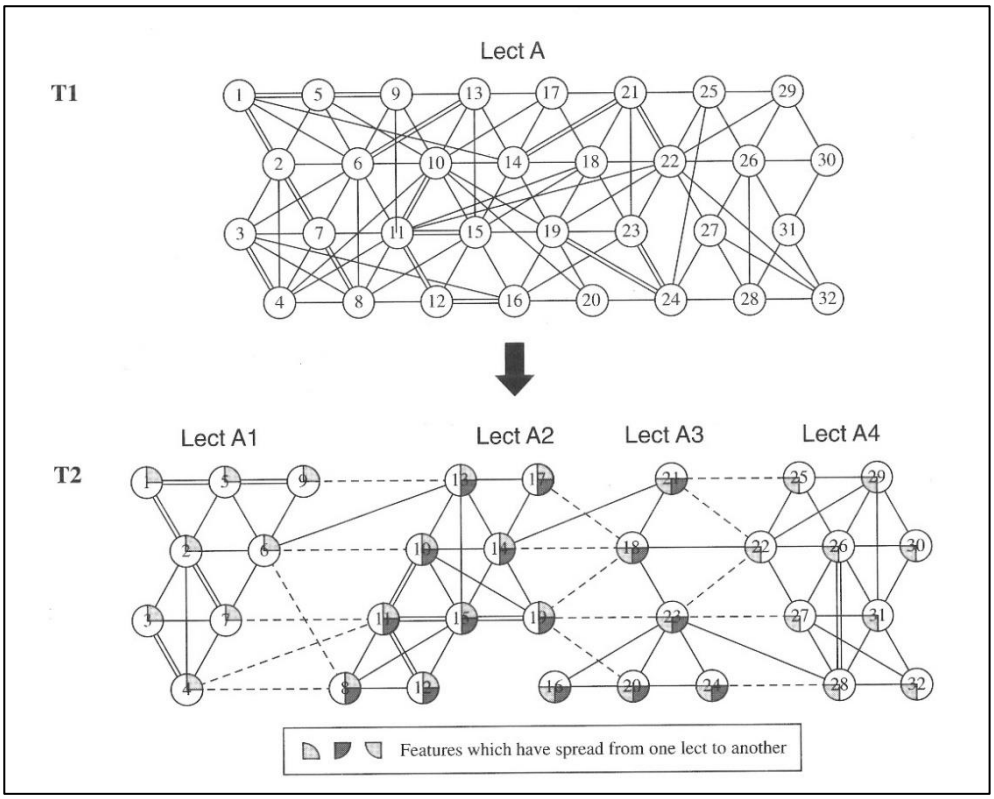

Figure 10. Lectal differentiation according to Ross (1997).

Ross (1997: 228) furthermore includes "language fusion" and "linkage rejoining" in his model, which refer to shared innovations affecting languages or dialects that have already undergone exclusive innovations. Such processes presuppose that some level of mutual intelligibility remains between the dialects in question, although this is notoriously difficult to measure, especially for reconstructed protolanguages. This resembles Gamkrelidze & Ivanov's (1995: 364) treatment of the satem branches, and suffers from the same theoretical problem: lack of falsifiability. However, Ross's example from Anejom̃ (Austronesian) is instructive. Anejom̃ shows two reflexes of the Proto-Oceanic article **na*, which seem to be lexically distributed without phonological or other conditioning. This situation may be explained by assuming that Anejom̃ is a fusion of two separate dialects.[32]

Yet, the Anejom̃ example is not comparable to satemization, since here the satem branches show internally consistent, but slightly variable, reflexes of the Proto-Indo-European velar series (cf. 1.3). It has been argued that the limited velar reflexes of Proto-Indo-European palatovelars in Balto-Slavic point to borrowings from a centum-dialect, which would also be compatible with Balto-Slavic originating in a fusion of two Indo-

[32] It may be argued that the conditioning factor simply has not been found yet, but assuming that the analysis is correct, the postulation of a fusion event is a possible explanation.

European dialects (Čekman 1974: 133), but these alleged centum forms are rather conditionally depalatalized palatovelars (cf. Kortlandt 1978b) that do not justify such a scenario. In any case, language fusion and linkage rejoining are processes that must be taken into account in subgrouping research, but perilous to incorporate into models of Indo-European diversification. Given the language family's time depth, it is difficult to estimate at which point the would-be branches of Indo-European were different enough that shared innovation would no longer be possible. Accordingly, language fusion or linkage rejoining should only be assumed when tree-like splits and linkage breakings fail to explain the data, i.e., as a tertiary hypothesis.

2.5. Linguistic palaeontology

Linguistic palaeontology (a term coined by Pictet 1859–1863) is an extension of the comparative method. While protolanguages may be reconstructed based on the comparative method, the basic idea of linguistic palaeontology is to infer aspects of the culture of the speakers of a certain protolanguage based on the contents of the reconstructed lexicon. Consequently, the focus lies on the semantics of reconstructed words. In most cases, the goal is to compare the culture of the protolanguage community to archaeological cultures, in order to formulate hypotheses on the timeframe and location of the protolanguage and its speakers, known as the *homeland* (Ger. *Urheimat*).

The first systematic application of linguistic palaeontology can probably be attributed to Kuhn (1845). Kuhn compiled reconstructed Proto-Indo-European words pertaining to the structure of the family (kinship terms), government, domesticated as well as wild animals, agriculture, and housing. Based on this, he argued that Proto-Indo-European speakers organized their society on the model of the family, which was patriarchal and patrilocal. Furthermore, Kuhn concluded that the Indo-Europeans practiced both animal husbandry and agriculture, and that they had transitioned from a nomadic to a sedentary lifestyle. However, he does not use these results to infer an Indo-European homeland; rather, Kuhn (1845: 1–2) asserts in his introduction that Proto-Indo-European was spoken in Asia. In a revised version, Kuhn (1850: 338) specifically rejects a steppe homeland, and instead places it on the Tibetan plateau.[33] Similarly, Pictet (1859–1863), who applied linguistic palaeontology to the Proto-Indo-European lexicon in great detail, nevertheless takes a Bactrian homeland as a given, and interprets the lexicon against this assumption.[34]

Schrader (1883) reversed this workflow, and (in the second, revised edition of his *magnum opus*, published in 1890) located the Indo-European homeland in the Pontic-Caspian steppe based on linguistic palaeontological considerations, thus formulating the Steppe hypothesis. The most important arguments were the existence of Proto-Indo-

[33] "[U]nsre gemeinsamen Vorfahren weideten ihre Heerden nicht in den kahlen Steppen, sondern auf den bewaldeten Bergen Hochasiens [Tibetan plateau, AP]…" (Kuhn 1850: 338).
[34] For example, Pictet (1859–1863: I, p. 382–86) argues that the Indo-Europeans must have known the camel, despite the fact that no Proto-Indo-European word for 'camel' can be reconstructed, since the two-humped camel originates in Bactria, which according to Pictet is the location of the homeland.

European nomadic pastoralist vocabulary (i.e., domesticated animals and wagon terminology) vs. the absence of shared Indo-European agricultural terminology, as well as the presence of a word for 'horse' vs. the absence of words for 'camel' and 'donkey'.[35] These arguments have stood the test of time and still feature in studies on the Indo-European homeland (e.g., Mallory 1989; Anthony 2007; Anthony & Ringe 2015). Apart from arguments bearing on technologies and subsistence strategies, recent scholarship has also revisited the relationship between kinship terms as evidence for social organization (cf. Kuhn 1845) and archaeological and genetic evidence for patrilocal families in Chalcolithic Europe (Sjögren et al. 2020), which show a striking correspondence.

Schrader (1883) discusses three important methodological principles of linguistic palaeontology. The first principle (Schrader 1883: 168ff) is that conclusions should in the first instance be based on positive rather than negative evidence; in other words, *absence of evidence is not evidence of absence*. The failure to reconstruct a shared Indo-European word for a particular concept does not imply that Proto-Indo-European speakers were unfamiliar with that concept. If, however, a whole semantic field is absent from the protolanguage, and contrasts with abundant positive evidence for another, such as the lack of terms for relatives on the wife's side of the family vs. the rich terminology for relatives on the husband's side, Schrader argues that this may be significant and therefore tentatively may be used as evidence of absence. Mallory (2021: 281) argues that arguments based on absence should not be rejected by default, since they at least potentially may correlate to real-world situations, and are often used in other historical disciplines, such as archaeology.

Another principle, which follows from the comparative method, is that only words that are actually reconstructable to a particular linguistic layer, e.g., Proto-Indo-European, may be used as evidence for reconstructing the culture of that speech community (Schrader 1883: 175ff). For example, words that are only found in one or a few branches should not be back-projected to the protolanguage without good arguments. This ties into the general discussion of which criteria an etymon needs to fulfil to be reconstructed to Proto-Indo-European (cf. 2.2.5 above). Similarly, root cognates with different derivatives in various branches do not license the reconstruction of the derived concepts to the protolanguage, if the root in question is still productive in the branches (Schrader 1883: 188ff). Great care must be taken to understand the derivational history of words within the branches, to exclude independent innovations from consideration. A related issue is loanwords between branches, which must be filtered out to avoid projecting them back to the protolanguage (Schrader 1883: 201ff).[36] However, Schrader argues that it may in some cases be impossible to differentiate between an inherited formation and an early borrowing between Indo-European dialects that had not yet diverged phonologically to the extent that a borrowing could be identified (similarly Hehn 1877: 487–88).

---

[35] As Schrader (1890) acknowledges in his preface, Benfey (1875) had already proposed a Pontic-Caspian steppe homeland, but merely as a comment on Hehn's (1873: 16–17) Asian homeland, which according to Benfey did not explain the existence of a Proto-Indo-European word for 'salt'. Tomaschek (1878: 862; 1883: 706) supported Benfey's hypothesis based on the language contact between early Indo-European and Uralic languages.

[36] On the other hand, the identification of early borrowings may point to the adoption of a foreign or novel concept, whereas language internal derivation (e.g., PIE *$k^{w}ek^{w}lo$- 'wheel') may suggest a "native" technological development.

A third principle (Schrader 1883: 194ff) is that attested semantics should not be back-projected to reconstructed forms without good arguments. For example, the reconstructed *$h_1ek^{\acute{}}uo$- 'horse' did not necessarily refer to a domesticated horse (Hehn 1877: 53–54; Renfrew 1989), and so does not by itself prove that Proto-Indo-Europeans kept and/or rode domesticated horses. In general, great care must be taken on a case-by-case basis not to overinterpret the semantic meaning of reconstructed lexemes.

The validity of linguistic palaeontology as a scientific methodology has been called into question (Pulgram 1958: 145; Anttila 1972: 373; Renfrew 1987: 77f; Zimmer 1990: 7; Clackson 2013). Heggarty, one of the most vocal critics of linguistic palaeontology, argues that the chance of parallel derivation and the uncertainty of semantic reconstruction makes it impossible to use reconstructed lexemes as evidence for extra-linguistic situations (Heggarty 2006: 189). While he accepts reconstruction of forms as reliable in principle, Heggarty (2006: 190; 2013: 162; 2014: 607–8; Heggarty et al. 2023: S20–21) argues that semantic reconstruction to the level of detail required by linguistic palaeontology is not possible (cf. Krell 1998: 279), since semantic change is not governed by laws in the same way as sound change. For example, Heggarty claims that there is no way to determine with any degree of certainty that PIE *k^wek^wlo- referred to the wheel of a wagon rather than some other circular object, with independent semantic shifts to 'wheel' in Indo-Iranian, Greek, Germanic, and Tocharian.

However, as we have seen above, the fact that both formal and semantic reconstruction suffers from a certain level of uncertainty has long been recognized by proponents of linguistic palaeontology; in any historical discipline, interpretations of the data are ultimately statements of likelihood. As Mallory (2021: 280) puts it, a reconstructed semantic meaning is an "inference to the best explanation […] for the current senses of a set of cognates". In the case of PIE *k^wek^wlo-, other than the fact that its descendants refer to the 'wheel (of a vehicle)',[37] meaning that assuming independent semantic shifts is uneconomical, we may point to the typological tendency of the directionality *concrete* > *abstract*, which makes it less likely that Gr. κύκλος 'circle, ring, wheel' preserves an original meaning 'circle', to which 'wheel' is secondary. Furthermore, PIE *k^wek^wlo- seems to be derived from *$k^wel(H)$- 'to move around, roam', suggesting that its original meaning was associated with transportation rather than simply rotation or a round shape. Coleman's (1988: 450) assertion, cited by Heggarty (2014: 608), that *k^wek^wlo- cannot be reconstructed for Proto-Indo-European because it is only attested in four branches, fails to take into account that the reduplicated stem of *k^wek^wlo- is not a productive derivational type in any of the relevant branches (or indeed any branch of Indo-European). Heggarty's criticism is of course not limited to *k^wek^wlo-, but concerns the methodology as a whole. However, this one example shows that he dismisses linguistic palaeontology without taking all the relevant facts into account.

Heggarty (2013: 163–64; 2018: 169) further criticizes linguistic palaeontology as being subjective, because it has been used as evidence for conflicting homeland hypotheses (cf. Bryant 2001: 123). However, as Mallory (2021: 279) points out, the existence of

[37] Or, in the case of ToA *kukäl*, ToB *kokale* 'chariot, wagon', to the wheeled vehicle itself.

competing hypotheses is the normal situation for all scientific frameworks. Heggarty does not take into account that other variables might be responsible for the varying results achieved by different scholars, such as different data sets, failure to correctly apply the comparative method and the methodological principles of linguistic palaeontology, or reliance on incorrect archaeological models.[38]

Thus, in the present work, linguistic palaeontology will be used following the methodological principles outlined above. The emphasis will lie on positive evidence and semantic fields that plausibly may be correlated to archaeological evidence.

[38] Heggarty (2018: 169) specifically mentions Gamkrelidze & Ivanov's (1995) Armenian hypothesis as opposed to the Steppe hypothesis. However, Gamkrelidze & Ivanov's (1995: 763–767) conclusions are based on a dataset that includes many lexemes that are highly unlikely to be Proto-Indo-European, e.g., **osono-* 'donkey' (based on the irregular correspondence of Gr. ὄνος 'donkey' and Lat. *asinus* 'donkey', cf. de Vaan 2008: 57), as well as controversial archaeological hypotheses, such as the Near Eastern origin of wheel/wagon technology (Schier 2015: 113).