Journal of Historical Linguistics
VOLUME 9 NUMBER 1 2019

ISSN 2210-2116 / E-ISSN 2210-2124
John Benjamins Publishing Company

Journal of Historical Linguistics VOLUME 9 NUMBER 1 2019

Journal of Historical Linguistics

SPECIAL ISSUE
Understanding language genealogy
Alternatives to the tree model

VOLUME 9 NUMBER 1

2019

John Benjamins Publishing Company

Save the trees

Why we need tree models in linguistic reconstruction (and when we should apply them)

Guillaume Jacques and Johann-Mattis List
Centre des Recherches Linguistiques sur l'Asie Orientale | Max Planck Institute for the Science of Human History

Skepticism regarding the tree model has a long tradition in historical linguistics. Although scholars have emphasized that the tree model and its long-standing counterpart, the wave theory, are not necessarily incompatible, the opinion that family trees are unrealistic and should be completely abandoned in the field of historical linguistics has always enjoyed a certain popularity. This skepticism has further increased with the advent of recently proposed techniques for data visualization which seem to confirm that we can study language history without trees. In this article, we show that the concrete arguments that have been brought up in favor of achronistic wave models do not hold. By comparing the phenomenon of incomplete lineage sorting in biology with processes in linguistics, we show that data which do not seem as though they can be explained using trees can indeed be explained without turning to diffusion as an explanation. At the same time, methodological limits in historical reconstruction might easily lead to an overestimation of regularity, which may in turn appear as conflicting patterns when the researcher is trying to reconstruct a coherent phylogeny. We illustrate how, in several instances, trees can benefit language comparison, although we also discuss their shortcomings in modeling mixed languages. While acknowledging that not all aspects of language history are tree-like, and that integrated models which capture both vertical and lateral language relations may depict language history more realistically than trees do, we conclude that all models claiming that vertical language relations can be completely ignored are essentially wrong: either they still tacitly draw upon family trees or they only provide a static display of data and thus fail to model temporal aspects of language history.

Keywords: tree model, Historical Glottometry, phylogenetic networks, incomplete lineage sorting

https://doi.org/10.1075/jhl.17008.mat
Journal of Historical Linguistics 9:1 (2019), pp. 128–166. issn 2210-2116 | e-issn 2210-2124

1. Introduction

All languages develop by descent with modification (Darwin 1859): linguistic material is transferred from generation to generation of speakers, with slight modifications in pronunciation, denotation, and grammar potentially summing up to changes large enough that when two or more linguistic varieties have been separated in some way, be it by geographical or political separation of their speakers, they may become mutually incomprehensible. It is true that not all linguistic material is necessarily inherited from the parent generation. Linguistic material can easily be transferred across linguistic boundaries or diffuse across similar speech varieties. This, however, does not change the fact that the primary process of language transmission is through childhood acquisition of a first language (Ringe, Warnow & Taylor 2002: 61; Hale 2007: 27–48). The fact that languages that are mutually incomprehensible and markedly different can share a common genetic origin was one of the great insights of 19th-century linguistics, and even if lateral forces of diffusion are able to drastically change the shape of languages, this does not invalidate the crucial role that vertical transmission plays in language history. We follow Labov (2007: 347) in strictly distinguishing transmission of language via first language acquisition from diffusion via contact as two distinct processes.

In the following, we will aim to substantiate this viewpoint. We will start from a brief overview of the historical debate between proponents of tree models and proponents of wave models in the history of linguistics (Section 2), then introduce the core arguments of the new debate regarding trees and waves (Section 3). After this, we will defend tree thinking in historical linguistics by showing that even patterns which do not look tree-like at first glance can still be explained using a branching tree model (Section 4.1); we will also show that, conversely, some patterns that appear to demonstrate common inheritance may in fact go back to processes of language contact, which can be readily incorporated into a rooted network model in which a family tree model serves as a backbone representing inheritance with horizontal edges representing borrowing events (Section 4.2).

After presenting several examples that illustrate the benefits of trees in historical language comparison (Section 5) and also pointing to examples of their obvious shortcomings (Section 6), we conclude that both tree- and non-tree-like processes need to be taken into account when trying to draw a realistic scenario of language history. The logical and practical necessity of using both models for treelike and non-tree-like evolution shows that we cannot simply abandon the tree model in historical linguistics, but should rather work on integrating vertical transmission and horizontal diffusion in a common framework.

2. Dendrophobia and dendrophilia in linguistics

In order to get a clearer picture of the major arguments brought up to support or to dismiss the family tree model, it is useful to have a closer look at the origins of the tree model and the discussions that it sparked. In the following, we will give a brief overview on the development of tree thinking ("dendrophilia") and tree skepticism ("dendrophobia") in linguistics, from its beginnings up to the present day.

2.1 Tree thinking in Schleicher's work

Although he was not the first to draw language trees,[1] it was August Schleicher (1821–1866) who popularized tree thinking in linguistics. In two early papers from 1853 (Schleicher 1853a, 1853b) and numerous studies published thereafter (e.g. Schleicher 1861, 1863), Schleicher propagated the idea that the assumptions about language history could be best "illustrated by the image of a branching tree" (Schleicher 1853a: 787).[2] It should be noted that there was no notable influence from Darwin's writings in his work. It is more likely that Schleicher was influenced by stemmatics (for manuscript comparison, see Hoenigswald 1963: 8). Even today, historical linguistics has certain tendencies that resemble tendencies found in the field of stemmatics much more closely than they do in evolutionary biology. It seems that Schleicher's enthusiasm for the drawing of language trees had quite an impact on Ernst Haeckel (1834–1919) (Sutrop 2012), since – as Schleicher himself pointed out (Schleicher 1863: 14) – linguistic trees by then were explicit, drawing on and visualizing concrete data from real-world languages; this contrasts to Darwin's Origins, in which the data were abstract and illustrated hypothetical taxa (Darwin 1859).

Despite his seemingly radical idea to model language history as a process of diversification exclusively via branching and splitting, it is important to note that Schleicher was not a careless proponent of tree thinking. In his work we find many examples that show that he was aware of potential problems resulting from the tree model. In his open letter to Haeckel, Schleicher explicitly pointed to problems of language mixing, using Latin and its descendants as an example; he compared this mixing to plant hybrids in biology, identifying this hybridization as a second factor leading to differentiation (Schleicher 1863: 18). In his earlier work, he explicitly mentioned language contact and borrowing of linguistic features as a process central in language history and development (Schleicher 1861: 6); he also emphasized the importance of distinguishing borrowed traits from inherited traits in language classification (Schleicher 1848: 30). Continuing the analogy between language development and species evolution, Schleicher also pointed to the difficulty in finding sharp borders between languages, dialects, and speech varieties ("Sprache, Dialekt, Mundarten und Untermundarten"); this fact finds a mirror in the difficulty in distinguishing between species and individuals (Schleicher 1863: 21). This last point in particular clearly demonstrates that Schleicher did not think that language splits were exclusively the product of abrupt separations of speakers and that he was aware of the idealizing aspect of the Stammbaum.

1. The first trees and networks depicting language development date back to at least the 17th century; for details, see List et al. (2016), Morrison (2016), and Sutrop (2012).

2. Our translation. Original text: "[Diese Annahmen, logisch folgend aus den Ergebnissen der bisherigen Forschung,] lassen sich am besten unter dem Bilde eines sich verästelnden Baumes anschaulich machen."

2.2 Tree skepticism in the work of Schmidt and Schuchardt

Schleicher's tree thinking, however, did not last long in the world of historical linguistics. By the beginning of the 1870s, Hugo Schuchardt (1842–1927) and Johannes Schmidt (1843–1901) published critical views, claiming that vertical descent was but one aspect of language evolution (Schmidt 1872; Schuchardt 1900). While Schmidt remained very vague in his criticism, Schuchardt was more concrete, pointing in particular to the problem of diffusion between very closely related languages: "We connect the branches and twigs of the family tree with countless horizontal lines and it ceases to be a tree" (Schuchardt 1900: 9).[3]

While Schuchardt's observations were based on his deep knowledge of the Romance languages, Schmidt drew his conclusions from a thorough investigation of shared cognate words in the major branches of Indo-European. In this investigation, he found patterns of words that were in a strong "patchy distribution" (see List et al. 2014) – that is, a distribution that showed many gaps across the languages under investigation, with only a few (if any) patterns that could be found across all languages. One seemingly surprising fact was, for example, that while Greek and Sanskrit shared about 39% of cognate vocabulary (according to Schmidt's count; see Geisler & List 2013) and Greek and Latin shared 53%, Latin and Sanskrit shared only 8%. Assuming that Greek and Latin had a common ancestor, Schmidt asserted, it was very difficult to explain the differences in the degree of vocabulary cognate between Greek and Sanskrit versus the vocabulary cognate between Latin and Sanskrit (Schmidt 1872: 24). Furthermore, this pattern

3. Our translation. Original text: "Wir verbinden die Äste und Zweige des Stammbaums durch zahllose horizontale Linien, und er hört auf ein Stammbaum zu sein."

of patchy distributions seemed to be repeated in all branches of Indo-European that Schmidt investigated. Schmidt thus concluded:

> No matter how we look at it, as long as we stick to the assumption that today's languages originated from their common proto-language via multiple furcation, we will never be able to explain all facts in a scientifically adequate way.[4]
>
> (Schmidt 1872:17)

Schmidt, however, did not stop with this conclusion but proposed another model of language divergence to take the place of the family tree model: "I want to replace [the tree] by the image of a wave that spreads out from the center in concentric circles, becoming weaker and weaker the farther they get away from the center"[5] (Schmidt 1872:27). Since then, this new model, the so-called "wave theory" (*Wellentheorie* in German), has been energetically discussed in articles and books on historical linguistics, sometimes being promoted as the missing complement to Schleicher's *Stammbaumtheorie* (Campbell 1999:187–200, Goebl 1983), sometimes being treated as its more realistic alternative (Gabelentz 1891:194–195). Despite the apparent simplicity of the wave theory as reflected in its succinct presentation in handbooks of historical linguistics, the theory is the center of much confusion, both among linguists and among those without training in historical linguistics. This confusion is reflected not only in the discussions among dendrophilists and dendrophobists but also in the various attempts to visualize the waves. While Schmidt did not give a visualization in his 1872 book, he gave one three years later (Schmidt 1875:199); this is shown in Figure 1 along with an English translation. It is difficult to interpret this figure, due not only to the scan quality (re-rendered from the original here) but also to its structure. It displays languages in a pie chart-like diagram in a quasi-geographic space. No information regarding the ancestral states of the relevant languages is given, and no temporal dynamics are shown. Being quasi-geographic, quasi-quantitative, and quasi-structured, the visualization is hard to understand, and the famous waves themselves are the last thing the figure brings to mind. Schmidt does not seem to ignore that evolution has a time dimension, but he does seem to deliberately neglect it when drawing his waves.

4. Our translation. Original text: "Man mag sich also drehen und wenden wie man will, so lange man an der anschauung fest hält, dass die in historisches Zeit erscheinenden sprachen durch merfache gabelungen aus der Ursprache hervorgegangen seien, d.h. so lange man einen stammbaum der indogermanischen Sprachen annimmt, wird man nie dazu gelangen alle die hier in frage stehenden tatsachen wissenschaftlich zu erklären."

5. Our translation, original text: 'Ich möchte an seine [*des Baumes*] stelle das bild der welle setzen, welche sich in concentrischen mit der entfernung vom mittelpunkte immer schwächer werdenden ringen ausbreitet.'

This confusion is also reflected in the scholarly literature. In the fifty years following Schmidt's publication, there was a wide range of various attempts to visualize the wave theory, ranging from Venn diagrams (Hirt 1905:93) to early networks (Bonfante 1931:174). The only publication known to us that retained Schmidt's pie-chart visualization was Meillet (1908:134), in which the author applied it to Indo-European languages (see Geisler & List 2013 for details on early visualizations of the wave theory). After Schleicher's initial, rather pictorial, tree drawings, language trees began quickly to be schematized in historical linguistics. In contrast, the correct way to draw a wave remains disputed up to today. Some scholars have adopted the influential isogloss-map representation by Bloomfield (1973:316) when they visualize the wave theory (Anttila 1972:305, Burlak & Starostin 2005:153–170, Holzer 1996:13–48). Many scholars, however, still use alternative visualizations (Lehmann 1969[1962]:124) or only mention the wave theory without further illustrations (Hock 1986). Visualization problems cannot be taken as primary arguments against a theory's validity. They may, however, reflect problems of internal coherence, and these problems of internal coherence are already reflected in the above-mentioned early interpretations of the Wellentheorie. It is therefore not surprising that Schmidt's wave theory provoked more negative than positive responses after its publication (Brugmann 1884; Hirt 1905).

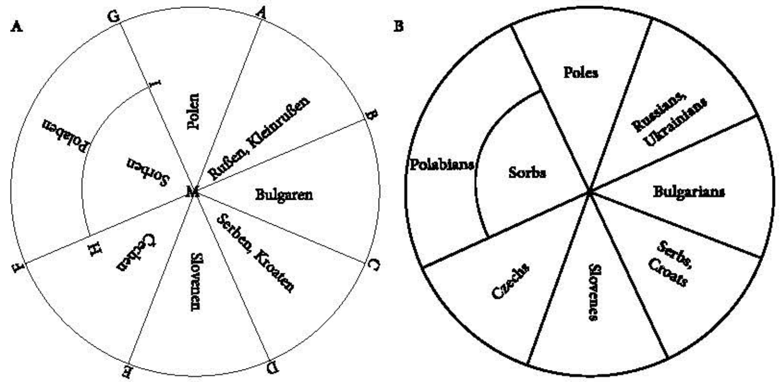

Figure 1. Schmidt's Wave Theory. A: Schmidt's visualization of the Wave Theory from 1875. B: English translation

2.3 Early arguments against the Stammbaumtheorie

Geisler & List (2013:118–120) distinguish three different kinds of criticisms that have been raised against the family tree model (and in favor of the wave theory):

(i) practicability problems, (ii) plausibility problems, and (iii) adequacy problems. "Practicability problems" refers to the problems involved in using the tree model to analyze a given set of languages. Critics such as Schmidt (1872), mentioned above, point out in particular the issue of conflicting evidence as something that the tree model cannot accommodate. "Plausibility problems" refers to the realism of the family tree model, which is reflected in the obvious simplifications that the tree model necessitates. Critics addressing this point emphasize that languages do not necessarily split abruptly but slowly diverge, accompanied by complex waves of diffusion (Schuchardt 1900; Schmidt 1872). Questions of "adequacy" grow out of debates over the purpose of writing language history in historical linguistics. Critics complain that family trees reduce all the vital aspects of language history that are the substance of diversity within a language family to nothing more than the process of vertical descent. A similar argument has been brought up in biology, where the "tree of life" has been labeled the "tree of one percent," due to the fact that only a minimal amount of the data seems to point to vertical descent (Dagan & Martin 2006).

Geisler & List (2013) emphasize that while all three types of criticism have been brought up against the family tree model, it is clear that their theoretical strength differs. Rejecting a model for reasons of practicability is straightforward, but this argument cannot be used to prove that a model is wrong or inadequate. An inability to find evidence for a tree in a given dataset is no proof that the family tree model is wrong, just like how an inability to distinguish borrowed from inherited traits (especially in deeper time depths) cannot be considered proof against the existence of tree-like divergence of languages. Geisler & List (2013) conclude that the stronger arguments against the family tree model are those that challenge its plausibility (particularly those that discuss the presumed split-process by which languages diverge) or its adequacy (particularly those that discuss its inability to provide a full picture of language history in all its complexity).

Putting adequacy to the side, the distinction between practicability and plausibility can be reframed as a distinction between the methodological and theoretical problems with the Stammbaumtheorie. Debates over practicability question the methodological possibility of inferring language trees from linguistic data (in essence, questioning the power of the methods available to us), while debates over plausibility question the adequacy of the model itself. While Schmidt's arguments were largely methodological in nature, pointing to conflicts in the data (which are mostly based on a misunderstanding of the nature of scientific inference and phylogenetic reconstruction, as pointed out in detail in Geisler & List 2013), Schuchardt's arguments are theoretical. He questions the process of divergence itself, claiming that languages do not split in an abrupt, binary fashion; rather, they slowly diverge, while at the same time exchanging material in a non-vertical manner.

An even greater problem, though one not addressed in Schuchardt (1900), is the possibility of convergence. When convergence leads to hybridization, it clearly contradicts the core of the tree model (see Section 6), as rooted trees can be mathematically defined as directed, acyclic graphs in which all nodes have no more than one parent. Interestingly, Schleicher was well aware of these problematic theoretical aspects of the tree model. He explicitly pointed to the possibility of hybridization (Schleicher 1848), and he emphasized the often gradient transition underlying language divergence (Schleicher 1863: 21). On the other hand, he also deliberately ignored these aspects in the family tree model, giving a strict preference to divergence and vertical inheritance.[6]

Proponents of the wave theory, on the other hand, were much less clear about the different processes they sought to model. Do wave-like processes of language change reflect borrowing among closely related languages, or are they intended to reflect language change in general? While Schuchardt (1900) seems to distinguish the two, pointing to horizontal lines ("horizontale Linien") that make a network out of a tree, Schmidt (1872) is much less explicit, although he often invokes the idea of gradual transitions between language borders (Schmidt 1875: 200), thus emphasizing the gradualness of diversification rather than the interference of vertical and lateral processes in language change. Given the diversity of opinions and the lack of concreteness, it is difficult to determine a core theory to which scholars refer when mentioning the wave theory; while some see the wave theory as the horizontal counterpart of the family tree (Baxter 2006: 74), others see the wave theory as a theory explaining linguistic divergence (Campbell 1999: 188–191).

3. The new debate on trees and waves

Accompanying the "quantitative turn" in historical linguistics in the beginning of the 21st century (List 2014: 209–210), the debate over trees and waves has been recently revived. While most textbooks had previously treated the two models as working together to provide a complementary view of external language history[7]

6. Yet he may have tried to visualize genetic closeness independently of elapsed time since separation, as can be seen from the tree in Schleicher (1861: 7), where he notes that the length of the lines indicated the divergence time, while the distance between the lines the degree of genetic closeness ("Die länge der linien deutet die zeitdauer an, die entfernung derselben von einander den verwantschaftsgrad"). This can be interpreted in such a way that Schleicher tried to include potential convergence after separation into his trees.

7. External language history is here used in the sense of Gabelentz (1891: 179–290), who distinguishes it from internal language history, which points to different stages of one and the same language.

(Lehmann 1992; Anttila 1972) or as models of two different aspects of language change (Campbell 1999), more and more linguists now discuss the models as two opposing perspectives on language change (Heggarty, Maguire & McMahon 2010; François 2014). One reason for the revival of the discussion can certainly be found in the prevalence of trees in recent phylogenetic studies in historical linguistics (Gray & Atkinson 2003; Atkinson & Gray 2006; Ringe, Warnow & Taylor 2002; Pagel 2009). Both trees and waves had been playing a less prominent role in the field for some time prior to the "quantitative turn";[8] however, biological methods for phylogenetic reconstruction applied to large linguistic datasets can utilize a tree diagram to analyze and display large amounts of language data in a way that is much more transparent than the classical method of identifying shared innovations, leading to a recent increase in the use of tree diagrams.

Yet not long after the first biological software packages began being used for phylogenetic tree reconstruction in linguistics, new techniques for visualizing splits networks[9] provided by the SplitsTree software package (Huson 1998) offered scholars a fresh view on conflicts in their data. Often propagated as a reconciliation of tree and wave theory (Bryant, Filimon & Gray 2005; Ben Hamed & Wang 2006; McMahon & McMahon 2005) and easy to apply to linguistic distance data, splits networks have quickly become a very popular tool in historical linguistics (Gray, Bryant & Greenhill 2010; Heggarty, Maguire & McMahon 2010; Ben Hamed 2005; Bowern 2010).

3.1 Phylogenetic tree reconstruction after the quantitative turn

Classical phylogenetic tree reconstruction in historical linguistics is very similar to the process of cladistics in biology (Hennig 1950; see also Lass 1997: 105–171), insofar as it makes use of a small set of characters which are inherently weighted and represent unique innovations in order to uncover the phylogeny of a language family. The idea of unique innovations – changes that define a subgroup – is very old in linguistics and can be found in work as early as that of Karl Brugmann (1849–1919), although it was later scholars such as Isidore Dyen who popularized the principle in historical linguistics (see Chrétien 1963 and Dyen 1953). Brugmann himself justified the use of shared innovations in subgrouping as follows:

8. Even Morris Swadesh was extremely wary of using his lexicostatistic method for producing family trees. Instead, he published a map on "interrelationships of American Indian languages" that came closer to the wave theory in its interpretation (Swadesh 1959: 23).

9. Most of these techniques have been based on the NeighborNet algorithm (Bryant & Moulton 2004), but see Hurles (2003) for the earliest example of splits networks in linguistics known to us.

> The only thing that can shed light on the relation among the individual language branches […] are the specific correspondences between two or more of them, the innovations, by which each time certain language branches have advanced in comparison with other branches in their development.[10]
>
> (Brugmann 1967[1886]: 24)

The reason why linguists put such a great emphasis on shared innovations in subgrouping is obvious: while related languages can easily share features they have retained from their common ancestor, features which separate them from other languages in the same family and which may be interpreted as a new development can provide a strong argument for subgrouping. The problem, however, which is often downplayed in this context, is how to *identify* these exclusively shared innovations. If languages share common features (apomorphies in cladistic terminology), this does not necessarily mean that these features qualify as innovations (synapomorphies), since they could likewise have (i) been borrowed (see Section 4.2.1), (ii) been retained from the common ancestor of all languages (symplesiomorphies), (iii) emerged independently (homoplasies), or (iv) been erroneously annotated as shared features. Furthermore, differential loss or further development of features in subgroups may easily mask shared innovations, and consequently an innovation that was originally shared by a group of languages may give the impression of being patchily distributed. This is further complicated by the fact that variation of linguistic features occurs in all languages and might very well be traced back to the ancestral language. If this is the case and variation is later resolved randomly across the lineages, what looks like a shared innovation is in fact a shared retention or an independent development, a combination of (ii) and (iii), a possibility that will be further discussed in Section 4.1. None of these problems are new to historical linguistics: we can already find all of these points apart from the problem of variation in the proto-language in Brugmann (1884), who concludes that proposed innovations must be *frequent* enough to reduce the possibility of chance in order to be applicable to subgrouping (see also Dyen 1953).

It is difficult to give a concrete definition of frequency in historical linguistics, as scholars often intuitively weight characters, assigning more importance to certain kinds of evidence (e.g., form similarities in morphological paradigms; see Nichols 1996) than to other types (e.g., isolated lexical items or frequent sound change patterns which are likely to recur independently), and most debates

10. Our translation. Original text: "Das einzige nun, was auf das Verhältnis der einzelnen Sprachzweige zu einander[, auf die Art des Hervorgangs der Einzelsprachen aus der idg. Ursprache] Licht werfen kann, sind die besonderen Übereinstimmungen zwischen je zwei oder mehreren von ihnen, die Neuerungen, durch die jedesmal gewisse Sprachzweige gegenüber den andern in der Entwicklung vorangeschritten erscheinen."

regarding subgrouping center around the question of how different types of evidence should be weighted or how data should be interpreted. As an example, see the discussion in Sagart (2015), in which the author proposes that the innovations presented in Blust (1999) are better interpreted as retentions.

Phylogenetic approaches that had originally been developed for evolutionary biology offer a different approach to the problem; they use a larger pool of characters and explicit models of character evolution to automatically find the phylogenetic tree that best explains the data according to different criteria (likelihood, parsimony) while simultaneously determining which characters have been retained and which have been innovated (Greenhill & Gray 2012). Classical linguists often mistrust these methods, criticizing their "black box" character.[11] While the criticism is justified to some extent, it should be kept in mind that it is not the methods themselves which are non-transparent or inaccurate, but rather their application and the data they are applied to. Methods for phylogenetic reconstruction, be they based on parsimony, maximum likelihood, or Bayesian inference, are not black box methods per se. As the methods are based on an explicit modelling of evolutionary processes,[12] the tree or forest of trees they infer is based on detailed historical scenarios in which the history of each character in the data is calculated.

3.2 Linguistic data and data-display networks

As mentioned before, splits networks enjoy a considerable popularity in recent quantitative approaches in historical linguistics. Unfortunately, many scholars have failed to understand that splits networks are merely a tool for data display (Morrison 2010) and not a tool that directly produces a phylogenetic analysis. Splits networks are very useful for exploratory data analysis, notably:

> (1) the automatic extraction of previously unknown patterns with regard to groups of objects, without using known structures in the data; (2) the detection of anomalous objects in the dataset; and (3) providing a compact representation of the dataset, which can be easily visualized as a connected graph.
> (Morrison 2014: 2)

11. This "character" is specifically reflected in the fact that multiple steps that lead to a certain conclusion (i.e. in phylogenetic reconstruction) are rarely shown to the users. Rather, users see only the aggregated results.

12. This includes parsimony, since we are not talking about statistical modelling, but process modelling, which is usually a simple birth-death process in parsimony as well as in maximum likelihood and Bayesian inference.

However, they do not produce a hypothesis regarding the way in which languages or biological species have diverged or recombined, and they must be strictly distinguished from explicit evolutionary networks, which display "evolutionary relationships between ancestors and descendants" (Morrison 2011: 43).

The claim that splits networks are not equivalent to phylogenetic trees or phylogenetic networks often leads to confusion among scholars, as it does not seem to be very clear precisely what phylogenetic trees and networks are intended to represent. For us, the crucial difference between data-display approaches and true phylogenetic accounts is the lack of or presence of an explicit time dimension displaying events of divergence (or recombination)[13] of lineages. Whenever we are dealing with attested divergence, as, for example, in the case of mutually unintelligible languages which are obviously genetically related, we are dealing with at least one ancestral variety from which the attested varieties developed. How many further ancestral languages we assume at different stages of the development of the language family depends on the power of our methods, the time depths involved, and language-family-specific factors. However, if we ignore the ancestral varieties in our analysis completely, as is done with splits networks, we lose all temporal dynamics, and as a result end up with nothing more than a representation of the data rather than a concrete hypothesis on the development of the languages under investigation.

3.3 Shared innovations and Historical Glottometry

A very recent approach that has also attempted to dismiss the tree model is the theory of Historical Glottometry (François 2014; Kalyan & François 2018). Glottometry results from dissatisfaction with conflicting data in historical linguistics, dissatisfaction similar to that expressed in Schmidt (1872). Additionally, glottometry follows Ross (1988) in assuming that language divergence can proceed in terms of both concrete separation ("social split" according to François 2014) and dialect divergence. While the former process involves the complete separation of the speakers of a given language, mostly based on geographic dislocation of parts of a population, the latter involves the slow divergence of language varieties into dialect areas which may later result in a complete split and the loss of mutual intelligibility. Essentially, this argument resembles that of Schuchardt (1900), as it attacks the process of concrete language split as it is visually suggested by the tree model. Given the high diffusibility of linguistic features across mutually intelligible varieties, reconstructing a fully resolved tree showing language divergence

13. When dealing with recombination of lineages (like under the assumption of language mixing; see Section 6), a tree model is not enough and a network has to be used.

in split processes may be difficult, if not impossible, in a scenario of language divergence. Ross (1988) uses the term "linkage" to refer to closely related language varieties that diffused rather than separated and uses specifically marked multifurcating nodes (polytomies) to highlight these varieties in his genetic subgrouping of Oceanic languages. Kalyan & François (2018) criticize this solution as unsatisfying, emphasizing that polytomies mask the fact that innovations can easily spread across dialect networks, thus creating intersecting, fuzzy subgroups. The solution proposed by Historical Glottometry is to use the classical comparative method to collect shared traits, intended to represent exclusively shared innovations, for the language family under investigation, then to display these traits as weighted isogloss maps in which weighting is represented by the thickness of a given isogloss. This is illustrated in Figure 2a, in which four hypothetical languages are given that are connected by three isoglosses, out of which two are in conflict with each other.

Three general problems with the method of Historical Glottometry need to be mentioned here. First, the resulting visualizations can by no means qualify as phylogenetic analyses, as they lack any time dimension. They are more similar to data display networks, and the fact that isoglosses are aggregated in numeric weights indicating isogloss strength makes them little more informative than splits networks produced with the NeighborNet algorithm. This does not mean that the measures proposed by glottometry do not have their specific value, but unlike the tree model, which displays a concrete evolutionary hypothesis, glottometric diagrams are mere tools for data visualization, as they do not allow ancestral languages to be included in the analysis.[14]

Second, the use of the term "innovation" in Historical Glottometry is logically problematic. According to the practice reported in François (2014), all instances in which a form in one language deviates in some respect from its reconstructed proto-form are interpreted as innovations. It seems to be further assumed that an innovation starts with its first introduction by a speaker and is diffused during the period of mutual intelligibility (François 2014:178). Parallel innovations – i.e., innovations which look similar but happened independently of each other – are acknowledged as such (François 2016:57), but when it comes to computing the diagrams, they are not distinguished from uniquely shared innovations; François

14. Mathematically, the isogloss model proposed by glottometry corresponds to a hypergraph, in which edges can connect more than one vertex (Newman 2010:122–123). Given that hypergraphs are equivalent to bipartite networks, it also seems that with the existing metrics applied in glottometry, not all mathematical possibilities are exhausted, and instead of weighting isoglosses using the cohesiveness value proposed in (François 2014), it might be interesting to look into different projections of bipartite into monopartite graphs (Newman 2010:124–125).

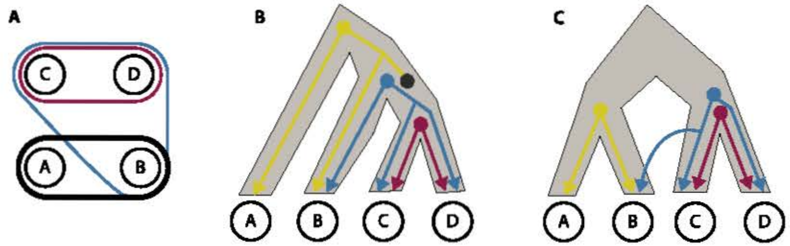

Figure 2. Historical glottometry, the family tree model, and an evolutionary network
Legend:

A: A glottometric diagram with weighted isoglosses drawn between languages sharing an innovation which are in apparent conflict with each other.
B: A demonstration of how the scenario in A can be explained with help of a family tree by assuming differential loss of the black isogloss.
C: An evolutionary network representing another possible explanation for the patterns; assumes that the blue innovation was borrowed from language C into language B.

(2014:177) does not even attempt to distinguish between the two. Thus, innovations in glottometry represent two different processes: namely, (i) cases of unique deviation from linguistic traits inherited from the Ursprache (true shared innovations in the cladistic sense); and (ii) cases of parallel development.

Leaving aside the fact that using a proto-language to identify innovative traits silently acknowledges a tree-like divergence from the beginning, even if it turns out to be a star-phylogeny from which all descendants separated at once, this broad notion of shared innovations in the practice of glottometry bears practical and theoretical problems, especially given that the identified shared and parallel innovations in glottometry are used as an argument against tree-like patterns of separation of ancestral languages. Shared innovations in the cladistic sense are never in conflict with a tree, since they are *defined* as those elements which constitute the tree. They are rigorously distinguished from shared retentions, lateral transfer, and parallel developments (Fleischhauer 2009). Scholars often overlook this, since they interpret the term "shared innovation" as a *descriptive* term when the term in fact is meant to be *explanatory*. When labeling certain features as shared innovations, these scholars seem to provide a mere description of the data, while the term additionally denotes a judgment – an explanation for a certain phenomenon. The descriptive use of explanative terminology can be seen as a general problem in linguistic terminology, as reflected in terms like "pronominalization" (see Jacques 2016:2), "polysemy" (see François 2008; List, Terhalle & Urban 2013), or "assimilation" (see List 2014:32). In all these cases, the terms do not only describe a phenomenon but also explain it. While descriptively, "assimilation" could be seen as a process by which a sound becomes more similar to a neighboring sound, in most definitions scholars further add that this process is due to the *influence* of the neighboring sound (Campbell & Mixco 2007:16). The term is thus not only used to describe a phenomenon, but also to explain it. The same applies to the term "shared innovation." On the one hand, scholars use it to denote a set of similarities shared by a certain group of languages; on the other hand, they use it to denote synapomorphies, namely shared inherited similarities which define a subgroup in the cladistic tree.

The specific confusion involving the term "shared innovation," however, is not restricted to linguistics; it also occurs in biology (De Laet 2005). A cladistic analysis seeks to identify which out of a large pool of possible characters could be used to define a subgroup and thus potentially reflect true shared innovations. If a supposed set of innovations shows internal conflict with possible tree topologies, this means, from a cladistic perspective, that some of these innovations have been wrongly proposed. This is illustrated in Figure 2b, in which the data from Figure 2a are explained by differential loss of a shared character in one clade of a tree. Given that we can often hardly distinguish whether homologous characters in languages are due to independent change or inheritance, a fact which is explicitly admitted by François (2014), conflicts with possible tree topologies can by no means be taken as rigorous proof that a substantial amount of the data cannot be explained by a tree.[15] Interestingly, this was emphasized much earlier in the history of linguistics when Brugmann (1884) criticized the wave theory by Schmidt (1872), because Schmidt had similarly assumed that all exclusively shared traits could have originated only once, ignoring the possibility of erroneous judgments, parallel development, borrowing, shared retention, and chance.

Third, given that Kalyan & François (2018) admit that innovations develop *somewhere*, their approach is by no means less agnostic than the use of multifurcating tree topologies by Ross (1988), as we would assume that an innovation first occurs in a small community from which it spreads outwards. Theoretically, it may thus be possible to draw explicit pathways of diffusion which could be rendered as horizontal edges in an evolutionary network, as illustrated in Figure 2c. Since Historical Glottometry refuses to increase the level of explicitness in data-display, its analyses remain unsatisfactory, as historical linguistics should have more to offer than vague statements regarding shared traits between language varieties.

15. A further problem, which is often ignored, is the possibility of erroneous annotations (see Wichmann 2017 for a more detailed account on false positives and false negatives in cladistic subgrouping).

4. Saving the trees from the critics

Given the logical necessity of allowing for divergence, a specific part of language history can be modeled with the help of a tree if specific processes like recombination (hybridization, creolization; see Section 6) can be excluded. That such a tree model does not necessarily represent all aspects of language history is obvious; even the strongest tree proponents would not deny this. Whether the amount of inheritance versus borrowing in language history is as low as it has been presumed for biology, a field in which tree critics have labeled the tree of life as the "tree of one percent" (Dagan & Martin 2006), is an interesting question worth being pursued further. However, given that we know that language varieties can diverge to such an extent that they lose mutual intelligibility, it is clear that a model for language history which handles divergence and splits of lineages is necessary. How these splits proceed in the end – whether they are best viewed as multifurcations after the split of a larger dialect continuum in several parts or as bifurcations – depends on our insights into the language family under investigation and into the processes of external language change in general.

When scholars point out that a given dataset lacks a tree-like signal, or that the tree-like signal for the subgrouping of a given language family is not strong, they often take this as direct evidence for large-scale language contact or linkage scenarios (Ross 1988). This, however, is by no means the only explanation for reticulations in datasets, and there are many other reasons why a given data selection may fail to reveal a tree (see the general overview in Morrison 2011:44–66). The most obvious, and in cases of large datasets most frequent, reason is erroneous codings, which occur particularly in those cases where the data have not been thoroughly checked by experts in the field (Geisler & List 2010) or where automatic analyses have introduced a strong bias. Another obvious reason for reticulation in a dataset is the selection of the data. Commonalities in sound change patterns and grammatical features, for example, often do not represent true shared innovations, but rather independent development. Additionally, it is often very hard to distinguish between synapomorphy and homoplasy, especially for sound changes (Chacon & List 2015:182–183); this is exacerbated by the fact that the majority of sound change patterns are extremely common, while rare sound changes are often very difficult to prove.

Apart from borrowing, dialect differentiation, data coding, and homoplasy, another often overlooked cause of conflicts in the data is the phenomenon of "incomplete lineage sorting" (Galtier & Daubin 2008). Incomplete lineage sorting is a well-known process in biology, in which polymorphisms (characters which are differently expressed in the same population, e.g., eye color) in the ancestral lineages are inherited by the descendant species when rapid divergence occurs

(Rogers & Gibbs 2014). Incomplete lineage sorting can explain, for instance, why 30% of the genes in a gorilla's genome have more similarity to the human or chimpanzee genome than the same genes in the human and chimpanzee genomes have to each other – surprising, given that human and chimpanzee are the closest relatives (Scally et al. 2012). In a recent study, List et al. (2016) proposed that incomplete lineage sorting may likewise occur in language history, given the multiple sources of polymorphisms in language change, ranging from near synonymy of lexical items via suppletive paradigms to word derivation.

Apart from these polymorphisms which may or may not be inherited across lineages before they are later randomly resolved, a further language-internal factor not mentioned by List et al. (2016) is that of sociolinguistic variation. This variation can occur in an entire population or even within a single speaker. The process of incomplete lineage sorting is further illustrated in Figure 3, where the two aspects – namely, sociolinguistic variation and language-internal variation – are contrasted. Note that in neither of these cases do we need to invoke strong language contact or situations of large-scale diffusion in dialect networks. Both patterns are perfectly compatible with a "social split" situation as invoked by François (2014), although they are based on fully resolved bifurcating trees. This shows that supposed reticulations or a lack of tree-like signal in the data do not necessarily prove the absence of tree-like patterns of divergence. They, rather, expose the weakness of our methods for finding the tree in the forest of individual histories of linguistic traits. In the following sections, we will illustrate this in more detail by showing how variation inherited from an ancestor language may be lost incompletely across lineages and by showing how the failure to identify true innovations may lead us astray when searching for convincing phylogenies.

4.1 Inherited variation and incomplete lineage sorting

Lexically-specific sound changes play an important role in Historical Glottometry, based on the assumption that they are "strongly indicative of genealogy, because they are unlikely to diffuse across separate languages" (François 2014:178). Out of 474 shared traits which are classified as innovations in François (2014), 116 (24%) belong to this type. In view of the low diffusibility of such traits,[16] overlapping isoglosses constitute a major problem for the tree model from the point of view of supporters of glottometry. Regardless of whether lexically-specific sound changes have more difficulty crossing language boundaries than other types of

16. This assertion remains to be demonstrated, but we accept it for the sake of argument.

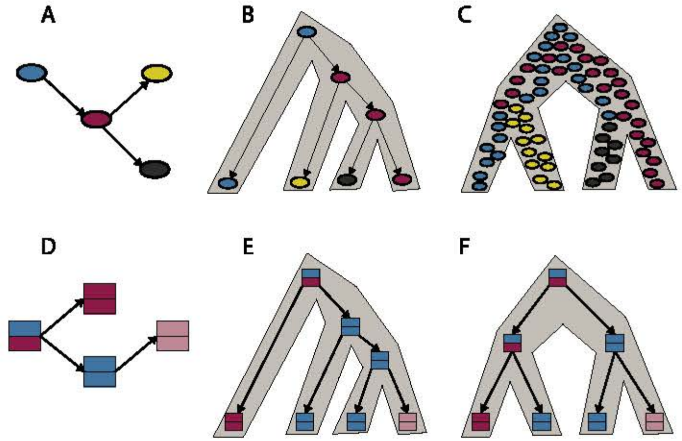

Figure 3. Incomplete lineage sorting due to sociolinguistic (A–C) and linguistic variation (D–F) and its impact on phylogenetic reconstruction and genetic subgrouping

Explanation:
A: Pattern of known directional evolution of a character (e.g., a sound change pattern)
B: A parsimonious tree resulting from the pattern in A
C: Alternative pattern assuming that the blue character already evolved in the ancestral language where it was used as a variant along with the original red character. Since the variation already occurred at the time of the ancestral language, it was inherited in the two descendant languages from which the character further developed. As a result, another tree topology can be reconstructed.
D: Example of a process of paradigm levelling
E: Parsimonious scenario of one tree topology, variant 1
F: Parsimonious scenario of one tree topology, variant 2

features, overlapping innovations can, as mentioned above, also be accounted for by assuming the existence of variation in the proto-language.[17]

Languages are never completely uniform, and fieldwork linguists working on unwritten languages commonly notice that even siblings can present significant differences in the pronunciation of certain words or even morphological paradigms (see, for instance, Genetti 2007:29–30). While some innovations can spread quickly to the entire community (or at least to all members of a specific

17. In this case, however, we can no longer speak of true innovations in the cladistic sense, given that, as mentioned above, the term "innovation" is explanatory and not descriptive and presupposes that a trait is uniquely shared by the subgroup that it defines.

age-group), in other cases it is possible for two competing forms (innovative vs. archaic) to remain used in the same speech community for a considerable period of time. This is observed in particular with sporadic changes, such as irregular metatheses, dissimilation, assimilation, or item-specific analogy.

When language differentiation occurs while forms are still competing, daughter languages can inherit the competing forms; subsequently, the innovative form may eventually prevail or disappear in a non-predictable way in each daughter language. If such a situation occurs, the distribution of the innovation will not directly match a particular node. This phenomenon is better illustrated by analogical levelling than by sporadic sound changes, as in the case of the former, the variation comes from well-understood morphological alternations that have been generalized in different ways in different language varieties, though the same account would be valid for the sporadic changes.

To illustrate how alternations and variation in the proto-language can blur the phylogeny, we take two examples from Germanic: the Proto-Germanic noun *knabō, *knappaz 'boy' and the dative second plural pronoun *izwiz or *iwiz.

4.1.1 *Alternations*

The reflexes of Proto-Germanic *knabō, knappaz 'boy',[18] an n-stem noun whose reflexes in the modern and ancient languages are particularly complex, can be found in Figure 4 (data from Kroonen 2011: 71, 128; Kroonen 2013: 294).

Using the attested ancient and modern forms with the known sound laws applied backwards, no fewer than four proto-forms have to be postulated: *knaban-, *knapan-, *knabban and *knappan-. Some languages have more than one reflex of this etymon (with diverging specialized meanings), and their distribution does not fit any accepted classification of the Germanic languages: for instance, while nearly all Germanicists agree on the existence of an Anglo-Frisian "Ingvaeonic" branch, we see that English sides with either German (in having a reflex of *knaban-) or with Dutch (the Old English reflex of *knapan-, lost in modern English) rather than with Frisian.

Unlike for most other language families, the detailed knowledge that has been accumulated regarding the history of Germanic languages allows us to go further than merely stating the presence of irregular correspondences: it is possible to account for them with a detailed model. It is now near-universally accepted that doublets such as these are due to the effect of Kluge's law (the change from *-Cn- to a geminate voiceless stop in pretonic position, *C being any pre-Germanic stop)

18. The reflexes of this proto-form have developed distinct meanings in the attested languages, including 'squire', but this aspect is not considered here.

on the endings of n-stem nouns in Pre-Proto-Germanic (stage 0) (Kluge 1884; Kroonen 2011).

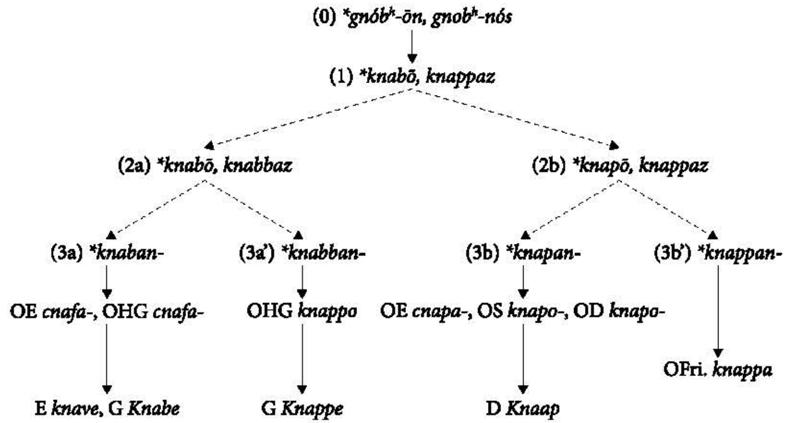

Figure 4. Several layers of variation: The etymon *knabō, knappaz 'boy' in Germanic

The paradigm of the noun 'boy' (as with all nouns of the same type) in Proto-Germanic (stage 1) had an alternation between *-b- and *-pp-. This complex alternation was variously levelled as *-b-/-bb- or *-p-/-pp- by stage 2; note that within a single language, not all items belonging to this declension class underwent levelling in the same way, and that some languages even have competing innovative (OE *cnapa* from *knapan-) and archaic (OE *cnafa* from *knaban-) forms for the same etymon (in this particular case, note that only the archaic form has been preserved with a different meaning in modern English *knave*). After simplification of the *-b-/-pp- alternation, all languages underwent a second wave of analogy, generalizing either the stem of the nominative (archaic *knaban- or innovative *knapan-) or that of the genitive (archaic *knappan- or innovative *knabban-), resulting in the four variants attested throughout the Germanic languages.

4.1.2 *Proto-variation*

Not all types of variations in the proto-language, however, can be straightforwardly accounted for by analogical levelings of paradigms, and in some cases alternative forms may have to be reconstructed back to the proto-language.

Germanic second person pronouns provide an example of this. The accusative and dative of the second person plural pronoun go back to two proto-forms: *izwiz (for Gothic *izwiz* and Old Norse *yðr* by dissimilation from *irwir; see Bugge 1855: 251) and *iwiz (Old English *ēow*, Old High German *iu*).

Some scholars argue that *iwiz is original and that *izwiz is an innovative form resulting from analogical leveling with the first plural accusative/dative *unsiz (through a stage *iwsiz followed by metathesis; see Kroonen 2013:275). Since Gothic is incontrovertibly a member of the first branch of Germanic, this shared isogloss with Old Norse is clearly a problem if this form is a single-event innovation.

However, there is no clear consensus on the origin of these forms. Brugmann (1890:804; see also Streitburg 1900:265) argues instead that *izwiz and *iwiz are both ancient, as Proto-Indo-European had both *wes (Sanskrit second plural accusative-genitive-dative *vas*) and *swes (Welsh *chwi*). A particle *e (Greek *e-kêi* 'there', Sanskrit *a-sau* 'this') was added to both of these alternative forms, resulting in *ewes → *iwiz and *eswes → *izwiz respectively.[19]

Brugmann's idea implies that two proto-forms co-existed in Proto-Germanic for the accusative/dative of the second person plural. This is by no means a cross-linguistically uncommon state of affairs,[20] and this type of situation may account for irregularities in pronominal systems in other parts of the world (cf. François 2016).

4.1.3 *Concluding remarks*

We do not deny the potential value of item-specific changes of this type as evidence for studying phylogeny. However, it is obvious that isoglosses based on item-specific analogical levelling and sporadic sound change will overlap with each other, since competing forms can be maintained within the same language variety and only later be incompletely sorted across different lineages.

4.2 The problem of identifying lexical innovations

In order to identify inherited lexical innovations and distinguish them from recent borrowings, the method of Historical Glottometry uses a fairly uncontroversial criterion: etyma whose reflexes follow regular sound correspondences are considered to be inherited (François 2014:176–178). Thus, whenever a common proto-form can be postulated for a particular set of words across several languages (which can thus be derived from this proto-form by the mechanical application of regular sound changes), it is considered in this model to be part of the inherited vocabulary and can be used, if applicable, as a common innovation.

19. Proto-Indo-European *e shifts to Germanic *i in unaccented syllables.

20. For instance, Japhug has several competing forms for the first and second person pronouns, as well as for the dative postposition, within a single variety and without counting dialectal variations (Jacques 2017:624).

This approach, however, neglects an important factor: while regular sound correspondences are *necessary* for analyzing forms in related languages as cognates, i.e. originating from the same etymon in their common ancestor,[21] they are not *sufficient* due to the existence of undetectable borrowings and nativized loanwords.

4.2.1 *Undetectable borrowings*

Sound changes are not always informative enough to allow the researcher to discriminate between inherited word and borrowing. When a form contains phonemes that remained unchanged or nearly unchanged from the proto-language in all daughter languages (because no sound change, or only trivial changes, affected them), there is no way to know whether it was inherited from the proto-language or whether it was borrowed at a later stage.

This type of situation is by no means exceptional and can be found in various language families. We present here two examples of borrowings undetectable by phonology alone: 'aluminum' in Tibetan languages and 'pig' in some Algonquian languages. Amdo Tibetan *hajaŋ* 'aluminum' and Lhasa *hájã* 'aluminum' look like they regularly originate from a Common Tibetan form *ha.jaŋ.[22] This is, of course, impossible for obvious historical reasons, as aluminum came into use in Tibetan areas in the twentieth century at a time when Amdo Tibetan and Lhasa Tibetan were already mutually unintelligible. This word is generally explained (Gong Xun, personal correspondence) as an abbreviated form of *ha.tɕaŋ jaŋ.po* 'very light', but this etymology is not transparent to native speakers of either Amdo or Lhasa Tibetan. This word was coined only once[23] and was then borrowed into other Tibetan languages[24] and neighboring minority languages under Tibetan influence (as for instance Japhug *χajaŋ* 'aluminum'). In this case, a phonetic borrowing from

21. Note, however, that cognacy is a more complex concept than is usually believed (List 2016) and that even forms originating from exactly the same etymon in the proto-language may present irregular correspondences due to analogy.

22. In Amdo Tibetan, Common Tibetan *h-*, *j-*, *-a* and *-aŋ* remain unchanged (Gong 2016). In Lhasa Tibetan, two sound changes relevant to this form occurred: a phonological high tone developed with the initial *h-*, and *-aŋ* became nasalized *ã*.

23. We are not aware of detailed historical research on the history of this particular word, but in any case, it matters little for our demonstration whether it was first coined in Central Tibetan or in Amdo.

24. In some Tibetan languages such as Cone, in which it is found as *hæ jǎː* (Jacques 2014:306), there is clear evidence that the word is borrowed from Amdo Tibetan and is not native (otherwise *yhæ jaː* would have been expected).

Amdo *hajaŋ* could only yield Lhasa *hájã*, since *h-* only occurs in high tone in Lhasa, and since final *-ŋ* has been transphonologized as vowel nasality.[25]

Several Algonquian languages share a word for 'pig' (Fox *koohkooša*, Miami *koohkooša* and Cree *kôhkôs*) that is ultimately of Dutch origin (Goddard 1974; Costa 2013). Hockett (1957: 266) pointed out that these forms must be considered to be loanwords "because of the clearly post-Columbian meaning; but if we did not have the extralinguistic information the agreement in shape (apart from M[enominee]) would lead us to reconstruct a [Proto-Central-Algonquian] prototype." The forms from these three languages could be regularly derived from Proto-Algonquian *koohkooša, a reconstruction identical to the attested Fox and Miami forms.[26]

Undetectable borrowings are also a pervasive phenomenon in Pama-Nyungan, where, with a few exceptions such as the Arandic and Paman groups, most languages present too few phonological innovations to allow easy discrimination of loanwords from cognates (Koch 2004: 46).

The same situation can be observed even if later sound changes apply to both borrowings and inherited words. Whenever borrowing takes place after the separation of two languages but before any diagnostic sound change has occurred in either the donor or the receiver language, or if the donor and the receiver languages have undergone identical sound changes up to the stage at which the borrowings occurred, phonology alone is not a sufficient criterion to distinguish between inherited words and loanwords.

A classic case is that of Persian borrowings in Armenian. As Hübschmann (1897: 16–17) put it, "in isolated cases, the Iranian and the genuine Armenian forms match each other phonetically, and the question whether borrowing [or common inheritance] has to be assumed must be decided from a non-linguistic point of view."[27] Table 1 presents a non-exhaustive list of such words, with the corresponding Proto-Iranian etyma.

The Armenian case shows that undetectable loans are not restricted to cases like those discussed above, in which a particular word contains only segments which have not been affected by sound changes at any point of the development

25. Likewise, in the case of borrowing from Lhasa into Amdo, the rhyme *-aŋ* would be the only reasonable match for Lhasa *-ã*.

26. However, it is true that, as shown in Taylor 1990, some Algonquian languages have forms that cannot regularly derive from a *koohkooša (for instance Ojibwe has *gookoozh* instead of expected †*gookoozh*), and that the ambiguity between cognate and loanword only exists with Fox, Miami and Cree. We wish to thank an anonymous reviewer for pointing out this fact.

27. Our translation. Original text: "In einzelnen Fällen kann allerdings das persische und echt armenische Wort sich lautlich decken und die Frage, ob Entlehnung anzunehmen ist oder nicht, muss dann nach andern als sprachlichen Gesichtspunkten entschieden werden."

Table 1. Armenian words which cannot be conclusively demonstrated to be either borrowings from Iranian or inherited words on the basis of phonetic analysis

Armenian	Meaning	Proto-Iranian	Reference
naw	boat	*nāu-	Hübschmann (1897: 16–17, 201), Martirosyan (2010: 466, 715)
mēg	mist	*maiga-	Hübschmann (1897: 474), Martirosyan (2010: 466, 715)
mēz	urine	*maiza-	Hübschmann (1897: 474), Martirosyan (2010: 466, 715)
sar	head	*sarah-	Hübschmann (1897: 236, 489), Martirosyan (2010: 571)
ayrem	burn	*Haid-	Hübschmann (1897: 418), Martzloff (2016: 145)

from the proto-language into its daughter languages. Undetectable loans are also possible when a particular word is borrowed before any sound change which could affect its phonetic material occurred in either the giver or recipient language (or if both languages have identical sound changes for words of this particular shape), even if numerous sound changes occurred after borrowing took place. It is possible that post-borrowing sound changes even remove phonetic clues which could have allowed us to distinguish between loanwords and inherited words.

The situations illustrated above can be seen as clear evidence that undetectable borrowings can occur even when two language varieties are mutually unintelligible. Neglecting the distinction between inherited words and undetectable borrowings, as in the approach propagated by adherents of Historical Glottometry, amounts to losing crucial historical information; it does not seem justified to blame the family tree model for a shortcoming in our methods of linguistic reconstruction.

4.2.2 *Nativization of loanwords*

In the previous section, we discussed cases in which borrowing took place prior to diagnostic sound changes, thus making it impossible to effectively use sound changes to distinguish between loanwords and inherited words. There is, however, evidence that even when diagnostic sound changes exist, they may not always be an absolutely reliable criterion.

When a particular language contains a sizeable layer of borrowings from another language, bilingual speakers can develop an intuition of the phonological correspondences between the two languages and apply these correspondences to newly borrowed words, a phenomenon known as "loan nativization."

The best documented case of loan nativization occurs between Saami and Finnish (the following discussion is based on Aikio 2006). Finnish and Saami are only remotely related within the Finno-Ugric branch of Uralic, but Saami has borrowed a considerable quantity of vocabulary from Finnish, with some words being borrowed at a stage before most characteristic sound changes had taken place and others more recently. Table 2 presents examples of cognates between Finnish and Northern Saami that illustrate some recurrent vowel and consonant correspondences.

Table 2. Examples of sound correspondences in inherited words between Finnish and Northern Saami (data from Aikio 2006:27)

Finnish	Northern Saami	Proto-Finno-Ugric	Meaning
käsi	*giehta*	*käti	hand
nimi	*namma*	*nimi	name
kala	*guolli*	*kala	fish
muna	*monni*	*muna	egg

The correspondence of final *-a* to *-i* and final *-i* to *-a* in disyllabic words found in the native vocabulary, as illustrated by the data in Table 2, is also observed in Saami words borrowed from Finnish. This includes recent borrowings, such as *mearka* from *merkki* 'sign, mark' and *báhppa* from *pappi* 'priest' (from Common Slavic *păpъ, itself of Greek origin), even though the sound change from Proto-Uralic to Saami leading to the correspondence *-a* : *-i* had already taken place at the time of contact. These correspondences are pervasive even in the most recent borrowings, to the extent that according to Aikio (2006:36), "examples of phonetically unmarked substitutions of the type F[innish] *-i* > Saa[mi] *-i* and F[innish] *-a* > Saa[mi] *-a* are practically nonexistent, young borrowings included."

In cases such as *báhppa* 'priest', the vowel correspondence in the first syllable *á* : *a* betrays its origin as a loanword, as the expected correspondence for a native word would be *uo* : *a* as in the word 'fish' in Table 2 (Aikio 2006:35 notes that this correspondence is never found in borrowed words).

However, there are cases in which recent loanwords from Finnish in Saami present correspondences indistinguishable from those of the inherited lexicon, as *barta* 'cabin' from Finnish *pirtti*, itself from dialectal Russian *pert'* 'a type of cabin'. These words show the same *CiCi* : *CaCa* vowel correspondence as the word 'name' in Table 2. Here, again, the foreign origin of this word is a clear indication that *barta* 'cabin' cannot have undergone the series of regular sound changes leading from Proto-Finno-Ugric *CiCi to Saami *CaCa*, and that instead the common vowel correspondence *CiCi* : *CaCa* was applied to Finnish *pirtti*.

Loan nativization can also occur between genetically unrelated languages. A clear example is provided by the case of Basque and Spanish (Trask 2000:53–54, Aikio 2006:21–23). A recurrent correspondence between Spanish and Basque is word-final *-ón* to *-oi*. Early Romance *-one (from Latin *-onem*) yields Spanish *-ón*. In Early Romance borrowings into Basque, however, this ending undergoes the regular loss of intervocalic *-n- (a Basque-internal sound change), and yields *-one → *-oe → *-oi*. An example of this correspondence is provided by Spanish *razón* and Basque *arrazoi* 'reason', both from Early Romance **ratsone* (from the Latin accusative form *ratiōnem*). This common correspondence has, however, been recently applied to recent borrowings from Spanish such as *kamioi* 'truck' and *abioi* 'plane' (from *camión* and *avión*). This adaptation has no phonetic motivation, since word-final *-on* is attested in Basque, and can only be accounted for as over-application of the *-oi* : *-ón* correspondence.

Nativization of loanwords is still a poorly investigated phenomenon and can only be detected in language groups whose historical phonology is already very well understood. While it has been documented for few languages as clearly as it has been for Saami and Basque, there is no reason to believe that this phenomenon is rare cross-linguistically. Its existence implies that sound laws cannot be used as an absolute criterion for distinguishing between inherited and borrowed common vocabulary (and thus between true shared innovations and post-innovation borrowings).

5. The benefit of trees in language comparison

In the previous section, we attempted to illustrate that not all patterns that look non-tree-like at first glance require a tree-free explanation, and that, at the same time, patterns that look like excellent examples of exclusively shared innovations may turn out to result from language contact. In addition, trees have several distinct advantages over more complex types of network representation, such as hybridization networks (Morrison 2011:139), which makes the tree model preferable in the absence of evidence of its inapplicability (for more on this, see Section 6).

5.1 Parallel innovations

Trees can be used to detect cases of parallel innovations or features spread through contact. A typical example of such a situation is provided by Semitic. As shown in Table 3, Hebrew and Akkadian share no less than four common innovative sound changes in the evolution of their consonantal systems:

- *θ → *ʃ* (merging with *ʃ)
- *ð → *z* (merging with *z)
- *θ' → *s'* (merging with *s')
- *ɬ' → *s'* (merging with *s')

Table 3. Reflexes of Proto-Semitic coronals in a selected set of Semitic languages (Huehnergard 1997); innovative features shared by Akkadian and other languages are indicated in grey

Proto-Semitic	Akkadian	Hebrew	Biblical Aramaic	Standard Arabic
*t	*t*	*t*	*t*	*t*
*d	*d*	*d*	*d*	*d*
*θ	*ʃ*	*ʃ*	*t*	*θ*
*ð	*z*	*z*	*d*	*ð*
*s	*s*	*s*	*s*	*s*
*z	*z*	*z*	*z*	*z*
*ʃ	*ʃ*	*ʃ*	*ʃ*	*s*
*ɬ	*ʃ*	*ɬ*	*ɬ*	*ʃ*
*t'	*t'*	*t'*	*t'*	*tˤ*
*θ'	*s'*	*s'*	*t'*	*ðˤ*
*s'	*s'*	*s'*	*s'*	*sˤ*
*ɬ'	*s'*	*s'*	*ʕ*	*dˤ*

While phonology could seem, at first glance, to support grouping Akkadian and Hebrew together while excluding Aramaic and Arabic, the bulk of morphological and lexical innovations incontrovertibly support Akkadian being the first branch of the family and Aramaic and Hebrew being closer to each other than either of them is to Arabic (see, for instance, Hetzron 1976; Huehnergard 2006); this subgrouping is summarized in Figure 5. Bayesian phylogenetic analyses that have been proposed for Semitic confirm this insight (see, for instance, Nicholls & Ryder 2011). Here, the tree reconstructed from overwhelming, independently-collected evidence provides us with the near certainty that the innovative features shared by Hebrew and Akkadian are either parallel innovations or isoglosses transmitted through contact, and cannot be common innovations of these two languages.

5.2 Reconstruction of the Ursprache

Trees can be used to determine which features are reconstructible to the Ursprache and which are more likely to be later innovations. To illustrate this specific benefit of family trees, let us take the case of Semitic prepositions. Akkadian differs from the rest of the family in that its spatial prepositions are *in* and *ana*, while the other languages have forms going back to *l- and *b-. Geez (an Ethio-Semitic language, belonging to a sub-branch of West Semitic), however, has a cognate of Akkadian *in* : the preposition *ən*, which appears in some expressions (Huehnergard 2006:16, Kogan 2015:119). Additionally, Akkadian does have a frozen trace of the preposition *b- (Rubin 2005:45–46). Since none of these four prepositions are the result of recent and obvious grammaticalization processes, there is no way without the tree model to decide which should be reconstructed to Proto-Semitic and which should not. Thanks to the Stammbaum in Figure 5, however, we know that since the prepositions *inV and *b- are attested (even as traces) in both Akkadian and West Semitic and are not recently grammaticalized, they can be safely reconstructed to Proto-Semitic.

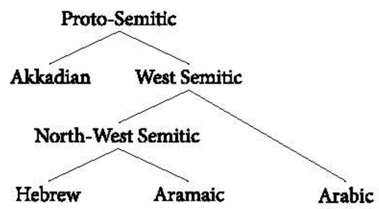

Figure 5. A simplified Stammbaum of Semitic languages

5.3 Directionality of change

As a byproduct of the reconstruction of particular features to the proto-language, trees can be used to determine the directionality of changes in ambiguous cases. While the directionality can sometimes be determined using the body of attested knowledge on sound changes (e.g. Kümmel 2007) or semantic changes (e.g., Urban 2011), there are still many isoglosses, particularly in inflectional morphology, whose interpretation as innovations or retentions is nearly impossible by direct comparison between languages.

As an example of the benefits of trees in determining the directionality of a semantic change, let us examine the root *ʔmr in Semitic (Kogan 2015:233, 331, 544). This root is attested in various languages with a slightly different meaning; Table 4 provides its reflexes in several languages. The meaning of this root is highly divergent across these languages: it is a perception verb ('see', 'look at') in some languages and a verb of speech ('say', 'command') in others. It is not obvious at first glance which of the different meanings was the original one.

The family tree of Semitic, however, provides a scenario of how the meaning of this root evolved across the family. The use of this root as a perception verb is found in both Akkadian and Northwest Semitic (Ugaritic): perception is thus most likely to be the archaic meaning. Ugaritic, in which the root means both 'to look at' and 'say', represented an intermediate stage, where both meanings were still in competition (this may be a preservation of the Proto-West Semitic stage).

In Hebrew and Arabic, the use of this root as a perception verb has disappeared, and Arabic has further narrowed down its meaning to 'to command'.

Table 4. Reflexes of the root *ʔmr in several Semitic languages (Kogan 2015: 233, 331, 544)

Language	Reflex	Meaning
Akkadian	*amārum*	to see
Hebrew	*ʔāmar*	to say, to declare, to command
Ugaritic	*ʔmr*	to say; to look at
Arabic	*ʔamara*	to order

The pathway of semantic change (1) is a possible account of the evolution of the meaning of this root in Semitic, which is compatible with the tree in Figure 5.

(1) 'see, look at' → 'address' → 'say' → 'command'

In this particular case, the tree model does not only help us to solve an ambiguous question in Proto-Semitic reconstruction; it also provides evidence for a semantic change that might otherwise not have been clearly attested.

5.4 Common tendencies of language change

Many processes of linguistic change are overwhelmingly frequent and widespread. However, apart from highly controversial attempts to find a universal constant of lexical replacement rates (Swadesh 1955), most of the knowledge regarding change preferences in language history – be they family-specific, areal, or global – has never been explicitly modelled, since most scholars work from intuition about common tendencies. Language phylogenies and modern phylogenetic approaches, however, allow us to quantify the processes in various ways, and although most currently applied models lack linguistic realism, they offer a promising starting point for future efforts. In addition to intuitive accounts of frequency and cross-linguistic studies, such as the one regarding sound change by Kümmel (2007), there are other promising approaches: for example, phylogenetic approaches, in which the evolution of linguistic characters (phonetic, morphological, semantic) is modelled by inferring how the characters evolved along a given phylogeny, may yield interesting insights into common tendencies of language change. These approaches allow us to process larger amounts of data, but at the same time, they are not able to handle uncertainty in their inferences. Even less sophisticated approaches, such as weighted parsimony, can provide interesting insights into sound change patterns which frequently occur independently of each other along different branches of a tree (Chacon & List 2015). Static models of shared commonalities, like the isogloss-maps of glottometry, do not provide insights into the dynamics and tendencies of common processes of language change.

5.5 Language change and migration history

Trees can be used to make sense of population prehistory and can help to enhance the comparison of linguistic and archaeological evidence. Clues regarding the history and the spread of a language family can be obtained using the reconstructible vocabulary for particular nodes. For instance, the presence of a reconstructible etymon *kasp- 'silver' (Akkadian *kaspum*, Ugaritic *ksp*, and Hebrew *késeφ*, among other languages; Huehnergard 2012: 14–16) suggests that silver smelting could have been known to the speakers of Proto-Semitic, an idea supported by the evidence of cupellation in Syria as early as the 4th millennium BC (Pernicka, Rehren & Schmitt-Strecker 1998). Other metals, however, are only reconstructible to lower branches of the family; for instance 'iron' does not occur earlier than Proto-Canaanite (*barðill-, Hebrew *barzel*; cf. Kogan 2015: 287),[28] an observation compatible with the much later spread of iron technology (Yahalom-Mack & Eliyahu-Behar 2015).

Of course, as shown in Section 4.1, words that are compatible with the sound laws of inherited vocabulary may nevertheless be diffused by contact (especially a form like *kasp-, which remained unchanged in most of the ancient attested languages). As a result, "linguistic paleontology" should always be used with great caution. By using the tree model to advance our knowledge regarding directionality preferences in semantic shift and morphological change, as well as our knowledge regarding the strength of certain tendencies, we may be able to consolidate paleolinguistic evidence and finally put this highly controversial field on more solid ground.

6. The limits of the tree model

While the tree model has undeniable advantages and remains the most powerful model for understanding the vertical history of most languages, there undoubtedly remains a residue of cases in which this model is not applicable, even taking incomplete lineage sorting into account. These cases consist of instances in which one language results from the merger of two previously unintelligible languages (whether or not the two varieties are demonstrably related or not).

28. Similar forms in other languages such as Akkadian *parzillum* 'iron' do not follow the regular correspondences and cannot be cognate.

The clearest and best documented example of this type is Michif, a contact language based on Canadian French and Plains Cree (Bakker 1997). Example (2), taken from Antonov (2015), illustrates the main features of this language (elements from French are in bold and those from Cree are underlined). Nearly all verbs and verbal morphology come from Plains Cree, except the verbs 'to be' and 'have'; these come from French, and retain the source language's complete irregular paradigms for these verbs, including French tense categories, as shown by (3). Most nouns and adjectives come from French. Some determiners are from French (the articles), but the demonstratives are from Cree; nouns can take the Cree obviative suffix -*(w)a*, and some nouns are compatible with possessive prefixes (like *o*- below).

(2) *o-pâpa-wa êtikwenn kî-wîkimê-yiw onhin* ***la***
3-father-OBV apparently PST-marry-3.OBV → 3.PROX this:AN:OBV DEF:FEM:SG
fâm*-a*
woman-OBV
'Her father apparently married that woman…' (1: 8–9)

(3) ***stit=enn pchit orfelin***
BE:3SG:PST=INDEF:F:SG little orphan
'She was a little orphan' (1: 2)

The descent of a language like Michif, and potentially also the descent of less extreme contact languages, cannot be represented by the tree model, as the representation would require two roots (from languages belonging to unrelated families). A more complex type of network, a directed network with multiple roots, would be necessary to represent a language of this type. This might be fruitful, as the near-perfect division of the French and the Cree components of this language might allow for a meaningful representation of the nature of language mixture.

The applicability of the tree model on a global scale crucially depends on the rarity of languages like Michif. If, as the data available to us seem to show, this language is truly exceptional (because its genesis occurred in a very special setting that is unlikely to have existed at earlier stages of history), there are few obstacles against accepting the tree model to represent the vertical descent of languages.

7. Conclusion

In this paper, we have tried to save Schleicher's family tree model from being cut into pieces by critics speaking prematurely. We have shown that Schleicher himself was far more aware of the obvious insufficiencies of his tree model than is usually

acknowledged in the literature, and that the wave theory by Schmidt, which is often praised as the alternative to the tree, never truly reached the level of sophistication necessary to depict the temporal dynamics of language history. After briefly introducing the new stage of the debate between proponents of trees and proponents of waves, we looked at Historical Glottometry, whose supporters stand as some of the strongest opponents of family trees in contemporary historical linguistics. We have shown, however, that Historical Glottometry suffers from the same problems as Schmidt's Wellentheorie, insofar as glottometry lacks temporal dynamics and is not capable of distinguishing true innovations from independently developed shared traits. We further substantiated this claim by illustrating how conflicts in linguistic data, which are taken as prima facie evidence against trees, can often be explained using a traditional family tree model, especially in cases where linguistic variation has been inherited from the ancestor language. On the other hand, we have shown how overlapping isoglosses, which are treated as evidence against tree-like evolution in language history, can likewise be explained by invoking classical processes of language contact. In order to further substantiate the claim that trees are worth being saved, we provided several examples of the usefulness of tree models in linguistic reconstruction, ranging from the detection of parallel innovations up to an enhancement of the methodology underlying linguistic paleography. We are aware that there are situations in language history, like language mixture, where trees cannot be used, but as long as these situations remain exceptional, we do not see any theoretical or practical justification for abandoning the family tree model as the standard to represent vertical aspects of language evolution.

Language history is incredibly complex and, even with more than 200 years of research into it, we have only seen the tip of the huge iceberg of possible processes in language evolution. No linguist would deny that not all aspects of language history are tree-like. Languages can split and branch when their speakers separate, but they do not necessarily do so; even after separation, languages may still easily exchange all kinds of linguistic material. We therefore agree with all tree skeptics that a language tree necessarily reduces linguistic reality, emphasizing only processes of vertical descent. On the other hand, however, we do not agree with the viewpoint that tree drawing per se is useless. Given our knowledge that we can, in theory, clearly distinguish processes of inheritance from processes of borrowing, we should make use of rooted phylogenies which distinguish vertical from lateral processes. While we explicitly acknowledge that integrated models which capture both vertical and lateral language relations may depict language history more realistically,[29] we do not accept the conclusion that vertical language change can be completely ignored. "Treeless" approaches, like Historical Glottom-

29. These would be true evolutionary networks in the sense of Morrison (2011).

etry or splits networks, either silently still use family trees or only provide a static display of data and thus fail to model temporal aspects of language history.

Acknowledgements

This research was supported by the DFG research fellowship grant 261553824 (JML) and the ERC Starting Grant 715618 CALC (JML). We thank Dmitry Nikolaev, Gong Xun, John Cowan, Juho Pystynen, Julien Dufour, Martin Kümmel, Mathieu Ségui, Mikhail Zhivlov, Nathan Hill, Rémy Viredaz, Simon Greenhill, Thiago Chacon, Tiago Tresoldi, and Yoram Meroz for a vivid discussion after we circulated the first draft of the paper. We thank Russell D. Gray for providing additional pointers to missing bibliographic items and Malcom Ross for additional comments to our first draft. We also thank the anonymous reviewers for a challenging discussion.

Abbreviations

F	feminine
INDEF	indefinite
ILS	Incomplete Lineage Sorting
OBV	obviative
PST	past
SG	singular

References

Aikio, Ante. 2006. Etymological Nativization of Loanwords: A Case Study of Saami and Finnish. *Saami Linguistics* ed. by Ida Toivonen & Diane Nelson, 17–52. Amsterdam: Benjamins.

Antonov, Anton. 2015. The Obviative as a Differential Object Marker in Michif. Paper presented at 11th Biennial Conference of the Association for Linguistic Typology, Albuquerque, USA, August 1–3, 2015.

Anttila, Raimo. 1972. *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan.

Atkinson, Quentin D. & Russell D. Gray. 2006. How Old is the Indo-European Language Family? Illumination or More Moths to the Flame? *Phylogenetic Methods and the Prehistory of Languages* ed. by Peter Forster & Colin Renfrew, 91–109. Cambridge/Oxford/Oakville: McDonald Institute for Archaeological Research.

Bakker, Peter. 1997. *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Metis*. Oxford: Oxford University Press.

Baxter, William H. 2006. Mandarin Dialect Phylogeny. *Cahiers de Linguistique – Asie Orientale* 35:1.71–114. https://doi.org/10.3406/clao.2006.1748

Ben Hamed, Mahé. 2005. Neighbour-Nets Portray the Chinese Dialect Continuum and the Linguistic Legacy of China's Demic History. *Proceedings of the Royal Society B* 272:1567.1015–1022. https://doi.org/10.1098/rspb.2004.3015

Ben Hamed, Mahé & Feng Wang. 2006. Stuck in the Forest: Trees, Networks, and Chinese Dialects. *Diachronica* 23:1.29–60. https://doi.org/10.1075/dia.23.1.04ham

Bloomfield, Leonard. 1973. *Language*. London: Allen & Unwin.

Blust, Robert. 1999. Subgrouping, Circularity, and Extinction: Some Issues in Austronesian Comparative Linguistics. *Selected Papers from the Eighth International Conference on Austronesian Linguistics* ed. by Elizabeth Zeitoun & Paul Jen-kuei Li, 31–94. Taipei: Academia Sinica.

Bonfante, Giuliano. 1931. I dialetti indoeuropei. *Annali del R. Istituto Orientale di Napoli* 4.69–185.

Bowern, Claire. 2010. Historical Linguistics in Australia: Trees, Networks and Their Implications. *Proceedings of the Royal Society B* 365:1559.3845–3854.

Brugmann, Karl. 1884. Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen. *Internationale Zeischrift für allgemeine Sprachewissenschaft* 1.228–256.

Brugmann, Karl. 1890. *Grundriß der vergleichenden grammatik der indogermanischen sprachen*, Vol. 2. Strassburg: Karl J. Trübner.

Brugmann, Karl. 1967[1886]. *Einleitung und Lautlehre: vergleichende Laut-, Stammbildungs- und Flexionslehre der indogermanischen Sprachen*. Berlin, Leipzig: Walter de Gruyter.

Bryant, David, Flavia Filimon & Russell D. Gray. 2005. Untangling Our Past: Languages, Trees, Splits and Networks. *The Evolution of Cultural Diversity: A Phylogenetic Approach* ed. by Ruth Mace, Clare J. Holden & Stephen Shennan, 67–84. London: UCL Press.

Bryant, David & Vincent Moulton. 2004. Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular Biology and Evolution* 21:2.255–265. https://doi.org/10.1093/molbev/msh018

Bugge, Sophus. 1855. Die formen der geschlechtslosen persönlichen pronomina in den germanischen sprachen. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 4:4.241–256.

Burlak, Svetlana Anatol'evna & Sergej Anatol'evič Starostin. 2005. *Sravnitel'no-istoričeskoe jazykoznanie*. Moscow: Akademia.

Campbell, Lyle. 1999. *Historical Linguistics: An Introduction (Second Edition)*. Edinburgh: Edinburgh University Press.

Campbell, Lyle & Mauricio Mixco. 2007. *A Glossary of Historical Linguistics*. Edinburgh: Edinburgh University Press.

Chacon, Thiago Costa & Johann-Mattis List. 2015. Improved Computational Models of Sound Change Shed Light on the History of the Tukanoan Languages. *Journal of Language Relationship* 13:3.177–204.

Chrétien, C. Douglas. 1963. Shared Innovations and Subgrouping. *International Journal of American Linguistics* 29:1.66–68. https://doi.org/10.1086/464714

Costa, David J. 2013. Borrowing in Southern Great Lakes Algonquian and the History of Potawatomi. *Anthropological Linguistics* 55:3.195–233. https://doi.org/10.1353/anl.2013.0013

Dagan, Tal & William Martin. 2006. The Tree of One Percent. *Genome Biology* 7:10.118. https://doi.org/10.1186/gb-2006-7-10-118

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

De Laet, Jan E. 2005. Parsimony and the Problem of Inapplicables in Sequence Data. *Parsimony, Phylogeny, and Genomics* ed. by Victor A. Albert, 81–116. Oxford: Oxford University Press.
Dyen, Isidore. 1953. Review of the book *Malgache et maanjan: Une comparaison linguistique* by Otto Chr. Dahl. *Language* 29:4.577–590. https://doi.org/10.2307/409983
Fleischhauer, Jens. 2009. A Phylogenetic Interpretation of the Comparative Method. *Journal of Language Relationship* 2.115–138.
François, Alexandre. 2008. Semantic Maps and the Typology of Colexification: Intertwining Polysemous Networks Across Languages. *From Polysemy to Semantic Change* ed. by Martine Vanhove, 163–215. (= *Studies in Language Companion Series*, 106.) Amsterdam: Benjamins. https://doi.org/10.1075/slcs.106.09fra
François, Alexandre. 2014. Trees, Waves and Linkages: Models of Language Diversification. *The Routledge Handbook of Historical Linguistics* ed. by Claire Bowern & Bethwyn Evans, 161–189. London: Routledge.
François, Alexandre. 2016. The Historical Morphology of Personal Pronouns in Northern Vanuatu. *Faits de Langues* 47.25–60.
Gabelentz, Hans Georg C. 1891. *Die Sprachwissenschaft*. Leipzig: T.O. Weigel.
Galtier, Nicolas & Vincent Daubin. 2008. Dealing With Incongruence in Phylogenomic Analyses. *Proceedings of the Royal Society B* 363:1512.4023–4029. https://doi.org/10.1098/rstb.2008.0144
Geisler, Hans & Johann-Mattis List. 2010. Beautiful Trees on Unstable Ground: Notes on the Data Problem in Lexicostatistics. *Die Ausbreitung des Indogermanischen. Thesen aus Sprachwissenschaft, Archäologie und Genetik* ed. by Heinrich Hettrich. Wiesbaden: Reichert.
Geisler, Hans & Johann-Mattis List. 2013. Do Languages Grow on Trees? The Tree Metaphor in the History of Linguistics. *Classification and Evolution in Biology, Linguistics and the History of Science: Concepts – Methods – Visualization* ed. by Heiner Fangerau, Hans Geisler, Thorsten Halling & William Martin, 111–124. Stuttgart: Franz Steiner Verlag.
Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110198812
Goddard, Ives. 1974. Dutch Loanwords in Delaware. *A Delaware Indian Symposium* ed. by Herbert C. Kraft, 153–160. Harrisburg: Pennsylvania Historical and Museum Commission.
Goebl, Hans. 1983. Stammbaum und Welle. *Zeitschrift für Sprachwissenschaft* 2:1.3–44.
Gong, Xun. 2016. A Phonological History of Amdo Tibetan Rhymes. *Bulletin of the School of Oriental and African Studies* 79:2.347–374. https://doi.org/10.1017/S0041977X16000070
Gray, Russell D., David Bryant & Simon J. Greenhill. 2010. On The Shape and Fabric of Human History. *Philosophical Transactions of the Royal Society B* 365:1559.3923–3933. https://doi.org/10.1098/rstb.2010.0162
Gray, Russell D. & Quentin D. Atkinson. 2003. Language-tree Divergence Times Support the Anatolian Theory of Indo-European Origin. *Nature* 426:6965.435–439. https://doi.org/10.1038/nature02029
Greenhill, Simon J. & Russell D. Gray. 2012. Basic Vocabulary and Bayesian Phylolinguistics. *Diachronica* 29:4.523–537. https://doi.org/10.1075/dia.29.4.05gre
Hale, Mark. 2007. *Historical Linguistics: Theory and Method*. Malden/Oxford/Victoria: Blackwell.

Heggarty, Paul, Warren Maguire & April McMahon. 2010. Splits or Waves? Trees or Webs? How Divergence Measures and Network Analysis Can Unravel Language Histories. *Philosophical Transactions of the Royal Society B* 365:1559.3829–3843. https://doi.org/10.1098/rstb.2010.0099
Hennig, Willi. 1950. *Grundzüge einer Theorie der phylogenetischen Systematik*. Berlin: Deutscher Zentralverlag.
Hetzron, Robert. 1976. Two Principles of Genetic Reconstruction. *Lingua* 38:2.89–108.
Hirt, Herman. 1905. *Die Indogermanen*, Vol. 1. Strassburg: Trübner.
Hock, Hans Henrich. 1986. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
Hockett, Charles F. 1957. Central Algonquian Vocabulary: Stems in k-. *International Journal of American Linguistics* 23:4.247–268. https://doi.org/10.1086/464417
Hoenigswald, Henry M. 1963. On the History of the Comparative Method. *Anthropological Linguistics* 5:1.1–11.
Holzer, Georg. 1996. *Das Erschließen unbelegter Sprachen*. Frankfurt am Main: Lang.
Huehnergard, John. 1997. *A Grammar of Akkadian*. Atlanta: Scholars Press.
Huehnergard, John. 2006. Proto-Semitic and Proto-Akkadian. *The Akkadian Language in its Semitic Context: Studies in the Akkadian of the Third and Second Millennium BC* ed. by G. Deutscher & N.J.C. Kouwenberg, 1–18. Leiden: Nederlands Instituut voor het Nabije Oosten.
Huehnergard, John. 2012. *An Introduction to Ugaritic*. Peabody: Hendrickson Publishers.
Hurles, Matthew E., Elizabeth Matisoo-Smith, Russell D. Gray & David Penny. 2003. Untangling Oceanic Settlement: The Edge of the Knowable. *Trends in Ecology and Evolution* 10:10.531–540. https://doi.org/10.1016/S0169-5347(03)00245-3
Huson, Daniel H. 1998. SplitsTree: Analyzing and Visualizing Evolutionary Data. *Bioinformatics* 14:1.68–73. https://doi.org/10.1093/bioinformatics/14.1.68
Hübschmann, Heinrich. 1897. *Armenische Grammatik, 1 Theil: Armenische Etymologie* Leipzig: Breitkopf und Härtel.
Jacques, Guillaume. 2014. Cone. *Phonological Profiles of Little-studied Tibetic Varieties* ed. by Jackson T.-S. Sun, 265–371. Taipei: Academia Sinica.
Jacques, Guillaume. 2016. Tangut, Gyalrongic, Kiranti and the Nature of Person Indexation in Sino-Tibetan/Trans-Himalayan. *Linguistics Vanguard* 2:1.1–13. https://doi.org/10.1515/lingvan-2015-0033
Jacques, Guillaume. 2017. Japhug. *The Sino-Tibetan Languages (Second Edition)* ed. by Graham Thurgood & Randy LaPolla, 614–634. London: Routledge.
Kalyan, Siva & Alexandre François. 2018. Freeing the Comparative Method from the tree model: A framework for Historical Glottometry. In Ritsuko Kikusawa & Lawrence Reid (eds), *Let's talk about trees: Genetic Relationships of Languages and Their Phylogenic Representation (Senri Ethnological Studies, 98)*. Ōsaka: National Museum of Ethnology. 59–89.
Kluge, F. 1884. Die germanische Consonantendehnung. *Beitrage der deutschen Sprache und Literatur* 9.149–186.
Koch, Harold. 2004. A Methodological History of Australian Linguistic Classification. *Australian Languages: Classification and the Comparative Method* ed. by Claire Bowern & Harold Koch, 17–60. Amsterdam: Benjamins. https://doi.org/10.1075/cilt.249.06koc
Kogan, Leonid. 2015. *Genealogical Classification of Semitic: The Lexical Isoglosses*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614515494

Kroonen, Guus. 2011. *The Proto-Germanic n-stems: A Study in Diachronic Morphophonology*. Amsterdam: Rodopi.
Kroonen, Guus. 2013. *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
Kümmel, Martin J. 2007. *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert Verlag.
Labov, William. 2007. Transmission and Diffusion. *Language* 83:2.344–387.
Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620928
Lehmann, Winfred Philipp. 1969[1962]. *Einführung in die historische Linguistik*. Heidelberg: Carl Winter.
Lehmann, Winfred Philipp. 1992. *Historical Linguistics (Third Edition)*. London: Routledge.
List, Johann-Mattis. 2014. *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.
List, Johann-Mattis. 2016. Beyond Cognacy: Historical Relations Between Words and Their Implication for Phylogenetic Reconstruction. *Journal of Language Evolution* 1:2.119–136. https://doi.org/10.1093/jole/lzw006
List, Johann-Mattis, Shijulal Nelson-Sathi, Hans Geisler & William Martin. 2014. Networks Of Lexical Borrowing and Lateral Gene Transfer in Language and Genome Evolution. *BioEssays* 36:2.141–150. https://doi.org/10.1002/bies.201300096
List, Johann-Mattis, Jananan Sylvestre Pathmanathan, Philippe Lopez & Eric Bapteste. 2016. Unity and Disunity In Evolutionary Sciences: Process-Based Analogies Open Common Research Avenues for Biology And Linguistics. *Biology Direct* 11:39.1–17.
List, Johann-Mattis, Anselm Terhalle & Matthias Urban. 2013. Using Network Approaches to Enhance the Analysis of Cross-Linguistic Polysemies. *Proceedings of the 10th International Conference on Computational Semantics – Short Papers*, 347–353. Stroudsburg: Association for Computational Linguistics.
Martirosyan, Hrach K. 2010. *Etymological Dictionary of the Armenian Inherited Lexicon*. (= *Leiden Indo-European Etymological Dictionary Series*, 8.) Leiden: Brill.
Martzloff, Vincent. 2016. Arménien geri 'captif': étymologie et phraséologie à la lumière d'un parallèle en albanien du Caucase. *Wékwos* 2.109–179.
McMahon, April & Robert McMahon. 2005. *Language Classification by Numbers*. Oxford: Oxford University Press.
Meillet, Antoine. 1908. *Les dialectes indo-européens*. Paris: Librairie Ancienne Honoré Champion.
Morrison, David A. 2010. Review of the book *Sequence Alignment: Methods, Models, Concepts, and Strategies*, ed. by Michael S. Rosenberg. *Systematic Biology* 359:3.363–365.
Morrison, David A. 2011. *An Introduction to Phylogenetic Networks*. Uppsala: RJR Productions.
Morrison, David A. 2014. Phylogenetic Networks: A New Form of Multivariate Data Summary for Data Mining and Exploratory Data Analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4:4.296–312. https://doi.org/10.1002/widm.1130
Morrison, David A. 2016. Genealogies: Pedigrees and Phylogenies Are Reticulating Networks Not Just Divergent Trees. *Evolutionary Biology* 43:4.456–473. https://doi.org/10.1007/s11692-016-9376-5
Newman, M. E. J. 2010. *Networks: An Introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199206650.001.0001
Nicholls, Geoff K. & Robin J. Ryder. 2011. Phylogenetic Models for Semitic Vocabulary. *Proceedings of the 26th International Workshop on Statistical Modelling* ed. by David Conesa, Anabel Forte, Antoni López-Quílez & Facundo Muñoz, 431–436.

Nichols, Johanna. 1996. The Comparative Method as Heuristic. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change* ed. by Mark Durie & Malcolm Ross, 39–71. New York: Oxford University Press.
Pagel, Mark. 2009. Human Language as a Culturally Transmitted Replicator. *Nature Reviews Genetics* 10.405–415. https://doi.org/10.1038/nrg2560
Pernicka, Ernst, Thilo Rehren & Sigrid Schmitt-Strecker. 1998. Late Uruk Silver Production by Cupellation at Habuba Kabira, Syria. *Metallurgica Antiqua: In Honour of Hans-Gert Bachmann and Robert Maddin* ed. by Thilo Rehren, Andreas Hauptmann & James D. Muhly 123–134. Bochum: Deutsches Bergbau Museum.
Ringe, Donald, Tandy Warnow & Ann Taylor. 2002. Indo-European and Computational Cladistics. *Transactions of the Philological Society* 100:1.59–129. https://doi.org/10.1111/1467-968X.00091
Rogers, Jeffrey & Richard A. Gibbs. 2014. Comparative Primate Genomics: Emerging Patterns of Genome Content and Dynamics. *Nature Reviews Genetics* 15:5.347–359. https://doi.org/10.1038/nrg3707
Ross, Malcolm. 1988. *Proto Oceanic and the Austronesian Languages of Western Melanesia*. (= *Pacific Linguistics C*, 98.) Canberra: Pacific Linguistics.
Rubin, Aaron D. 2005. *Studies in Semitic Grammaticalization*. Winona Lake: Eisenbrauns.
Sagart, Laurent. 2015. "East Formosan" and the PAN Palatals. Talk presented at the 13th International Conference on Austronesian Linguistics, Taipei, Taiwan, July 18–22, 2015. Available at https://www.academia.edu/22871609/
Scally, Aylwyn, Julien Y. Dutheil, LaDeana W. Hillier, Gregory E. Jordan, Ian Goodhead, Javier Herrero, Asger Hobolth, Tuuli Lappalainen, Thomas Mailund, Tomas Marques-Bonet, Shane McCarthy, Stephen H. Montgomery, Petra C. Schwalie, Y. Amy Tang, Michelle C. Ward, Yali Xue, Bryndis Yngvadottir, Can Alkan, Lars N. Andersen, Qasim Ayub, Edward V. Ball, Kathryn Beal, Brenda J. Bradley, Yuan Chen, Chris M. Clee, Stephen Fitzgerald, Tina A. Graves, Yong Gu, Paul Heath, Andreas Heger, Emre Karakoc, Anja Kolb-Kokocinski, Gavin K. Laird, Gerton Lunter, Stephen Meader, Matthew Mort, James C. Mullikin, Kasper Munch, Timothy D. O'Connor, Andrew D. Phillips, Javier Prado-Martinez, Anthony S. Rogers, Saba Sajjadian, Dominic Schmidt, Katy Shaw, Jared T. Simpson, Peter D. Stenson, Daniel J. Turner, Linda Vigilant, Albert J. Vilella, Weldon Whitener, Baoli Zhu, David N. Cooper, Pieter de Jong, Emmanouil T. Dermitzakis, Evan E. Eichler, Paul Flicek, Nick Goldman, Nicholas I. Mundy, Zemin Ning, Duncan T. Odom, Chris P. Ponting, Michael A. Quail, Oliver A. Ryder, Stephen M. Searle, Wesley C. Warren, Richard K. Wilson, Mikkel H. Schierup, Jane Rogers, Chris Tyler-Smith & Richard Durbin. 2012. Insights into Hominid Evolution from the Gorilla Genome Sequence. *Nature* 483:7388.169–175. https://doi.org/10.1038/nature10842
Schleicher, August. 1848. *Zur vergleichenden Sprachengeschichte*. Bonn: König.
Schleicher, August. 1853a. Die ersten Spaltungen des indogermanischen Urvolkes. *Allgemeine Monatsschrift für Wissenschaft und Literatur* 3.786–787.
Schleicher, August. 1853b. O jazyku litevském, zvláště ohledem na slovanský června. *Časopis Českého Museum* 27.320–334.
Schleicher, August. 1861. *Kurzer Abriss einer Lautlehre der indogermanischen Ursprache*. Weimar: Böhlau.
Schleicher, August. 1863. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: Hermann Böhlau.

Schmidt, Johannes. 1872. *Die Verwantschaftsverhältnisse der indogermanischen Sprachen.* Weimar: Hermann Böhlau.
Schmidt, Johannes. 1875. *Zur Geschichte des indogermanischen Vocalismus. Zweite Abteilung.* Weimar: Hermann Böhlau.
Schuchardt, Hugo. 1900. *Über die Klassifikation der romanischen Mundarten: Probe-Vorlesung, gehalten zu Leipzig am 30 April 1870.* Graz.
Streitberg, Wilhelm. 1900. *Urgermanische grammatik.* Heidelberg: Carl Winter's Universitätsbuchhandlung.
Sutrop, Urmas. 2012. Estonian Traces in the Tree of Life Concept and in the Language Family Tree Theory. *Journal of Estonian and Finno-Ugric Linguistics* 3.297–326.
Swadesh, Morris. 1955. Towards Greater Accuracy in Lexicostatistic Dating. *International Journal of American Linguistics* 21:2.121–137. https://doi.org/10.1086/464321
Swadesh, Morris. 1959. The Mesh Principle in Comparative Linguistics. *Anthropological Linguistics* 1:2.7–14.
Taylor, Allan R. 1990. A European Loanword of Early Date in Eastern North America. *Anthropological Linguistics* 32:3/4.187–210.
Trask, Larry. 2000. Some Issues in Relative Chronology. *Time Depth in Historical Linguistics, Vol. 1* ed. by Colin Renfrew, April McMahon & Larry Trask, 45–58. Cambridge: The McDonald Institute for Archaeological Research.
Urban, Matthias. 2011. Asymmetries in Overt Marking and Directionality in Semantic Change. *Journal of Historical Linguistics* 1:1.3–47. https://doi.org/10.1075/jhl.1.1.02urb
Wichmann, Søren. 2017. Genealogical Classification in Historical Linguistics. *Oxford Research Encyclopedia of Linguistics*, 1–19. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.78
Yahalom-Mack, Naama & Adi Eliyahu-Behar. 2015. The Transition from Bronze to Iron in Canaan: Chronology, Technology, and Context. *Radiocarbon* 57:2.285–305. https://doi.org/10.2458/azu_rc.57.18563

Address for correspondence

Johann-Mattis List
Department of Linguistic and Cultural Evolution
Max Planck Institute for the Science of Human History
Kahlaische Str. 10
07745 Jena
mattis.list@shh.mpg.de

Co-author information

Guillaume Jacques
INALCO-CRLAO
Centre des Recherches Linguistiques sur
l'Asie Orientale