

Save the Trees: Why We Need Tree Models in Linguistic Reconstruction (and When We Should Apply Them)

Guillaume Jacques¹ and Johann-Mattis List^{2*}

¹Centre des Recherches Linguistiques sur l'Asie Orientale, Paris

²Max Planck Institute for the Science of Human History, Jena,

*corresponding author: mattis.list@shh.mpg.de

Abstract

Skepticism against 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 from historical linguistics has always enjoyed a certain popularity. This skepticism has further increased with recently proposed techniques for data visualization which seem to confirm that we can study language history without trees. We show that the concrete arguments brought up in favor of anachronistic wave models do not hold. In comparing the phenomenon of *incomplete lineage sorting* in biology with processes in linguistics, we show that data which does not seem to be resolvable in trees may well be explained without turning to diffusion as an explanation. At the same time, methodological limits in historical reconstruction may easily lead to an overestimation of regularity, which may in turn surface as conflicting patterns when trying to reconstruct a coherent phylogeny. We illustrate how trees can benefit language comparison in several examples, but we also point to their drawbacks 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, we conclude that all models which claim that vertical language relations can be completely ignored are essentially wrong: Either they silently still use 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

1 Introduction

All languages develop by *descent with modification* (Darwin 1859): linguistic material is transferred from generation to generation of speakers, and slight modifications in pronunciation, denotation, and grammar may sum up to changes which are so large 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. Not all linguistic material is necessarily inherited from the parent generation. Linguistic material may easily be transferred across linguistic boundaries or diffuse across similar speech varieties. This does, however, not change the fact that the primary process by which languages are transmitted is the acquisition of a first language by children (Ringe et al. 2002:61, Hale 2007:27-48). That largely incomprehensible and different languages may share a common genetic origin was one of the great insights of 19th century linguistics, and even if lateral forces of diffusion may drastically change the shape of languages, this does not invalidate the crucial role that vertical transmission plays in language history, and 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 want to substantiate this viewpoint. Starting from a brief overview on the historical debate between trees and waves in the history of linguistics, we will introduce the core arguments of the new debate about trees and waves, and then defend tree thinking in historical linguistics by showing that patterns which do not look tree-like from a first sight may still be explained by a branching tree model, while on the other hand patterns that look like common inheritance may go back to processes of language contact, which can be readily included into a rooted network model in which the family tree model serves as a backbone representing inheritance, and horizontal edges represent borrowing events.

After illustrating the benefits of trees in historical language comparison in several examples, also pointing to their obvious shortcomings, 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 tree-like 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 instigated. 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 today.

2.1 Tree Thinking in Schleicher's Work

Although 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,b), and numerous studies published thereafter (see, for example, Schleicher 1861, 1863), he propagated the idea that the assumptions about language history could be best 'illustrated by the image of a branching tree' (Schleicher 1853a:787).² Note that there was no notable influence by Darwin here. It is more likely that Schleicher was influenced by *stemmatics* (manuscript comparison, see Hoenigswald 1963:8). Even today, historical linguistics has certain features that resemble stemmatics much more closely than evolutionary biology. It seems that Schleicher's enthusiasm for the drawing of language trees had quite an impact on Ernst Haeckel (1834-1919, see Sutrop 2012), since – as Schleicher pointed out himself (Schleicher 1863:14) – linguistic trees by then were explicit, pointing to concrete languages, and not abstract, pointing to hypothetical taxa, like the one Darwin showed in his *Origins* (Darwin 1859).

Despite the seemingly radical idea to model language history as process of diversification exclusively via branching and splitting, it is important to note that Schleicher was not a careless proponent of tree thinking. Judging from his work we find many examples showing that he was aware of potential problems resulting from the tree model. Thus, in his open letter to Haeckel, Schleicher, taking Latin and its descendants as an example, explicitly pointed to problems of language mixing which he compared to plant hybrids in biology, identifying it as a second factor leading to differentiation (Schleicher 1863:18). In his earlier work, he mentioned language contact and borrowing of linguistic features explicitly as a process characteristic for language history (Schleicher 1861:6), emphasizing the importance of distinguishing borrowed from inherited traits in language classification (Schleicher 1848:30). Following up the analogy with species evolution, Schleicher also pointed to the problem of finding sharp borders between languages, dialects, and speech varieties ('Sprache, Dialekt, Mundarten und Untermundarten'), which finds a counter-part in the distinction between species and individuals (Schleicher 1863:21). Especially this last point clearly reflects that Schleicher did not exclusively think that language splits were a product of abrupt separation of speakers, and that he was aware of the idealizing aspect of the *Stammbaum*.

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, especially pointing to the problem of diffusion between very closely related languages:

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

²Our 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'.

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. What he found were patterns of words that were in a strong *patchy distribution* (see List et al. 2014), that is, showing many gaps across the languages, with only a few (if at all) patterns that could be found across all languages. One seemingly surprising fact was, for example, that Greek and Sanskrit shared about 39% of cognates (according to Schmidt's count, see Geisler & List 2013), Greek and Latin shared 53%, but Latin and Sanskrit only 8%. Assuming that Greek and Latin had a common ancestor, Schmidt found it very difficult to explain how the similarities between the two languages with Sanskrit could be so different (Schmidt 1872:24). Furthermore, this pattern of patchy distributions seemed to be repeated in all branches of Indo-European that Schmidt compared in his investigation. 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. (Schmidt 1872:17)⁴

Schmidt, however, did not stop with this conclusion but proposed another model of language divergence instead 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. (Schmidt 1872:27)⁵

Since then, this new model, the so-called *wave theory* (*Wellentheorie* in German) has been vividly discussed in articles and textbooks in historical linguistics, sometimes being promoted as the missing complement of Schleicher's *Stammbaumtheorie* (Campbell 1999:187-200, Goebel 1983), sometimes being treated as its more realistic alternative (Gabelentz 1891:194f). Despite the apparent simplicity of the wave theory reflected in succinct presentation in historical linguistics handbooks, the theory is the center of much confusion, not only among linguists themselves, but also among those who are not primarily trained in historical linguistics. This confusion is not only reflected in the discussions between dendrophilists and dendrophobists, but also in the various attempts that have been made to visualize the waves. While Schmidt did not give a visualization in his book from 1872, he gave one three years later (Schmidt 1875:199), as shown in Figure 1 along with an English translation. It is difficult to interpret this Figure, not only due to

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

⁴Our 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.'

⁵Our 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.'

the scan quality, but also due to its structure, which is hard to understand intuitively: It displays languages in a pie-chart-like diagram in a quasi-geographic space. No information regarding ancestral states of the 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 least thing one thinks about when inspecting it. Schmidt does not seem to ignore that evolution has a time dimension, but he seems to deliberately neglect it when drawing his waves.

The confusion is also reflected in the scholarly literature. In the fifty years following Schmidt's publication, we can find a wide range of different attempts to visualize the wave theory, ranging from Venn diagrams (Hirt 1905:93) to early networks (Bonfante 1931:174). The only publication that retained Schmidt's pie-chart visualization known to us was Meillet (1908:134), who 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. 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 visualizing 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 to discredit a theory. 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).

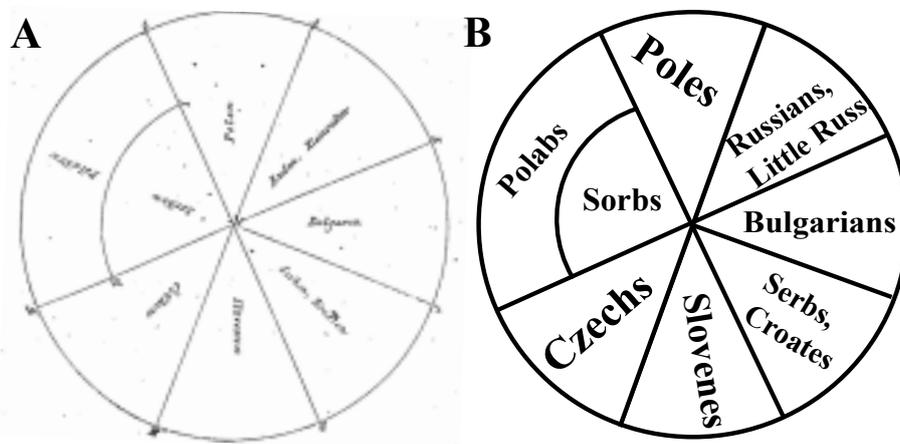


Figure 1: Schmidt's Wave Theory. A: Schmidt's visualization of the Wave Theory from 1875. B: English translation.

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): (1) practicability prob-

lems, (2) plausibility problems, and (3) adequacy problems. Practicability problems refer to the problems in applying the tree model to analyse a given set of languages. Critics, such as Schmidt (1872) mentioned above, point out in particular the issue of conflicting evidence. Plausibility problems refer to the realism of the family tree model and are reflected in obvious simplifications provoked by the tree model. They are reflected in critics emphasizing that languages do not necessarily split abruptly but slowly diverge accompanied by complex waves of diffusion (Schuchardt 1900, Schmidt 1872). Adequacy refers to the purpose of writing language history in historical linguistics. Critics complain that family trees break down all vivid aspects that are substantial for the diversification of a language family to processes 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* given 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. Refusing a model for reasons of practicability is straightforward, but it 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, in the same way as the inability to distinguish borrowed from inherited traits (especially in deeper time depths) can be considered proof against the existence of tree-like divergence of languages. Geisler & List (2013) conclude, that stronger arguments against the family tree model are those that challenge its plausibility, with respect to the presumed split-process by which languages diverge, or its adequacy, with respect to its ability 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 redrawn as a distinction between *methodological* and *theoretical* problems of the Stammbaumtheorie. The former question the *methodological possibility* to infer language trees from linguistic data (referring to the power of the methods available to us), and the latter 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, but rather slowly diverge, while at the same time exchanging material in a non-vertical manner. An even greater problem, not addressed in Schuchardt (1900) is the possibility of *convergence*. When leading to *hybridization*, it clearly contradicts the tree model in its core (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 was explicitly pointing 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 deliberately ignored these aspects in the family tree model, giving a strict preference to divergence and vertical inheritance.⁶

⁶Yet 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

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, and 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

Along with the “quantitative turn” in historical linguistics in the beginning of the 21st century (List 2014:209f), the debate on trees and waves has been recently revived. While most textbooks had treated both models as a complementary view on *external language history*⁷ (Lehmann 1992, Anttila 1972) or as treating two completely different aspects of language change (Campbell 1999), more and more linguists now discuss the models as opposing perspectives (Heggarty et al. 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 et al. 2002, Pagel 2009). While both trees and waves had been playing a less prominent role for a long time,⁸ biological methods for phylogenetic reconstruction applied to large linguistic datasets now offer a much more transparent way to analyse and display large amounts of language data in a tree diagram than the classical method of identifying shared innovations.

Yet not long after the first biological software packages were used for phylogenetic tree reconstruction in linguistics, new visualization techniques for *splits networks* (most of them based on the NeighborNet algorithm, Bryant & Moulton 2004, but see Hurles et al. 2003 for the earliest example of splits networks in linguistics known to us), as 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 et al. 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 et al. 2010, Heggarty et al. 2010, Ben Hamed 2005, Bowern 2010).

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.

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

⁸Even Morris Swadesh was extremely careful to use his lexicostatistic method to produce family trees. Instead, he published a map on “interrelationships of American Indian languages” that comes closer to an interpretation in terms of the wave theory (Swadesh 1959:23).

3.1 Phylogenetic Tree Reconstruction after the Quantitative Turn

Classical phylogenetic tree reconstruction in historical linguistics is very similar to the basic ideas of cladistics in biology (Hennig 1950, see also Lass 1997:105-171), in so far 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, that is, changes that define a subgroup, is very old in linguistics, and can be found already in the work of Karl Brugman (1849-1919), although it was later scholars like 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:

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. (Brugmann 1967[1886]:24)⁹

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, which is often downplayed in this context, however, 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 have likewise (a) been borrowed (see Section 4.2.1), (b) been retained from the common ancestor of all languages (*symplesiomorphies*), (c) been independently emerged (*homoplasies*), or (d) been erroneously annotated as *shared features*. Furthermore, *differential loss* or further development of features in subgroups may easily mask shared innovations, and an innovation which was originally shared by a group of languages may give the impression to be patchily distributed. This is further complicated by the fact that variation of linguistic features occurs in all languages and may as 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 would in fact rather represent a shared retention or an independent development, thus, a combination of (b) and (c), as will be further discussed in Section 4.1. None of these problems is new to historical linguistics, and we can find all of these points, apart from the problem of variation in the proto-language, already in Brugmann (1884), who concludes that in order to be applicable to subgrouping, proposed innovations must be *frequent* enough to reduce the possibility of chance (see also Dyen 1953).

How frequency is defined in concrete in historical linguistics is difficult to say, as scholars often intuitively weight characters, assigning more importance to certain kinds of evidence (for example form similarities in morphological paradigms, see Nichols 1996) than to other types (isolated lexical items, or frequent sound change patterns which are likely to recur independently), and most debates regarding subgrouping center around the question of how the

⁹Our 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.’

different types of evidence should be weighted or the data interpreted. As an example, compare the discussion in Sagart (2015) who proposes that the innovations presented in Blust (1999) are better interpreted as retentions.

Phylogenetic approaches which were originally developed for application in evolutionary biology offer a different approach to the problem by using a larger pool of characters and explicit models of character evolution to automatically find the phylogenetic tree that explains the data according to different criteria (likelihood, parsimony) while at the same time determining which characters have been retained and which have been innovated (Greenhill & Gray 2012). Classical linguists often mistrust these methods, criticizing their *blackbox* character.¹⁰ While the criticism is justified to some extent, it should be kept in mind that it is not the methods itself which are non-transparent or inaccurate, but 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 blackbox methods per se. As the methods are based on an explicit modelling of evolutionary processes,¹¹ the tree or the forest of trees they infer are 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 misunderstand that splits networks are a mere tool for data display (Morrison 2010) not a tool that directly produces a phylogenetic analysis. While 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),

they do not produce any hypothesis on how languages or biological species diverged or recombined, and they must be strictly distinguished from explicit *evolutionary networks*, displaying ‘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 what phylogenetic trees and networks are supposed to represent. For us, the crucial difference between data-display approaches and true phylogenetic accounts is the lack of an explicit time dimension displaying events of divergence (or recombination)¹² 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

¹⁰This 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. What users see are only the aggregated results.

¹¹This includes parsimony, as 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.

¹²When dealing with recombination of lineages (like under the assumption of language mixing, see 6), a tree model is not enough, and a network has to be used.

time depths, and language-family-specific factors, but if we ignore the ancestral varieties in our analysis completely, as in the case of splits networks, we lose all temporal dynamics, and as a result, we end up with a mere 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 which also attempts to dismiss the tree model is the theory of *historical glottometry* (François 2014, Kalyan & François forthcoming). Similar to the arguments brought up by Schmidt (1872), glottometry results from the dissatisfaction of conflicting data in historical linguistics. Furthermore, glottometry follows Ross (1988) in assuming that language divergence may proceed in form of 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 comes close to the one of Schuchardt (1900), as it attacks the process of concrete language split as it is visually suggested by the tree model. Given the high diffusability of linguistic features across mutually intelligible varieties, a fully resolved tree reconstruction showing language divergence in split processes may become difficult if not impossible in such 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 them in his genetic subgrouping of Oceanic languages. Kalyan & François (forthcoming) criticize this solution as unsatisfying, emphasizing that polytomies mask 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, supposed to represent exclusively shared innovations, for the language family under investigation, and 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, where four hypothetical languages are given which are connected by three isoglosses of which two are in conflict with each other.

Three general problems with the method of historical glottometry need to be mentioned in this context: First, the resulting visualizations can by no means qualify as phylogenetic analyses, as they lack the 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 not much 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 for ancestral languages to be included in the analysis.¹³

¹³Mathematically, the isogloss model proposed by glottometry corresponds to a *hypergraph*, in which edges can connect more than one vertex (Newman 2010:122f). 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 2010:124f).

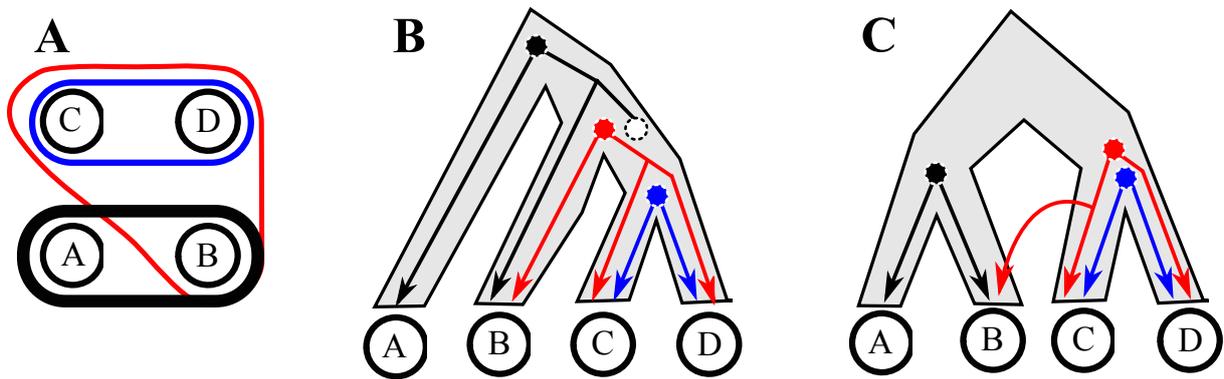


Figure 2: Historical glottometry, the family tree model, and an evolutionary network. A shows a glottometric diagram with weighted isoglosses drawn between languages sharing an innovation which are in apparent conflict with each other. B shows how the same scenario can be explained with help of a family tree by assuming differential loss of the black isogloss. C shows an evolutionary network representing another possible explanation for the patterns by assuming that the red innovation was borrowed from language C into language B.

Second, the use of the term *innovation* in historical glottometry is logically problematic. According to the practice reported in (François 2014), all instances where 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 was diffused in times 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, and (François 2014:177) even does not make an attempt to distinguish between the two. Thus, innovations in the sense of glottometry represent two different processes, namely (a) cases of unique deviation of linguistic traits from the Ursprache (true shared innovations in the cladistic sense), and (b) cases of parallel development.

Leaving aside the fact that using a proto-language in order 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 since 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, while the term in fact is meant to be *explanative*. When labeling certain features as shared innovations, they 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), or ‘polysemy’ (see François

2008, List et al. 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, and 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 but also occurs in biology (De Laet 2005). A cladistic analysis seeks to identify which of a large pool of *possible characters* could qualify to define a subgroup and thus reflect *potentially* 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, where the data from Figure 2A is 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, 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.¹⁴ Interestingly, this was emphasized much earlier in the history of linguistics when Brugmann (1884) criticized the *wave theory* by Schmidt (1872), as 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 (forthcoming) 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 where it spreads. 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 on shared traits between language varieties.

4 Saving the Trees from the Critics

Given the logical necessity to allow for divergence, a specific part of language history can be modeled with help of a tree if specific processes like recombination (hybridization, creolization, see 6) can be excluded. That such a tree model does not necessarily represent all aspects of language history is obvious, and even the strongest tree proponents would not deny it. Whether the amount of inheritance versus borrowing in language history is as low as it was supposed for biology, where tree critics have labeled the tree of life as the “tree of one percent” (Dagan &

¹⁴A 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).

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 *necessitates* a model for language history which handles divergence and splits of lineages. 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 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 are erroneous codings which occur especially in those cases where the data has 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 independent development, and especially for sound changes it is often very hard to distinguish between synapomorphy and homoplasy (Chacon & List 2015:182f), which 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, during 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 why 30% of the genes in a Gorilla's genome are more similar to the human genome or to the Chimpanzee genome although 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 language-internal factors of polymorphisms which may be inherited across lineages before they are later randomly resolved, a further factor not mentioned by List et al. (2016) is *variation in the population of speakers*, or *sociolinguistic variation*, which may even occur in one and the same 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 the 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 in the data, or 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 to find the tree in the forest of individual histories of linguistic traits. In the following, we will illustrate this in more detail by showing

how variation *inherited* from the 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.

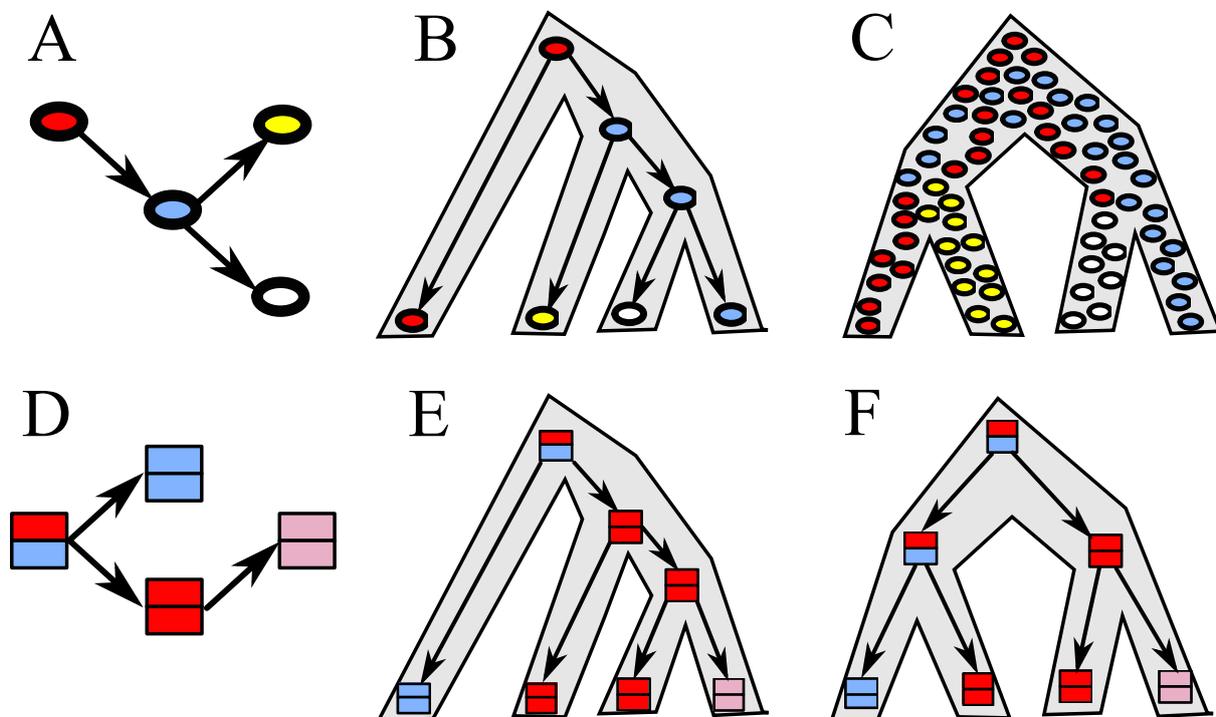


Figure 3: Incomplete lineage sorting due to sociolinguistic (A-C) and linguistic variation (D-F) and its impact on phylogenetic reconstruction and genetic subgrouping. A shows a pattern of known directional evolution of a character (e.g., a sound change pattern), and B shows one of the most parsimonious trees resulting from the pattern. C shows an alternative pattern by 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 languages, it was inherited in the two descendant languages from which the character further developed. As a result, another tree topology can be reconstructed. D gives an example for a process of paradigm levelling, and E and F show two possible equally parsimonious scenarios invoking different tree topologies each.

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,¹⁵ overlapping isoglosses constitute a major problem for the tree model in the

¹⁵This assertion remains to be demonstrated, but we accept it for the sake of argument.

view of glottometry. Regardless of whether lexically-specific sound changes have more difficulty to cross language boundaries than other types of features, overlapping innovations can, as mentioned above, also be accounted for by assuming the existence of variation in the proto-language.¹⁶

Languages are never completely uniform, and fieldwork linguists working on unwritten languages commonly notice that even siblings may 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 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; then 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 rather 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 of 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’ (Figure 4),¹⁷ an n-stem noun whose reflexes in the modern and ancient languages are particularly complex (data from Kroonen 2011:71,128, Kroonen 2013:294).

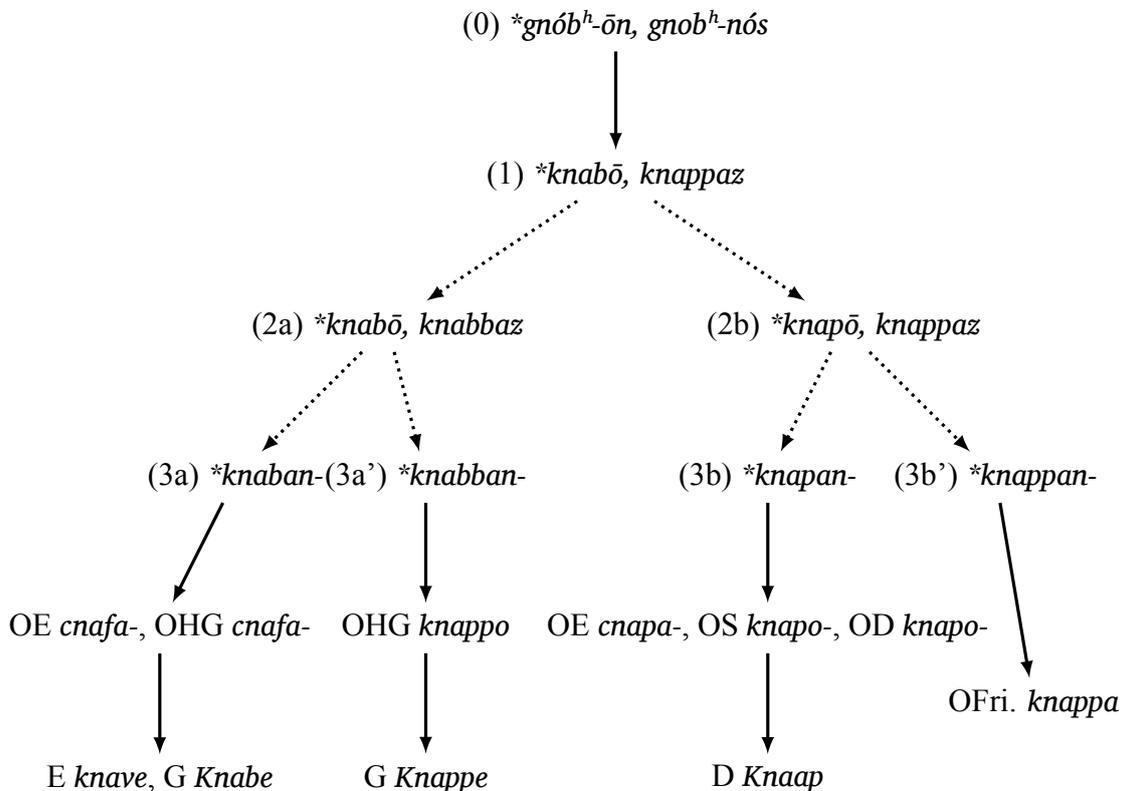
From the attested ancient and modern forms (if the known sound laws are applied backwards), no fewer than four protoforms have to be postulated: **knaban-*, **knapan-*, **knabban-* and **knappan-*. Some languages have more than one reflex of this etymon (with diverging specialized meanings), and their 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 ‘Ingvaemonic’ 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 Frisian.

Unlike most other language families, the detailed knowledge that has accumulated on the history of Germanic languages allows to go further than stating the presence of irregular correspondences: it is possible to account for them with a detailed model. It is now near-universally

¹⁶In this case, however, we can no longer speak of true innovations in the cladistic sense, as, as mentioned above, the term *innovation* is explanatory and not descriptive, and presupposes that a trait is uniquely shared by the subgroup which it defines.

¹⁷The reflexes of this proto-form have developed distinct meanings in the attested languages, including ‘squire’, but this aspect is not considered here.

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



accepted that doublets such as those 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) on the endings of n-stem nouns in pre-proto-Germanic (stage 0, Kluge 1884, Kroonen 2011).

The paradigm of the noun ‘boy’ (and 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 Germanic languages.

4.1.2 Proto-Variation

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

Germanic second person pronouns provide an example of a case of this type. 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 through **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 for analogical levelling with the first plural accusative/dative **unsiz* (through a stage **iwsiz* followed by metathesis, see Kroonen 2013:275). Since Gothic is uncontroversibly 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 Streitberg 1900:265) argues instead that **izwiz* and **iwiz* are both ancient; as proto-Indo-European had both **wes* (Sanskrit 2pl 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.¹⁸

Brugmann’s idea implies that two proto-forms co-existed in proto-Germanic for the accusative/dative of the 2pl. This is by no means a crosslinguistically uncommon state of affairs,¹⁹ 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 2014:176-8). 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.

This approach however neglects an important factor: while regular sound correspondences are a necessary condition for analyzing forms in related languages as cognates, i.e. originating

¹⁸Proto-Indo-European **e* shifts to Germanic **i* in unaccented syllables.

¹⁹For instance, Japhug has several competing forms for the first and second person pronouns, as well as form the dative postposition (Jacques 2017:624 within a single varieties (without counting dialectal variations).

from the same etymon in their common ancestor,²⁰ they are not a *sufficient* condition 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 *hájā* ‘aluminum’ look like they regularly originate from a Common Tibetan form **ha.jaŋ*.²¹ 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, p.c.) 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 has been coined only once,²² and was then borrowed into other Tibetan languages²³ and neighboring minority languages under Tibetan influence (as for instance Japhug *χajaŋ* ‘aluminum’). In this case a phonetic borrowing from Amdo *hajaŋ* could only yield Lhasa *hájā*, since *h-* only occurs in high tone in Lhasa, and since final *-ŋ* has been transphonologized as vowel nasality.²⁴

Several Algonquian languages, share a word for ‘pig’ (Fox *koohkooša*, Miami *koohkooša* and Cree *kōhkōs*) 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.²⁵

²⁰Note however that cognacy is a more complex concept that 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.

²¹In 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*.

²²We are not aware of a 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.

²³In some Tibetan languages such as Cone *hàjā*, Jacques (2014:306), there is clear evidence that the word is borrowed from Amdo Tibetan and is not native (otherwise †*hàjā* would have been expected).

²⁴Likewise, in the case of borrowing from Lhasa into Amdo, the rhyme *-aŋ* would be the only reasonable match for Lhasa *-ã*.

²⁵However, 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 *gookoosh* instead of expected †*gookoozh*), and that the ambiguity between cognate and loanword only exist with Fox, Miami and Cree (I wish to thank an anonymous reviewer for pointing out this fact).

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 for 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 occurred in either the donor or the receiver language, or if the donor and the receiver languages underwent identical sound changes up to the stage when the borrowings occurred, phonology alone is not a sufficient criterion to distinguish between inherited words and loanwords.

A classical 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.’²⁶ Table 1 presents a non-exhaustive list of such words, with the corresponding proto-Iranian etyma.

Table 1: Armenian words which cannot be conclusively demonstrated to be either borrowings from Iranian or inherited words from a phonetic point of view

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

The Armenian case shows that undetectable loans are not restricted to cases like those studied above, when a particular word only contains segments which have not been affected by sound changes from the proto-language to all its daughter languages. Undetectable loans are 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 to distinguish between loanwords and inherited words.

²⁶Our 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.’

What has been 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 in historical glottometry, amounts to losing crucial historical information, and it does not seem justified to blame the family tree model for an insufficiency of our linguistic reconstruction methodology.

4.2.2 Nativization of Loanwords

In the previous section, we discussed cases when borrowing took place before 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 is that 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, some at a stage before most characteristic sound changes had taken place, other more recently. Table 2 presents examples of cognates between Finnish and Northern Saami illustrating 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
<i>käsi</i>	<i>giehta</i>	* <i>käti</i>	‘hand’
<i>nimi</i>	<i>namma</i>	* <i>nimi</i>	‘name’
<i>kala</i>	<i>guolli</i>	* <i>kala</i>	‘fish’
<i>muna</i>	<i>monni</i>	* <i>muna</i>	‘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, including recent borrowings, such as *mearka* from *merkki* ‘sign, mark’ and *báhppa* from *pappi* ‘priest’ (from Common Slavic **pǎpъ*, 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 where 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 *непть* ‘a type of cabin’, which 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 2006:21-3). 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 phonetical 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 not yet been documented as clearly as in 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 have tried to illustrate that not all patterns which look non-tree-like on first sight require a tree-free explanation, while at the same time patterns that look like excellent examples for 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 (on which 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:

- $*\theta \rightarrow f$ (merging with $*f$)

- *ð → z (merging with *z)
- *θ' → s' (merging with *s')
- *ɫ' → 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 another language are indicated in grey)

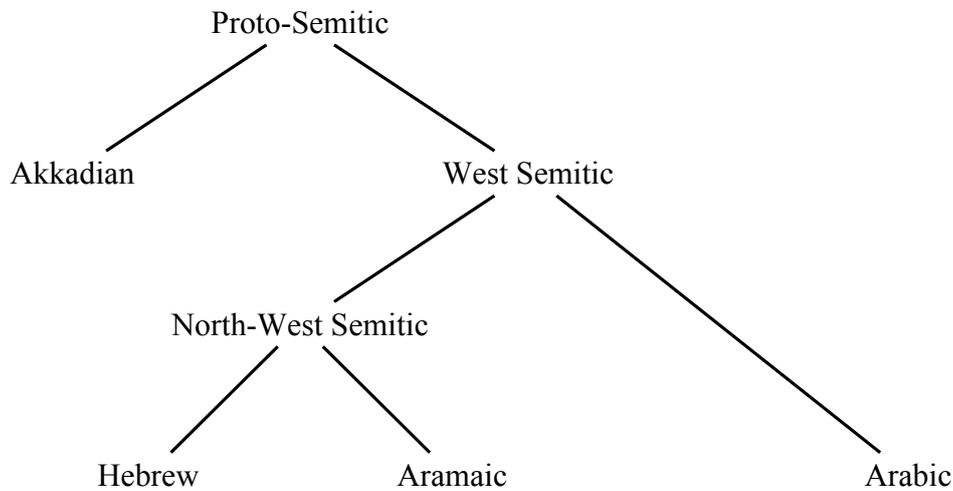
Proto-Semitic	Akkadian	Hebrew	Biblical Aramaic	Standard Arabic
*t	t	t	t	t
*d	d	d	d	d
*θ	ʃ	ʃ	t	θ
*ð	z	z	d	ð
*s	s	s	s	s
*z	z	z	z	z
*ʃ	ʃ	ʃ	ʃ	s
*ɫ	ʃ	ɫ	ɫ	ʃ
*t'	t'	t'	t'	t ^ʕ
*θ'	s'	s'	t'	θ ^ʕ
*s'	s'	s'	s'	s ^ʕ
*ɫ'	s'	s'	ʕ	d ^ʕ

While phonology could seem at first glance to support grouping Akkadian and Hebrew together in a group excluding Aramaic and Arabic, the bulk of morphological and lexical innovations incontrovertibly support that Akkadian is the first branch of the family, and that Aramaic and Hebrew are closer to each other than either of them is to Arabic (see for instance Hetzron 1976, Huehnergard 2006), and Bayesian phylogenetic analyses which 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), and Akkadian does have a frozen trace of the preposition

Figure 5: A simplified Stammbaum of Semitic languages



**b-* (Rubin 2005:45-6). Since none of the 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 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.

5.3 Directionality of Change

As a by-product 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 (in particular Kümmel 2007) or semantic changes (see for instance Urban 2011), there are still many isoglosses, in particular in inflectional morphology, whose interpretation as innovations or retentions is nearly impossible by direct comparison between languages.

As an example for 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 the languages; it is a perception verb (‘see’, ‘look at’) in some 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 again 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 North-West Semitic (Ugaritic): it is thus most likely to be the archaic meaning. Ugaritic, where the root means both ‘to look at’ and ‘say’, represents 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 ‘command’.

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

Language	Reflex	Meaning
Akkadian	<i>amārum</i>	to see
Hebrew	<i>ʔāmar</i>	to say, to declare, to command
Ugaritic	<i>ʔmr</i>	to say; to look at
Arabic	<i>ʔamara</i>	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. Apart from highly controversial attempts to find a universal constant of lexical replacement rates (Swadesh 1955), however, 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 have a mere 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 on frequency or cross-linguistic studies as the one on sound change by Kümmel (2007), 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, since they allow to process larger amounts of data, while at the same time not being able to handle uncertainty in the inferences. Even less sophisticated approaches, such as weighted parsimony, can provide interesting insights into sound change patterns which occur frequently independent of each other, along different branches of a tree (Chacon & List 2015). Static models of shared commonalities, as the isogloss-maps of glottometry, do not allow us to get access to 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 can be obtained regarding the history and the spread of a language family using the vocabulary reconstructible for particular nodes. For instance, the presence of a reconstructible etymon **kasp-* ‘silver’ in Akkadian (*kaspum*), Ugaritic *ksp* and Hebrew (*késeḫ*) among other languages (Huehnergard 2012:14-6), suggests that silver

smelting could have been possibly 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 et al. 1998). Other metals however are only reconstructible to lower branches of the family; for instance ‘iron’ does not occur earlier than proto-Cananean (**barðill-*, Hebrew *barzel*, cf. Kogan 2015:287),²⁷ 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 advancing our knowledge regarding directionality preferences in semantic shift and morphological change as well as regarding the strength of certain tendencies with help of the family tree model, 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 where it is not applicable even taking incomplete lineage sorting into account, when one language results from the merger of two previously unintelligible languages (whether or not the two varieties are demonstrably related or not).

The clearest and best documented example of this type is Michif, a contact language based on Canadian French and Plains Cree (Bakker 1997). Example (1) (taken from Antonov 2015) illustrates the main features of this language (elements from French are in bold, and those from Cree in italics). Nearly all verbs and verbal morphology come from Plains Cree, except the verbs ‘to be’ and ‘have’ which are from French with their complete irregular paradigms (including French tense categories, as shown by example 2). Most nouns and adjectives come from French. Some determiners are from French (the articles) but the demonstrative are from Cree, and 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** **fâm-a**
 3-father-obv apparently pst-marry-3obv→3prox this:an:obv def:fem:sg woman-obv
 ‘Her father apparently married that woman...’ (1:8-9)

- (3) **stît=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 that of less extreme contact languages, cannot be represented by the tree model, as it requires two roots (from languages belonging to unrelated families). A more complex type of network, a directed network with multiple roots, is necessary to represent a language of the type, and the near-perfect division of the

²⁷Similar forms in other languages such as Akkadian *parzillum* ‘iron’ do not follow the regular correspondences and cannot be cognate.

French and the Cree components of this language may 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 premature critics. 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 really reached the level of sophistication to depict the temporal dynamics of language history. After briefly introducing the new debate of trees and waves, with the method of historical glottometry as one of the strongest opponents of family trees in contemporary historical linguistics, we have shown that this method suffers from the same problems as Schmidt's *Wellentheorie*, in so far 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 that conflicts in linguistic data, which are taken as *prima facie* evidence against trees, may often be explained in 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, may 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 for 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 to abandon the family tree model as the standard to represent vertical aspects of language evolution.

Language history is incredibly complex and despite more than 200 years of research 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, and 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, however, necessitates the 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 lan-

guage history more realistically,²⁸ we do not accept the conclusion that vertical language change can be completely ignored. “Treeless” approaches, like historical glottometry or splits networks, either silently still use family trees, or they 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, Yoram Meroz for a vivid discussion after we circulated the first draft of the paper. We thank Russell D. Gray for providing additional hints to missing bibliographic items, and Malcom Ross for additional comments to our first draft. We also thank the anonymous reviewers for a challenging discussion.

²⁸This would be true evolutionary networks in the sense of Morrison (2011).

References

- Aikio, Ante. 2006. Etymological nativization of loanwords: A case study of Saami and Finnish. In Ida Toivonen & Diane Nelson (eds.), *Saami Linguistics*, 17–52. Amsterdam: Benjamins.
- Antonov, Anton. 2015. The obviative as a differential object marker in Michif. In *Association for Linguistic Typology 11th Biennial Conference*, University of New Mexico, Albuquerque.
- 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? In Peter Forster & Colin Renfrew (eds.), *Phylogenetic methods and the prehistory of languages*, 91–109. Cambridge and Oxford and 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, W. H. 2006. Mandarin dialect phylogeny. *Cahiers de Linguistique – Asie Orientale* 35(1). 71–114.
- 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 of London, Series B: Biological Sciences* 272(1567). 1015–1022.
- Ben Hamed, Mahe & Feng Wang. 2006. Stuck in the forest: Trees, networks and Chinese dialects. *Diachronica* 23. 29–60.
- Bloomfield, Leonard. 1973. *Language*. London: Allen & Unwin.
- Blust, Robert. 1999. Subgrouping, circularity, and extinction: Some issues in Austronesian comparative linguistics. In *Selected Papers from the 8th International Conference on Austronesian Linguistics*, 31–94. Taipei: Academica Sinica.
- Bonfante, G. 1931. I dialetti indoeuropei [The Indo-European dialects]. *Annali del R. Istituto Orientale di Napoli* 4. 69–185.
- Bowern, C. 2010. Historical linguistics in Australia: trees, networks and their implications. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 365(1559). 3845–3854.
- Brugmann, Karl. 1884. Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen [On the relationships of the Indo-European languages]. *Internationale Zeitschrift für allgemeine Sprachwissenschaft* 1. 228–256.
- Brugmann, Karl. 1890. *Grundriß der vergleichenden grammatik der indogermanischen sprachen*, vol. II 2.1. Strassburg: Karl J. Trübner.
- Brugmann, Karl. 1967[1886]. *Einleitung und Lautlehre: vergleichende Laut-, Stammbildungs- und Flexionslehre der indogermanischen Sprachen* [Introduction and phonetics: comparative studies on sounds, stems, and inflexion of Indo-European languages]. Berlin, Leipzig: Walter de Gruyter. Unmodified reprint of the reprint of the second edition of the unmodified new print from 1930.
- Bryant, David, Flavia Filimon & Russell D. Gray. 2005. Untangling our past: Languages, Trees, Splits and Networks. In Ruth Mace, Clare J. Holden & Stephen Shennan (eds.), *The evolution of cultural diversity: A phylogenetic approach*, 67–84. London: UCL Press.
- Bryant, David & V. Moulton. 2004. Neighbor-Net. *Molecular Biology and Evolution* 21(2). 255–265.
- 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'evic Starostin. 2005. *Sravnitel'no-istoričeskoe jazykoznanie* [Comparative-historical linguistics]. Moscow: Akademia.
- Campbell, Lyle. 1999. *Historical linguistics. An introduction*. Edinburgh: Edinburgh University Press 2nd edn.
- 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.
- Costa, David J. 2013. Borrowing in Southern Great Lakes Algonquian and the history of Potawatomi. *Anthropological Linguistics* 55(3). 195–233.
- Dagan, Tal & William Martin. 2006. The tree of one percent. *Genome Biology* 7(118). 1–7.

- 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. In Victor A. Albert (ed.), *Parsimony, phylogeny, and genomics*, 81–116. Oxford: Oxford University Press.
- Dyen, Isidore. 1953. [Review] Malgache et maanjan: Une comparaison linguistique by Otto Chr. Dahl. *Language* 29(4). 577–590.
- 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. In Martine Vanhove (ed.), *From polysemy to semantic change*, 163–215. Amsterdam: Benjamins.
- François, Alexandre. 2014. Trees, waves and linkages: models of language diversification. In Claire Bowerman & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 161–189. 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* [Linguistics]. Leipzig: T. O. Weigel.
- Galtier, Nicolas & Vincent Daubin. 2008. Dealing with incongruence in phylogenomic analyses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 363(1512). 4023–4029. doi:10.1098/rstb.2008.0144.
- Geisler, H. & J.-M. List. 2013. Do languages grow on trees? The tree metaphor in the history of linguistics. In Heiner Fangerau, Hans Geisler, Thorsten Halling & William Martin (eds.), *Classification and evolution in biology, linguistics and the history of science. Concepts – methods – visualization*, 111–124. Stuttgart: Franz Steiner Verlag.
- Geisler, Hans & Johann-Mattis List. 2010. Beautiful trees on unstable ground. Notes on the data problem in lexicostatistics. In Heinrich Hettrich (ed.), *Die Ausbreitung des Indogermanischen. Thesen aus Sprachwissenschaft, Archäologie und Genetik*, Wiesbaden: Reichert. Document has been submitted in 2010 and is still waiting for publication.
- Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. Berlin: Mouton de Gruyter.
- Goddard, Ives. 1974. Dutch loanwords in Delaware. In Herbert C. Kraft (ed.), *A Delaware Indian Symposium*, 153–60. Harrisburg: Pennsylvania Historical and Museum Commission.
- Goebel, 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 African and Oriental Studies* 79(2). 347–374.
- Gray, R. D., D. Bryant & S. J. Greenhill. 2010. On the shape and fabric of human history. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 365(1559). 3923–3933.
- Gray, Russell D. & Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426(6965). 435–439.
- Greenhill, Simon J. & Russell D. Gray. 2012. Basic vocabulary and Bayesian phylolinguistics. *Diachronica* 29(4). 523–537.
- Hale, Mark. 2007. *Historical linguistics: Theory and method*. Malden and Oxford and Victoria: Blackwell.
- Heggarty, P., W. Maguire & A. McMahon. 2010. Splits or waves? Trees or webs? How divergence measures and network analysis can unravel language histories. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 365(1559). 3829–3843.
- Hennig, Willi. 1950. *Grundzüge einer Theorie der phylogenetischen Systematik* [Foundations of the theory of phylogenetic systematics]. Berlin: Deutscher Zentralverlag.
- Hetzron, Robert. 1976. Two principles of genetic reconstruction. *Lingua* 38(2). 89–108.
- Hirt, Herman. 1905. *Die Indogermanen* [The Indo-Europeans], 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.
- Hoeningwald, Henry M. 1963. On the history of the comparative method. *Anthropological Linguistics* 5(1). pp.1–11.
- Holzer, Georg. 1996. *Das Erschließen unbelegter Sprachen* [The reconstruction of unattested languages]. Frankfurt am Main: Lang.

- Huehnergard, John. 1997. *A grammar of Akkadian*. Atlanta: Scholars Press.
- Huehnergard, John. 2006. Proto-Semitic and Proto-Akkadian. In G. Deutscher & N.J.C Kouwenberg (eds.), *The Akkadian Language in its Semitic Context: Studies in the Akkadian of the Third and Second Millennium BC*, 118. Leiden: NINO.
- Huehnergard, John. 2012. *An Introduction to Ugaritic*. Peabody, MA: 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.
- Huson, Daniel H. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 14(1). 68–73.
- Hübschmann, Heinrich. 1897. *Armenische Grammatik. I. Theil: Armenische Etymologie* [Armenian Grammar. First part. Armenian etymology]. Leipzig: Breitkopf und Härtel.
- Jacques, Guillaume. 2014. Cone. In Jackson T.-S. Sun (ed.), *Phonological profiles of little-studied Tibetic varieties*, 265–371. Taipei: Academia Sinica.
- Jacques, Guillaume. 2016. Tangut, Gyalrongic, Kiranti and the nature of person indexation in Sino-Tibetan/Trans-Himalayan. *Linguistic Vanguard* 1–13. doi:10.1515/lingvan-2015-0033.
- Jacques, Guillaume. 2017. Japhug. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages, second edition*, 614–634. London: Routledge.
- Kalyan, Siva & Alexandre François. forthcoming. Freeing the comparative method from the tree model: A framework for Historical Glottometry. In Ritsuko Kikusawa & Lawrence Reid (eds.), *Let's talk about trees: Tackling problems in representing phylogenetic relationships among languages*, Osaka: National Museum of Ethnology.
- Kluge, F. 1884. Die germanische Consonantendehnung. *Beiträge der deutschen Sprache und Literatur* 9. 149–186.
- Koch, Harold. 2004. A methodological history of Australian linguistic classification. In Claire Bowerman & Harold Koch (eds.), *Australian Languages: Classification and the Comparative Method*, 17–60. Amsterdam: Benjamins.
- Kogan, Leonid. 2015. *Genealogical Classification of Semitic: The Lexical Isoglosses*. Berlin: Mouton de Gruyter.
- 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.
- Lehmann, Winfred Philipp. 1969[1962]. *Einführung in die historische Linguistik* [Introduction to historical linguistics]. Heidelberg: Carl Winter. Autorisierte, vom Verfasser durchgesehene Übersetzung von Rudolf Freudenberg.
- Lehmann, Winfred Philipp. 1992. *Historical linguistics*. London: Routledge 3rd edn.
- List, J.-M. 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. doi: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.
- List, Johann-Mattis, Jananan Sylvestre Pathmanathan, Philippe Lopez & Eric Baptiste. 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. In *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: Brill.

- Martzloff, Vincent. 2016. Arménien geri ‘captif’ : étymologie et phraséologie à la lumière d’un parallèle en albanien du Caucase [Armenian geri ‘captive’: etymology and phraseology in the light of a parallel in the Albanian of the Caucasus]. *Wékwos* 2. 109–179.
- McMahon, April & Robert McMahon. 2005. *Language classification by numbers*. Oxford: Oxford University Press.
- Meillet, Antoine. 1908. *Les dialectes Indo-Européens* [The Indo-European dialects]. Paris: Librairie Ancienne Honoré Champion.
- Morrison, D. A. 2011. *An introduction to phylogenetic networks*. Uppsala: RJR Productions.
- Morrison, David A. 2010. Review of *sequence alignment: methods, models, concepts, and strategies* ed. by M. S. Rosenberg (2009). *Systematic Biology* 359(3). 363–365.
- Morrison, David A. 2014. Phylogenetic networks: a new form of multivariate data summary for data mining and exploratory data analysis. *WIREs Data Mining and Knowledge Discovery* doi:doi:10.1002/widm.1130.
- Morrison, David A. 2016. Genealogies: Pedigrees and phylogenies are reticulating networks not just divergent trees. *Evolutionary Biology* doi:10.1007/s11692-016-9376-5. Published online before print.
- Newman, M. E. J. 2010. *Networks. An Introduction*. Oxford: Oxford University Press.
- Nicholls, Geoff K. & Robin J. Ryder. 2011. Phylogenetic models for Semitic vocabulary. In *Proceedings of the 26th International Workshop on Statistical Modelling*, .
- Nichols, Johanna. 1996. The comparative method as heuristic. In Mark Durie (ed.), *The comparative method reviewed. Regularity and irregularity in language change*, 39–71. New York: Oxford University Press.
- Pagel, Mark. 2009. Human language as a culturally transmitted replicator. *Nature Reviews. Genetics* 10. 405–415.
- Pernicka, E., Th. Rehren & S. Schmitt-Strecker. 1998. Late Uruk silver production by cupellation at Habuba Kabira, Syria. In Th. Rehren, A. Hauptmann & J. D. Muhly (eds.), *Metallurgica Antiqua. In Honour of Hans-Gert Bachmann and Robert Maddin*, 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.
- Rogers, J. & R. A. Gibbs. 2014. Comparative primate genomics: emerging patterns of genome content and dynamics. *Nature Reviews Genetics* 15(5). 347–359.
- Ross, Malcom D. 1988. *Proto-Oceanic and the Austronesian languages of Western Melanesia*. Canberra: Pacific Linguistics.
- Rubin, Aaron. 2005. *Studies in Semitic Grammaticalization*. Winona Lake.
- Sagart, Laurent. 2015. ‘East Formosan’ and the PAN palatals. Talk, held at the 13th International Conference on Austronesian Linguistics, Taipei, July 18–22. URL: <https://www.academia.edu/22871609/>.
- Scally, A., J. Y. Duthell, L. W. Hillier, G. E. Jordan, I. Goodhead, J. Herrero, A. Hobolth, T. Lappalainen, T. Mailund, T. Marques-Bonet, S. McCarthy, S. H. Montgomery, P. C. Schwalie, Y. A. Tang, M. C. Ward, Y. Xue, B. Yngvadottir, C. Alkan, L. N. Andersen, Q. Ayub, E. V. Ball, K. Beal, B. J. Bradley, Y. Chen, C. M. Clee, S. Fitzgerald, T. A. Graves, Y. Gu, P. Heath, A. Heger, E. Karakoc, A. Kolb-Kokocinski, G. K. Laird, G. Lunter, S. Meader, M. Mort, J. C. Mullikin, K. Munch, T. D. O’Connor, A. D. Phillips, J. Prado-Martinez, A. S. Rogers, S. Sajjadian, D. Schmidt, K. Shaw, J. T. Simpson, P. D. Stenson, D. J. Turner, L. Vigilant, A. J. Vilella, W. Whitener, B. Zhu, D. N. Cooper, P. de Jong, E. T. Dermitzakis, E. E. Eichler, P. Flicek, N. Goldman, N. I. Mundy, Z. Ning, D. T. Odom, C. P. Ponting, M. A. Quail, O. A. Ryder, S. M. Searle, W. C. Warren, R. K. Wilson, M. H. Schierup, J. Rogers, C. Tyler-Smith & R. Durbin. 2012. Insights into hominid evolution from the gorilla genome sequence. *Nature* 483(7388). 169–175.
- Schleicher, August. 1848. *Zur vergleichenden Sprachengeschichte* [On comparative language history]. Bonn: König.
- Schleicher, August. 1853a. Die ersten Spaltungen des indogermanischen Urvolkes [The first splits of the Indo-European people]. *Allgemeine Monatsschrift für Wissenschaft und Literatur* 3. 786–787.
- Schleicher, August. 1853b. O jazyku litevském, zvláště ohledem na slovanský června 1853 [On the Lithuanian language, with a special focus on Slavic]. *Časopis Českého Museum* 27. 320–334.
- Schleicher, August. 1861. *Kurzer Abriss einer Lautlehre der indogermanischen Ursprache* [A brief sketch of the sound system of the Indo-European proto-languages]. Weimar: Böhlau.
- Schleicher, August. 1863. *Die Darwinsche Theorie und die Sprachwissenschaft* [Darwinian theory and linguistics]. Weimar: Hermann Böhlau.

- Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen* [On the genetic relations among the Indo-European languages]. Hermann Böhlau.
- Schmidt, Johannes. 1875. *Zur Geschichte des indogermanischen Vocalismus. Zweite Abteilung* [On the history of Indo-European vocalism. Second part]. 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.
- 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, Vol. 32, No. 3/4 (Fall - Winter, 1990)*, pp. 187-210 32(3/4). 187–210.
- Trask, Larry. 2000. Some issues in relative chronology. In Colin Renfrew, April McMahon & Larry Trask (eds.), *Time Depth in Historical Linguistics, vol. 1*, 45–58. Cambridge, UK: 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.
- Wichmann, Søren. 2017. Genealogical Classification in Historical Linguistics. In *Oxford Research Encyclopedia of Linguistics*, 1–19. Oxford: Oxford University Press. doi: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.