

## 2. PHYLOGENETIC CLASSIFICATIONS AND NETWORK APPROACHES IN LINGUISTICS



# DO LANGUAGES GROW ON TREES? THE TREE METAPHOR IN THE HISTORY OF LINGUISTICS

*Hans Geisler and Johann-Mattis List*

## THE RETURN OF THE TREES

Among biologist as well as linguists, it is now widely accepted that there are many striking parallels between the evolution of life forms and the history of languages. Starting from the rise of language studies as a scientific discipline in the early 19<sup>th</sup> century up to today's recent "quantitative turn" in historical linguistics, scholars from both disciplines have repeatedly pointed to similarities between the respective research objects in biology and linguistics. Of all these parallels, the use of *family trees* to model the differentiation of species (genomes and languages) is surely the most striking one. Methodically speaking, genealogical relations between linguistic and biological species are visualized with help of bifurcating trees which indicate the splitting of ancestral into descendant taxa. Being developed independently in linguistics and biology (Hoenigswald 1963), the tree model suffered different fates in both disciplines: While the reconstruction of phylogenetic trees successively became one of the key objectives in evolutionary biology, the tree model was controversially disputed in linguistics and – although to no time completely abandoned – never became a true part of the consensus.

Although linguists always had certain reservations regarding the tree model, it recently experienced a surprising revival. While earlier linguistic work on phylogenetic reconstruction was almost exclusively based on the intuitive weighting of features from very small samples of well-studied ancient languages, the integration of stochastic methods originally designed for biological applications made it possible to analyze large quantitative datasets automatically (Gray and Atkinson 2003; Atkinson and Gray 2006). Whereas tree construction has played for some time a minor role in historical linguistics, it has again become a specific field of historical linguistic endeavor in the last two decades (Pagel 2009: 414).

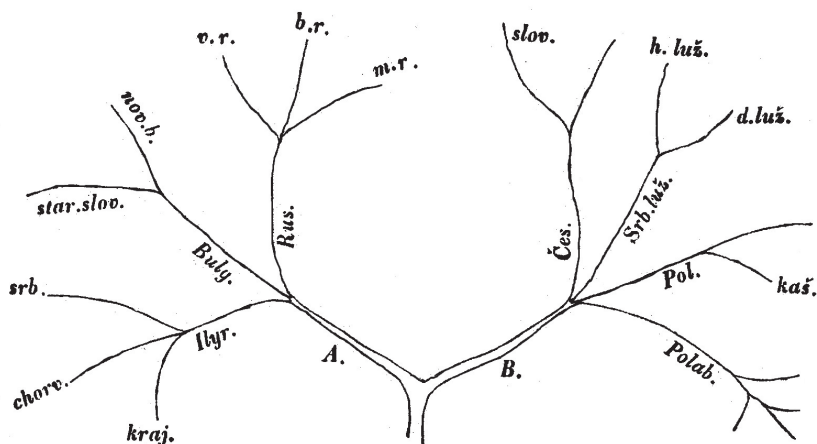


Figure 1: An early family tree of the Slavic languages by Čelakovský (1853)

In the following, we will give a short overview how the tree model emerged and made his way into 19th and 20th century linguistics. Though nearly ousted by competing alternative models like the wave theory for quite a long time, bifurcating trees has prevailed recently because they make it possible to represent logically dichotomous relations between species and languages which in turn allow for easy phylogenetic algorithmization. Unfortunately, the tree model's logical simplicity masks the complexity of biological and linguistic research objects in many respects. The most important surely is the wide neglect of language contact and ensuing borrowing. Thus, we claim that only combined approaches which describe both the vertical and the horizontal components of language relations are apt to depict the intrinsically distorted character of language change adequately.

## THE ORIGIN OF THE TREES

According to the current view in historical linguistics, one can roughly distinguish two different kinds of language relations: *genealogical language relations*, i.e. relations which are due to the common descent from an ancestor language, and *non-genealogical language relations*, i.e. relations which are a result of language contact. One of the key tasks of historical linguistic research is to find out whether resemblances between languages are a result of the former or the latter kind of language relationship. Otherwise, no language history could be drawn.

To infer whether specific resemblances between languages are due to contact or due to inheritance, however, is a complicated task, and in many cases there is no clear-cut procedure to discriminate between the two. The deeper one goes back in time, the greater becomes the problem of inference. Thus, German *Kopf* “head” and English *cup* “cup” probably go back to Proto-Germanic *\*kuppa-* “cup” (Orel 2003), yet whether the word was borrowed from Latin into Proto-Germanic (Kluge and

Seebold 2002) or inherited from Proto-Indo-European (Orel 2003) cannot be resolved with full confidence (see Figure 2). It is therefore not surprising that the idea that language relations can be divided into genealogical and non-genealogical ones was developed considerably late. Before the 19<sup>th</sup> century, the dominant view on language relations was non-genealogical. Discussions regarding the origination of languages were restricted to the biblical myth of the *Tower of Babel*.

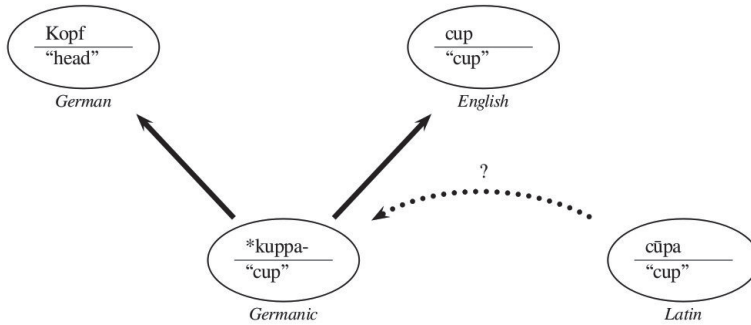


Figure 2: German *Kopf* and english *cup*: inheritance or borrowing?

### Early views on language relations

*Catastrophism* as opposed to *gradualism* (or *evolutionarism*) was a leading paradigm in scholarly thinking up to the early 19<sup>th</sup> century. While in fields like geology and biology the biblical creation myth accounted for the origin of the earth and the species (Christy 1983), the origin of all languages was explained by the biblical myth of the *Tower of Babel* according to which all languages originated as aberrations of a single language after the *confusion of tongues*. As in geology and biology the origin of a given diversity was explained as the result of an ad-hoc *catastrophic* event. All languages were assumed to be derived from the mysterious “Adamic Language” which became later directly identified with Hebrew. Consequently, the *Hebrew Paradigm* heavily influenced the way scholars would investigate language relations (Klein 2004). Since the monophyletic origin of all languages was apparently already proven by the religious dictum and linguistic change was characterized as an abrupt process of decay, the scholars mostly restricted their research to speculative etymological studies trying to show that all languages had inherited at least some words of Hebrew (cf., e.g., the work of Münster 1523, Reuchlin 1506, or Cruciger 1616).

While – held back by the Hebrew Paradigm – the genealogical perspective was only sporadically adopted and investigated. Scholars were well aware of the fact that languages can influence each other in many different ways. The non-genealogical perspective on language relations was the prevailing one in pre-19<sup>th</sup> century linguistics (Allen 1953: 55–7), and there are many examples in the literature, where scholars explicitly make use of non-genealogical explanations in order to explain specific resemblances between certain languages (see, e.g., Cratylus, *Institutio Ora-*

toria, Webb 1787). This is obviously due to the fact that language contact is fairly easy to recognize, not only for those who show a special interest in languages but also for “normal” speakers who are in contact to people who speak in different tongues.

### The discovery of tree-likeness

August Schleicher (1821–1868) is often regarded as the founding father of historical linguistics, being the one who established it as a real science (Fox 1995: 23–7). His two main contributions to historical linguistics were the method of linguistic reconstruction (Schleicher 1861) and the development of the tree model to visualize genealogical language relations.

Schleicher’s tree model (Schleicher 1853a and 1853b) is the cumulation of several findings which were made during the early 19<sup>th</sup> century. In this time, scholars such as Jacob Grimm (1785–1863) and Rasmus Rask (1787–1832), had detected that – in contrast to previous opinions – certain aspects of languages, namely their sound systems, did not change *chaotically*, but *regularly*, making it possible to compare different languages systematically for common traits (see Rask 1818, Grimm 1822). Along with Franz Bopp’s (1791–1867) independent detection of many grammatical resemblances between Sanskrit and many European languages such as Latin, Greek, and Gothic (see Bopp 1816), it seemed, for the first time, no longer possible to explain these similarities by accident or derivation, but only by a common origin of these specific languages. Abandoning the Hebrew Paradigm, and adopting the hypothesis that sound change was a regular process, scholars apparently had finally found a method by which it was possible to distinguish vertical from horizontal language relations.

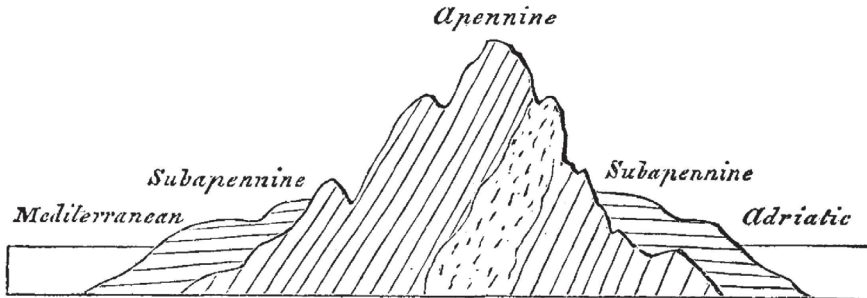
The *regularity hypothesis* upon which the new vertical thinking in linguistics was built summarizes three major characteristics of sound change which are already explicitly mentioned in Schleicher’s early work. According to these characteristics, sound change is a *universal*, a *gradual*, and a *law-like* process (cf. Schleicher 1848: 25). *Universality* implies that the process is independent of time and space, *graduality* implies that the process is neither abrupt nor chaotic, and *law-likeness* implies that the process is (to a great degree) exceptionless.

Perspective	Hebrew paradigm	Schleicher
heredity	sporadic	systematic
differentiation	singular	recurrent
change	chaotic	regular

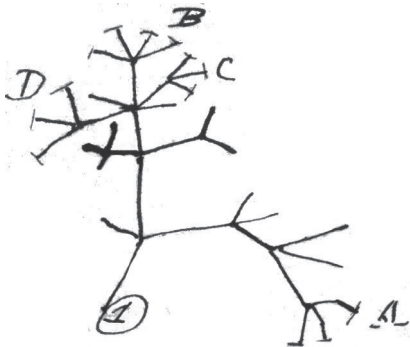
Table 1: The hebrew paradigm and Schleicher’s ‘Tree Model’

The new theory of vertical language relations which is directly reflected in the tree model, radically differs from the earlier conception of language relations within the Hebrew Paradigm: *change* is no longer seen as a *chaotic*, but as a *regular* process,

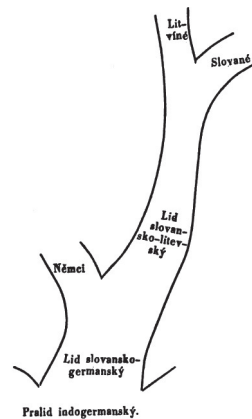
heredity is no longer believed to be a *sporadic* but a *systematic* phenomenon, and the *origination* of new languages is no longer identified with a *singular* event but as a process which repeatedly occurs during all times (see Table 1).



a) Lyell (1830)



b) Darwin (1837)



c) Schleicher (1853b)

Figure 3: Illustrations by Lyell (a), Darwin (b), and Schleicher (c)

### Common paradigms in Geology, Biology and Linguistics

At about the same time when linguists realized that the apparently chaotic and sporadic phenomena of language change were indeed universal and gradual, geologists and biologists came to similar conclusions in their own fields. Between 1830 and 1833 the English geologists Charles Lyell (1797–1875) published his multi-volume book *Principles of Geology* (Lyell 1830–1833) in which he substantiated the claim – first brought forward by James Hutton (1726–1797) – that the shape of the earth was the result of slow-moving and gradually operating forces which were acting independently of times and places. In 1859 the English biologist Charles

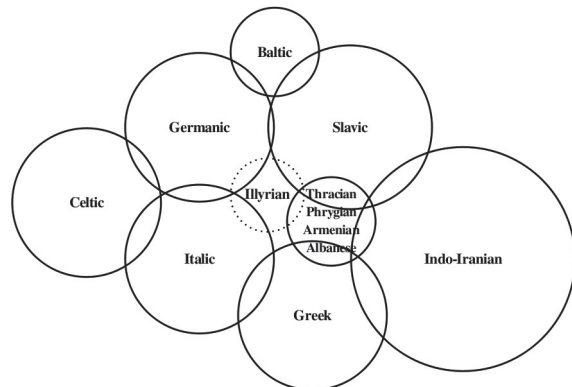
Darwin (1809–1882) published the famous book *On the Origin of Species* in which he first introduced the idea that the diversity of life was due to the universally and gradually operating force of natural selection (Darwin 1859). Whether these new ideas regarding the universality and graduality of certain processes in different disciplines were due to mutual influence or due to the spirit of the age: The new paradigm of *uniformitarianism* made it possible to reconstruct the prehistory of regions, species, and languages under the common slogan “The present is a key to the past”, and scholars from all three disciplines noticed and discussed the nature and the implication of the parallels they found in the different fields of research.

### Back to Dendrophobia

In contrast to evolutionary biology, where family trees became the leading paradigm for the description of species differentiation, the popularity of language trees soon began to fade in the newly established discipline of historical-comparative linguistics. In 1872 Johannes Schmidt (1843–1901) published the book *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen* (Schmidt 1872) in which he pointed to various problems regarding the applicability and the adequacy of the tree model. He pointed out that the data of the Indo-European languages did not suggest a simple tree-like differentiation. In order to account for his findings, he proposed the so-called *Wave Theory* according to which certain changes spread like waves in concentric circles over neighboring speech communities. Even two years earlier Hugo Schuchardt (1842–1927) criticized the assumption that languages simply split and then evolve independently (as suggested by the tree model), emphasizing that languages usually diverge gradually while at the same time mutually influencing each other: ‘We connect the branches and twigs of the tree with countless horizontal lines and it ceases to be a tree’ (Schuchardt 1870 [1900] 11)<sup>1</sup>.



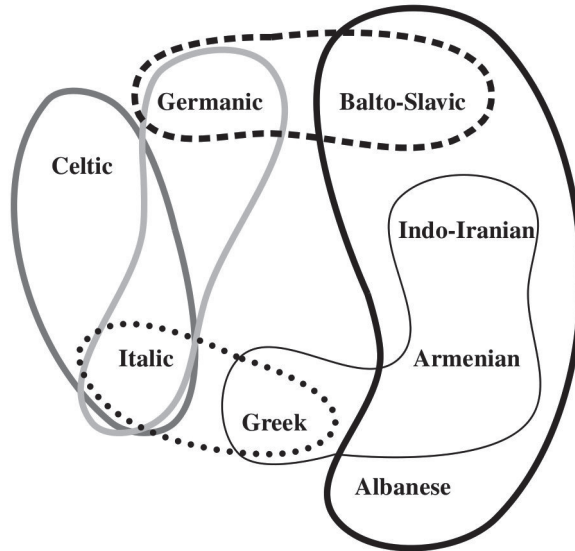
a) After Meillet (1908:134)



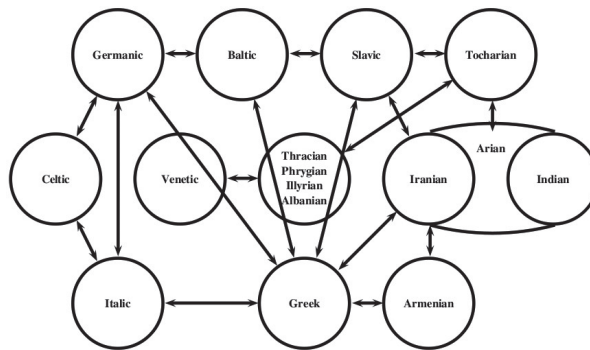
b) After Hirt (1905:93)

1 Translation of the authors, original text: ‘Wir verbinden die Äste und Zweige des Stammbaums durch zahllose horizontale Linien, und er hört auf ein Stammbaum zu sein.’





c) After Bloomfield (1933:316)



d) After Bonfante (1931:174)

Figure 4: The 'Wave Theory' in different visualizations

While most of the scholars were well aware of the inadequacy of the family-tree model, they had great difficulties in coming up with a conclusive alternative model which would describe and depict the complex reality of the phylogenetic history of languages in an equally simple and straightforward way. The fruitless quest for new metaphors is reflected in numerous different visualizations of Schmidt's Wave Theory ranging from simple geographical maps (Schmidt 1875: 199; Meillet 1908: 134), via overlapping circles (Hirt 1905: 93) or alternating boundaries (Bloomfield 1933: 316), up to networks (Bonfante 1931: 174), as illustrated in Figure 5.

What all these visualizations have in common is that they emphasize the spatial extension of languages which is neglected within the tree-model. At the same time,

however, the time dimension is sacrificed: Languages are arranged on a map and relations between the languages are marked, yet all relations are displayed as static differences, not as dynamic processes of differentiation. Surely, this lack of dynamicity was one of the reasons, why linguists never abandoned the family tree completely, but rather used both models in dependence of the respective problems they were dealing with.

### PROBLEMATIC ASPECTS OF THE ‘TREE MODEL’

The tree model can be criticized by questioning its practicability, its plausibility, or its adequacy. Although criticism regarding plausibility and adequacy seems to be stronger than criticism regarding practicability, most of the arguments which have been brought forward against the tree model belong to the latter kind.

#### Practicability of the model

Many of the early opponents of Schleicher’s *Stammbaum* disfavored the tree model because they experienced problems when trying to apply it. Most of these cases were due to conflicts in the data: Apparently, the tree model could not account for the distribution of common features in the descendant languages. Thus, applying a quasi-quantitative account, Schmidt (1872) listed words which were patchily distributed over the major Indo-European subgroups in support for his Wave Theory. In his counts, for example, there are 132 words which are reflected in both Latin and Old Greek but not in Old Indian, 99 words which are reflected in both Old Indian and Old Greek but not in Latin, yet only 20 words which occur in Latin and Old Indian but not in Old Greek. Following Schmidt’s line of thought, these counts contradict the tree model, since they suggest a strange pattern of closeness between the three languages where Old Greek is close to both Latin and Old Indian while Old Indian and Latin are only close to Old Greek (see Figure 5a).

However, this argumentation has a striking shortcoming, in so far as it ignores the *temporary status* of knowledge in the historical sciences. Thus, Schmidt’s estimations for common roots between Latin and Old Indian are considerably low. According to estimates drawn from Nicolaev (2007), there are 364 cognate words in Old Greek and Old Indian which are not reflected in Latin, 199 between Old Indian and Latin which are not reflected in Old Greek, and 379 between Latin and Old Greek which are not reflected in Old Indian (see Figure 5b). This shows that drawing conclusions from historical data is always preliminary. If a current state of knowledge disfavors the tree model, this doesn’t need to hold for future states.

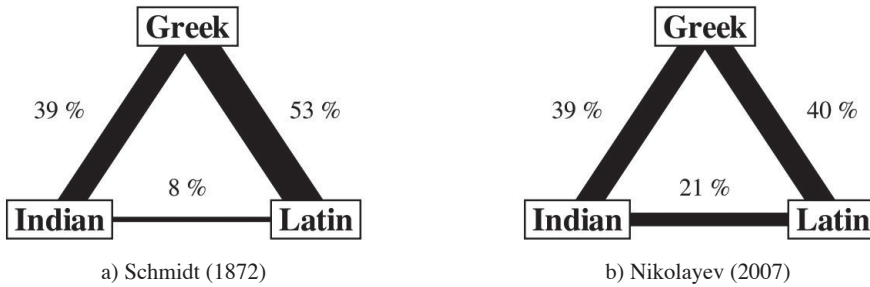


Figure 5: Cognate counts in old Greek, Latin, and old Indian

A further point which is mentioned by Schmidt himself is the impact of hidden borrowings. As it was mentioned before, the deeper one goes back in time, the more difficult it becomes to distinguish clearly between similarities due to genealogical and similarities due to non-genealogical relations. If the high amount of shared cognates between Old Greek and Latin turned out to be the result of the close contact between the two languages, this would, however, not contradict the tree model, it would only show that it is of crucial importance to disentangle vertical and horizontal relations before the reconstruction of family trees can be applied faithfully.

#### Plausibility of the model

From the above-mentioned one can conclude that pointing to the impracticability of the tree model cannot consistently prove its inadequacy, since the practicability of a given model can be overcome by advanced methods. Objections regarding plausibility, on the other hand, are much stronger, since they question the model itself.

The arguments raised by the opponents of the tree model come along with its obvious simplifications: When mapped onto a family tree, languages are reified and treated as discrete objects located in space and time. Language divergence is necessarily characterized as an abrupt event, and no reverse process of convergence is allowed. In reality such a situation is met only under rare circumstances when the speakers of a language separate geographically. Under normal circumstances, however, languages form areal continua of slightly diverging varieties. A strict separation of languages does only hold for distantly related, standardized, written languages. Lacking the geographical dimension, the family tree can neither model language divergence in all its complexity, nor can it account for the opposite process of convergence which eventually may even lead to hybridization. Thus, from what is known from studies on dialect geography and language divergence, there are obvious plausibility issues when trying to model language history with the help of trees only.

### Adequacy of the model

Apart from the apparent plausibility issues arising from Schleicher’s Stammbaum, the discontent of most linguists with family trees surely also results from their lack of adequacy. If the goal of historical linguistics is to describe realistically how languages evolve, it is surely not enough to simply point to their vertical history, since the horizontal aspects of language history are surely at least as – if not even more – characteristic for language history as the vertical ones. The tree model’s lack of expressiveness is surely one of the most important reasons for the general reluctance of linguists to draw phylogenetic trees: If family-trees are simply not realistic enough to depict what linguists know about the history of the languages they investigate, why should one even make the effort to reconstruct them?

### SPECIES EVOLUTION AND LANGUAGE CHANGE

The key assumption of the new approaches in historical linguistics is that the characteristic processes of language change and biological evolution are so similar that the methods designed for one discipline may also be used in the other one, despite the fact that the domains differ (Croft 2008: 225). The use of biological methods requires certain analogies to be made between linguistic and biological processes and entities. Table 2 lists some of the most common ones which can often be found in the literature. Thus, regarding the *unit of heredity*, the biological *gene* is usually set in analogy with the linguistic *word*, both being ‘discrete heritable units’ (Pagel 2009: 406). *Replication* of the heritable units is achieved via concrete mechanisms of *reproduction* in biological evolution and via *learning* in language history. From the perspective of *origination*, *cladogenesis* in biology is identified with *language splitting* in linguistics (ibid.). From the perspective of *change*, the driving forces of biological evolution such as *natural selection* and *genetic drift* are compared with *social selection* and *trends* eventually leading to language change (ibid.). Last not least, differentiation is usually assumed to be treelike, and the impact of “horizontal forces” on evolution is considered to be rather low in both cases.

Aspects	Species	Languages
unit of heredity	gene	word
replication	(asexual and sexual) reproduction	learning
origination	cladogenesis	language splitting
forces of change	natural selection and genetic drift	social selection, trends
differentiation	treelike	treelike (?)

Table 2: Some apparent parallels between species and languages

Assuming that these parallels hold, it seems perfectly plausible to use the methods developed for the application in one discipline in the other. However, it is important to be aware not only of the parallels but also of the differences between the research objects of both disciplines. The most striking difference between languages and genomes is that biological evolution manifests itself *substantially* while language history does not. In terms of Popper (1978), genome evolution and language evolution take place in different *worlds*: While biological organisms are part of *world 1*, the ‘world that consists of physical bodies’ (ibid. 143), languages belong to *world 3*, the ‘world of the products of the human mind, such as languages; tales and stories and religious myths’ (ibid. 144) which are replicated by learning.

Since we are dealing with very different domains here, the processes dominating in biological and linguistic evolution may also differ quite significantly. Thus, the *unit of heredity* in biology, the gene, is built from a set of universal characters which can be found in all organisms. The *unit of heredity* in linguistics, the word, however, is built from a set of sounds which are distinctive only with respect to the language they belong to. Unlike genes, words are not drawn from a universal alphabet, but from alphabets which themselves are subject to change. Therefore, biological methods which only work on global similarity, such as traditional alignment algorithms, necessarily fail to detect these specific similarities which are of interest to historical linguistics.

When using biological methods in linguistic applications, it is therefore important to be very cautious, and to check whether the parallels really hold or whether one is simply led astray by some first-glance lookalikes. It seems that the use of biological methods in historical linguistics is not always based on a thorough reflection regarding the question of comparability. Phylogenetic reconstruction, for example, is usually based on cognate-sets extracted from lexicostatistical wordlists. In these analyses, the wordlists, reflecting the so-called *basic vocabulary* of the languages under investigation (Swadesh 1955), are usually compared with the *core genome* in biology, i.e. they are supposed to represent the most stable, slow-changing, and least borrowing-prone part of a language’s lexicon. However, given the independence of word form and meaning, which does *not* hold for biology, there is no objective procedure to determine a language’s basic items. As a result, the creation of basic lists in linguistics is based on a manual procedure which could be shown to be very prone to errors in item translation (Geisler & List 2013) and the identification of borrowings (Nelson-Sathi et al. 2011).

## ALTERNATIVES TO THE 'TREE MODEL'

Given that the family tree is not sufficient to model language history in all its complexity, while the Wave Theory lacks the dynamicity of the tree model, remaining a mere static, map-like visualization of shared similarities, one may raise the question whether there are any other possibilities to display both genealogical and non-genealogical relationships between languages. Given that both tree and wave reflect certain aspects of language relations, the most straightforward alternative would be to combine both models in a network approach where both horizontal and vertical language relations are displayed (see Nelson-Sathi et al. this volume). Such a model has the advantage that it preserves of the tree model's dichotomous logic with clear-cut categorizations, but further allows fine-graded mapping of language contact. Although the idea of combining trees and waves has been developed very early in the history of linguistics, there are only a few attempts to visualize or formalize it (Southworth 1964; Holzer 1995), and it was only recently that a quantitative approach for the reconstruction of phylogenetic networks based on lexicostatistical wordlists has been proposed (Nelson-Sathi et al. 2011). Nevertheless, given the complexity of language history, combined networks of horizontal and vertical language relations seem to offer a promising alternative to both trees and waves in historical linguistics.

## CONCLUSION

In this paper we showed that during the history of linguistics the family tree never played a major role. Soon after the model was first introduced, scholars criticized the concept for its obvious shortcomings and proposed various other ways to model language history, none of which gained broad acceptance. The recent quantitative turn in historical linguistics which was initiated by the adaptation of new automatic methods initially designed for evolutionary biology led to an unexpected revival of the tree model in historical linguistics. Although the new methods doubtlessly decrease the amount of subjectivity inherent in the traditional intuitive approaches to phylogenetic reconstruction, they do not cope for the simplifying character of the tree model *per se*. In order to model language history in a realistic way, combining approaches which reflect the vertical as well as the horizontal aspects of language relations are needed.

## REFERENCES

- Allen, W.S. (1953) 'Relationship in comparative linguistics', in Transactions of the Philological Society, 52–108.
- Atkinson, Q.D. & R.D. Gray (2006) 'How old is the Indo-European language family? Illumination or more moths to the flame?' in P. Forster and C. Renfrew (eds), Phylogenetic methods and the prehistory of languages. McDonald Institute monographs (Cambridge/Oxford/Oakville: McDonald Institute for Archaeological Research): 91–109.

- Bloomfield, L. (1933 [1973]) *Language* (London: Allen & Unwin).
- Bonfante, G. (1931) 'I dialetti indoeuropei', in *Annali del R. Istituto Orientale di Napoli* 4: 69–185.
- Bopp, F. (1816) *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache. Nebst Episoden des Ramajan und Mahabharas in genauen metrischen Uebersetzungen aus dem Originaltexte und einigen Abschnitten aus den Veda's* (Frankfurt am Main: Andreäische Buchhandlung).
- Čelakovský, F.L. (1853) *Čtení o srovnací mluvnici slovanské* (Prague: Universitě pražské).
- Christy, C. (1983) *Uniformitarianism in linguistics. Studies in the history of linguistics* (Amsterdam/Philadelphia: John Benjamins).
- Cratylus: Plato (ca. 399 BC), English Translation, in Fowler, H.N., *Plato in Twelve Volumes* (London: William Heinemann).
- Croft, W. (2008) 'Evolutionary Linguistics', *Annual Review of Anthropology* 37, 219–34.
- Cruciger, G. (1616) *Harmonia linguarum quatuor cardinalium: Hebraicae Graecae Latinae & Germanicae: In Qua Praeter Summum Earum Consensum, acceptionumque propriarum ab impropris distinctionem, perpetua unius ab altera, origo perspicue deducitur* (Frankfurt, Tampach & Bringer).
- Darwin, C. (1837) *Notebook on Transmutation of Species*.
- Darwin, C. (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).
- Fox, A. (1995) *Linguistic reconstruction. An introduction to theory and method* (Oxford: University Press).
- Geisler, H. & J.-M. List (2013) 'Beautiful trees on unstable ground. Notes on the data problem in lexicostatistics', in H. Hettrich (ed) *Die Ausbreitung des Indogermanischen. Thesen aus Sprachwissenschaft, Archäologie und Genetik* (Wiesbaden: Reichert).
- Gray, R.D. & Q.D. Atkinson (2003) 'Language-tree divergence times support the Anatolian theory of Indo-European origin', *Nature* 426.6965: 435–9.
- Grimm, J. (1822) *Deutsche Grammatik*. 2nd ed. Vol. 1 (Göttingen: Dieterichsche Buchhandlung).
- Hoeningswald, H.M. (1963) 'On the History of the Comparative Method', *Anthropological Linguistics* 5.1: 1–11.
- Holzer, G. (1995) *Das Erschließen unbelegter Sprachen. Zu den theoretischen Grundlagen der genetischen Linguistik* (Frankfurt am Main: Lang).
- Hirt, H. (1905) *Die Indogermanen. Ihre Verbreitung, ihre Urheimat und ihre Kultur*. Vol. 1 (Strassburg: Trübner).
- Institutio Oratoria: Quintilian (ca. 50 AD) *Institutio Oratoria*. English Translation: J.S. Watson, (Iowa: Institutes of Oratory).
- Klein, W.P. (2004) 'Was wurde aus den Wörtern der hebräischen Ursprache? Zur Entstehung der komparativen Linguistik aus dem Geist etymologischer Spekulation', in G. Veltri & G. Necker (eds), *Gottes Sprache in der philologischen Werkstatt. Hebraistik vom 15. bis zum 19. Jahrhundert*, Proceedings of the Symposium "Die Geburt der Philologie aus dem Geist der Hebraistik" (Wittenberg, Oct. 6–6, 2002), *Studies in European Judaism* 11 (Leiden: Brill): 3–23.
- Kluge, F. (ed) (2002) *Etymologisches Wörterbuch der deutschen Sprache*. Cont. by E. Seebold. 24th ed. (Berlin: de Gruyter).
- Lyell, C. (1830) *Principles of geology, being an attempt to explain the former changes of the Earth's surface, by reference to causes now in operation*. Vol. 1. (London: John Murray).
- Meillet, A. (1922 [1908]) *Les dialectes Indo-Européens* (Paris: Librairie Ancienne Honoré Champion).
- Münster, S. (1523) *Dictionarium Hebraicum, nunc primum editum et typis excusum, adiectis Chaldaicis vocabulis non parum multis*. (Basileae: Apud Frobeniu).
- Nelson-Sathi, S.; List, J.-M.; Geisler, H.; Fangerau, H.; Gray, R.D.; Martin, W. & T. Dagan (2011) 'Networks uncover hidden lexical borrowing in Indo-European language evolution', *Proceedings of the Royal Society B*. 1713.278: 1794–803.
- Nikolayev, S. (2007) *Indo-European etymology*. available online: <http://starling.rinet.ru/cgi-bin/main.cgi>.

- Orel, V. (2003) *A handbook of Germanic etymology* (Leiden: Brill).
- Pagel, M. (2009) 'Human language as a culturally transmitted replicator', *Nature Reviews* 10: 405–15.
- Popper, K.R. (1978) 'Three worlds', *The Tanner Lectures on Human Values*: 143–67.
- Rask, R.K. (1818) *Undersøgelse om det gamle Nordiske eller Islandske sprogs oprindelse* (Copenhagen: Gyldendalske Boghandlings Forlag).
- Reuchlin, J. (1506) *Ioannis Reuchlin Phorcensis LL. Doc. ad Dionysium fratrem suum Germanum de rudimentis Hebraicis libri III* (Pforzheim: Thomas Anshelm).
- Schleicher, A. (1848) *Zur vergleichenden Sprachengeschichte* (Bonn: König).
- Schleicher, A. (1853a) 'Die ersten Spaltungen des indogermanischen Urvolkes', *Allgemeine Monatsschrift für Wissenschaft und Literatur*: 786–7.
- Schleicher, A. (1853b) 'O jazyku litevském, zvláště na slovanský. Čteno v posezení sekci filologické král. České Společnosti Nauk dne 6. června 1853', *Časopis Českého Museum* 27: 320–34.
- Schleicher, A. (1861) *Compendium der vergleichenden Grammatik der indogermanischen Sprache*. Vol. 1: *Kurzer Abriss einer Lautlehre der indogermanischen Ursprache, des Altindischen (Sanskrit), Alteranischen (Altbaktrischen), Altgriechischen, Altitalischen (Lateinischen, Umbrischen, Oskischen), Altkeltischen (Altirischen), Altslawischen (Altbulgarischen), Litauischen und Altdeutschen (Gotischen)* (Weimar: Böhlau).
- Schmidt, J. (1872) *Die Verwandtschaftsverhältnisse [sic!] der indogermanischen Sprachen* (Weimar: Böhlau).
- Schmidt, J. (1875) *Zur Geschichte des indogermanischen Vocalismus* (Weimar: Böhlau).
- Schuchardt, H. (1900[1870]) 'Über die Klassifikation der romanischen Mundarten', *Probe-Vorlesung, gehalten zu Leipzig am 30. April 1870* (Graz: Brevier): 166–88.
- Swadesh, M. (1955) 'Towards greater accuracy in linguistic dating', *International Journal of American Linguistics* 21.1: 121–37.
- Webb, D. (1787) 'Some reasons for thinking, that the Greek language was borrowed from the Chinese', in *Fourmont, Notes on the Grammatica Sinica of Mons* (London: Dodsley).