

# Form and formalism in linguistics

Edited by

James McElvenny

History and Philosophy of the Language  
Sciences 1



Series information: ./langsci/seriesinfo/hpls-info.tex not found!

# Form and formalism in linguistics

Edited by

James McElvenny



McElvenny, James (ed.). 2019. *Form and formalism in linguistics* (History and Philosophy of the Language Sciences 1). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/214>

© 2019, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/>

ISBN: 978-0-000000-00-0 (Digital)

no print ISBNs!

no DOI

Source code available from [www.github.com/langsci/214](http://www.github.com/langsci/214)

Collaborative reading: [paperhive.org/documents/remote?type=langsci&id=214](http://paperhive.org/documents/remote?type=langsci&id=214)

Cover and concept of design: Ulrike Harbort

Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono

Typesetting software: Xe<sub>La</sub>TeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

[langsci-press.org](http://langsci-press.org)

Storage and cataloguing done by FU Berlin

# Contents

<b>Preface</b> James McElvenny	<b>iii</b>
<b>1 Visual formalisms in comparative-historical linguistics</b> Judith Kaplan	<b>1</b>
<b>Index</b>	<b>35</b>



# Preface

James McElvenny

University of Edinburgh

Notions of “form” have a long history in Western thought on language. When linguistics emerged as an institutionalized discipline in the early decades of the nineteenth century, its practitioners could look back on a multitude of senses and uses of “form”, embedded in a variety of conceptual schemes. Even though many nineteenth-century linguists sought to emphasize the novelty of their work and imagined a radical break with the “pre-scientific” past (see [Morpurgo Davies 1998](#): chap. 1), both their everyday practice and their theoretical views were permeated by an intellectual inheritance stretching back over centuries, in which “form” occupied a central place.

On a practical level, “form” has long been employed in a general sense to refer to the perceptible outer appearances of linguistic expressions, especially in connection with the inflectional variants of words. On a deeper theoretical level, there has often been an effort to find underlying motivations for these appearances and so conceive of “form” in senses loaded with metaphysical and epistemological significance. This was the path taken by such movements as the medieval Scholastics and the Enlightenment-era General Grammarians (see [Law 2003](#): chaps. 8 and 11), whose successors in the nineteenth century – despite often disavowing their predecessors – were similarly engaged in a search for the cognitive, biological or aesthetic bases of linguistic form.

A particularly prominent figure in nineteenth-century discussions of form in language was Wilhelm von Humboldt (1767–1835), whose writings served as the point of departure for many later scholars. For Humboldt and his followers, there is a sense in which all language is form and nothing else, in that language is the representation we make of the world which, in Kantian fashion, we shape according to our perceptive faculties. “The essence of language”, writes [Humboldt \(1905 \[1820\]](#): 17), “consists in pouring the material of the phenomenal world into



the form of thoughts.” (*Das Wesen der Sprache besteht darin, die Materie der Erscheinungswelt in die Form der Gedanken zu giessen.*) A commonplace among the Humboldtians was to claim that each language has its own characteristic form of representation discernible in the form of its expressions. The task of the linguist is to capture these forms and analyse them for what they reveal about the mental, cultural and physical life of language speakers (see Morpurgo Davies 1998: chap. 5; Trabant 1986; McElvenny 2016).

The centrality of form to linguistic scholarship continued into the structuralist era. The *Cours de linguistique générale* of Ferdinand de Saussure (1857–1913) famously contains the assertion that “language is a form and not a substance” (*la langue est une forme et non une substance*) (Saussure 1922 [1916]: 169). Following on from the earlier Humboldtian position, a fundamental tenet of structuralism is to conceive of languages as self-contained structures imposed on the material substrate of the world. In describing phonological, grammatical and semantic apparatuses of languages, the structuralist is engaged in an investigation of linguistic form (for a classical structuralist account couched in these terms, see Lyons 1968: 54–70).

In the generativist era, Noam Chomsky’s (b. 1928) efforts to construct an intellectual genealogy for his work involved an attempted appropriation of Humboldtian “form”, which rekindled awareness of these ideas in mainstream linguistics. In his *Cartesian linguistics*, Chomsky (2009 [1966]: 69–77) sought to assimilate Humboldtian form to his own innovation of generative rules as the underlying system that allows for the creative use of finite means to produce an infinite array of expressions.

The fecundity of “form” is visible not only in its polysemy, but also in the family of derivatives it has brought into the world, including such terms as “formal”, “formalized” and “formalist/formalism”. Like their parent, these terms defy concise definition, although when applied as labels to directions in linguistic research they generally imply concentration on internal systematicity to the exclusion of external explanatory factors alongside an inclination to abstraction and axiomatization – two tendencies that may in fact manifest independently of one another (cf. Newmeyer 1998). As is explored in several contributions to this volume, formalism as a research mindset is at home in many fields – such as logic, mathematics, aesthetics and literary studies – and represents an area of rich historical cross-pollination between linguistics and other disciplines.

In a separate but related sense, “formalism” as a count noun refers to the devices employed in the representation and analysis of phenomena. Various formalisms in this sense, along with the theoretical views to which they are tied,



are also examined in the following chapters.

In composing this volume, we have come together as historians of science and philosophers of language and linguistics to take a critical look at notions of form and their derivatives, and the role they have played in the study of language over the past two centuries. We investigate how these notions have been understood and used, and what this reveals about the way of thinking, temperament and daily practice of linguists.

The first contribution to our volume is Judith Kaplan's examination in **chapter 1** of the role of visual formalisms in representing genealogical relationships between languages. Engaging with some of the latest literature on material culture in the history of science, Kaplan explores how visual diagrams and metaphors helped in grasping relationships between languages in comparative-historical grammar, from the nineteenth century up to the present day. She finds that the tensions between the dominant models of language relationship – “tree” versus “wave” models – were typically conceived in a visual mode, whether this was explicitly represented in a diagram or initially described only as a visual metaphor. She observes shifting commitments to the realism of representations and mutual influences between linguists and those working in neighbouring sciences.

In chapter 2, James McElvenny compares competing nineteenth-century accounts of “alternating sounds” – a cover term for the apparent unstable phonological variation found in “exotic” languages – for the different attitudes towards linguistic form that they reveal. The traditional view took alternating sounds to be a feature of “primitive” languages, which were assumed to have not attained the levels of formal arbitrariness characteristic of European languages. Franz Boas (1858–1942) famously refuted this view by insisting that all languages have fully developed phonologies and ascribed alternating sounds to perceptual error on the part of outside observers. Georg von der Gabelentz (1840–1893), on the other hand, embraced the phenomenon and wielded it against Neogrammarian doctrine, the leading formal theory of his day. Both Boas and Gabelentz' positions can claim a measure of theoretical sophistication and at the same time contain obvious faults. McElvenny places these positions in their historical context and considers why Boas' view was so well received in linguistics while Gabelentz' was not.

Chapter 3 turns to the links between linguistic, psychological and, above all, aesthetic theory in the work of Edward Sapir (1884–1939). In this chapter, Jean-Michel Fortis provides a detailed exposition of Sapir's writings on form in language, concentrating in particular on Sapir's notion of “form-feeling” and following the trail – in some places explicitly marked by Sapir himself and in others

reconstructed by Fortis through terminological and conceptual detective work – to identify his sources of inspiration. Fortis places Sapir in a finely interlaced intellectual network spanning across contemporary *Gestalt* psychology and German art theory, with a heritage extending at least as far back as the Romantic period around the turn of the eighteenth to the nineteenth century.

The focus on Sapir continues in chapter 4, where Els Elffers critically compares Sapir's philosophy of science to that of Jerry Fodor (1935–2017) and examines the implications of their views for the treatment of linguistic form. Looking at Sapir's arguments against the "superorganic" in language scholarship and Fodor's proposal for "token physicalism", she finds striking similarities between the two, despite their very different intellectual contexts: Sapir was responding to ideas in anthropology emerging from debates about the nature of the *Geisteswissenschaften* in contrast to the *Naturwissenschaften*, whereas Fodor was responding to logical positivism. Both scholars, however, concerned themselves with how best to demarcate the individual sciences, with the specific example of linguistics in mind, and settled on the principle of demarcating the sciences not according to their subject matter but the way in which that subject matter is conceived.

In chapter 5, Bart Karstens undertakes a re-examination of the genesis of linguistic structuralism and its early interaction with Russian Formalism, a school of literary analysis from the early twentieth century. Karstens engages in a detailed investigation of the scholarly network around Roman Jakobson (1896–1982) and his role as a vector for the transmission of Russian Formalism first to the Prague School of structuralism in the 1920s and then later in the United States. While formalist doctrine was often heavily criticized by the early structuralists, Karstens shows that various formalist views informed elements of early structuralism.

A similar story of "resistant embrace" is told in chapter 6, where John Joseph reconsiders the place of structuralism in French linguistics of the mid-twentieth century, before the onset of the "post-structuralist" period. Focusing on such figures as Emile Benveniste (1902–1976), Henri Meschonnic (1932–2000), Aurélien Sauvageot (1897–1988) and their closest contemporaries, Joseph demonstrates that each of these figures has a complex relationship to structuralism: at times criticizing the apparent premises of the approach while employing recognizably structuralist forms of analysis, or publicly avowing structuralism while straying away from its principles in their own work.

In chapter 7, Ryan Nefdt surveys some of the radical changes in theory that generative linguistics has undergone in its short history and derives from them positive lessons for the philosophy of science. Amid the turbulence and insta-

bility that has characterized generative theory, he identifies one constant: the formal structures in language that generative linguists describe. With the durability of this constant in mind, he advocates for a position of structural realism in the philosophy of linguistics. Such a position, he argues, would allow linguists to escape pessimistic meta-induction – that is, the notion that we must necessarily expect our theories to one day be refuted and superseded – and allows them to step away from the ontology of natural languages, thereby securing the epistemological basis of the formal approach to language.

The gaze of the last two chapters in our volume is largely directed towards current questions in the philosophy of linguistics, specifically the role of normativity and authority in language description. After first tracing the origins of generative grammar in formalist approaches to logic, Geoffrey Pullum, in chapter 8, develops a new perspective on the classical distinction between prescriptivism and descriptivism. He contends that the value of a grammatical description lies in the precise, formalized account it provides of a particular set of linguistic practices, which can guide those who may wish to participate in those practices. In serving as a guide, every grammar has normative force, but is not necessarily prescriptive: the grammar-reader may follow its advice but is not compelled to do so.

In chapter 9, Nick Riemer identifies the ideologies of language he sees embodied in the “unique form hypothesis”, the assumption that every linguistic expression can be reduced to a single, universally agreed underlying representation. While linguists might seek to distance themselves from this hypothesis and its implications, it is, argues Riemer, a recurring motif in linguistics, especially prominent in the teaching of the discipline. Its effects in education are particularly pernicious, since teachers, due to the exigencies of pedagogy, can usually offer no justification for the unique forms they present to their students other than arbitrary authority, a practice that reinforces unreflective submission to authority of all kinds, both at university and in life. Acknowledging that most linguists would shudder at such consequences, Riemer pleads for greater open-mindedness among linguists towards critique of the discipline’s foundations.

Although dealing with a broad range of topics from diverse perspectives and in different styles, this volume is the product of concerted collective effort. Each of us came to this project with existing ideas about form and formalism in linguistics. These ideas we set out in draft chapters, which we discussed in person at a meeting in Edinburgh in August 2018. After our meeting, we revised the chapters to reflect the insights gained through our discussion. It is these revised chapters, shaped and harmonized by our dialogue, that are contained in this volume.

## References

- Chomsky, Noam. 2009 [1966]. *Cartesian linguistics: a chapter in the history of rationalist thought*. James McGilvray (ed.). Cambridge: Cambridge University Press.
- Humboldt, Wilhelm von. 1905 [1820]. Über das vergleichende Sprachstudium in Beziehung auf die verschiedenen Epochen der Sprachentwicklung. In Albert Leitzmann (ed.), *Wilhelm von Humboldts gesammelte Schriften*, vol. IV, vol. IV. Berlin: Behr.
- Law, Vivien. 2003. *The history of linguistics in Europe from Plato to 1600*. Cambridge: Cambridge University Press.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- McElvenny, James. 2016. The fate of form in the Humboldtian tradition: The *Formungstrieb* of Georg von der Gabelentz. *Language and Communication* 47. 30–42. <http://dx.doi.org/10.1016/j.langcom.2015.12.004>.
- Morpurgo Davies, Anna. 1998. *History of linguistics: Nineteenth-century linguistics*. London: Longman.
- Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, Massachusetts: MIT Press.
- Saussure, Ferdinand de. 1922 [1916]. *Cours de linguistique générale*. Charles Bally & Albert Sechehaye (eds.). Paris: Payot.
- Trabant, Jürgen. 1986. *Apeliotes, oder der Sinn der Sprache: Wilhelm von Humboldts Sprach-Bild*. München: Fink.

## Chapter 1

# Visual formalisms in comparative-historical linguistics

Judith Kaplan

University of Pennsylvania

This paper examines visual formalisms in comparative-historical linguistics from the perspective of the history of science. It shows that visual aids representing key understandings of language relationship have followed on pre-existing visual metaphors. Using this observation to pry open canonical metaphors of language relationship, it traces the ways in which these “visualizations” have both consolidated existing research programs and opened up new lines of inquiry for students and recent advocates of phylogenetic methods.

## 1 Introduction

From the interiority of brain atlases to the distant topography of Mars, from the intimate realm of nano-images to the global modelling of climate data, a recent swell in computerized visualization techniques is transforming the face of scientific research, pedagogy, and generalist publications. Commenting on this trend in 2014, Lorraine Daston judged computer simulations to be “the greatest revolution in scientific empiricism since the canonization of observation and experiment in the late seventeenth century” (Daston 2014: 321). These developments, moreover, have had a profound impact on scholarship in Science and Technology Studies: many have hailed the growing sophistication of digital visual culture as an opportunity to re-think classical theories of scientific representation. Crucially, their efforts have emphasized the *materiality* of scientific representation in a turn away from questions of truth-as-correspondence and social infrastructure. For this new generation, the “material enactments” of scientific images are to be taken just as seriously as the embodied practices and community norms



surrounding them, and a good deal more seriously than any quest to faithfully represent the natural world (Coopmans et al. 2014: 3; see also Kusakawa 2016).

How might these conversations relate to the formalisms of comparative-historical linguistics?<sup>1</sup> Like economics, chemistry, and molecular biology, to name a few arenas, diachronic linguistics has embraced and disseminated a raft of colourful and complex data visualizations since the early 2000s (see, e.g., Gray et al. 2009). Thoughtful critics have entertained the possibility that a belated turn to phylogenetic modeling (an iterative statistical approach to genealogical classification thought to have revolutionized biological systematics during the 1970s) has made “tree thinking” viable again – vindicating the model for purposes of linguistic inquiry (Lopez et al. 2013). But we may equally well use the opportunity occasioned by this surge in tree thinking to reflect on the status of visual culture and epistemology in the language sciences more generally. Looking at canonical visual topoi for understanding language relationship over roughly the last 150 years, my chapter attempts to do just that.

Two points emerge from this line of questioning. First, the paper shows that well-known diagrams of language relationship derive from pre-existing verbal descriptions: words came first and were subsequently elaborated by pictures.<sup>2</sup> This chronology illuminates George Lakoff’s distinction between “conceptual” and “image” metaphors in complicated ways (Lakoff 1987), where “metaphor” itself is defined basically as a way of “understanding and experiencing one kind of thing [...] in terms of another” (Lakoff & Johnson 1980: 455). First, it shows that conceptual metaphors – understood to be systematic, quotidian, and extendable – and image metaphors – more limited in scope, hewn from conventional mental images, and characterized by “one-shot mapping” – often interact and define one another. This helps to explain the layering of arboreal and genealogical metaphors of language relationship, for instance, in comparative-historical linguistics. This double representation shows an inclination to capture the relatively abstract (the genealogical) with the relatively concrete (the arboreal). Furthermore, it helps us to understand that the fundamental metaphor here for conceptualizing relationship is LANGUAGE IS A LIVING THING. But even beyond this framework, the example suggests that image metaphors can become progressively conventional over time, to the point where they *are* systematic, quotidian,

---

<sup>1</sup>See James McElvenny’s preface to this volume. This chapter takes formalisms to be those “devices employed in the representation and analysis of phenomena” (p. iv), as he elaborates.

<sup>2</sup>This point generalizes to the history of biology. Here, Peter Simon Pallas (1741–1811) is credited with originating the tree of life (see his *Elenchus Zoophytorum*, 1766), though the visual was at that point purely descriptive, not diagrammatic. The diagrammatic rendering of Pallas’ idea came some 63 years later (Eichwald 1829).

and extendable over a wide-range of phenomena. This point takes on extra significance in the context of science, where the extension of conventional mental images has contributed to the development and communication of theories about the way the world actually works (see [Boyd 1979](#): 357).

In these respects, historical linguistics looks very much like molecular biology. Natasha Myers, tackling the latter tradition, traces the mechanical interventions of structural biology and bio-engineering back to the circulation of machinic metaphors for life during the 1870s ([Myers 2015](#), “Introduction”). They provided Thomas Henry Huxley, analogizing between the “protoplasmic theory of life” and the “‘horology’ of a clock”, with a bridge from the “visible tangible and manipulable world” of everyday life to an “invisible, intractable world of biological molecules” ([Huxley 1880](#); [Myers 2014](#): 157). Myers ultimately presents the idea of the “molecular machine” as a powerful “material-semiotic actor” capable of directing practitioners – initiates, especially – to travel certain lines of experimental inquiry ([Haraway 1991](#); [Myers 2014](#): 165–168).

The development and persistence of foundational metaphors for vertical (LANGUAGE IS A LIVING THING, entailing a genealogical concept of relationship) and horizontal (LANGUAGE IS A PHYSICAL THING, entailing a proximity theory of relationship) transfer in linguistics exhibit characteristics that are like the ones Meyers describes. They are hybrid in nature (both verbal and visual) and they feature prominently in texts that have served the consolidation of disciplinary knowledge. Attending to these similarities points up the enduring significance of texts, alongside material culture, in the history of science.

To focus on trees alone – rumoured to be the “most universally widespread of all great cultural symbols” ([Pietsch 2012](#): 1) – might sustain conclusions on the specific conceptual ramifications of “biosystematic iconography” (see, e.g., [Pulgram 1953](#): 69) and illuminate large scale patterns of change over time – from realism to anti-realism, and lately back again. Indeed, numerous studies of these phenomena already exist (see, e.g., [Southworth 1964](#); [Hoenigswald & Wiener 1987](#)). Instead, by drawing together trees and their alternatives in what follows, I hope to show how such representations highlight certain notions of relationship while removing others from view.

This, then, is the second point of the paper: visual metaphors and visual aids of language relationship matter a great deal because they constrain objects and programmes of research. I aim to establish this point by showing how difficult it has been for linguists to visualize vertical and historical relationships at the same time. While the advent of algorithms like NeighborNet in 2003 purported to give researchers the tools needed to see variation within hierarchy – forests

and trees – there is a much longer history of failed attempts to capture both kinds of relationship in a single visualization (Bryant & Moulton 2004). Select examples of this tension are woven throughout my discussion of canonical types in what follows.

These points are developed over the course of five sections. First, I consider classificatory diagrams lacking figural elaboration, in other words, lists and tables (on diagrams, see Bigg 2016). My hope is that this starting point will denaturalize the turn to tree thinking in the second half of the nineteenth century. The following section then introduces the dominant visualizations for understanding historical relationship – trees and waves – manifest in works that have gone on to have canonical status in pedagogy and historiography. Complementing previous studies of these texts, this part of the paper emphasizes the relationship between words and images and highlights conceptual problems that were encountered during the 1870s in bringing trees and waves meaningfully together. Next, I look at the uptake of these visualizations in twentieth-century textbooks. How were they introduced, drawn, and qualified? What attempts, if any, were made to see “the wave process and the splitting process” simultaneously? Section four entertains the possibility that computational models offered a new bifocal lens on these processes of linguistic differentiation. The conclusion offers a few ideas about the benefits of integrating the historiography of science and linguistics when it comes to specifically *visual* formalizations.

## 2 From tables to trees

MultiTree is a “digital library of language relationships” that was launched in 2006, funded by the United States National Science Foundation in 2012, and hosted by LINGUIST List as recently as 2018 (<http://new.multitree.org>). The stated aim of the project is to facilitate research in historical linguistics, “representing the most complete collection of language relationship hypotheses in a user-friendly, visually-appealing, and interactive format”. As with many such projects, it is a resource with ambitions vis-a-vis expert, interdisciplinary, and public audiences alike. While the visualizations presented on the site may be young (among other innovations, users can “climb” branches to view trees from individual nodes in rectangular and radial layouts), the data is often rather old. A search on “Mayan”, for instance, retrieves a potentially interactive visualization of a classificatory note composed by the Swiss-American ethnologist A. S. Gatschet in the mid-1890s, shown in Figure 1. The entry is quite simple: it pictures Mayan as a root node linked to six sub-groups (Huastec, Maya proper, Tzental, Mam, Quiché, and



Pokom).

What is curious about this presentation is how unmotivated it makes the tree actually seem. Whereas science studies scholars have directed painstaking attention to the implications of rooted (versus unrooted) trees, top-to-bottom (versus left-to-right) orientation, branching patterns, and the like, the manipulability of MultiTree undercuts all authorial intentionality on such fronts. Moreover, the original publication of Gatschet's account of the "Maya Linguistic Family" holds nary a tree – a hierarchical outline format using Roman and Arabic numerals sufficed just as well for his classificatory purposes (Gatschet & Campbell 1973: 250–251). As Pietsch and others have noted, early arboreal representations merely translated tables into trees, perhaps explaining some authors' preference for their growth from the left-hand margin of a printed page (Wells 1987: 51; Pietsch 2012: 7–10; Archibald 2014: 57).<sup>3</sup>

Why bother layering the biosystematic metaphor on top of the familial (see Wells 1987: 49 on "mixed metaphors", 53–54 on biological imports)? When it comes to MultiTree, this choice not only fosters comparability across the database, it also reflects architects' stated commitments to aesthetics, access, and "fun". In other words, the visual is second nature to those already familiar with the techniques of comparison and sub-grouping, and it recruits potential newcomers to those particular methodological approaches.

This brief example is meant to suggest that tree thinking did not have a necessary or inevitable trajectory in comparative-historical linguistics. Trees were not the only means available for the organization of information on ancestor-descendant relations.<sup>4</sup> Rather, they have served additional rhetorical purposes, taken up in the next section.

---

<sup>3</sup>This is to say nothing of the local influence from stemmatics in linguistics, dating back to the sixteenth century, where trees typically drop down in a branching pattern from an original manuscript positioned at the top of the page (see e.g. Maher 1966; Hoenigswald 1975; Cameron 1987). Setting questions of priority aside, Müller (1913 [1891]: vol. I, 537) encourages reflection on the visual culture of linguistics in the late nineteenth-century, as his *Lectures on the Science of Language* includes both genealogical trees and tables. In this example, the table gives Müller more space for textual elaboration – allowing him to differentiate, for instance, between "living" and "dead" languages, and to layer vertical groupings, reflecting geography, on top of horizontal brackets, reflecting genetic affiliation.

<sup>4</sup>The case in biology on this point is somewhat different. In the case of Lamarck's diagrams, for example, the shift from tables (e.g. 1778) to trees (e.g. 1809) coincided with a definite conceptual shift (re. species mutability). The recognition of variation and change over time did not correspond with the visual in linguistics (Archibald 2014: ch. 3).



1853 was a pivotal year with regard to the visualization of language relationship

<sup>5</sup>In fact, the “Arbre Généologique” by Felix Galet (ca. 1807) is often cited as the first tree of language relationship (Hellström 2012: 242). This chronology roughly aligns with the biological context, where the first known tree diagram of relationship was published by Augustin Augier in 1801.

weighty arboreal realism (see Maher 1966; Hoenigswald 1975; Koerner 1987). Significantly, Geisler and List suggest that this diagram was a formalization, after the fact, of Schleicher's first published reflections on the comparative history of languages some five years earlier (Schleicher 1848). Identifying relationship with descent, and pinning that conception on the tree, they assert, Schleicher's "new theory of vertical language relations [was] directly reflected in the tree model" that has since become so familiar (Geisler & List 2013: 114).

What more did Schleicher invest in his "schema" than a procedure for classification by two-way splits? Historians have emphasized notions of parsimony, regularity (Geisler & List 2013: 117, 114–115), organicism (Wells 1987: 56), and programmatic ambition (Koerner 1975: 755). More concretely, Schleicher told readers directly that branch length served as an indicator of "duration" and that the distance between branches was meant to indicate "degrees of relationship", left otherwise undefined (Schleicher 1853: 8). Going further, Schleicher emphasized the importance of his trees for training newcomers to the field, and their projected departure from older philological traditions:

In the present work an attempt is made to set forth the inferred Indo-European original language side by side with all extant derived languages. Besides the advantages offered by such a plan, *in setting immediately before the eyes of the student the final results of the investigation in a more concrete form, and thereby rendering easier his insight into the nature of a particular Indo-European language*, there is, I think, another advantage of no less importance, namely that it shows the baselessness of the assumption that the non-Indian Indo-European languages were derived from Old-Indian (Sanskrit), an assumption which has not yet entirely disappeared. (Schleicher 1967 [31871]: 94)

My emphasis on the first "advantage" described in this passage, the pedagogical advantage, presents the tree diagram as tool for summing up and disseminating research findings to those just entering the field — for Schleicher it was decidedly *not* a means to new linguistic knowledge. References to the "eyes" and "insight" of the student recall Daston's (2008) depiction of the "all-at-once-ness" of disciplined perception, seen here to be very much in the making through the association between pedagogy and disciplinary differentiation. That said, phylogeny does not appear to be a primary goal, in and of itself. Rather, it is celebrated as a means to better understand a "particular" language under investigation. Schleicher's philosophy of science did not necessarily demand knowledge of a general sort (Nyhart 2012). His visual epistemology appears to have involved

a kind of inward tendency, from sight to insight, both in the cultivation of the student and the discipline.

By the early 1870s, the outlines of the Indo-European family had been drawn, giving comparativists considerable cause for celebration. Nevertheless, exceptions persisted. As a young professor of German and Slavic at the University of Bonn, Johannes Schmidt (1843–1901), a student of Schleicher's, tackled these difficulties head-on. In a 31-page monograph on *The Relationships of the Indo-Germanic Languages* [*Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*], he demonstrated that unique resemblances can be identified between any two Indo-European branches, and that these tend to increase with geographic proximity. In light of this observation, he argued that linguistic changes spread laterally like waves on a pool of geographically distributed speech, rather than vertically, through a process of strict cleavage and differentiation. With each change distributed individually, he projected an image of successive waves moving out and interacting from a variety of centers – a network of linguistic features differentiated through space.

Though Schmidt did not give readers a diagram of his *Wellentheorie* in 1872, he did picture it in words.<sup>6</sup> With the following passage, Schmidt invited readers to join him in an image metaphor:

If we want now to represent the relationships of the Indo-Germanic languages in a picture that illustrates the origin of their diversity, then we must completely abandon the idea of a family tree. I would like to *put a picture of a wave in its place*, which diffuses concentrically with the distance from the mid-point in ever weaker rings. It does not matter that our language area makes no circle, rather a circle-sector at best, with the most primitive language at one end, not the centre [...] There were not initially any boundaries between languages within this domain, two arbitrarily distant dialects, A and X, were connected to each other by continuous varieties, B, C, D etc. [...] (Schmidt 1872: 27–28)<sup>7</sup>

<sup>6</sup>The critique of Schleicher's tree thinking through alternative visual metaphors came even earlier in Hugo Schuchardt's 1870 lecture "On the Classification of the Romance Dialects", where he speaks of killing the tree by binding together numerous branches and twigs with "horizontal lines" (Schuchardt 1900 [1870]: 11).

<sup>7</sup>"Wollen wir nun die verwandtschaftsverhältnisse der indogermanischen sprachen in einem bilde darstellen, welches die entstehung irer verschiedenheiten veranschaulicht, so müssen wir die idee des stammbaumes gänzlich aufgeben. Ich möchte an seine stele das bild der welle setzen, welche sich in concentrischen mit der entfernung vom mittelpunkte immer schwächer werdenden ringen ausbreitet. Dass unser sprach-gebiet keinen kreis bildet, sondern höchstens einen kreissektor, dass die ursprünglichste sprache nicht im mittelpunkte, sondern an dem

With references to “pictures”, geometry, and dialect labels, this passage reads as though it were captioning a printed diagram, though Schmidt did not provide one at first to accompany the text. Indeed, he went on to challenge the visual altogether – asserting the priority of linguistic data over any such formalization. As far as he was concerned, “[p]ictures have only marginal value in science, and if the one chosen here is displeasing to someone, he can replace it at will with something better without changing the results of the foregoing analysis”.

Perhaps this attitude partly explains why his first attempt to provide a visual aid lagged some three years behind his introduction of the image metaphor. Perhaps this reluctance derived from problems inherent to the visualization of horizontal relationship. Geisler and List suggest as much through their side-by-side presentation of several “fruitless” attempts to draw an alternative to Schleicher’s trees – from overlapping circles (Hirt 1905), to the spokes of a wheel (Meillet 1908), to early networks (Bonfante 1931), and alternating boundaries (Bloomfield 1933). I turn now to the challenging case of another visual metaphor that attempted to capture vertical and horizontal relationship simultaneously.

J. Heinrich Hübschmann (1848–1908) heeded Schmidt’s call to data-driven analysis in his comparative work on Armenian. Like Schmidt, Hübschmann studied with Schleicher at the University of Jena, though he later completed his degree under the Iranian philologist Martin Haug (1827–1876) at the University of Munich. Having defended a dissertation on Avestan and Old Persian philology in 1872, he turned in the next three years to an investigation of the relationship between Iranian and Armenian in pursuit of his Habilitation at the University of Leipzig. Initially, Hübschmann was only interested in Armenian insofar as it contributed to an understanding of the internal phylogeny of the Iranian family of languages. In line with the general consensus of the day, Hübschmann had, up to this point, taken a high degree of shared vocabulary as evidence of the fact that Armenian was part of Iranian (Schmitt 1976).

His postdoctoral studies culminated in the 1967 [1875] paper “On the Position of Armenian in the Circle of the Indo-Germanic Languages” [*Ueber die Stellung des Armenischen im Kreise der indogermanischen Sprachen*]. Here Hübschmann sorted out non-native parts of the lexicon and analysed the remaining “native” words, which allowed him to identify strata in the historical development of the language. His findings compelled his colleagues to recognize Armenian as an independent branch of the Indo-European family – not a sub-group of Iranian

---

einen ende des gebietes ligt, tut nichts zur sache [...] Sprachgrenzen innerhalb dises gebietes gab es ursprünglich nicht, zwei von einander beliebig weit entfernte dialekte des selben A und X waren durch continuierliche varietäten B, C, D, u.s.w. mit einander vermittelt.”

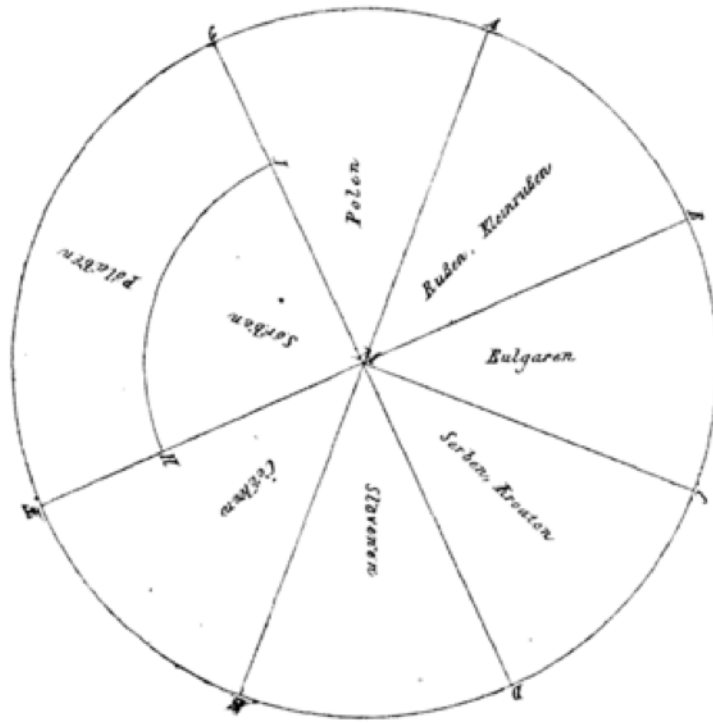


Figure 2: Schmidt (1875: 199). The text goes on to tell readers that the radia between lettered points should be read as isogloss lines, carving out dialects like pieces of a pie.

after all.

In the 1967 [1875] essay, Hübschmann demonstrated an extensive political history of contact and exchange between native Armenian-speaking and Iranian-speaking groups, suggesting that the similarities traditionally invoked in support of the “prevailing” view of common descent were in fact derived from the sector of borrowed, not inherited, vocabulary. Working back and forth between etymological and phonological evidence, Hübschmann next established provisional sound laws unique to Armenian, which undermined long-standing confidence in what were thought to be cognates between the languages in question. From the lexicon, Hübschmann then moved to grammatical considerations – the inflectional morphology of Iranian and Armenian, which exhibited many surface similarities. He attributed these to processes of analogical change – a psychological process of association tending to regularize words with similar meanings or inflectional paradigms, a mechanism of convergence.

Adding further phonological evidence to the balance against family relationship in this case, Hübschmann came to a fairly radical position, one that prioritized horizontal over vertical relationship:

Through the last part of our investigation, such a tight bond has without question been constructed between Armenian and European that it would be easier to tear Armenian from Aryan than from European. Among the European languages it stands closest to Balto-Slavic [...] In this situation, friends of the family tree [...] will certainly be inclined to separate Armenian completely from Aryan and make it a purely European language. Against this view I might first refer to the fact that Armenian does not take part completely in the split of *a* and *r* [...]

If further research makes this conclusion definitive, then the impossibility of setting up a family tree of the Indo-European languages would be strikingly demonstrated. For Armenian would be the connecting ring of both parts in the chain of the Aryan-Balto-Slavic languages, not a branch between two branches. And then too the family tree, which Johannes Schmidt's vigorous might has overturned, would remain lying forever [...]

But if Armenian is to be the connecting member between Iranian and Balto-Slavic, between Aryan and European, then in my opinion it must have played the role of an intermediary at a time when they were still very similar to one another, when the historical period had not yet drawn the present sharp boundary between them, but when they were still related to one another as dialects. (Hübschmann 1967 [1875]: 183)

Like all the examples encountered thus far, Hübschmann appealed first to a visual metaphor in this passage rather than a visual aid. Further, his discussion highlights inherent difficulties in drawing vertical and horizontal relationship together. To see Armenian as a link in the chain between Aryan and European, it was necessary to focus in on “*a time* when they were still very similar”. If trees lacked, fundamentally, a feeling for spatially distributed variation, waves were completely without a sense of timing.

Though this painstaking work secured Hübschmann's reputation as the “father” of modern Armenian linguistics – a doubly genealogical claim – previous accounts have not had much to say about the degree to which his positivism paradoxically threatened to topple his faith in the comparative method. In a paper “On the pronunciation and transcription of Old Armenian”, published the following year, Hübschmann pushed Schmidt's visual still further:

[...] it seems that languages can have similar sound systems without being related to one another, that the sound system of a language by others, i.e. local influences can be assumed, leading one to infer the congruencies between the sound systems of two languages less from their origin as from their local gathering. This statement seems to me for the determination of the genealogy of languages to be important and in linguistics to reward further success than heretofore was the case [...] If Iranian languages on the border of India show Indic sound similarities, must one therefore believe that they stand nearer to the Indic than the other Iranian languages? (Hübschmann 1876: 66)<sup>8</sup>

Evidently, the dictates of historical fidelity required taking vertical *and* horizontal relationship into consideration. But this proved remarkably difficult to capture visually. Hübschmann's best practice was to toggle back and forth between the two.

## 4 Metaphors and visual aids in twentieth-century textbooks

The canonical topoi just considered enjoyed a hearty afterlife in the *Disziplin-geschichte* of the late nineteenth century, and its review in textbooks thereafter. This section looks at the deployment of visual aids in that genre, building on previous studies of print and pedagogy in the history of chemistry, physics, and biology (Bertomeu-Sánchez et al. 2002; Kuhn 1962; Hopwood 2015). This literature has shown how scientific textbooks specialized from the late eighteenth century on, emphasizing their “use in formal teaching and their pedagogical and scientific authority”; their significance for disciplinary self-fashioning; and their “major role” in the making of *interactional expertise*, that is “the worldviews of citizens, what they know, what they do, what they are” (Simon 2016: 475, 479;

---

<sup>8</sup>“Aus alledem ergibt sich, dass Sprachen das gleiche Lautsystem haben können, ohne miteinander verwandt zu sein, dass das Lautsystem einer Sprache von äusseren, d. h. lokalen Einflüssen bedingt sein kann, und man aus der Gleichheit des Lautsystems zweier Sprachen weniger auf ihre Verwandtschaft als auf ihr locales Beisammensein zu schliessen hat. Dieser Satz scheint mir für die Beurtheilung der Verwandtschaftsverhältnisse der Sprachen wichtig zu sein und in der Linguistik mehr Beachtung zu verdienen als es bisher der Fall war. Also, um unsern Satz auf einen Fall anzuwenden: wenn iranische Sprachen an der Grenze Indiens indische Lauteigen-thümlichkeiten (durch den Gebrauch von Lingualen und Aspiraten) zeigen, hat man darum zu glauben, dass sie dem Indischen näher als die andern iranischen Sprachen stehen?”



Johns 1998: 406–408).<sup>9</sup>

Leonard Bloomfield’s *Language* met all of these criteria: it served as a provocative introduction to descriptive linguistics in 1933, asserting a new program while disciplining perception (Bloomfield & Hoijer 1965: v-vi). Comparative-historical material, notably, bookends the text: it appears first as a kind of “preface history”, recounting progress toward the modern “scientific” study of language, and in the second half of the book, aligning with Bloomfield’s priorities and programmatic vision. Far from a straightforward reproduction of earlier works, the presentation of historical research in the later part was designed for American students – those just beginning linguistics “who often d[o] not have the background in Indo-European languages” necessary to “understand texts that present methodology very largely in terms of concrete problems drawn from the older Indo-European languages” (Bloomfield & Hoijer 1965: vi). Put differently, the update shifted from exemplars to models, in line with what could reasonably be assumed of a new generation of students.

What, then, was textually self-evident, and what did Bloomfield think needed elaboration through the use of visual aids? By far the most common diagram in *Language* is the table, followed by maps (eight), and only then abstract visualizations of the sort laid out in the previous section. Interestingly, Bloomfield identifies visual metaphors and visual aids in teasing out the implications of the former. Students read:

The comparative method assumes that each branch or language bears independent witness to the forms of the parent language, and that identities or correspondences among the related languages reveal features of the parent speech. This is the same thing as assuming, firstly, that the parent community was completely uniform as to language, and secondly, that this parent community split suddenly and sharply into two or more daughter communities, which lost all contact with each other. Often enough, the comparative method assumes successive splittings of this sort in the history of a language [...] The comparative method thus shows us the ancestry of languages in the form of a family tree, with successive branchings [...]

---

<sup>9</sup>Josep Simon (2016) contends that “textbook” had come to mean a book conceived for instructional purposes within formal education by the middle of the nineteenth century, picking up on the earlier convention of designating canonical works, excerpted with spaces for students’ interlineal notes, as *texts*. Thus, he implies a direct connection between the history of note-taking practices and the development of formal, printed textbooks. This contextualizes John Joseph’s compelling 2017 discussion of the ambiguous relationship between pictures and words in Saussure’s *Cours de linguistique générale* within a broader history of science and education.

(Bloomfield 1933: 311)

This passage rehearses standard criticisms of the tree model — namely, ancestral uniformity and clean two-way splits — showing it to have heuristic power despite being unrealistic. Thus, it signals an advance over nineteenth-century understanding. “The earlier students of Indo-European did not realize that the family-tree diagram was merely a statement of their method; they accepted the uniform parent languages and their sudden and clear-cut splitting as historical realities” (Bloomfield 1933: 311). In this way, Bloomfield subordinated Schleicher’s visual aid to a method of inference. Translating this into Lakoff’s terminology, an “image metaphor” moved toward the “conceptual” register.

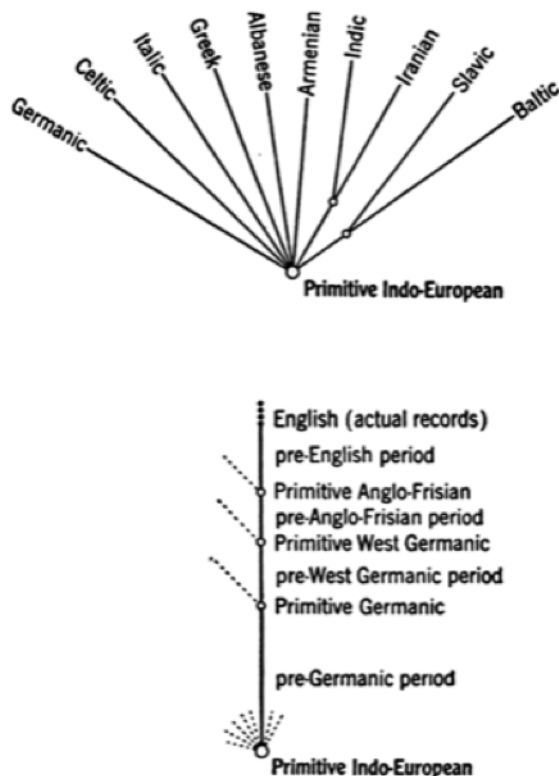


Figure 3: Bloomfield (1933: 312)

This shift was reflected in the highly idealized visual that accompanied the text, shown in Figure 3. The lengths and distances between branches are not particularly measured, and the labels refer to groupings and periods rather than

specific language entities. Bloomfield was an anti-realist tree thinker, to be sure: the diagrams above depict relations but not relatives.

In the text, Bloomfield persistently refers tree thinking to “older scholars”, setting off his positive variationist approach. To complement a series of examples highlighting exceptions to the assumption of clean two-way splits, he adapted a visual aid from the Germanist and linguistic paleontologist Otto Schrader (1855–1919), shown in Figure 4.

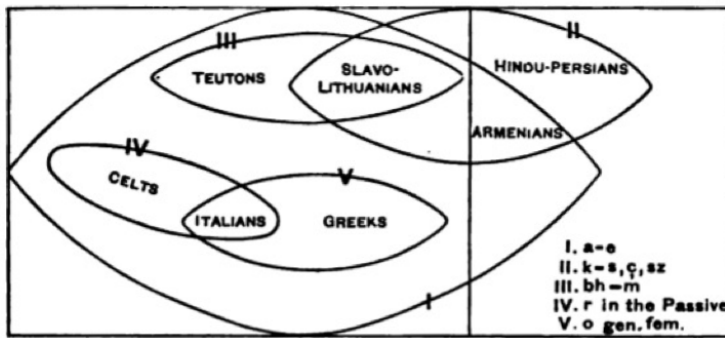


FIG. 6.

Figure 4: Schrader (1890: 65)

The citation to Schrader may at first seem surprising, given that the authors differ in their selection of “special resemblances”, hence, group assignments (Bloomfield 1933: 317). However, both authors allow that the groups could be drawn differently depending on the forms taken into consideration. Bloomfield explained his image in Figure 5 – containing elements of uniformity and variation – in terms of Schmidt’s “wave-hypothesis”, which he endorsed. “Indeed”, he wrote favourably, it is “the picture presented by the local dialects in the areas we can observe” (Bloomfield 1933: 317). Charles Hockett (1916–2000), writing in a more richly visual idiom some twenty-five years later, would refer a reprint of the same image to overlaid notions of vertical and horizontal relationship (cf. Hüb-schmann 1967 [1875]).

Bloomfield’s influence can be seen throughout the pages of Hockett’s textbook, *A Course in Modern Linguistics* (1959 [1958]). Hockett introduced his subject in rigorous terms: “Linguistic research can accomplish nothing unless it is strictly inductive” (Hockett 1959 [1958]: 7). Such primacy of “actual usage, as determined by observation” was born out in the sequence of chapters, which proceed from the smallest units of synchronic observation – defined through examples and pre-

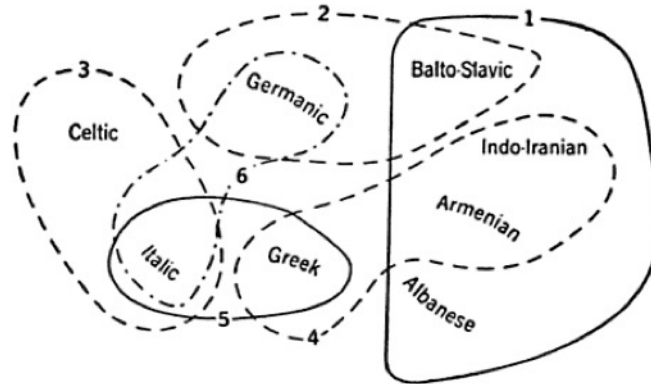


FIGURE 3. Some overlapping features of special resemblance among the Indo-European languages, conflicting with the family-tree diagram.—Adapted from Schrader.

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

Figure 5: Bloomfield (1933: 316)

sented with rules for exacting description – to the more complex, with chapters on language diachrony and other un-observables saved for the end. Indeed, he did not even mention the distinction between synchronic and diachronic linguistics in the book until Chapter 36.

There are clues to Hockett's visual epistemology throughout the text, with bearing on the way he called upon diagrams of language relationship. First, the text shows a remarkable tolerance for the kind of idealization any visualization of language relationship would require. Elaborating on the "design of a language" through its five subsystems, he allowed, for instance, that no description "can claim more than a kind of by-and-large accuracy" (Hockett 1959 [1958]: 139). He similarly flagged the underdetermined and heuristic nature of grammatical description in connection with immediate constituents a few pages later. The following passage is perhaps unexpected behaviourist fare:

[...] grammatical analysis is still, to a surprising extent, an art: the best and clearest descriptions of languages are achieved not by investigators who follow some rigid set of rules, but by those who through some accident of life-history have developed a flair for it [...] Consequently, the reader will find in these sections many an example which the writer has handled

in one way, but which might also be handled in some other way [...] Indeed, the reader should be alert for possible instances where conciseness of statement has unintentionally concealed uncertainty. (Hockett 1959 [1958]: 147)

Reflections like these contextualize his use of visual metaphors (e.g., the persistent reference to phonemic “shape”) and visual aids in the text (Hockett 1959 [1958]: 130–132). With respect to the latter, Hockett’s use of two-dimensional abstract representations involved explicit pedagogical aims. For example, in the notes to his chapter on “Canonical Forms and Economy”, Hockett taught students *how to see* the three diagrams in Figure 6 which were most often used to represent complex morphological systems — the “maze”, “freightyard”, and “rollercoaster” (Hockett 1959 [1958]: 290–292; Harris 1951; Hoenigswald 1950).

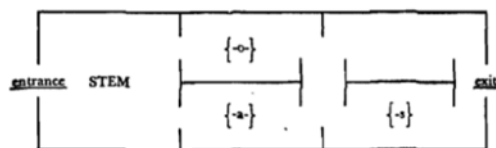


FIGURE 34.1. THE MAZE



FIGURE 34.2. THE FREIGHTYARD

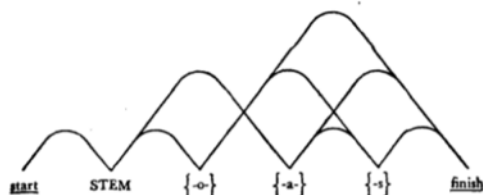


FIGURE 34.3. THE ROLLERCOASTER

Figure 6: Hockett (1959 [1958]: 291)

Hockett told readers that the example provided was derived from the inflection of gendered Spanish adjectives, a pattern “too simple to need diagrammatic display”, thus a “good one with which to demonstrate the diagramming techniques”. The words accompanying these images train students in the techniques of visual analysis – navigating the maze, for example, one proceeds “from left to right, never crossing any lines”; once in the freightyard, there is “no turning back”; and while riding the rollercoaster, “one can turn down wherever there is a curved top,

but not an angle”. Thus, Hockett formalized morphological rules, rendered them exhaustive, and made them intelligible for beginners. The penultimate paragraph on the matter provides lessons in *critical* visual analysis. Here Hockett pointed out that vertical alignment in the first two diagrams can be “read” as indicating “a single positional class”, whereas the “rollercoaster has the advantage of listing all the inflectional affixes along the bottom for ease of checking against inadvertent duplications”. The discussion concludes with an exercise that recruits students to the practice of diagrammatic visualization, suggesting that this was not merely a means of summing up, but rather an active part of research practice.

How was this visual epistemology brought to bear on questions of language relationship? Hockett defined relationship in terms of common origin: divergence being the key factor, not time. Accordingly, he made a sharp distinction between linguistic and biological phylogeny – “languages do not ‘reproduce’ either sexually or asexually”, instead, they “simply *continue*” (Hockett 1959 [1958]: 369). For this reason, familial metaphors to do with parenthood and ancestry appeared “shaky” in his estimation. “At a given point in time,” he concluded,

a set of related languages is merely what would be a set of dialects of a single language except that the links between the dialects have become very tenuous or have been broken [...] The mere fact of relationship thus becomes of secondary importance. More important is the *degree* of relationship. (Hockett 1959 [1958]: 369)

The first tree diagrams to appear in the text serve methodological rather than representational ends. Grouping together Proto-Germanic and its descendants without concern for sub-grouping, two sets of arrows in Figure 7 instead illustrate the logic of traditional and inverted reconstruction.

A more realistic tree diagram, reproduced in Figure 8, appears a few pages later, qualified, with tips on interpretation.

Of this canonical image, Hockett wrote:

The vertical dimension represents time, increasing as one goes from bottom to top. G has been placed earlier than the other four languages because our records of it date from an earlier century. Read literally, the diagram would suggest that, after Proto-Germanic, first the speakers of what was to become G split off from the rest [...], that somewhat later the speakers of what was to become OI moved away from the rest, and that, finally, the remaining group split three ways – the splits, in each case, being more or

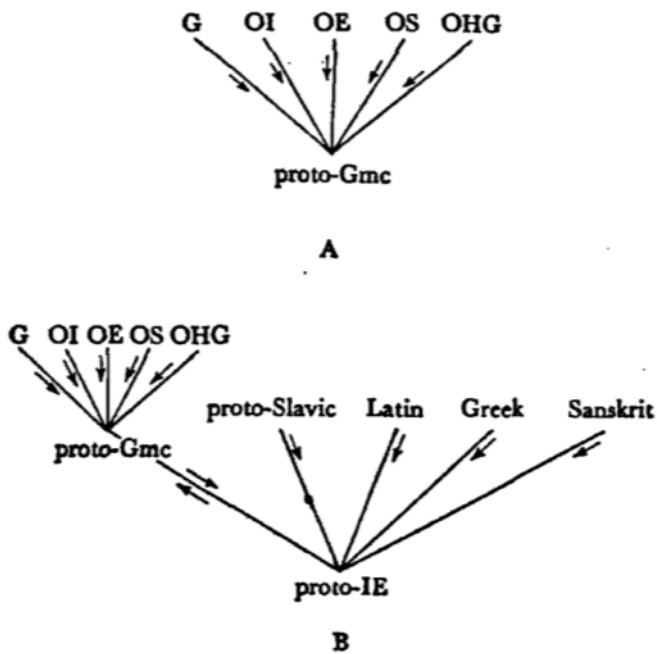


Figure 7: Hockett (1959 [1958]: 514)

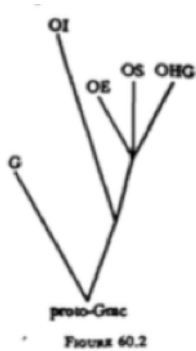


Figure 8: Hockett (1959 [1958]: 519)

less sudden. Now such literal interpretation is not contrary to what sometimes happens in history. But it is dangerous to assume that this is always what happened, because there are other ways in which divergence can come about. (Hockett 1959 [1958]: 519–521)

This passage begins with extremely rudimentary visual directives, a reminder of the pedagogical aim of the work. It also says something — through reference to “our records” — about the research labour that goes into “cooking” the data summarized by the family tree. From there, readers are directed through a “literal” reading of the tree. The important point here is that insights on historical *process* are being extracted from a given *pattern* of historical relationship. While a bracketed table might capture a similar classification of language relationship, it would be harder to read in such a richly narrative way. This is because it would lack the “entailments” of the underlying image and conceptual metaphors we have been tracing. Ultimately, Hockett concluded, the tree was a possible, but unrealistic way of representing relationship. “It is imperative for us to remember that our reconstruction wears a disguise of greater preciseness than can validly be ascribed to it, but to throw it out for this reason would be folly” (Hockett 1959 [1958]: 523). He proposed the alternative illustrated in Figure 9 instead.

The diagram in Figure 9, in effect, zooms in on the base of the previous tree, looking at it through a series of cross-sectional slices like a flip-book. A succession of slices was Hockett’s best effort to visualize relationship simultaneously in space and time. It culminated in the recapitulation of Bloomfield’s diagram of dialect geography, formulated here in the context of a methodological argument rather than a list of discrepant data. This reflects another step away from exemplary towards formal instruction.

In his overview of scientific textbooks, Josep Simon calls for a transdisciplinary exploration of the genre, beyond a traditional emphasis on disciplines and discipline-formation in connection with this site of “normal science” (Simon 2016: 475). In partial fulfilment of that call, the final example in this section pivots from general introductions to a textbook devoted specifically to the sub-discipline in question, Theodora Bynon’s *Historical Linguistics* (1977). Significantly, Bynon introduces the twentieth-century organization of linguistic knowledge with an extended image metaphor:

The representation of the evolution of a language as consisting in a succession of discrete states is no more a true reflection of the situation than is the representation of a circle by a number of straight lines connecting successive points around its circumference. For, however large a number



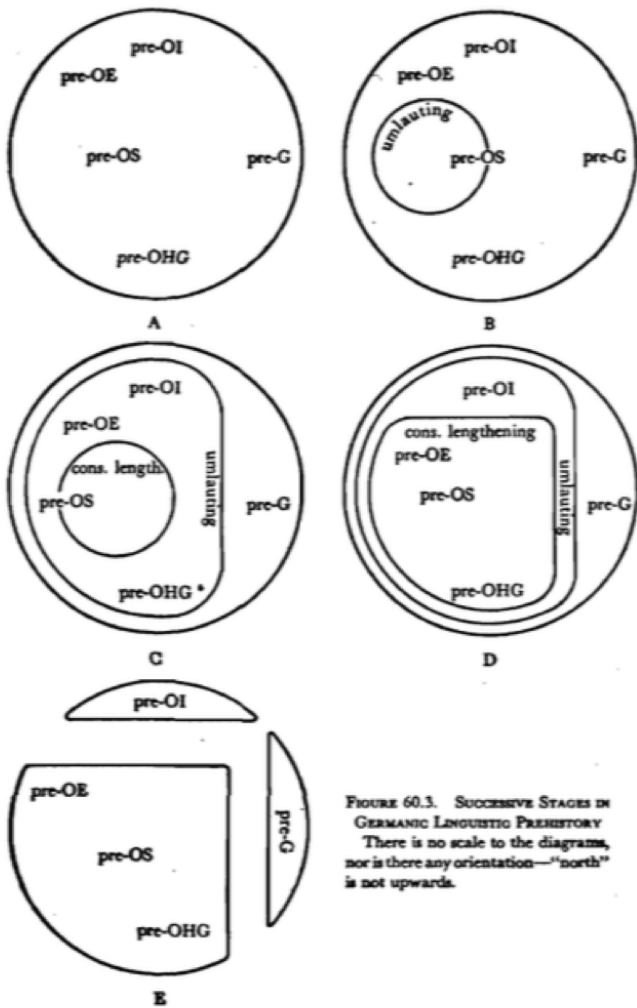


FIGURE 60.3. SUCCESSION STAGES IN GERMANIC LINGUISTIC PREHISTORY  
There is no scale to the diagrams, nor is there any orientation—"north" is not upwards.

520

Figure 9: Hockett (1959 [1958]: 520)

of such points are taken the resulting figure will never be a genuine circle and, in the same way, however many language states are considered over a given period their succession will never provide a true picture of the unbroken *continuity* of a language in time. It is thus due to the limitations of our methodology that we are faced with the rather absurd situation that language evolution, although observable retrospectively in its *results*, appears to totally elude observation as a *process* while it is actually taking place. (Bynon 1977: 2)

So much for any attempt to comprehend, let alone visualize or represent, historical products and processes realistically. Accepting this fate, Bynon accordingly opts for a “two-fold strategy”. First, she presents models of linguistic development from the neogrammarians to the transformational-generative school. “We must study [the] results [of language change] as abstracted from the grammatical descriptions of successive language states and [...] of related languages” (Bynon 1977: 6). Second, she turns to the “question of the connection between language change and social and geographical space” (Bynon 1977: 6). Rather than worry about the historical fidelity of either approach to the study of relationship and differentiation, this text holds them apart schematically.



Figure 10: Bynon (1977: 173)

According to this overall scheme, questions of the linear development of language through time – and, with them, trees – are isolated from those pertaining to internal variation and contact. Trees appear in her narrative as a bridge from the consideration of change within individual languages to changes between them. For Bynon, trees are not primary, they do the work of summing up the “rules” of differentiation – a sign of her times. That said, she emphasizes their visual interpretation more than either Bloomfield or Hockett. Describing a downward branching tree linking English, German, the Greek dialects, Persian and

Sanskrit back to Proto Indo-European through a series of innovations, students read:

In the tree diagram the horizontal dimension [...] represents “space” in a much idealized form – not in an absolute geographical sense but rather in terms of contact or absence of contact between speech communities – whereas the vertical dimension represents time. The branches of the tree then represent channels of transmission, that is the paths along which innovations have been transmitted, and whenever a branch divides into two or more this implies the splitting up of a speech community indicated by the fact that subsequent innovations are no longer shared. (Bynon 1977: 66)

Clearly, Bynon embraced anti-realistic tree thinking, though here she invests more words in training students to see this model than her structuralist predecessors. Notably, she resists the urge to conclude her discussion with an all-inclusive diagram of the Indo-European family, opting for a two-page chart instead, shown in Figure 11.

This diagram, she concludes, has the advantage of not overstating putative relationships, designed, as it was, to bring “together loosely according to branches and periods the main languages of the Indo-European family for which actual material survives.” As for the tree model, Bynon states, it should be reserved for “the display of rules relating successive systems” (Bynon 1977: 70). At roughly the same time as phylogenetic modelling was taking off in biology, Bynon was relinquishing the analogy between languages and biosystematics entirely.

## 5 From anti-realism to realism in the digital era

At a 1973 conference on the topic of “Lexicostatistics in Genetic Linguistics”, convened in Montreal, Paul Black presented a paper on the adaption of Multidimensional Scaling (MDS) to historical linguistics, which built on a prior collaboration with Isidore Dyen and Joseph Kruskal of Bell Labs. His work was an early attempt to model hybrid transfer computationally using data collected by others. Black endorsed the controversial use of statistics in comparative-historical linguistics, stating by way of introduction that metrical analysis of linguistic distance “permits a multidimensional recognition of relations” (Black 1973: 43–92). Thus, Black adapted canonical metaphors of language relationship – forged on the basis of Indo-European data in a two-dimensional environment – to a new geographic and conceptual space.

	2000 B.C.	1000 B.C.	B.C. A.D.	1000 A.D.
Albanian				Albanian
Greek	Mycenaean Greek	Homeric Greek Classical Greek	Koine	Greek
Anatolian	Hittite/ Luwian/ Palaic/ Hieroglyphic Hittite/			
Tocharian			Tocharian A/ Tocharian B/	
Indo-Iranian	Indic (Indo-Aryan)	Vedic Sanskrit	Classical Sanskrit Prakrit Pali	Hindi, Urdu Panjabi Gujarati Bengali Assamese Sindhi Sinhalese
			Avestan/ Old Persian	Persian Kurdish Pashto Balochi Ossetic
Armenian			Classical Armenian	Armenian

Figure 11: Bynon (1977: 68). This is half of the chart, which stretches over two pages.

Carried over from the world of marketing, psychology, and political science, Black described MDS as a way to see continuous variation (“cline structure”) within a hierarchy of discrete classes (the evolutionary tree). The objective was hybrid in nature:

While a “family tree” diagram or some other representation of a hierarchical subgrouping is an obviously appropriate way of describing the temporal hierarchy of linguistic splits through which a group of languages may have evolved from a common ancestral protolanguage, multidimensional scaling can be used to investigate and describe the spatial variation which originates in the wave-like spread of linguistic innovations within a single language, and which may also persist within the evolutionary tree to an extent sufficient to hamper the correct inference of this tree. (Black 1973: 43)

According to Black’s discussion, the method was new in that it began by *testing* – rather than assuming or imposing – the fit between tree classifications, wave models, and the actual language data (in this case, pertaining to Bikol, Lower Niger dialects, Konsoid, and Salish) they were meant to represent. From there, distances between each of the entities under consideration (dialects or languages) were scaled so as to approximate “actual physical distances”. Looking at Bikol dialects, for example, one might figure percentages of lexicostatistical retention, subtract each from one hundred percent, and map each percentage point of difference as a distance of one tenth of an inch. A common retention of 79%, as in the case of Sorsogon and Masbate, for instance, might yield a one-dimensional distance of 2.1 inches according to this method. Black continued,

Oas might then be added to the picture by placing it 3.1 inches (corresponding to 69%) from Sorsogon and 4.2 inches (corresponding to 58%) from Masbate; these relationships would then be well represented in two-dimensional space as a triangle. (Black 1973: 52)

This was reasonably straightforward. But, as Black pointed out, the procedure becomes increasingly unwieldy as more dimensions are added to the mix, such that the representation “might prove to be difficult to visualize and interpret” should dimensionality not be “restricted to some very small number” (Black 1973: 53). With these words, Black was confronting the difficulty of reconciling language data with formalized relationships, fidelity with the “all-at-onceness” of disciplined perception. Even if MDS was escaping the constraints of the printed

page as a research tool, Black was still bound to two dimensions when it came to the communication of research findings. In order to flatten a full set of distance measures into a two-dimensional representation with some degree of intelligibility, it was necessary to adjust the original percentages in a rationalized way that might be traced back to the original data. Electronic computers were thought to have the power needed to pull this off. Using the KYST program,<sup>10</sup> Black specified the range of possible dimensions, a rule for scaling, and the lexicostatistic data, ultimately yielding images like those in Figure 12.

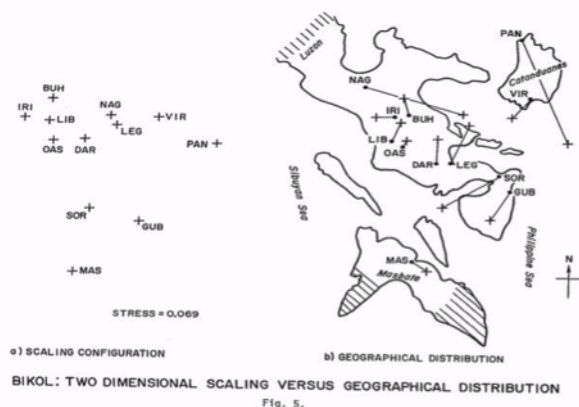


Figure 12: Black (1973)

On the left-hand side of Figure 12, we have the distances (inverse percentage agreements) between twelve dialects scaled down to two dimensions from a potential total of eleven. The stress index given (.069) is a measure of how far these adjusted values differ from the original data fed into the program. The right-hand side shows these values projected into geographic space, with arrows indicating differences between scaled values and places where the dialects were actually observed.

This last point highlights the extent to which representational validity was of concern to those working with MDS. The iterative nature of this method – tinkering with dimensions and scaling to preserve fidelity to the data – foregrounded the issue of realism in a way that pre-computational diagramming did not. If earlier examples primarily served rhetorical, didactic, or programmatic functions,

<sup>10</sup> KYST, pronounced “kissed”, was one of several MDS programs available in the 1970s. The name derives from those of its architects: Kruskal, Young, Shepard and Torgerson (see Kruskal et al. 1973).

MDS thus can be seen to align with the advent of a new period of experimental visualization in comparative-historical linguistics.

In this respect, MDS looks like a forerunner of the use of phylogenetic networks in linguistics (Stevens 2013). Russell Gray, for instance, a self-described “evolutionist” and the newly appointed director of the Max Planck Institute for the Science of Human History in Jena, has zealously promoted the use of such probabilistic modelling techniques in linguistics, linking their powers of discovery to new scientific frontiers (e.g., <http://www.mpi.nl/events/nijmegen-lectures-2014/lecture-videos>, accessed 2 August 2018). Cheerfully labelling this work *lexomics*, Gray has sought to model and test tacit assumptions about comparative-historical methods, in addition to reconstructing family trees.

One of the programs commonly used in such work is SplitsTree, a tool for producing visualizations of phylogenetic networks (Greenhill et al. 2010). The idea, according to architects Daniel Huson (a specialist in bioinformatics at Tübingen University) and David Bryant (a mathematician from the University of Auckland) is to “use split networks, which are not trees, to represent phylogenetic signals that, for the most part, originate from trees” (Huson & Bryant 2006: 254–267). Through an iterative modelling process, this program prioritizes fidelity to the givens – acknowledging the realistic complexity of historical relationship while revealing the presence of latent trees to “plain sight”.

## 6 Conclusion

This paper has surveyed visual metaphors and visual aids of language relationship and divergence from the mid-nineteenth century to the early 2000s. It has shown that, until recently, visual metaphors took priority over visual aids. In some cases – Schmidt’s most notably – we have seen the very utility of visual aids called fundamentally into question. The complicated overlay of mental images and concepts in this history suggests that there may be movement between these categories of metaphoric expression, accounting for both disciplinary consolidation and the open-endedness of linguistic practices. This emphasis on metaphor provides an interesting counterpoint to the recent material turn in STS scholarship on practices of scientific representation. Highlighting the difficulties linguists have encountered in their efforts to comprehend the “all-at-once-ness” of language relationship in both time and space, I have suggested that canonical visualizations (whether presented in words or pictures) matter a great deal in pedagogy and practice. Phylogenetic modelling has offered new hope to those in pursuit of a “natural classification” insofar as it offers this kind of dual vision.

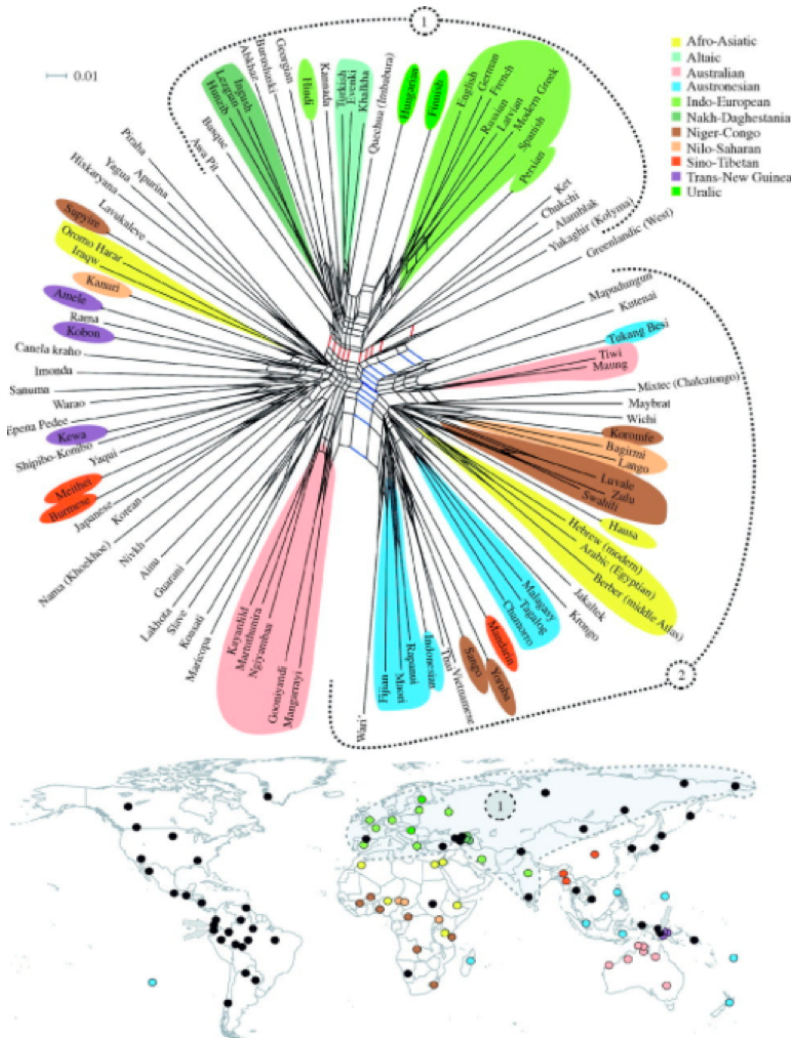


Figure 13: Greenhill et al. (2010: 2445)



## References

- Archibald, J. David. 2014. *Aristotle's ladder, Darwin's tree: The evolution of visual metaphors for biological order*. New York: Columbia University Press.
- Bertomeu-Sánchez, José-Ramon, Antonio García-Belmar & Bernadette Bensaude-Vincent. 2002. Looking for an order of things: Textbooks and chemical classifications in nineteenth century france. *Ambix* 49(3). 227–250.
- Bigg, Charlotte. 2016. Diagrams. In Bernard Lightman (ed.), *A companion to the history of science*. Malden, MA: Wiley, Blackwell.
- Black, Paul. 1973. Multidimensional scaling applied to linguistic relationships. In Isidore Dyen & Jucquois (eds.), *Lexicostatistics in genetic linguistics ii: Proceedings of the Montreal conference*.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Bloomfield, Leonard & Henry Hoiyer. 1965. *Language history: From Language (1933 ed.)* New York: Holt, Rinehart & Winston.
- Bonfante, Giuliano. 1931. I dialetti indoeuropei. *Annali del R. Istituto Orientale di Napoli* 4. 69–185.
- Boyd, Richard. 1979. Metaphor and theory change: What is “metaphor” a metaphor for? In Andrew Ortony (ed.), *Metaphor and thought*, 356–408. Cambridge: Cambridge University Press.
- Bryant, David & Vincent Moulton. 2004. Neighbor-net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2). 255–265.
- Bynon, Theodora. 1977. *Historical linguistics*. Cambridge: Cambridge Textbooks in Linguistics.
- Cameron, H. Don. 1987. The upside-down cladogram: Problems in manuscript affiliation. In Henry Hoenigswald & Linda Wiener (eds.), *Biological metaphor and cladistic classification: An interdisciplinary perspective*, 227–242. Philadelphia: University of Pennsylvania Press.
- Čelakovský. 1853. *Čtení o srovnávací mluvnici slovanské na universitě pražské*. Prague: F. Řivnáče.
- Coopmans, Catelijne, Janet Vertesi, Michael Lynch & Steve Woolgar. 2014. *Representation in scientific practice revisited*. Cambridge, MA: MIT Press.
- Daston, Lorraine. 2008. On scientific observation. *Isis* 99(1). 97–110.
- Daston, Lorraine. 2014. Beyond representation. In Coopmans et al (ed.), *Representation in scientific practice revisited*. Cambridge, MA: MIT Press.

- Eichwald, Carl Eduard von. 1829. *Zoologia specialis quam expositis animalibus tum vivis, tum fossilibus potissimum rossiae in universum....* Vilnius: Josephus Zawadzki.
- Gatschet, Albert S. & Lyle Campbell. 1973. Gatschet's classification of Mayan languages. *International Journal of American Linguistics* 39(4). 250–252.
- Geisler, Hans & Johann-Mattis 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.), *A companion to the history of science*, 111–124. Stuttgart: Franz Steiner.
- Gray, Russell, Alexei Drummond & Simon J. Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in pacific settlement. *Science* 323(5913). 479–483.
- Greenhill, Simon J., Quentin D. Atkinson, Adam Meade & Russell D. Gray. 2010. The shape and tempo of language evolution. *Proceedings of the Royal Society B* 277. 2443–2450.
- Haraway, Donna. 1991. *Simians, cyborgs, and women: The reinvention of nature*. New York: Routledge.
- Harris, Zellig. 1951. *Methods in structural linguistics*. Chicago: Chicago University Press.
- Hellström, Peter. 2012. Darwin and the tree of life: The roots of the evolutionary tree. *Archives of Natural History* 39(2). 234–252.
- Hirt, Herman. 1905. *Die Indogermanen: Ihre Verbreitung, ihre Urheimat und ihre Kultur I*. Strassburg: Trübner.
- Hockett, Charles. 1959 [1958]. *A course in modern linguistics*. New York: Macmillan.
- Hoenigswald, Henry. 1950. Morpheme order diagrams. *SiL* 8. 79–81.
- Hoenigswald, Henry. 1975. Schleicher's tree and its trunk. In Werner Abraham (ed.), *Ut videam: Contributions to an understanding of linguistics: For Pieter A. Verburg*, 157–160. Lisse & Holland: P. de Ridder.
- Hoenigswald, Henry & Linda Wiener (eds.). 1987. *Biological metaphor and cladistic classification: An interdisciplinary perspective*. Philadelphia: University of Pennsylvania Press.
- Hopwood, Nick. 2015. *Haeckel's embryos: Images, evolution, and fraud*. Chicago: Chicago University Press.
- Hübschmann, J. Heinrich. 1876. Ueber Aussprache und Umschreibung des Armenischen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 30. 53–73.
- Hübschmann, J. Heinrich. 1967 [1875]. On the position of Armenian in the sphere of the indo-european languages. In Winfred P. Lehmann (ed.), *A reader in nine-*

- teenth century historical Indo-European linguistics*, 164–189. Bloomington: Indiana University Press. Original: “Ueber die Stellung des Armenischen im Kreise der indogermanischen Sprachen”, *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen*, 23: 5–49.
- Huson, Daniel & David Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biological Evolution* 23(2). 354–367.
- Huxley, Thomas H. 1880. Biology. In *Encyclopaedia Britannica*.
- Johns, Adrian. 1998. *The nature of the book: Print and knowledge in the making*. Chicago: Chicago University Press.
- Joseph, John. 2017. The arbre-tree sign: Pictures and words in counterpoint in the *Cours de Linguistique Générale*. *Semiotica* 217. 147–171.
- Koerner, E. F. Konrad. 1975. European structuralism: Early beginnings. In Thomas Sebeok (ed.), *Current trends in linguistics 13: Historiography of linguistics*, 717–827. The Hague: Mouton.
- Koerner, E. F. Konrad. 1987. On Schleicher and trees. In Henry Hoenigswald & Linda Wiener (eds.), *Biological metaphor and cladistic classification: An interdisciplinary perspective*, 109–114. Philadelphia: University of Pennsylvania Press.
- Kruskal, Joseph, Forrest Young & Judith Seery. 1973. *How to Use KYST: A Very Flexible Program to Do Multidimensional Scaling and Unfolding*. [http://www.netlib.org/mds/kyst2a\\_manual.ps](http://www.netlib.org/mds/kyst2a_manual.ps). Bell Laboratories Unpublished Manuscript.
- Kuhn, Thomas. 1962. *The structure of scientific revolutions*. Chicago: Chicago University Press.
- Kusukawa, Sachiko. 2016. Classics from this journal: Martin Rudwick’s “The emergence of a visual language for geological science, 1760–1840”, *History of Science*, xiv: 3, 1976, pp. 149–95. *History of Science* 54(1). 98–104.
- Lakoff, George. 1987. Image metaphors. *Metaphor and Symbolic Activity* 2(3). 219–222.
- Lakoff, George & Mark Johnson. 1980. Conceptual metaphor in everyday language. *The Journal of Philosophy* 77(8). 453–486.
- Lopez, Philippe, Johann-Mattis List & Eric Baptiste. 2013. A preliminary case for exploratory networks in biology and linguistics: The phonetic network of chinese words as a case study. In Heiner Fangerau, Hans Geisler, Thorsten Halling & William Martin (eds.), *Classification and evolution in biology, linguistics, and the history of science*, 181–196. Stuttgart: Franz Steiner.
- Maher, John P. 1966. More on the history of the comparative method: The tradition of Darwinism in August Schleicher’s work. *Anthropological Linguistics* 8(3). 1–11.

- Meillet, Antoine. 1908. *Les dialectes indo-européens*. Paris: La Société de Linguistique de Paris.
- Müller, F. Max. 1913 [1891]. *The science of language*. New York: Charles Scribner's Sons.
- Myers, Natasha. 2014. Rendering machinic life. In Coopmans et al. (ed.), *Representation in scientific practice revisited*, 153–175. Cambridge, MA: MIT Press.
- Myers, Natasha. 2015. *Rendering life molecular: Models, modelers, and excitable matter*. Durham: Duke University Press.
- Nyhart, Lynn. 2012. Wissenschaft and Kunde: The general and the special in modern science. *Osiris* 27(1). 250–275.
- Pietsch, Theodore. 2012. *Trees of life: a visual history of evolution*. Baltimore: Johns Hopkins University Press.
- Priestly, Tom. 1975. Schleicher, Čelakovský, and the family-tree diagram: a puzzle in the history of linguistics. *Historiographia Linguistica* 2. 299–333.
- Pulgram, Ernst. 1953. Family tree, wave theory, and dialectology. *Orbis: Bulletin International de Documentation Linguistique* 2(1).
- Schleicher, August. 1967 [31871]. Introduction to a *Compendium of the comparative grammar of the Indo-European, Sanskrit, Greek and Latin languages*. In Winfred P. Lehmann (ed.), *A reader in nineteenth century historical Indo-European linguistics*, 87–96. Bloomington: Indiana University Press. Original: *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Weimar: Böhlau.
- Schleicher, August. 1848. *Sprachvergleichende Untersuchungen*. Bonn: H. B. Koenig.
- Schleicher, August. 1853. Die ersten Spaltungen des indogermanischen Urvolkes. *Allgemeine Zeitung fuer Wissenschaft und Literatur*.
- Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Böhlau.
- Schmidt, Johannes. 1875. *Zur Geschichte des indogermanischen Vocalismus II*. Weimar: Böhlau.
- Schmitt, Rüdiger. 1976. Einführung. In Rüdiger Schmitt (ed.), *Heinrich Hübschmann: kleine Schriften zum Armenischen*. Hildesheim & New York: Georg Olms.
- Schrader, Otto. 1890. *Prehistoric antiquities of the Aryan peoples: a manual of comparative philology and the earliest culture*. Trans. by F. B. Jevons. London: Charles Griffin.

- Schuchardt, Hugo. 1900 [1870]. Über die Klassifikation der romanischen Mundarten: Probe-Vorlesung, gehalten zu Leipzig am 30. April 1870. In Leo Spitzer (ed.), *Hugo Schuchardt Brevier*, 166–188.
- Simon, Josep. 2016. Textbooks. In Bernard Lightman (ed.), *A companion to the history of science*. Malden, MA: Wiley, Blackwell.
- Southworth, Franklin. 1964. Family-tree diagrams in comparative linguistics. *Language* 40. 557–565.
- Stevens, Hallam. 2013. *Life out of sequence: a data-driven history of bioinformatics*. Chicago: University of Chicago Press.
- Wells, Rulon. 1987. The life and growth of language: Metaphors in biology and linguistics. In Henry Hoenigswald & Linda Wiener (eds.), *Biological metaphor and cladistic classification: An interdisciplinary perspective*, 39–80. Philadelphia: University of Pennsylvania Press.





# Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit <http://www.langsci-press.org/donate> to provide financial support or register as a community proofreader or typesetter at <http://www.langsci-press.org/register>.







# Form and formalism in linguistics

“Form” and “formalism” are a pair of highly productive and polysemous terms that occupy a central place in much linguistic scholarship. Diverse notions of “form” – embedded in biological, cognitive and aesthetic discourses – have been employed in accounts of language structure and relationship, while “formalism” harbours a family of senses referring to particular approaches to the study of language as well as representations of linguistic phenomena. This volume brings together a series of contributions from historians of science and philosophers of language that explore some of the key meanings and uses that these multifaceted terms and their derivatives have found in linguistics, and what these reveal about the mindset, temperament and daily practice of linguists, from the nineteenth century up to the present day.

