

Beyond Edit Distances: Comparing Linguistic Reconstruction Systems

Johann-Mattis List

Max Planck Institute for the Science of Human History, Jena, list@shh.mpg.de

1 Introduction

Jäger's target article illustrates how a fully automated workflow of the comparative method, starting from the proof of genetic relationship up to the reconstruction of proto-forms can be applied to a small set of Romance languages. There are quite a few points deserving further discussion in this very interesting and very well-written overview on the current state-of-the-art in computational historical linguistics. For example, I do not completely agree with Jäger's claim that methods that make use of orthographical string representations would not qualify as part of NLP-near computational historical linguistics. I think that the current practice in NLP is much less rigorous than Jäger assumes here, with many of the first but also the more recent algorithms that addressed quantitative tasks in historical linguistics making exclusive use of orthographic language data (Hauer and Kondrak 2011; Serva and Petroni 2008; Ciobanu and Dinu 2018).

I feel, however, that a detailed discussion of these smaller issues can probably be better done in personal communication than in a comment on a target paper and colleagues might also comment on this. Instead, I would like to use the space that was given to me, in order to discuss Jäger's use of the *normalized edit distance* (an extension of the classical edit distance by Levenshtein 1965) as an evaluation measure for the differences in reconstruction systems. While this may seem to be a very minor point of the full workflow, the proper comparison of reconstruction systems has far more implications, not only for the field of computational linguistics, but also for the field of historical linguistics in general. Furthermore, as I will try to show, it could open interesting and new possibilities to measure, among others, the intelligibility among dialects.

In the following, I will quickly discuss the so-called abstractionalist-realist debate in historical linguistics and illustrate its importance to understand the nature of reconstruction systems (Section 2). I will then comment on phonological reconstruction as a specific task in historical linguistics, and show how it differs from Jäger's workflow (Section 3). I will then summarize my view on the nature of the proto-language in linguistic reconstruction (Section 4), and emphasize the need to use a measure for the structural comparison of reconstruction systems, as opposed to substantial comparison using methods such as the normalized edit distance (Section 5). Developing a first method to measure the structural similarity of reconstruction systems in Section 6 and applying this measure to Jäger's data illustrates that the automated reconstructions proposed by Jäger come much closer to the gold standard than the use of edit distance as an evaluation measure may suggest. In the conclusion in Section 7, I make clear that the new measure I propose can only partially account for the abstract nature of reconstruction systems and that more powerful systems are needed for a realistic comparison.

2 The abstractionist-realist debate in historical linguistics

Ever since Schleicher proposed the first proto-forms for Proto-Indo-European (Schleicher [1861] 1866), there has been great disagreement among scholars how to interpret the reconstruction systems (see Lass 2017 for a recent account). In Schleicher's language theory, which compared languages with organisms (Ross and Durie 1996, 14), the quality of the proto-language was clear-cut and concrete. It represented the highest level of development of Proto-Indo-European in its "vorhistorischer Periode", which in its "historischer Periode" had gradually disintegrated into the various Indo-European languages (cf. Schleicher [1861] 1866, 1:4). This concrete view of the proto-language certainly gave reasons for Schleicher's realism, leading to the composition of his famous fable (Schleicher 1868). Later on, scholars exhibited a more critical attitude towards the reconstruction systems, emphasizing their abstract character. Schmidt's *Wellentheorie* (1872) can be understood as a reaction to Schleicher's realism, revealing the proto-language as a "wissenschaftliche fiction" (Schmidt 1872, 31). The wave theory, however, raised new problems for linguistic reconstruction, since the proto-language was now deprived of its uniform character and the question what the reconstruction systems actually represent had to be evaluated again. This can be seen in Brugmann's scepticism that the reconstruction systems could ever reflect a linguistic reality (Brugmann, 1904, 25), and even more in Saussure's work, who completely rejected a positive specification of the sound values of reconstruction systems (Saussure 1916, 303).

The question about the nature of the proto-language, the reconstruction systems, and their relationship remained unsolved, leading to various proposals in favour of the realists or the abstractionists. Hall's reconstruction of Proto-Romance, which was actually testable against an attested proto-language, could not convince all scholars, "that the degree of realism attained is reasonably high" (Hall 1960, 204). This became obvious in the discussion that was induced by the glottalic theory. While the supporters of the theory based their claims on a realistic conception of the reconstruction systems (cf. e.g. Salmons 1993, 69f), the opponents took a rather abstract position (cf. e.g. Schrodtt 1989, 147, who suggested to replace the phonetic differentiation within the PIE plosive series by index numbers).

It is not important what one actually thinks about the nature of the proto-language. The past has shown that people can perfectly agree on particular reconstructions even if they disagree on the question of whether a reconstruction should be typologically plausible or not (as demanded by Jakobson 1971). When working on approaches to automate certain aspects of the comparative method, however, it is important to keep the abstractionalist-realist debate in mind, since this debate finds its reflection in the way in which linguists practice historical language comparison.

3 Phonological reconstruction

Phonological reconstruction is one of the specific subtasks of the comparative method which Jäger attempts to automate. The classical workflow in which scholars proceed here is to start

from previously identified *sound correspondence patterns* and to assign each distinct pattern a certain proto-sound. Sound correspondence patterns can be seen as a specific partitioning analysis of the *alignment sites* of a set of cognate words identified for a group of related languages (List 2019). While two distinct sound correspondence patterns can reflect the same proto-sound, one sound correspondence pattern is usually thought to always reflect the same proto-sound (unless scholars make use of internal reconstruction). Phonological reconstruction can in theory be made at two points of the workflow. One can first carry out a *lexical reconstruction* to identify the actual words that were used to express a certain meaning in the ancestral language and then do the phonological reconstruction, or one can project proto-forms for all aligned cognate sets in the data and later decide which of these proto-forms are likely to have belonged to the proto-language.

Jäger's workflow uses ancestral state reconstruction methods to address the task of lexical reconstruction before working on phonological reconstruction. This procedure is straightforward. In the traditional literature in comparative linguistics, however, scholars usually start with phonological reconstruction and then proceed to lexical reconstruction. Lexical reconstruction itself is not given much thought in classical historical language comparison, and scholars working on South-East Asian languages often even ignore lexical reconstruction completely (for a rare exception, see Mann 1998). Instead, they pick the homologous morphemes from their word comparisons, assign some rough meaning to them, and then propose an ancient pronunciation based on the correspondence patterns they observe (Hill and List 2017, 52f).

While Jäger's procedure is equally justified and specifically useful, as it emphasizes the importance of lexical reconstruction (or *onomasiological reconstruction*, see Jäger and List 2018) while most classical approaches ignore it, the fact that linguists tend to start with phonological reconstruction reflects the importance of *abstractionalist* thinking in classical studies on historical language comparison. Even if most linguists agree that linguistic reconstruction should try to propose a language that can be pronounced and that resembles a natural language, their practice accepts a certain degree of abstractness. This abstractness is often also used to cope with the problem of uncertainty, which has never been convincingly addressed in the classical framework of the comparative method. As an example, compare the usage of *H* to denote *any* of the three laryngeals in Proto-Indo-European in cases where evidence from Greek is lacking.

4 The nature of reconstruction systems

According to scholarly practice, reconstruction systems in historical linguistics are neither abstract nor realistic. They are not abstract because scholars pay attention to phonological or phonetic characteristics when proposing their proto-forms and also discuss differences in reconstruction systems among colleagues in the light of the proto-values they propose. They are not realistic because scholars fail to provide all due detail which they would need in order to reconstruct a proto-language up to the point where they could use it for a conversation. Since the reconstruction methodology is based on the core assumption that observed variation can be

explained by former conformity, reconstruction systems tend to be deprived of a temporal or a spatial dimension.

Attempts to reconstruct variation in the proto-language have been made (Joseph 2006). We also know that variation in the proto-language may account for cases where data appears less tree-like (List et al. 2016; Jacques and List, forthcoming). Nevertheless, most work on linguistic reconstruction favors abstract over realistic solutions. Even if scholars might see themselves as representatives of the realist-camp, the core methodology of phonological reconstruction demands a certain amount of abstractionism. Variation as explanation should only be invoked as a last resort, when no further way to explain it can be found. Had Saussure concluded that the variation among verbal roots in Sanskrit was due to real variation in the proto-language, he would never have proposed the *coefficients sonantiques*, presumed vocalic elements of unknown pronunciation, which can explain the variation as a former regularity (Saussure 1879). These elements now build the core of the modern laryngeal theory in Proto-Indo-European, that states that these alternations go back to sounds lost in almost all Indo-European languages (Meier-Brügger 2002, 236–55). In order to minimize the amount of speculation in linguistic reconstruction, it is important to limit the range of explanations that can be invoked to account for variation, and theories that increase the freedom of scholars to provide just-so-stories (such as, for example, *lexical diffusion*, as shown by Hill 2016) need to be avoided when it comes to linguistic reconstruction.

5 Comparing reconstruction systems

Given that most if not all reconstruction systems in historical linguistics are abstract in their nature (even if scholars occasionally use them to write fables), no reconstruction system should be taken *literally*, i.e., interpreted in strict phonetic terms. In order to compare two or more reconstruction systems for the same proto-language, it is not enough to compare the proposals sequence by sequence. Since the reconstruction practice allows for a considerable amount of freedom in the choice of symbols, any literal comparison of reconstruction systems, i.e., any comparison which compares the substance of the segments in the proto-sequences proposed by different scholars will drastically exaggerate the differences which different reconstruction systems exhibit.

Number	System A	System B	System C
1	k a ŋ	tʃ u k	1 2 3
2	p a ŋ	t u k	4 2 3
3	t a ŋ	p u k	5 2 3
4	k i k	tʃ a tʃ	1 6 1
5	p i k	t a tʃ	4 6 1
6	t i k	p a tʃ	5 6 1

Table 1: Structurally identical reconstruction systems.

To illustrate this point, consider the three fictive reconstruction systems given in Table 1. While the first two systems appear “realistic” and different in their substance, the third system is “abstract”, following the above-mentioned idea of Saussure to use numbers instead of letters to represent proto-forms. Although the systems may appear to be quite different upon first sight, a closer inspection quickly reveals that they are in fact *identical* with respect to their structure.

This is proven in Table 2, where all sound correspondence patterns are listed and we can see that each distinct sound in each of the three systems *always* corresponds to the same distinct sound in the other system. We could thus simply define a substitution cipher and convert each of the systems in any of the other systems, by simply applying this substitution cipher letter by letter. Although different in their *substance*, all three systems are *structurally identical*.

Pattern	System A	System B	System C
1	k	tʃ	1
2	a	u	2
3	ŋ	k	3
4	p	t	4
5	t	p	5
6	i	a	6

Table 2: Sound correspondence patterns for the three systems.

If we used the edit distance to calculate the difference between the three systems here, we would always end up with 1, i.e., a suggested complete difference. Since the symbols differ almost completely in all three cases, the edit distance, which assumes a segment-by-segment conversion of one sequence into another, without taking any form of abstract structural identity into account, has to *edit* each of the sequences completely. This should illustrate why one should not use edit distance to compare reconstruction systems. Given the systematic nature of sound change and the abstract nature of linguistic reconstruction, it is not the substantial differences that should count for the evaluation, but the structural differences (see List 2014 for details on structural as opposed to substantial differences).

One might argue that Jäger compares his reconstruction system with Latin, a language for which we have a realistic system, given that it is attested in written form. But since sound change is a systematic process, the differences we observe between a proposed reconstruction system and a gold standard will still need to be measured in the form of structural differences rather than substantial differences. This can be easily illustrated when looking at Jäger’s automated reconstructions. In three cases (“come”, “new”, and “see”), we find that Jäger’s system proposes the consonant *v* while the Latin form shows a *w* (*wenire*: *venir*, *nowus*: *novo*, *widere*: *veder*). Structurally, Jäger’s system proposes the correct sound, but substantially, the edit distance counts this proposal as an error. If we replace Jäger’s *v* manually by a *w* and compute the edit distance, the average distance shrinks by one percent.

6 Measuring differences in reconstruction systems

As a first step towards a realistic comparison of reconstruction systems, we need a measure of what I called *structural identity* above. Luckily, this measure is already available in form of B-Cubed scores (Amigó et al. 2009). B-Cubed scores offer a straightforward way to compare *partitioning analyses* (or *cluster analyses*) with each other. In the task of *automatic cognate detection* in computational historical linguistics, for example, B-Cubed scores are frequently used to compare how well an algorithm performs in comparison with a gold standard (Hauer and Kondrak 2011; Jäger, List, and Sofroniev 2017; List, Greenhill, and Gray 2017).

B-Cubed scores indicate if two partitioning analyses that may be labelled differently are essentially identical by measuring how well one partitioning analysis predicts the other. In the same way in which they can be used to measure the similarity of partitioning analyses, they can be used to measure the structural similarity of (aligned) sequences (for a first study in this regard, see List 2018), since a sequence can similarly be treated as a partitioning analysis: if we have four entities A, B, C, and D and partition those into two groups, AB and CD, we can likewise represent this partition of the sequence A B C D as a labelled sequence $A^1 B^1 C^2 D^2$ (with labels indicating the partition), or as a sequence of labels 1 1 2 2. In the same way, we can treat any sequence of sounds (in a word, a sentence, or a text) of a given language as a partitioning of the positions in the sound sequence into groups, with the sounds representing their labels. Thus, German [m a m a] “mother” could be seen as a clustering of 1 2 3 4 (representing the position of a sound in the sound sequence) into the two groups [m] (1 and 3) and [a] (2 and 4), and German [p a p a] would be *structurally identical*, given that it breaks the positions of the word in the same way.

If we can treat a sound sequence as a partition of positions, we can compare two sound sequences that are aligned with each other with help of B-Cubed scores in the same way in which we compare different partitioning analyses. If the B-Cubed scores yield 1, the sequences are structurally identical, and if they are 0, they are structurally different. In order to illustrate this, I use the implementation of B-Cubed scores as provided by the LingPy software package (<http://lingpy.org>, List et al. 2018) to compare Jäger’s reconstructed proto-language with the Latin gold standard with respect to the structural differences.

The new method which I propose to evaluate the degree of structural identity of reconstruction systems proceeds in two steps. In a first step, the automated reconstruction system and the gold standard, or alternatively two expert reconstruction systems that we want to compare for differences, are being aligned with each other. This step can be carried out automatically, but it may be useful to manually adjust it, specifically in those cases where we expect a high degree of *substantial differences*. In a second step, the B-Cubed scores for the aligned sequences are calculated. B-Cubed scores are presented in form of three values, the *precision*, the *recall*, and the *F-Score* (the harmonic mean of precision and recall).

Given that B-Cubed scores measure how well two systems *predict* each other, B-Cubed precision and B-Cubed recall can be directly interpreted in this way. Precision measures how

well the first system predicts the second system, recall predicts how well the second system predicts the first system, and the F-Score summarizes the results. Note that a system that predicts another system well does not necessarily show a high structural similarity to the system it predicts. B-Cubed precision for the comparison of a sequence A B C D with a sequence A A B B is 1.0, i.e., the first sequence predicts the second sequence perfectly, since we can find a substitution cipher that converts the first into the second sequence without errors. However, the second sequence cannot be converted to the first sequence, and therefore, the B-Cubed recall will be 0.5, and the F-Score will be 0.67.

Language	Words	Edit Dist.	Precision	Recall	F-Score
ITALIAN_GROSSETO_TUSCAN	40	0.562	0.389	0.374	0.381
SARDINIAN_CAMPIDANESE	40	0.563	0.379	0.388	0.383
EMILIANO_FERRARESE	40	0.587	0.387	0.386	0.386
ROMANIAN	39	0.579	0.393	0.39	0.391
ITALIAN	40	0.543	0.388	0.397	0.392
SARDINIAN_LOGUDARESE	40	0.508	0.43	0.38	0.403
TURIA_ROMANIAN	23	0.585	0.425	0.383	0.403
SICILIAN_UnnamedInSource	38	0.527	0.445	0.421	0.433
SARDINIAN	39	0.489	0.458	0.422	0.439
Jäger (this volume)	40	0.48	0.441	0.481	0.46

Table 3: Applying B-Cubed scores to evaluate Jäger’s automated reconstruction.

In Table 3, the results for the top-ten comparisons of Latin with the other Romance languages and Jäger’s reconstructions are shown, ordered in increasing number by their B-Cubed F-Scores. While the differences in normalized edit distances in Jäger’s system are very small, amounting to less than 0.01 between Sardinian and Jäger’s system, the differences in B-Cubed F-Scores show that Jäger’s algorithm is structurally much more similar to Latin than the extant languages. What is furthermore interesting is that Jäger’s system is best in *predicting* Latin, as we can see from the recall, while it is less well *predicted* by Latin than, for example, Sardinian or Sicilian. All in all, this sums up to a considerable advantage of 0.021 points in the B-Cubed Scores in comparison with Sicilian.

7 Outlook

While the new measure for the comparison of reconstruction systems proposed here has several advantages over edit distance as it was used in the previous literature (Bouchard-Côté et al. 2013), it cannot fully account for the abstract nature of reconstruction systems. The problem is that what I called *structural identity* is not the only structural difference we can observe when dealing with different reconstruction systems. Another type is what one could call *structural equivalence*, as illustrated in Table 4.

Number	System A	System A'	System B	System B'
1	k a η	1 2 3	k ə η	1 2 3
2	k u η	1 4 3	k ^w ə η	4 2 3
3	k i η	1 5 3	k ⁱ ə η	5 2 3
4	p a k	6 2 1	p ə k	6 2 1
5	p u k	6 4 1	p ^w ə k	7 2 1
6	p i k	6 5 1	p ⁱ ə k	8 2 1

Table 4: Structurally equivalent reconstruction systems which are not structurally identical.

The two reconstruction systems differ not only substantially but also structurally from each other. This can be seen from their numerical representations, which would be identical if the systems were structurally identical. But both systems encode the same information, and we can convert one system into the other and vice versa with help of a substitution cipher that takes *context* into account. In such a model, *k* in System A would correspond to *k* in System B when followed by *a*, and to *k^w* and *kⁱ* when followed by *u* or *i*, respectively, and *ə* in System B can be converted to *a* when preceded by *k* or *p*, or to *u* or *i* when preceded by a labialized or a palatalized consonant, respectively. The fact that we can convert one system into the other and vice versa without loss of information shows that they can serve the same purpose as proto-languages and predict the daughter languages equally well. To identify their structural equivalence, however, B-Cubed scores are not sufficient, since they can only identify structural identity, and more research will be needed to account for structural equivalence in addition to structural identity when dealing with reconstruction systems.

Acknowledgments

This research was funded by the ERC Starting Grant 715618 “Computer-Assisted Language Comparison” (<http://calc.digling.org>).

Supplementary Material

The code and the data to calculate the B-Cubed scores for different reconstruction systems is available for download from Zenodo (<https://doi.org/10.5281/zenodo.3242122>).

References

- Amigó, E., J. Gonzalo, J. Artiles, and F. Verdejo. 2009. “A Comparison of Extrinsic Clustering Evaluation Metrics Based on Formal Constraints.” *Information Retrieval* 12 (4): 461–86.
- Bouchard-Côté, A., D. Hall, T. L. Griffiths, and D. Klein. 2013. “Automated Reconstruction of Ancient Languages Using Probabilistic Models of Sound Change.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (11): 4224–9.

- Brugmann, K.. 1904. *Kurze Vergleichende Grammatik Der Indogermanischen Sprachen. Auf Grund Des Fünfbändigen grundrisses Der Vergleichenden Grammatik Der Indogermanischen Sprachen von K. Brugmann Und B. Delbrück' verfasst.* Photomechanischer Nachdruck 1970. Walter de Gruyter & Co., Berlin. Strassburg: Karl J. Trübner.
- Ciobanu, A. M., and L. P. Dinu. 2018. "Simulating Language Evolution: A Tool for Historical Linguistics." In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, 68–72. Association of Computational Linguistics.
- Hall, R. A. 1960. "On Realism in Reconstruction." *Language* 36 (2): 203–6.
- Hauer, B., and G. Kondrak. 2011. "Clustering Semantically Equivalent Words into Cognate Sets in Multilingual Lists." In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, 865–73. AFNLP.
- Hill, N. W. 2016. "A Refutation of Song's (2014) Explanation of the 'Stop Coda Problem' in Old Chinese." *International Journal of Chinese Linguistic* 2 (2): 270–81.
- Hill, N. W., and J.-M. List. 2017. "Challenges of Annotation and Analysis in Computer-Assisted Language Comparison: A Case Study on Burmish Languages." *Yearbook of the Poznań Linguistic Meeting* 3 (1): 47–76.
- Jacques, G., and J.-M. List. Forthcoming. "Save the Trees: Why We Need Tree Models in Linguistic Reconstruction (and When We Should Apply Them)." *Journal of Historical Linguistics* 9 (1): 128–67.
- Jäger, G. and J.-M. List. 2018. "Using ancestral state reconstruction methods for onomasiological reconstruction in multilingual word lists." *Language Dynamics and Change* 8 (1): 22-54.
- Jäger, G., J.-M. List, and P. Sofroniev. 2017. "Using Support Vector Machines and State-of-the-Art Algorithms for Phonetic Alignment to Identify Cognates in Multi-Lingual Wordlists." In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Long Papers*, 1204–15. Valencia: Association for Computational Linguistics.
- Jakobson, R.. 1971. "Typological Studies and Their Contribution to Historical Comparative Linguistics." In *Selected Writings*, by Roman Jakobson, 1:523–32. The Hague: Mouton.
- Joseph, B. D. 2006. "On Projecting Variation Back into a Proto-Language. With Particular Attention to Germanic Evidence and Some Thoughts on "Drift"." In *Variation and Reconstruction*, edited by T. D. Cravens, 103–18. Amsterdam; Philadelphia: John Benjamins.
- Lass, R.. 2017. "Reality in a Soft Science: The Metaphonology of Historical Reconstruction." *Papers in Historical Phonology* 2 (1): 152–63.
- Levenshtein, V. I. 1965. "Dvoičnye Kody S Ispravleniem Vypadenij, Vstavok I Zameščenij Simvolov." *Doklady Akademij Nauk SSSR* 163 (4): 845–48.
- List, J.-M. 2014. *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.
- . 2018. "More on Network Approaches in Historical Chinese Phonology (音韵学)." In *The*

- 2nd Li Fang-Kuei Society Young Scholars Symposium*, 157–74. Taipei: Li Fang-Kuei Society for Chinese Linguistics.
- . 2019. “Automatic Inference of Sound Correspondence Patterns Across Multiple Languages.” *Computational Linguistics* 1 (45): 137–61.
- List, J.-M., S. J. Greenhill, and R. D. Gray. 2017. “The Potential of Automatic Word Comparison for Historical Linguistics.” *PLOS ONE* 12 (1): 1–18.
- List, J.-M., S. Greenhill, T. Tresoldi, and R. Forkel. 2018. “LingPy. A Python Library for Quantitative Tasks in Historical Linguistics.” Jena: Max Planck Institute for the Science of Human History. 2018. <http://lingpy.org>.
- List, J.-M., J. S. Pathmanathan, P. Lopez, and E. Baptiste. 2016. “Unity and Disunity in Evolutionary Sciences: Process-Based Analogies Open Common Research Avenues for Biology and Linguistics.” *Biology Direct* 11 (39): 1–17.
- Mann, N. W. 1998. “A Phonological Reconstruction of Proto Northern Burmic.” PhD, Arlington: The University of Texas.
- Meier-Brügger, M. 2002. *Indogermanische Sprachwissenschaft*. 8th ed. Berlin; New York: de Gruyter.
- Ross, M., and M. Durie. 1996. “Introduction.” In *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*, edited by M. Durie, 3–38. New York: Oxford University Press.
- Salmons, J. C. 1993. *The Glottalic Theory: Survey and Synthesis*. Washington: Inst for the Study of Man.
- de Saussure, F. 1879. *Mémoire Sur Le Système Primitif Des Voyelles Dans Les Langues Indo-Européennes*. Leipzig: Teubner.
- de Saussure, F. 1916. *Cours de Linguistique Générale*. Edited by C. Bally. Lausanne: Payot.
- Schleicher, A. (1861) 1866. *Compendium Der Vergleichenden Grammatik Der Indogermanischen Sprache*. 2nd ed. Vol. 1. Weimar: Böhlau.
- . 1868. “Eine Fabel in Indogermanischer Sprache.” In *Beiträge Zur Vergleichenden Sprachforschung Auf Dem Gebiete Der Arischen, Celtischen Und Slawischen Sprachen*, edited by A. Kuhn and A. Schleicher, 206–8. 5. Berlin: Ferdinand Dümmler.
- Schmidt, J. 1872. *Die Verwandtschaftsverhältnisse Der Indogermanischen Sprachen*. Weimar: Hermann Böhlau.
- Schrodt, R. 1989. “Neue Forschungen Zur Germanischen Lautverschiebung – Ein Fall von Paradigmenwechsel?” In *The New Sound of Indo-European: Essays in Phonological Reconstruction*, edited by T. Vennemann, 137–52. Berlin: de Gruyter.
- Serva, M., and F. Petroni. 2008. “Indo-European Languages Tree by Levenshtein Distance.” *EPL* 81 (6).