

Pragmatics of Language Evolution

Johann-Mattis List
mattis.list@shh.mpg.de

DGfS Summer School on Experimental Pragmatics
Zentrum für Allgemeine Sprachwissenschaft, Berlin

Summer Semester 2019

Contents

1 Introduction	3
Introduction	3
Empirical Research on Language Evolution	12
2 Language Change	21
Empirical Studies on Sound Change	21
Cross-Linguistic Studies on Semantic Change	30
Empirical Approaches to Studying Language Contact	38
3 Beyond Language Change	47
Modeling Language Change	47
Evolution of Speech Acts	56
Evolution of Poetry	64

1 Introduction

The two introductory sessions discuss general aspects of historical language comparison, historical linguistics, and language evolution.

Introduction

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

The fact that “all languages evolve, as long as they exist” (Schleicher 1863: 18f) has been long known to linguists and does not surprise us anymore. The reasons why all language change constantly, however, is still not fully understood. What we know, however, is that language usage must be at the core of language evolution. It is the dynamics among speakers, who want to be understood and understand what others say, while at the same time trying to be efficient, convincing, or poetic when communicating with others. If the dynamics of language use are indeed one of the driving forces of language evolution, it is evident that the phenomena of language change need to be studied from the perspective of pragmatics. In times of constantly increasing amounts of digital language data, in various forms, ranging from wordlists via results of laboratory experiments to large historical corpora, it is clear that every attempt to understand the specific dynamics of language evolution must be carried out in an empirical framework. In the course, I will try to give a rather broad (but nevertheless eclectic) introduction into topics in historical linguistics in which pragmatics play a crucial role for the study of language change and its driving forces. In this context, we will look into empirical aspects of research on language evolution, empirical studies on sound change, and the pragmatics of language contact. In addition, we will also learn how language change can be modeled, and how we can study pragmatic phenomena themselves from an evolutionary perspective by investigating how speech acts and poetic traditions evolve.

1 Introduction

1.1 Nothing in linguistics makes sense except ...

In the famous essay *Nothing in biologists makes sense except in the light of evolution*, originally published in 1973, Theodosius Dobzhansky tried to defend evolutionary thinking in biology against its rival theories of creationism by emphasizing that evolutionary science would not necessarily be a threat to religious beliefs. Linguists feel the threat of creationism less strongly, and it happens rarely that new findings on the ancestry of our world’s largest language families are actively discussed in creationists’ circles¹ But given the importance that evolutionary thinking plays throughout all branches of biology, including those that do not necessarily deal with evolution *per se*, such as clinical studies, morphological studies, or bacteria, it is surprising that the historical (or evolutionary) aspect of language is so far only investigated by specialists, while the linguistic “mainstream”, be it fieldworkers, syntacticians, semanticists, or typologists, rarely resort to historical explanations in their research.

[Q] *Why would biologists bother about evolutionary questions when studying diseases?*

1.2 A tradition of ignoring history

People at times blame Ferdinand de Saussure’s (1857-1913) distinction between diachrony and synchrony (Saussure 1916) for the current lack of interest in historical questions in mainstream linguistics. Another reason certainly also lies in the problematic closeness of at least some proponents of historical linguistics to the racist theories culminating in the Nazi regime in Germany (Behr 2019). But given that

¹For a rare exception, see the discussion of a recent analysis of Sino-Tibetan divergence times by Sagart et al. (2019) at uncommondescent.com.

every language is beyond doubt a *product* of its history, that all facts about a language in a given stage go back to its specific history, it is still surprising to me to which degree the evolutionary perspective has been banned from the agenda in the field of mainstream linguistics.

[Q] *What importance does evolutionary thinking or historical thinking play in the specific linguistic research you pursue?*

1.3 Historical linguistics after the quantitative turn

By the end of the 20th century, not only mainstream linguistics had largely abandoned historical and evolutionary thinking, but also the field of historical linguistics itself had largely marginalized itself. The biggest questions concerning the history of the Indo-European languages seemed to have been solved or turned out to be unsolvable, public outreach could not be guaranteed for small-scale studies, and institutes began to be closed. When I started to study Indo-European linguistics at the Freie Universität Berlin in 2003, I knew already by then that my institute would be closed in 2013, and I grew up, scientifically, with the idea in mind that – as much as language change and the methodology of historical linguistics were fascinating me by then – it would be rather unlikely to find a job as a scientist in this dying field.

Interestingly, however, this changed quickly already during the time when I began to study historical linguistics. The *quantitative turn* (Geisler and List 2013), which started around the second millennium and reflected in a steadily growing amount of studies in historical linguistics devoted to quantitative topics, including phylogenetic reconstruction (Gray and Atkinson 2003), automated sequence comparison (Kondrak 2000), and the creation of large-scale online databases (Dryer and Haspelmath 2013), revived not only the interest of scholars in language evolution, but also led to a drastic increase with respect to public awareness, which up to now has also led to a drastic increase in public funding.²

[Q] *Despite the popularity of the new approaches in historical linguistics and linguistic typology, which is specifically also reflected in the growing amount of younger scholars and teams, scholars repeatedly raise critics, such as Dixon (2019), claiming that the new quantitative methods have “no relevance with respect to the established discipline of comparative-historical linguistics” (Chap. 11, Fn. 7). What are – in your opinion – the major aspects of the “established discipline of comparative-historical linguistics”?*

2 Levels of analysis in the historical sciences

2.1 Historiography

In order to understand more clearly the objectives of historical and typological language comparison, we need to first get a clearer picture of historical research and research in the historical sciences in general. As a starting point, let's have a look at the following three example sentences, which all reflect – in the one or the other way – different levels of analysis in the field of historiography.

- (A) “Julius Caesar was murdered by Junius Brutus with a considerable amount of stabs.”
- (B) “The assassination of senators was a frequent phenomenon in the history of ancient Rome.”
- (C) “*The structure of scientific revolutions* by Kuhn (1996) discusses the evolution of new models of scientific description.”

²This is not only reflected in the newly established Department of Linguistic and Cultural Evolution (Director: Russell D. Gray) at the Max Planck Institute for the Science of Human History in Jena, but also in numerous research grants that pursue historical and typological language comparison of different areas of the world.

1 Introduction

These three examples present us with three different aspects of historiography. The first example represents the description of an *individual event*. The second example describes the characteristics of a certain range of *specific events in a certain context*. The third example, finally, describes the characteristics of specific events in a *general context*. From these three examples, we can thus derive that historiography deals (at least) with three distinct topics, namely

1. the description of individual events,
2. the description of event types in a specific context, and
3. the description of event types in a general context.

Whether this account of historiography is exhaustive, is difficult to say, and historians or philosophers of science may have a much more detailed viewpoint on this, but I consider the distinction between the three levels of analysis (*individual, specific, general*) as a useful starting point to discuss the different levels of analysis important for research in historical linguistics.

[Q] Georg Wilhelm Friedrich Hegel (1770-1831) the German philosopher, who often wrote in a completely incomprehensible manner, presented in his *Vorlesungen über die Philosophie der Geschichte* (Lectures on the philosophy of history), posthumously published in 1837 three forms of historiography, which he called “*ursprüngliche Geschichte*” (“original history”), “*reflectirte Geschichte*” (“reflected history”), and “*philosophische Geschichte*” (“philosophical history”). How well do these terms match with the distinction of individual, specific, and general descriptions made above?

2.2 Linguistic Historiography (Sprachgeschichtsschreibung)

How can we compare the three levels of historiographic analysis with the levels of analysis we find in historical linguistics and *linguistic historiography*? We can again start from three examples.

- (A) “In Proto-Indo-European, the word for ‘father’ was **ph₂tér*.”
- (B) “English was heavily influence by French throughout its history.”
- (C) “Labial plosive consonants often change into labial fricatives.”

It should not be difficult to compare the three examples with the examples on historiography above: the first example presents an *individual fact* about a certain language; the second example describes *specific characteristics of the development of a specific language*; and the third example, finally, describes *processes and general tendencies in language evolution*.

From these examples, we can again, derive fundamental tasks of historical language comparison (and also for linguistic typology, as we will see). The first task is the *description of events and states of a given language*, which explicitly includes the description of different historical stages of a given language, be they reconstructed or still reflected in literature. The second task is the *description of evolutionary processes within one specific language or a group of languages*. The third task, finally, is the *description and analysis of general aspects of language evolution*.

[Q] The famous linguist Georg von der Gabelentz (1840–1893) made an interesting distinction between *internal and the external language history* (Gabelentz 2016). Where could we allocate this distinction in our three levels of linguistic historiography above?

2.3 From historical linguistics to diversity linguistics

One may wonder why I keep talking about historical linguistics *and* linguistic typology in this summary. The reason is that the phenomena we typically deal with in historical linguistics barely belong to the field of historical linguistics alone, but include also the field of linguistic typology. Martin Haspelmath has coined the term *diversity linguistics* to address those fields that deal with language diversity, including explicitly historical linguistics and linguistic typology.³ What seems important to me in this context is the fact that we can see from the examples for the different levels of linguistic analysis, that it is not easy to make a strong distinction between typology and history. Typology is just the consequence of individual and specific description, it is, what deals with *general* questions, reflecting what Gabelentz (2016) called *general linguistics* (*Allgemeine Sprachwissenschaft*).

This scientific discipline [i.e., “general linguistics”, JML] has the human language capacity itself as its research object. It wants to understand this capacity, not only with respect to the specific abilities of spirit and body, of which it is composed, but also, and as much as this is achievable, in the whole extend of its developments. (ibid.: 317)⁴

In the light of the three levels of research in linguistics, mentioned above, the traditional distinction between historical linguistics and linguistic typology (as well as areal linguistics) is basically wrong. What is important is the perspective of a given research. Here Haspelmath (2019) proposes to distinguish *p(articular)-linguistics* (which would be “Einzelsprachenlinguistik” in German) from *g(eneral)-linguistics* (which would be “allgemeine Sprachwissenschaft” in the sense of Gabelentz). Following Gabelentz, we should include *c(omparative)-linguistics* into this model. The distinction between *particular*, *comparative*, and *general* linguistics should make clear why the distinction between historical linguistics and linguistic typology (and at times also areal linguistics) is problematic: scholars who are called “historical linguists” often discuss both questions of p-linguistics, e.g., when describing an ancient state of a given language, and they may also discuss g-linguistics, as they will usually try to derive general trends and tendencies from the comparative data they accumulated. In the same sense, scholars working on areal linguistics can investigate the state of a particular language with respect to its location, they can compare languages in specific areas, and they can search for tendencies of convergence across different linguistic areas of the world.

[Q] When discussing problems of linguistic change, Eugenio Coseriu (1921-2002) suggests to distinguish three basic problems in linguistics, (a) the rational problem of change (“problema racional del cambio”), (b) the general problem of change events (“problema general de los cambios”), and (c) the historical problem of a given change (“problema histórico de tal cambio determinado”, Coseriu 1973: 65f). To which degree do Coseriu’s distinctions complement or contradict the distinction in p-, c-, and g-linguistics made above?

2.4 Language evolution and language history

While I pointed to theoretical problems of making a clear-cut distinction between historical linguistics and linguistic typology, it is obvious that the terms are still useful in practice, not only because of their long tradition, but also because historical linguistics explicitly deals with *change*, i.e., with *evolution*, while the aspect of evolution is not necessarily regarded as essential and important by all typologists (recall that linguists do not necessarily explain all of language in the light of evolution). The distinction into the different levels of analysis is therefore more important for those who describe themselves as “historical

³Unfortunately, I do not know of any publication in which Haspelmath explained this term, but his blogpost Diversity Linguistic Comment finds many examples for the usage of this term (apart from the very title of the blog).

⁴My translation, original text: Diese Wissenschaft hat das menschliche Sprachvermögen selbst zum Gegenstande. Sie will dies Vermögen begreifen, nicht nur in Rücksicht auf die geistlichen Kräfte und Anlagen, aus denen es sich zusammensetzt, sondern auch, soweit dies erreichbar ist, dem ganzen Umfange seiner Entfaltungen.

1 Introduction

linguists”, since it is *in* the field of historical linguistics, where scholars often confuse the different levels of particular, comparative, and general linguistics. While we can easily see (and I have tried to show) that the field of historical linguistics usually covers all three levels of analysis, it is obvious that the first level is a very specific case that is not “historical” *per se*. When dealing with the pragmatics of language evolution, our key objective is to solve questions of *general linguistics*, but since we make extensive use of *comparative linguistics* to feed our data on what we consider should reflect general linguistics, we will have to look at both comparative and general historical linguistics in due detail in this course. In contrast to modern syntax, semantics, and much of the research on linguistic universals, historical linguistics is hereby understood as the discipline that looks at the evolution of the phenomena under question, and that tries to explain these phenomena by investigating their development (or evolution).

[Q] *Is it possible or useful to speak of a sub-discipline of “particular historical linguistics”?*

3 Research workflows in the historical sciences

3.1 Three stages of research

When investigating the research practice not only in historical linguistics, but in the historical sciences in general, we can identify a specific research workflow that consists of three essential stages: *modeling*, *inference*, and *analysis*. This triad is inspired by (Dehmer et al. 2011: XVII) and expresses the general idea that (1) each scientific investigation necessitates a certain *model* or *idea* about the object of investigation (the stage of *modeling*) and that (2) we use different *methods* to *infer* more examples for the phenomena that our model predicts (stage of *inference*). Thanks to the inferred examples, we can finally (3) carry out *analyses* which ultimately help us to refine our model (stage of *analysis*). While the bulk of linguistic research that gets into the spotlight of media, funders, and colleagues belongs to the third stage of research, the first two stages are of crucial importance in order arrive at a point where an analysis is possible. For this reason, all aspects will be discussed in brief detail in the following, and I will try to introduce examples for all aspects throughout the course.

[Q] *While I am convinced that the distinction of modeling, inference, and analysis is useful for historical linguistics, I am less convinced of research in other areas. What do you think about the three aspects in the light of your specific experience and research in your specific field of linguistics?*

3.2 Modeling

Modeling is important for the description of a given problem. Without basic models we could not make any conclusions about historical events that constitute a specific development in the internal history of a given language or the external history of a given language family. For example, if we wanted to study different phenomena of *language contact*, we could not do this without having at least some idea of the phenomena involved in it, such as, for example, *lexical borrowing*, *syntactic inference*, or other processes which are characteristic for language contact (Weinreich 1953 [1974]). Similarly, we could not investigate the phenomenon of sound change without having a clear idea of phonetic (or phonological) change.

An example for the importance of modeling can be found in the detection of Verner (1877) that all those cases which Grimm (1822) had listed as exceptions to the typical change of plosives turning into fricatives (among others) in Germanic, first identified by Rask (1818), can in fact be explained when refining the Proto-Germanic reconstruction by adding stress patterns still reflected in the Vedic language. Without the idea that sound change should be inherently *regular*, Verner would never have felt triggered

to try to find the reason that caused the apparent exception. If Verner's model of sound change had been the one proposed by lexical diffusion theory (Chen 1972, Wang 1969), he might have never detected the explanation for the exception, since lexical diffusion allows and predicts a certain level of unexplainable exceptions, which renders the model useless for rigorous research in historical linguistics (Hill 2016).

[Q] *In an early paper dealing with phylogenetic approaches in historical linguistics, Atkinson and Gray (2006) emphasize that "models are lies that lead us to the truth" (94). What do they mean with this sentence?*

3.3 Inference

The stage of inference provides us with the data we need for our analysis. Without data, we could not understand a given problem, we could also not test our hypotheses or models. As in probably all scientific research, whenever scholars start to work on a given problem, their original data are insufficient and small in size. They are enough to help them to establish a first model, but not enough to test this model for its suitability. This iterative procedure is very typical for those historical linguistics who work in the framework of the so-called *comparative method* (Ross and Durie 1996), which can be seen as paradigmatic for an iterative procedure based on modeling, inference, and analysis. According to this framework, linguists start with a certain theory (they assume a set of languages to be related). They then collect first-order evidence in favor of this theory, which they again test against their original model. If the inferred data contradicts the model, either the model is refined or parts of the data discarded. In this fashion, the problem of circularity is circumvented by making use of an iterative fashion which accumulates evidence step by step, until a sufficient amount of inferences has been made to prove the original theory of relatedness.

In the field of comparative-historical linguistics, inference plays a crucial role. Unfortunately, the original techniques or "methods" for inference are rarely well-described, and it is expected that those who want to become historical linguists themselves should study from examples until they reach the level of proficiency (Schwink 1991). Although the methodological basis of historical linguistics is poorly established if not non-existent in many cases, practicing historical linguists tend to know very well what they are doing and why. While this holds specifically for the classical tasks of establishing cognates and regular sound correspondences for a given set of languages, there are other aspects of historical linguistics where scholars have to resort to ad-hoc assumptions or negative criteria. Here, one should probably not blame the field for providing insufficient descriptions of their methodology, but rather the phenomenon of language evolution, which is in the end responsible for the complexity. An example for a rather *hard* problem is the inference of borrowings, for which clear-cut methods hardly exist (List forthcoming).

[Q] *Why should it be more difficult to detect borrowings than genetically related words?*

3.4 Analysis

Any investigation that wants to exceed the stage of pure description needs to conduct an analysis in some form. Analyses are not only important for generalizations, but also for explanations in the historical sciences. Despite what one might think when reading research papers, an analysis does not necessarily constitute the end-point of an investigation, but it may often help to enhance the models which were used to carry out the inference.

While the majority of analyses in classical historical linguistics are done in a qualitative fashion, with scholars searching through data and testing their intuitive assumptions by investigating the data item by item, the quantitative turn in historical linguistics has led to an enormous increase in quantitative analyses, ranging from individual phylogenetic reconstructions for particular language families (Greenhill

1 Introduction

et al. 2009, Kolipakam et al. 2018, Sagart et al. 2019), via the investigation of cross-linguistic tendencies (Brown et al. 2013, List et al. 2018), up to attempts to obtain general insights into the dynamics of language evolution (Greenhill et al. 2017, Hua et al. 2019).

[Q] *What are the advantages and disadvantages of qualitative and quantitative analyses? What kinds of analyses do you perform in your research?*

4 Outlook

4.1 Goal of the course

The goal of this session was to introduce the participants to some basic aspects of thinking that I think are important when discussing how research is done in historical linguistics. Judging from the growing number of publications, the growing number of grants awarded to individual researchers, and a generally growing public interest in the results of our research, historical linguistics is currently gaining popularity. As a result, there are more and more researchers who want to complement their primarily synchronic research by taking an evolutionary viewpoint or investigating data that was produced by means of comparative analyses. Given the lack of systematicity in historical linguistics, however, and the unspoken rule that those who want to understand the principles of language change, it is still very difficult to obtain insights into the field by relying on one handbook alone. When discussing different aspects of language change in which pragmatics plays a role, my major goal is *not* to provide the participants with new insights into the pragmatic aspects of language evolution. Instead, my goal is to provide an overview on interesting topics in historical linguistics where pragmatic aspects are of a certain relevance for modeling and analysis, and to contrast these with the traditional and more recent quantitative methods for historical language comparison. In this way, I hope to equip the participants with the basic knowledge of historical linguistics methodology that they can themselves use their knowledge about pragmatics to investigate their own research questions.

4.2 Topics

The topics which I selected are a compromise of topics that I find interesting and topics that I know well enough to teach them. There will be some topics which I just find interesting, and it is possible that participants will have more insights into some of the questions. In all cases, however, I think that a specific historical (or evolutionary) perspective will provide the chance to create and test fresh ideas. In many cases, I will start from introducing our purely linguistic perspective on the phenomena in question and then try to elaborate with the participant to which degree a pragmatic perspective could be interesting to deepen our understanding of the problems. I will give a definite preference to the recently developed, empirical, and quantitative approaches and try to avoid approaches in historical linguistics that are exclusively based on qualitative analyses presented in prose. In doing so, I do not want to negate the importance and usefulness of qualitative approaches and insight, but rather reflect my personal strengths and insights.

The following table gives a rough overview on the content planned for each of the upcoming sections, contrasting the original title with some notes that further specify how I plan the content. Since I originally thought that I would teach nine sessions, I have merged the conclusion session with the session on the evolution of poetry.

No.	Topic	Content
1	Introduction	Try to set the basis for an understanding of the <i>historical (or evolutionary)</i> questions in historical linguistics.
2	Empirical research on language evolution	Discuss the basic aspects of the research that has been carried out in the field of quantitative historical linguistics since the quantitative turn.
3	Empirical studies on sound change	Given an overview of the phenomenon of sound change and discuss recent approaches and attempts to explain its actuation.
4	Cross-linguistic studies on semantic change	Discuss the phenomenon of semantic change and show how it can be studied cross-linguistically.
5	Empirical approaches to studying language contact	Present basic aspects of (primarily lexical) language contact situations and discuss new data-driven approaches.
6	Modeling language change	Discuss the general problem of modeling language change and provide an overview on recent attempts to solve it.
7	The evolution of speech acts	Introduce the handling of large-scale linguistic databases in quantitative historical linguistics and discuss how certain aspects of speech acts or linguistics rituals could be investigated.
8	The evolution of poetry and conclusion	Present some general ideas of how to study the evolution of poetry and make a final wrap-up.

References

- Atkinson, Q. D. and R. D. Gray (2006). "How old is the Indo-European language family? Illumination or more moths to the flame?" In: *Phylogenetic methods and the prehistory of languages*. Ed. by P. Forster and C. Renfrew. Cambridge, Oxford, and Oakville: McDonald Institute for Archaeological Research, 91–109.
- Behr, W. (2019). "«Urheimat» der Chinesen. Die Sprachwissenschaft und die Suche nach «Wurzeln» [Homeland of the Chinese people. Linguistics and the search for roots]". *Geschichte der Gegenwart* 4.6, 1–6. eprint: <https://geschichtedergegenwart.ch/urheimat-der-chinesen-die-sprachwissenschaft-und-die-suche-nach-wurzeln/>.
- Brown, C. H., E. W. Holman, and S. Wichmann (2013). "Sound correspondences in the world's languages". *Language* 89.1, 4–29.
- Chen, M. (1972). "The time dimension. Contribution toward a theory of sound change". *Foundations of Language* 8.4, 457–498. JSTOR: 25000618.
- Coseriu, E. (1973). *Sincronia, diacronia e historia. El problema del cambio lingüístico* [Synchrony, diachrony, and history. The problem of linguistic change]. Madrid: Biblioteca Románica Hispánica.
- Dehmer, M., F. Emmert-Streib, A. Graber, and A. Salvador, eds. and introd. (2011). Weinheim: Wiley-Blackwell.
- Dixon, R. M. W. (2019). *Australia's original languages. An introduction*. Sydney: Allen & Unwin.
- Dobzhansky, T. (1973). "Nothing in biology makes sense except in the light of evolution". *The American Biology Teacher* 35.3, 125–129.
- Dryer, M. S. and M. Haspelmath, eds. (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Gabelentz, G. von der (2016). *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Ed. by M. Ringmacher and J. McElvenny. Berlin: Language Science Press.
- Geisler, H. and J.-M. List (2013). "Do languages grow on trees? The tree metaphor in the history of linguistics". In: *Classification and evolution in biology, linguistics and the history of science. Concepts – methods – visualization*. Ed. by H. Fangerau, H. Geisler, T. Halling, and W. Martin. Stuttgart: Franz Steiner Verlag, 111–124.
- Gray, R. D. and Q. D. Atkinson (2003). "Language-tree divergence times support the Anatolian theory of Indo-European origin". *Nature* 426.6965, 435–439.
- Greenhill, S. J., T. E. Currie, and R. D. Gray (2009). "Does horizontal transmission invalidate cultural phylogenies?" *Proc. Biol. Sci.* 276.1665, 2299–2306.
- Greenhill, S. J., C. H. Wu, X. Hua, M. Dunn, S. C. Levinson, and R. D. Gray (2017). "Evolutionary dynamics of language systems". *Proceedings of the National Academy of Sciences of the United States of America* 114.42, E8822–E8829.
- Grimm, J. (1822). *Deutsche Grammatik*. 2nd ed. Vol. 1. Göttingen: Dieterichsche Buchhandlung. Google Books: MnsKAAAAIAAJ.
- Haspelmath, M. (2019). "Confusing p-linguistics and g-linguistics: Philosopher Ludlow on «framework-free theory»". *Diversity Linguistics Comment* 7.6. eprint: <https://dic.hypotheses.org/1801>.
- Hegel, G. W. F. (1837). *Georg Wilhelm Friedrich Hegels Vorlesungen über die Philosophie der Geschichte*. Ed. by E. Gans. Georg Wilhelm Friedrich Hegels Werke 9. Berlin: Duncker und Humblot. Google Books: y_EGAAAAcAAJ.
- Hill, N. (2016). "A refutation of Song's (2014) explanation of the 'stop coda problem' in Old Chinese". *International Journal of Chinese Linguistics* 2.2, 270–281.
- Hua, X., S. J. Greenhill, M. Cardillo, H. Schneemann, and L. Bromham (2019). "The ecological drivers of variation in global language diversity". *Nature Communications* 10.2047, 1–10.
- Kolipakam, V., F. M. Jordan, M. Dunn, S. J. Greenhill, R. Bouckaert, R. D. Gray, and A. Verkerk (2018). "A Bayesian phylogenetic study of the Dravidian language family". *Royal Society Open Science* 5.171504, 1–17.
- Kondrak, G. (2000). "A new algorithm for the alignment of phonetic sequences". In: *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*. (Seattle, 04/29–05/03/2000), 288–295.
- Kuhn, T. S. (1996). *The structure of scientific revolutions*. 3rd ed. Chicago: University of Chicago Press.
- List, J.-M. (forthcoming). "Automated methods for the investigation of language contact situations". *Language and Linguistics Compass* ????, 1–19.
- List, J.-M., S. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2018). "CLICS². An improved database of cross-linguistic colexifications assembling lexical data with help of cross-linguistic data formats". *Linguistic Typology* 22.2, 277–306.
- Rask, R. K. (1818). *Undersøgelse om det gamle Nordiske eller Islandske sprogs oprindelse* [Investigation of the origin of the Old Norse or Icelandic language]. Copenhagen: Gyldendalske Boghandlings Forlag. GoogleBooks: cHqJAAAAQAAJ; English translation: – (1993). *Investigation of the origin of the Old Norse or Icelandic language*. Trans. by N. Ege. Travaux du Cercle Linguistique de Copenhague 26. Copenhagen: The Linguistic Circle of Copenhagen.
- Ross, M. and M. Durie (1996). "Introduction". In: *The comparative method reviewed. Regularity and irregularity in language change*. Ed. by M. Durie. New York: Oxford University Press, 3–38.
- Sagart, L., G. Jacques, Y. Lai, R. Ryder, V. Thouzeau, S. J. Greenhill, and J.-M. List (2019). "Dated language phylogenies shed light on the ancestry of Sino-Tibetan". *Proceedings of the National Academy of Science of the United States of America* 116 (21), 10317–10322.
- Saussure, F. de (1916). *Cours de linguistique générale*. Ed. by C. Bally. Lausanne: Payot; German translation: – (1967). *Grundfragen der allgemeinen Sprachwissenschaft*. Trans. from the French by H. Lommel. 2nd ed. Berlin: Walter de Gruyter & Co.
- Schleicher, A. (1863). *Die Darwinsche Theorie und die Sprachwissenschaft. Offenes Sendschreiben an Herrn Dr. Ernst Haeckel*. Weimar: Hermann Böhlau. ZVDD: urn:nbn:de:vbv:12-bsb10588615-5.
- Schwink, F. (1991). *Linguistic typology, universality and the realism of reconstruction*. Washington: Institute for the Study of Man.
- Verner, K. A. (1877). "Eine Ausnahme der ersten Lautverschiebung [An exception to the first sound shift]". *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 23.2, 97–130.
- Wang, W. S.-Y. (1969). "Competing changes as a cause of residue". *Language* 45.1, 9–25. JSTOR: 4117748.
- Weinreich, U. (1974). *Languages in contact. With a preface by André Martinet*. 8th ed. The Hague and Paris: Mouton.

Empirical Research on Language Evolution

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

The goal of this session is to discuss briefly the basic aspects of the research that has been carried out in the field of quantitative (or computational) historical linguistics since the quantitative turn.

1 The quantitative turn in historical linguistics

1.1 Background

In the early 1950s, Morris Swadesh (1909–1967) presented a method to measure the genetic closeness between languages on the basis of a statistical formula that was ultimately based on counting the amount of shared cognates across standardized wordlists of different languages (Swadesh 1950). Although it seemed at first that the methods could revive the discipline of historical linguistics, which had past its prime after the structuralist turn in the begin of the 1920s, and had not seen any major methodological or analytical improvement since the begin of the 20th century.¹ Unfortunately, the original interest in the new ideas did not last long, and soon after it was first published, the new method was heavily criticized (Bergsland and Vogt 1962), and went out of vogue some 10 years later.

In the begin of the second millennium, Gray and Atkinson (2003) used similar data but different statistical methods to date the age of the Indo-European language family. They caused a similar stir as Swadesh had done almost half a century ago. But while Swadesh's method was filed away soon after it had been proposed, the method of Gray and Atkinson was part of a general *quantitative turn in historical linguistics*, which started at the begin of the second millennium. This quantitative turn is reflected in a large bunch of literature on such different topics as phonetic alignment (Kondrak 2000, Prokić et al. 2009), automated cognate detection (List 2014), and phylogenetic reconstruction (Atkinson and Gray 2006).

[Q] What may have been the reasons why Swadesh's approach was abandoned so quickly by historical linguists?

1.2 New studies on language evolution

We can distinguish four different aspects of research approaches in the course of the quantitative turn. As a first and most prominent aspect, we have research dealing with questions of *phylogenetic reconstruction* which usually involved *dating* as well. Language data are not only analyzed to yield a topology of the branching structure of the language family in question, but in addition, absolute branch lengths are often also inferred, which allow to estimate when a given language family has originated. The software and methods used for these studies are usually taken or inspired from approaches developed first in evolutionary biology. As of now, quite a few different language families have been analyzed in this way, including Indo-European (Chang et al. 2015, Gray and Atkinson 2003), Austronesian (Gray et al. 2009), Dravidian (Kolipakam et al. 2018), Bantu (Grollemund et al. 2015), Pama-Nyungan (Bowerman et al. 2011), Japonic (Lee and Hasegawa 2011), and Sino-Tibetan (Sagart et al. 2019). In addition, scholars have

¹The last major improvement, the decipherment of Hittite, which also helped to proof that it was an Indo-European language dated back to Hrozný (1915).

also attempted to provide unified methods that could be applied in a completely automated fashion to all languages of the world (Holman et al. 2011).

Another strand of research deals with the computation of inference procedures which were traditionally only carried out manually. Most prominently, we find here various attempts to automate different aspects of the general workflow of the traditional *comparative method* for historical language comparison (Weiss 2015). Breaking down the workflow into some of its major parts, we thus find (1) automated methods for the comparison of words, as reflected in methods for phonetic alignment (Kondrak 2000, Prokić et al. 2009) and automated cognate detection (Hauer and Kondrak 2011, List et al. 2016b, Turchin et al. 2010), (2) automated approaches for the detection of borrowings (List 2015, Mennecier et al. 2016, Nelson-Sathi et al. 2011),² (3) automated approaches for linguistic reconstruction (Bouchard-Côté et al. 2013, Jäger forthcoming), and (4) automated approaches for the detection of sound correspondences (List 2019).

While the second strand deals mostly with questions of inference, a third strand organizes inferred data in form of large-scale online databases that aggregate different kinds of information on the world's languages. The most prominent of these databases is beyond doubt the *World Atlas of Language Structures* (Dryer and Haspelmath 2013), but in addition we also find attempts to aggregate cross-linguistic information on phoneme inventories (Maddieson et al. 2013, Moran et al. 2014), polysemies (List et al. 2018), phonotactics (Donohue et al. 2013), borrowings (Haspelmath and Tadmor 2009), as well as datasets like D-Place, that compare cultural, environmental, and linguistic diversity (Kirby et al. 2016).

While the popular phylogenetic approaches deal with c-linguistics (or p-linguistics in a wider sense of the term), insofar as they deal with concrete languages in concrete times, trying to answer very specific (or *particular*) questions about their past, a fourth strand of research makes use of the new cross-linguistic databases along with results drawn from the phylogenetic approaches to investigate general aspects of language change, including questions like the rate of linguistic change and its correlates (Calude and Pagel 2011, Greenhill et al. 2017), the question to which degree environmental factors might have an impact on language evolution (Everett et al. 2015), or how language structures converge independent of contact or inheritance (Blasi et al. 2016).

[Q] *Why is the aspect of dating, i.e., the inference of absolute phylogenies, so important for the new methods in historical linguistics?*

1.3 Benefits of computational historical linguistics

Apart from the obvious benefit that the new quantitative methods have drastically revived the interest of scholars in historical linguistics, which also resulted in an increased amount of funding and a new generation of young scholars who are highly collaborative in their research and well trained in computational methods, the quantitative turn has also led to a considerable amount of rethinking in the field of historical linguistics, which offers new perspectives on the subject which have been ignored so far. First, we can see that the new methods shift the focus from *internal* to *external language* history, while at the same time turning away from the traditional focus on Indo-European alone.³ We can also see that the new methods lead to the raise of new questions, specifically addressing *general* questions of language history.

This is also reflected in new research approaches, which are more explicitly *data-centered* nowadays and often based on statistical or stochastic modeling. While research in historical linguistics has always been data-centered, the new methods have shown that the classical approaches to deal with

²See List (forthcoming) for an overview on these approaches.

³Compare classical handbooks such as the *Einführung in die vergleichende Sprachwissenschaft* by Szemerényi (1970), where the term *comparative linguistics* (which should be a general discipline) is seen as a synonym for *Indo-European linguistics*.

1 Introduction

data – namely the individual collection of extensive personal notes from the literature, and the publication of new insights from these personal collections in form of extensive prose – are reaching their limits in times where the amount of data is constantly increasing. Although the attempts to automate the classical methods have so far not yet led to a situation where computers could beat the experts,⁴ we have won many important and new insights into the methods and the practice of historical language comparison, specifically also because the new methods challenged classical (traditional) linguists to revise the methods they use and to increase the degree of explicitness by which they apply them.

[Q] *That languages interact with different factors is evident. What are the aspects that make it so difficult to study language change with help of computational frameworks?*

1.4 Problems and criticisms

Not all linguists have enthusiastically welcomed the new methods. While the various critics range from justified criticism, via exaggerations, up to complete ignorance for the initial goals of the computational approaches, and at times rather reflect the insulted ego of those who consider themselves as indisputable experts, the new field faces a couple of serious problems that are worth being criticized and rigorously analyzed. Among the most important of these are (1) problems with the data that is used in quantitative analyses, (2) problems of applicability of the computational approaches, and (3) problems of transparency and (4) comparability with respect to the results and methods which scholars report, and (5) problems of the general accuracy of the computational methods in comparison with experts.

The data problems related to the way in which data are compiled and curated, and what judgments they are based upon. The general problem here is that most of the phylogenetic approaches still make use of human-annotated data, trusting the expertise of only a small amount of experts to be enough to annotated data for at times more than 100 different languages. The danger of this procedure (which is to some degree difficult to avoid) are potential problems of inter-annotator-agreement, which may themselves, of course, impact the results (Geisler and List 2010). The problem of applicability and transparency is reflected in large amounts of software solutions and datasets that are only discussed in the literature, but have not been openly shared (List et al. 2017). As a result, there are quite a few methods out there that could provide valid solutions, but which have only been tested on one dataset and never officially been published, which comes close to a crisis of irreproducibility as it has been noted in many branches of science since the beginning of this millennium (Nature 2013).⁵

The problem of comparability results from missing standards in our field, which make it difficult to compare results across datasets, since it is often very tedious to lift the data used by different scholars to a level where they could be easily compared. The problem of accuracy, finally, is probably the hardest problem to address, since the problems of historical linguistics are often quite hard to solve automatically, specifically also because – as a rule – data is sparse, while most computational methods have been built based on the assumption that data to test and train algorithms would be abundantly available.

[Q] *What solutions can you think of to overcome the problems of transparency and comparability, which were mentioned above?*

⁴This is also not to be expected shortly, given that the only areas in which machines outperform humans so far are restricted fields, such as chess, or the go-game (Silver et al. 2016), and not in problems that need to be solved in open worlds.

⁵Luckily, this picture is slowly changing, thanks to extensive efforts to propagate free data and free code. At our department, for example, we have now decided to refuse to review papers where we are not given code and data, if they are needed for replication, following the idea of referee's rights as expressed by the editorial board of the journal Nature in 2018.

2 Towards a qualitative turn in diversity linguistics

2.1 Reconciling classical and computational research

The use of computer applications in historical linguistics is steadily increasing. With more and more data available, the classical methods reach their practical limits. At the same time, computer applications are not capable of replacing experts' experience and intuition, especially when data are sparse. If computers cannot replace experts and experts do not have enough time to analyse the massive amounts of data, a new framework is needed, neither completely computer-driven, nor ignorant of the assistance computers afford. Such computer-assisted frameworks are well-established in biology and translation. Current machine translation systems, for example, are efficient and consistent, but they are by no means accurate, and no one would use them in place of a trained expert. Trained experts, on the other hand, do not necessarily work consistently and efficiently. In order to enhance both the quality of machine translation and the efficiency and consistency of human translation, a new paradigm of computer-assisted translation has emerged (Barrachina et al. 2008: 3).

[Q] *Do you have experience with computer-assisted translation? If not, what role do computers and computer tools play for your research?*

2.2 Computer-assisted language comparison

Following the idea of computer-assisted frameworks in translation and biology, a framework for computer-assisted language comparison (CALC) is the key to reconcile classical and computational approaches in historical linguistics. Computational approaches may still not be able to compete with human experts, but when used to pre-process the data with human experts systematically correcting the results, they can drastically increase the efficiency of the classical comparative method and make up for the insufficiencies of current computational solutions. At the same time, bringing experts closer to computational and formal approaches will also help to increase the consistency of classical research, forcing experts to annotated their specific findings and corrections in due detail, without resorting to texts in prose and ad-hoc explanations.

[Q] *Classical linguists working on etymological research often emphasize the importance of looking into all details of language history, invoking the slogan "chaque mot a son histoire", which is, according to Campbell (1999: 189) traditionally attributed to Jules Gilliéron (1854-1926). Even if this was completely true, how can we still defend the recent attempts of computer-assisted and computer-based strategies in historical linguistics to work on a more formal and more quantitative handling of linguistic data?*

2.3 Data, Software, and Interfaces

In the framework of computer-assisted language comparison, data are constantly passed back and forth between computational and classical linguists. Three different aspects are essential for this workflow: Specific *software* allows for the application of transparent methods which increase the accuracy and the application range of current methods in historical linguistics and linguistic typology. Interactive *interfaces* serve as a bridge between human and machine, allowing experts to correct errors and to inspect the automatically produced results in detail. To guarantee that software and interfaces can interact directly, *data* need to be available in human- and machine-readable form.

1 Introduction

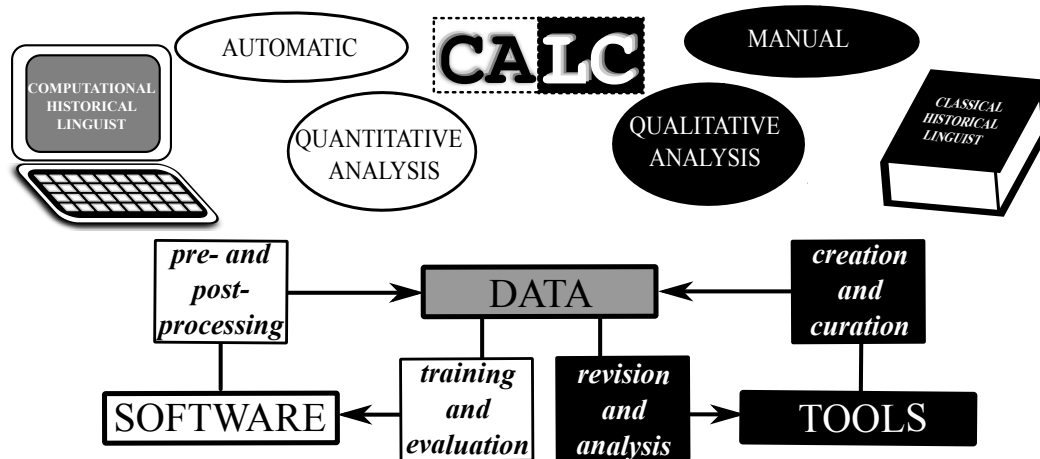


Fig. 1: Interplay of data, software, and interfaces in computer-assisted language comparison.

[Q] How exactly should one imagine data that are human- and machine-readable at the same time?

2.4 CALC project at the MPI-SHH in Jena

In the ERC-funded research project CALC (Computer-Assisted Language Comparison, List 2016), we try to establish a computer-assisted framework for historical linguistics. We pursue an interdisciplinary approach that adapts methods from computer science and bioinformatics for the use in historical linguistics. While purely computational approaches are common today, the project focuses on the communication between classical and computational linguists, developing interfaces that allow historical linguists to produce their data in machine readable formats while at the same time presenting the results of computational analyses in a transparent and human-readable way.

As a litmus test which proves the suitability of the new framework, the project attempts to create an etymological database of Sino-Tibetan languages (see Sagart et al. 2019 for initial attempts and results). The abundance of language contact and the peculiarity of complex processes of language change in which sporadic patterns of morphological change mask regular patterns of sound change make the Sino-Tibetan language family an ideal test case for a new overarching framework that combines the best of two worlds: the experience of experts and the consistency of computational models.

[Q] What may be the reason for choosing an interdisciplinary approach, and what are the most likely disciplines from which the project could take inspiration?

3 Important aspects of computational historical linguistics

To get a better understanding of the state of the art, the potential, and the limitations of computational approaches in historical linguistics after the quantitative turn, it is important to have a closer look at the problems, as they were outlined before, and how scholars try to address them today. Even more important, however, is to understand the basic ideas that underlie the new methods, and the topics that the methods deal with. To provide a short overview on these different aspects, we will follow the triad of *modeling*, *inference*, and *analysis*, as outlined in the session before. In this context, however, it is important to note explicitly that the division into the three aspects has its limits in practice, since what

counts as inference in a given research framework may at times count as analysis in another one and vice versa.

3.1 Modeling

The models that are used so far in computational historical linguistics are all rather simple. While this may at times be surprising for classical linguists, who have a very complex idea of change process and also very detailed knowledge of the complex range of what is possible in language change, reducing the complexity of models is a necessary step in all scientific research. Rather than trying to establish the most complex models before we start to infer something, we should investigate how far we can go with a simplifying model and where its specific limits lie.

Crucial aspects for the models in diversity linguistics are the concept of *language*, *word* (or linguistic sign), *word form*, and *word meaning*. Higher dimensions relevant for questions of language use, such as the speaker-listener interaction, are usually disregarded in the initial stages of investigation. The most common model for a language is to treat a given language as a bag of words (or a bag of linguistic signs). Depending on the perspective, one can invoke a set of grammatical rules by which these signs are combined to form sentences. The linguistic sign itself follows the basic idea of Saussure (1916) with the modification that the sign is not seen as a duplet of *form* and *meaning*, but a triplet of form, meaning, and the *language* to which the sign belongs (List 2014).

The sign form is usually modeled as a *sequence of sounds*, which implies that we can *segment* each word into a certain number of sounds. The sequences are constructed or constrained by *phonotactic rules*. If needed, one can add an additional layer of segmentation, dependent on the research question (e.g., one could look at a word consisting of morphemes consisting of sound segments, or a word consisting of syllables consisting of sound segments). These *secondary sequence structures* are of a certain importance in modern approaches for sequence comparison (List 2014, List et al. 2016b), but they are often also deliberately disregarded. While the sign form is best treated as a sequence of sounds, the sign meaning is usually handled as a *network of senses*.

While this model of language as a bag of words may seem very simple, it is effectively the model that was underlying most of the phylogenetic analyses that have been published so far. Additionally one should say, that even classical historical linguists tend to use this model in their analyses. When needed, throughout this course, we will discuss more complex models in due time.

To address the problem that we face a drastic lack of comparability with respect to the data that has been produced in diversity linguistics, the Cross-Linguistic Data Initiative (<https://clldf.clld.org>, Forkel et al. 2018) has published a set of recommendations for unified data standards in diversity linguistics, which are now gaining more and more popularity among scholars. These recommendations build more or less directly on the above-mentioned language model, and the current plan is to expand these further, based on the need and the availability of more complex models. As a very important aspect of standardization, CLDF comes along with *reference catalogs*, which are basically meta-datasets, that offer standards for the handling of languages (Glottolog, <https://glottolog.org>, Hammarström et al. 2018), concepts (Concepticon, <https://concepticon.clld.org>, List et al. 2016a), and sounds in transcription (CLTS, <https://clts.clld.org>, Anderson et al. 2018).

[Q] In addition to the modeling of the data, the modeling of the processes, which has been not mentioned here, is of great importance. What models can you think of that would explain, for example, the process of sound change, or the process of lexical change?

1 Introduction

3.2 Inference

As mentioned before, the inference of dated language phylogenies is by far the most popular of the computational methods proposed so far in the field of computational historical linguistics. Discussing the details of these approaches would, unfortunately, go beyond the scope of this session, but good review literature that provides some basic insights is now readily available (Greenhill 2015). What seems important to mention in this context is that the bag-of-words model mentioned before can be seen as the standard model that is essentially used to search for a language phylogeny. When discussing the simulation of language change in a later session, we will discuss more complex ways to simulate language change, which in theory also allow to handle the interaction between speaker and listener.

Second in popularity are methods for automated sequence comparison, which are very popular in dialectology, where methods for phonetic alignment are used to compute aggregate distances between dialect varieties, based on pronunciation distances derived from pre-selected lists of words (Nerbonne et al. 2011). In addition, methods for phonetic alignments are also used for the task of automated cognate detection, which tries to infer which words in a multi-lingual wordlist go back to the same ancestor. Techniques for automated cognate detection are quite well-developed by now, and have been shown to work surprisingly well, with accuracy scores of up to 90% on shallower language families (List et al. 2017), while the accuracy usually drops to around 60%-70% when dealing with larger datasets (Jäger et al. 2017). Further aspects of inference include automated borrowing detection (Mennecier et al. 2016), the detection of sound correspondences and sound correspondence patterns (List 2019), and also the automated prediction of so far unobserved words (Bodt and List 2019), which is specifically useful to support fieldworkers working on small groups of related languages.

[Q] *How can automated word prediction be useful for linguistic field work?*

3.3 Analysis

As it was mentioned briefly before, the distinction between what counts as inference and what counts as analysis are not always easy to draw. Intuitively, analysis should involve g-linguistic questions in the sense discussed in the first session, but it is clear that there is no formal justification for it, and it seems to depend more on the workflow, whether a certain step (such as – for example – phylogenetic inference) is labeled as part of the inference or the analysis step. An example for such a borderline case is the *Database of Cross-Linguistic Colexifications* (CLICS, <https://clics.clld.org>, List et al. 2018), which offers cross-linguistic accounts on polysemies, which are displayed in form of a network analysis that provides information on the relative cross-linguistic closeness of more than 1500 different concepts, reflected in more than 1000 of the world's languages. While CLICS is offering an analysis that shows – similar to Youn et al. (2016) – that lexical structure is surprisingly similar across languages, the analysis itself could be treated as some kind of inference, and analysed to answer bigger questions related to human cognition. The more classical analyses which are usually presented, however, try to test certain theories (which can relate to both p- and g-linguistic questions) by analysing the data which has been inferred previously. In these cases, the large-scale cross-linguistic databases, which are increasingly produced, play an important role, as they allow scholars to test their hypotheses on a global scale, allowing them, for example, to test hypotheses regarding the transmission of Creole languages (Blasi et al. 2017), the evolution of syntax (Widmer et al. 2017), or the impact of our diet on evolution of our speech sounds (Blasi et al. 2019).

[Q] *What hypotheses can be derived from current pragmatic theory that could be tested with help of cross-linguistic approaches?*

References

- Anderson, C., T. Tresoldi, T. C. Chacon, A.-M. Fehn, M. Walworth, R. Forkel, and J.-M. List (2018). "A Cross-Linguistic Database of Phonetic Transcription Systems" . *Yearbook of the Poznań Linguistic Meeting* 4.1, 21–53.
- Atkinson, Q. D. and R. D. Gray (2006). "How old is the Indo-European language family? Illumination or more moths to the flame?" In: *Phylogenetic methods and the prehistory of languages*. Ed. by P. Forster and C. Renfrew. Cambridge, Oxford, and Oakville: McDonald Institute for Archaeological Research, 91–109.
- Barrachina, S. et al. (2008). "Statistical approaches to computer-assisted translation" . *Computational Linguistics* 35.1, 3–28.
- Bergsland, K. and H. Vogt (1962). "On the validity of glottochronology" . *Current Anthropology* 3.2, 115–153. JSTOR: 2739527.
- Blasi, D. E., S. Wichmann, H. Hammarström, P. Stadler, and M. H. Christiansen (2016). "Sound-meaning association biases evidenced across thousands of languages" . *Proceedings of the National Academy of Science of the United States of America* 113.39, 10818–10823.
- Blasi, D. E., S. M. Michaelis, and M. Haspelmath (2017). *Nature Human Behaviour* 1, 723–729.
- Blasi, D. E., S. Moran, S. R. Moisiak, P. Widmer, D. Dediu, and B. Bickel (2019). "Human sound systems are shaped by post-Neolithic changes in bite configuration" . *Science* 363.1192, 1–10.
- Bodt, T. A. and J.-M. List (2019). "Testing the predictive strength of the comparative method: An ongoing experiment on unattested words in Western Kho-Bwa languages" . *Papers in Historical Phonology* 4.1, 22–44.
- Boucharde-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, 4224–4229.
- Bowern, C., P. Epps, R. Gray, J. Hill, K. Hunley, P. McConwell, and J. Zentz (2011). "Does Lateral Transmission Obscure Inheritance in Hunter-Gatherer Languages?" *PLoS ONE* 6.9, e25195.
- Calude, A. S. and M. Pagel (2011). "How do we use language? Shared patterns in the frequency of word use across 17 world languages" . *Philosophical Transactions of the Royal Society B* 366, 1101–1107.
- Campbell, L. (1999). *Historical linguistics. An introduction*. 2nd ed. Edinburgh: Edinburgh Univ. Press.
- Chang, W., C. Cathcart, D. Hall, and A. Garret (2015). "Ancestry-constrained phylogenetic analysis support the Indo-European steppe hypothesis" . *Language* 91.1, 194–244.
- Donohue, M., R. Hetherington, J. McElvenny, and V. Dawson (2013). *World phonotactics database*. Canberra: Department of Linguistics. The Australian National University.
- Dryer, M. S. and M. Haspelmath, eds. (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Everett, C., D. E. Blasi, and S. G. Roberts (2015). "Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots" . *Proceedings of the National Academy of Sciences of the United States of America* 112.5, 1322–1327.
- Forkel, R., J.-M. List, S. J. Greenhill, C. Rzymiski, S. Bank, M. Cysouw, H. Hammarström, M. Haspelmath, G. A. Kaiping, and R. D. Gray (2018). "Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics" . *Scientific Data* 5.180205, 1–10.
- Geisler, H. and J.-M. List (2010). "Beautiful trees on unstable ground. Notes on the data problem in lexicostatistics" . In: *Die Ausbreitung des Indogermanischen. Thesen aus Sprachwissenschaft, Archäologie und Genetik*. Ed. by H. Hettrich. Document has been submitted in 2010 and is still waiting for publication. Wiesbaden: Reichert.
- Gray, R. D. and Q. D. Atkinson (2003). "Language-tree divergence times support the Anatolian theory of Indo-European origin" . *Nature* 426.6965, 435–439.
- Gray, R. D., A. J. Drummond, and S. J. Greenhill (2009). "Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement" . *Science* 323.5913, 479–483.
- Greenhill, S. J., C. H. Wu, X. Hua, M. Dunn, S. C. Levinson, and R. D. Gray (2017). "Evolutionary dynamics of language systems" . *Proceedings of the National Academy of Sciences of the United States of America* 114.42, E8822–E8829.
- Greenhill, S. (2015). "Evolution and Language: Phylogenetic Analyses" . In: *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Ed. by J. D. Wright. Second Edition. Oxford: Elsevier, 370–377.
- Grollemund, R., S. Branford, K. Bostoen, A. Meade, C. Venditti, and M. Pagel (2015). "Bantu expansion shows that habitat alters the route and pace of human dispersals" . *Proceedings of the National Academy of Sciences of the United States of America* 112.43, 13296–13301.
- Hammarström, H., R. Forkel, and M. Haspelmath (2018). *Glottolog*. Version 3.3. URL: <http://glottolog.org>.
- Haspelmath, M. and U. Tadmor, eds. (2009). Berlin and New York: de Gruyter.
- 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*. (Chiang Mai, Thailand, 11/08–11/13/2011). AFNLP, 865–873.
- Holman, E. W. et al. (2011). "Automated dating of the world's language families based on lexical similarity" . *Current Anthropology* 52.6, 841–875. JSTOR: 10.1086/662127.
- Hrozny, B. (1915). "Die Lösung des hethitischen Problems [The solution of the Hittite problem]" . *Mitteilungen der Deutschen Orient-Gesellschaft* 56, 17–50.
- Jäger, G. (forthcoming). "Computational historical linguistics" . *Theoretical Linguistics*.
- 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*. "EACL 2017". Valencia: Association for Computational Linguistics, 1204–1215.
- Kirby, K. R. et al. (2016). "D-PLACE: A Global Database of Cultural, Linguistic and Environmental Diversity" . *PLOS ONE* 11.7, 1–14.
- Kolipakam, V., F. M. Jordan, M. Dunn, S. J. Greenhill, R. Bouckaert, R. D. Gray, and A. Verkerk (2018). "A Bayesian phylogenetic study of the Dravidian language family" . *Royal Society Open Science* 5.171504, 1–17.
- Kondrak, G. (2000). "A new algorithm for the alignment of phonetic sequences" . In: *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*. (Seattle, 04/29–05/03/2000), 288–295.
- Lee, S. and T. Hasegawa (2011). "Bayesian phylogenetic analysis supports an agricultural origin of Japonic languages" . *Proc. Biol. Sci.* 278.1725, 3662–3669.
- List, J.-M. (2014). *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press.
- (2015). "Network perspectives on Chinese dialect history" . *Bulletin of Chinese Linguistics* 8, 42–67.
- (2016). *Computer-Assisted Language Comparison: Reconciling Computational and Classical Approaches in Historical Linguistics*. Jena: Max Planck Institute for the Science of Human History.
- (2019). "Automatic inference of sound correspondence patterns across multiple languages" . *Computational Linguistics* 1.45, 137–161.
- (forthcoming). "Automated methods for the investigation of language contact situations" . *Language and Linguistics Compass* ????, 1–19.
- List, J.-M., M. Cysouw, and R. Forkel (2016a). "Concepticon. A resource for the linking of concept lists" . In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation*. "LREC 2016" (Portorož, 05/23–05/28/2016). Ed. by N. C. C. Chair), K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Ojijk, and S. Piperidis. European Language Resources Association (ELRA), 2393–2400.
- List, J.-M., P. Lopez, and E. Baptiste (2016b). "Using sequence similarity networks to identify partial cognates in multilingual wordlists" . In: *Proceedings of the Association of Computational Linguistics 2016 (Volume 2: Short Papers)*. Association of Computational Linguistics. Berlin, 599–605.
- 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. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2018). "CLICS². An improved database of cross-linguistic colexifications assembling lexical data with help of cross-linguistic data formats" . *Linguistic Typology* 22.2, 277–306.
- Maddieson, I., S. Flavier, E. Marsico, C. Coupé, and F. Pellegrino. (2013). "LAPSyD: Lyon-Albuquerque Phonological Systems Database" . In: *Proceedings of Interspeech*. (Lyon, 08/25–08/29/2013).
- Mennecier, P., J. Nerbonne, E. Heyer, and F. Manni (2016). "A Central Asian language survey" . *Language Dynamics and Change* 6.1, 57–98.
- Moran, S., D. McCloy, and R. Wright, eds. (2014). *PHOIBLE Online*. URL: <http://phoible.org/>.
- Nature, E. B. (2013). "Reducing our irreproducibility" . *Nature* 496.4, 398.
- (2018). "Referees' rights" . *Nature* 560, 409.
- Nelson-Sathi, S., J.-M. List, H. Geisler, H. Fangerau, R. D. Gray, W. Martin, and T. Dagan (2011). "Networks uncover hidden lexical borrowing in Indo-European language evolution" . *Proceedings of the Royal Society of London B: Biological Sciences* 278.1713, 1794–1803.
- Nerbonne, J., R. Colen, C. Gooskens, P. Kleiweg, and T. Leinonen (2011). "Gabmap - A web application for dialectology" . *Dialectologia* Special Issue II, 65–89.
- Prokić, J., M. Wieling, and J. Nerbonne (2009). "Multiple sequence alignments in linguistics" . In: *Proceedings of the EACL 2009 Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education*. "LaTeCH-SHET&R 2009" (Athens, 03/30/2009), 18–25. acm: 1642052.
- Sagart, L., G. Jacques, Y. Lai, R. Ryder, V. Thouzeau, S. J. Greenhill, and J.-M. List (2019). "Dated language phylogenies shed light on the ancestry of Sino-Tibetan" . *Proceedings of the National Academy of Science of the United States of America* 116 (21), 10317–10322.
- Saussure, F. de (1916). *Cours de linguistique générale*. Ed. by C. Bally. Lausanne: Payot.
- Silver, D. et al. (2016). "Mastering the game of Go with deep neural networks and tree search" . *Nature* 529.7587, 484–489.
- Swadesh, M. (1950). "Salish internal relationships" . *International Journal of American Linguistics* 16.4, 157–167. JSTOR: 1262898.
- Szemerényi, O. (1970). *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Turchin, P., I. Peiros, and M. Gell-Mann (2010). "Analyzing genetic connections between languages by matching consonant classes" . *Journal of Language Relationship* 3, 117–126.
- Weiss, M. (2015). "The comparative method" . In: *The Routledge Handbook of Historical Linguistics*. Ed. by C. Bowern and N. Evans. New York: Routledge, 127–145.
- Widmer, M., S. Auderset, J. Nichols, P. Widmer, and B. Bickel (2017). "NP recursion over time: Evidence from Indo-European" . *ACM Transactions on Asian Language Information Processing* 93.4, 799–826.
- Youn, H., L. Sutton, E. Smith, C. Moore, J. F. Wilkins, I. Maddieson, W. Croft, and T. Bhattacharya (2016). "On the universal structure of human lexical semantics" . *Proceedings of the National Academy of Sciences of the United States of America*.

2 Language Change

The three sections on language change deal with sound change, semantic change, and contact-induced change.

Empirical Studies on Sound Change

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we look at the phenomenon of regular sound change from different perspectives, discussing the historical context of how it was detected, the techniques of how it is studied, and the open questions that could not be solved so far. We conclude by discussing future challenges.

1 The detection of regular sound change

One of the most fundamental insights of historical linguistics, which has its origins in the very origins of the discipline, is the detection that sound change proceeds in what seems to be a mostly *regular* manner. What this exactly means has been subject of lengthy discussions. The notion of *regularity* also created a lot of confusion among those linguists and non-linguists without a detailed training in historical linguistics. This is in part reflected in computational approaches that seek to automate the classical approaches to linguistic reconstruction, which often quite naively ignore the fundamental aspect of regularity. Before we start to look into the techniques that linguists use in order to study sound change in detail, it is important to go back in history in order to review more closely how the concept of regularity evolved, and how it has been constantly challenged.

[Q] We will later spend more time on discussing what is actually meant by regularity, but judging from what you know about linguistics by now, what could regularity reflect in this context?

1.1 Rask, Grimm, and the detection of sound shifts

The early detection that sound change may follow general tendencies in a given language family is usually attributed to Rasmus Rask (1787–1832), who pointed to what he thought were frequent transitions (of sounds) from Greek and Latin to Icelandic (Rask 1818: 169). When reading about these findings, Jacob Grimm (1785–1863) further investigated these systematic similarities between Greek, Latin, and Germanic languages (specifically Gothic), and expanded the second version of his *Deutsche Grammatik* considerably. Grimm identified regular correspondences between consonants in Greek, Gothic, and Old High German, as shown in Table 1 below.

gr.	goth.	alth.	gr.	goth.	alth.	gr.	goth.	alth.
P	F	B(V)	T	TH	D	K	..	G
B	P	F	D	T	Z	G	K	CH
F	B	P	TH	D	T	CH	G	K

Table 1 Correspondences identified by Grimm (1822: 584).

What these formulas shown in the table meant that there were essentially many words of comparable meaning in the three languages in which the consonants formed patterns. If a word had a *p* in Greek (such as in *ποδ*- “foot”), it would reflect as *f* in Gothic (*fōtus*), and as *v* in Old High German (*vuoz*), yet not only in these three words, but in many more examples (see the detailed evidence in *ibid.*: 585). The crucial conclusion that Grimm drew from these observed patterns was that the *identity* of sounds (or letters) would not justify a comparison of words sharing a common origin. What would justify it instead was that the correspondence turned out to follow the rules he had detected while words with similar forms would result from coincidence or borrowing (*ibid.*: 588).

[Q] In the very same book, we also find him saying that the German word *Schrift* expressed “eight sounds in seven signs, since *f* stands for *p*” (Grimm 1822: 3, my translation). What does this statement tell us about the historical context in which Grimm worked?

1.2 Verner and the dawn of regularity

Grimm’s detection had pointed to an interesting tendency with respect to sound change, namely, that one could find patterns of corresponding sounds when comparing genetically related languages. The problem, however, was that these patterns did not seem to work in all cases. Grimm himself noted this, emphasizing that this consonant shift (he thought in terms of a change *from* Greek *to* Germanic) “proceeds in the majority but will never be pure in particular cases” (ibid.: 590, my translation). Some reasons for certain exceptions had been mentioned already by Grimm himself. He noted, for example, that the patterns were influenced by the presence of liquids or sibilants (ibid.). Only later, however, linguists managed to find a proper explanation for these exceptions by refining the formulation of what they started to call *sound laws*. After exceptions of the basic correspondence pattern established by Grimm were listed systematically by Lottner (1862), Grassmann (1863) could explain the first class of exceptions by pointing to systematic assimilation processes in Greek and Sanskrit (Meier-Brügger 2002: L 348), by which of two aspirated sounds which follow each other, the first loses its aspiration (cf. Sanskrit **dhá-dhā-mi* > *dádhāmi*, “I put”). The second class of exceptions, finally, could be shown by (Verner 1877) to reflect a regular process in Germanic languages, during which the correspondence patterns varied in strict correlation with presumed stress patterns in the Proto-Germanic language (which themselves were still reflected in Vedic Sanskrit).

[Q] In Verner’s original, he emphasizes that “Indo-European *k, t, p* first changed in all places to *h, þ, f*; these voiceless fricatives along with the voiceless fricative *s*, [...] became then voiced inside a word, when being in voiced neighborhood, but stayed voiceless when following after a stressed syllable” (ibid.). As an example, scholars often quote Gothic *broþar* in contrast to Vedic Sanskrit *bhrātar-* and Gothic *fadar* in contrast to Vedic Sanskrit *pitar-*. How should the stress be distributed in the Sanskrit words?

1.3 The Neogrammarian Manifesto

The fact that what formerly was thought to be a mere tendency could now, with help of refined rules that would allow to explain sound change as a process without exception lead to a great euphoria in the field and culminated in the so-called *Neogrammarian manifesto*:

All sound change, as long as it proceeds mechanically, follows exceptionless laws, i.e., the direction of the sound shift is the same with all members of a language community except from those cases in which the dialect split occurs, and all words in which the sound occurs in the same context are transformed without exception. (Osthoff and Brugmann 1878: XIII, my translation)

The principle assumption, that sound change proceeds without exceptions and that all apparent exceptions which one might observe can be regularly explained, be it by showing that the words under question are not cognate in the end, or that secondary processes have masked the former regularity, is still the working principle of classical historical linguistics and the first think historical linguists learn during their training.

[Q] What are the two central aspects that can be found in the quote from the Neogrammarian manifesto?

2 Language Change

1.4 Wang, Chen and the postulation of irregularity

Not all linguists would follow the opinion of the Neogrammarians. Especially dialectologists would often prefer to follow the famous slogan that “chaque mot a son histoire”, usually attributed to Jules Gilliéron (1854–1926, see Campbell 1999: 189). The doubts of the dialectologists were, however, not in direct contradiction to the Neogrammarian hypothesis of regularity, given that their theory did not state that *all words* in a given language change regularly, but rather emphasized that irregularities “could be accounted for [...] by certain less obvious mechanisms of borrowing and analogy” (Kiparsky 1988: 368).

In the 1960s, the situation changed drastically, when new research, which was almost exclusively based on the Chinese dialects, led to the postulation of a new mechanism of sound change which was in strict opposition to the hypothesis of the Neogrammarians.

Regarding the lexicon [they assumed] that a change always affects the whole lexicon, and can therefore be seen as an abrupt change. Regarding the sounds [they assumed] that the change proceeded step by step, and can therefore be seen as a gradual change. (Wang 2006: 109)¹

The results of the analyses of the Chinese dialectologists, however, suggested that a certain mechanism of sound change, which they later called *lexical diffusion*, proceeds in the exact opposite way, namely, in “a manner that is phonetically abrupt but lexically gradual. As the change diffuses across the lexicon, it may not reach all the morphemes to which it is applicable. If there is another change competing for part of the lexicon, residue may result” (Wang 1969: 9). Examples were specifically drawn from cases where words with exactly the same pronunciation in Middle Chinese, the ancestor of most Chinese dialects, turned out to develop two different readings, which led the scholars conclude that “[when] a phonological innovation enters a language it begins as a minor rule, affecting a small number of words” which later “gradually spreads across the lexicon” (Chen 1972).

Character	Pinyin	Meaning	Middle Chinese	Shuāngfēng
步	<i>bù</i>	„to walk”	bo ³	bu ³³
捕	<i>bǔ</i>	„to grasp”	bo ³	p ^h u ²¹
刨	<i>páo</i>	„to dig”	bæw ¹	bə ³³
跑	<i>páo</i>	„to scrape”	bæw ¹	p ^h ə ²¹
盜	<i>dào</i>	„to rob”	daw ³	də ³³
導	<i>dǎo</i>	„to lead”	daw ³	t ^h ə ³⁵

Table 2: Examples for irregularities in the readings of Shuāngfēng (ZIHUI).

[Q] Some basic examples for the idea of Chen (1972) that homophonous readings in Middle Chinese developing different readings in a given dialect are given in Table 2. Do these examples provide an undisputable proof for the existence of lexical diffusion as an alternative mechanism of sound change?

1.5 Labov and the study of sound change in progress

The theory of lexical diffusion is not only contrary to the inherent model of sound change underlying, but also tackles its most important implication: If sound change is by and large regular, it means we can reconstruct ancient stages of languages not reflected in written sources. But if considerable parts of our evidence turn out to reflect sound change processes that were not completely finished, this would make it much more difficult to carry out linguistic reconstruction. It seems, however, that the theory of lexical diffusion is not entirely correct. Firstly, Labov (1981) could show by investigating sound change in progress that there were two basic mechanisms of sound change, one mechanism that diffuses across the lexicon, and one in which a change captures all the words at the same time.

¹My translation, original text: 作為詞彙,要變就都變,因而是一種突變。作為語音,變化是逐漸的,因而是一種漸變。

There is no basis for contending that lexical diffusion is somehow more fundamental than regular, phonetically motivated sound change. On the contrary, if we were to decide the issue by counting cases, there appear to be far more substantially documented cases of Neogrammarian sound change than of lexical diffusion. (Labov 1994: 471)

Secondly, what is even more important in this context, it would even open the door for speculations, if one would treat sound change as a process that could be in principle irregular, as it is to be expected that nobody would have tried to resolve Grimm’s exceptions if one had thought that these were anyway impossible to be explained by means of regular “sound laws” (Hill 2016). Nevertheless, even when accepting the Neogrammarian idea of regularity, the question remains to which degree this regularity is persistent, given that we know well that processes like borrowing and analogy can mask it.

[Q] In this context, the term “mechanism” was used in order to distinguish lexical diffusion from Neogrammarian sound change. Would it not be possible to just use the term “process” instead of “mechanism”?

2 Techniques for the investigation of regular sound change

In order to accumulate the data needed to investigate how sound change proceeds, one needs specific techniques for inference. The most prominent method employed by scholars is traditionally called the *comparative method* (Meillet 1925 [1954]), which is essentially a bunch of techniques which are eclectically employed by linguists embarking on historical language comparison. If one asks different linguists, they will often differ with respect to what they think represents the comparative method best, and for this reason, this method is better treated as some kind of an *overarching framework* that scholars use in order to compare languages (Fox 1995, Jarceva 1990, Klimov 1990).

[Q] Why would linguists still talk of the comparative method, even if they know from their practice themselves that it is not a unified procedure?

2.1 Classical approaches in the framework of the comparative method

The comparative method is an overarching framework that historical linguists use to study language history. The application of the framework is tedious, involving many iterative steps. Scholars start by comparing words from different languages in order to identify sets of potentially related words (*cognates*). They then set up lists of sound correspondences and use this information to revise their initial list of cognates (see Table 3). This new information is again used to revise the list of corresponding segments, and so on, until the results can no longer be refined. By applying this method to two or more languages, linguists assemble *cognate words* and *correspondence patterns*, which are then used to infer change scenarios that explain the different correspondence patterns by invoking an ancestral language from which the sounds in the descendant languages (the reflex sounds) can be derived in the most convincing fashion.

Cognate List		Alignment				Correspondence List			
English	<i>foot</i>	f	υ	t		Eng.	Grk.	Freq.	
Ancient Greek	ποδ-	p	ɔ	d		f	p	3 x	
English	<i>father</i>	f	α:	θ	ə	ɪ	p ^h	1 x	
Ancient Greek	πατέρ-	p	a	t	ε	r	ɪ	r	2 x
English	<i>fear</i>	f	ɪə	ɪ	-	θ	t	1 x	
Ancient Greek	φοβέ-	p ^h	ɔ	b	e	t	d	1 x	
English	<i>fire</i>	f	αιə	ɪ					
Ancient Greek	πυρ-	p	y	r					

↑
irregular match!

Table 3: Detecting regular sound correspondences in classical historical language comparison.

2 Language Change

[Q] Table 3 gives an example with respect to the detection of sound correspondences between English and Ancient Greek. How can the principle be handled for more than one language?

2.2 Computer-assisted approaches

While traditional accounts on the inference of sound correspondences (and consecutively also accounts on the inference of sound change patterns) are still the predominant way in which linguists analyze the history of the world's languages, computational methods, specifically those that help linguists in their work rather than threatening to replace them, are constantly gaining ground. Among the most important techniques in this context are (1) techniques for *automated phonetic alignment*, which are needed as a basis for identifying corresponding sounds (Kondrak 2000, List 2014), (2) extended techniques for *automated cognate detection* (Arnaud et al. 2017, List et al. 2017), which make use of alignment techniques in order to search for the most likely candidates of related words across languages, and (3) relatively recent techniques for *automated correspondence pattern inference* (List 2019), which infer sound correspondences across multiple languages, offering a first starting point for phonological reconstruction.

The methods for phonetic alignments, cognate detection, and sound correspondence inference are quite advanced until now, and they start providing real help to linguists who investigate so far less thoroughly investigated language families (Chen 2019, Hill and List 2017, Kolipakam et al. 2018). With LingPy (<http://lingpy.org>, List et al. 2018a), a stable software package offers basic algorithms for phonetic alignment analyses and cognate detection. Furthermore, the data processed with LingPy can be directly inspected with help of web-based tools, such as the Etymological Dictionary Editor (EDICTOR, <http://edictor.digling.org>, List et al. 2017), allowing linguists to quickly modify their data, correcting the errors made by the algorithm, or converting it to formats needed for the further analysis with help of phylogenetic software. Online tutorials (e.g., <https://calc.hypotheses.org>) along with print tutorials (List et al. 2018b) run newcomers through the new techniques.

Another benefit of the methods that may be less evident from the first sight has been presented in a recent experiment on *word prediction*. Since scholars in fieldwork usually do not have time to elicit all words relevant for their study at ones, they can make use of the comparative method (either in a classical or a computer-assisted form) to predict how certain words would sound from the correspondence patterns they observe for the languages under investigation. This was in fact already mentioned by Grimm (1822: 589), who thought there would be a limited possibility to predict the consonantal shape of Germanic words if they were missing. In a recent experiment, we tested the usefulness of computer-assisted word prediction techniques (Bodt and List 2019), the so far unpublished results indicate that the expert fieldworker was able to predict missing words in the data with an accuracy of about 75%. More studies and experiments will be needed to further test and enhance the suitability of the procedure which was laid out in this pilot study.

[Q] If we manage to predict words with an accuracy of 75% (by an expert who made use of computational pre-processing), what does this tell us with respect to the question of the regularity of language change? Is it now regular after all or not?

3 Open questions on sound change

In their very influential paper titled *Empirical foundations for a theory of language change*, Weinreich et al. (1968) proposed a set of problems for future work in historical linguistics. That Campbell (1999: 194f) repeats these actual problems, shows that there has not been much success in increasing our understanding with respect to these problems. Five major problems are summarized by Campbell: the problem of the (1) *constraints* of language change, the problem of the (2) *transition* of different change

processes, the problem of the (2) *embedding* of change in linguistic and social relations, the problem of the (4) *evaluation* of change with respect to the speakers of a given language, and the problem of the (5) *actuation* of change, i.e., the question of why particular changes occur at particular times and places. Coseriu (1973: 65f) lists only three problems, namely “(a) el problema racional del cambio” (*why do languages change at all?*), “(b) el problema general de los cambios” (*under which circumstances do languages change?*), and “(c) el problema histórico de tal cambio determinado” (*why do particular changes take place*). In the following, we will discuss these three problems in more detail, specifically concentrating on sound change, and compare them with the ones mentioned by Cambell.

[Q] In this list of problems, actuation of sound change refers to the occurrence of particular changes in particular languages. Could one also understand the problem in a broader context?

3.1 The rational problem of sound change

The question of why sound change happens after all is difficult to answer, since it is difficult to find a direct benefit resulting from the process for a given language system (Anttila 1976). In an evolutionary framework, we would thus say that there is no apparent selective pressure that would favor the modification of sounds. On the contrary, it is known that sound change may increase the amount of grammatical irregularities.² That not all changes in evolution need to yield a direct benefit, however, is nothing new for biologists, who have been investigating what they call phenomena of *drift* already for a long time. In order to explain why sound change happens, most scholars nowadays assume variation at the synchronic level as its starting point (Ohala 1989, Kümmel 2008: 22, Paul 1880 [1886]: 30). It is further assumed that language systems are *robust* enough to tolerate a certain amount of sound change (Hockett 1965: 203f). Robustness itself results from the redundancy of speech (*ibid.*),³ which can be seen as an important feature of language, as it guarantees its functioning as a communication system. While these neutral theories of sound change seem to be obvious (and have been mentioned already quite early in the linguistic literature), it is less clear to which degree certain selective aspects could not also play a role in sound change. Blasi et al. (2019), for example, assume that the pronunciation of labiodentals was greatly facilitated along with changes in the diet of early humans. Everett et al. (2015) claim that tone languages evolve more frequently in humid climates. While neutral theories of evolution can in principle explain why sound change should be possible, we are still far away from being able to draw a conclusive picture of all the factors that may influence it.

[Q] Labov (2001: 15) emphasizes that “the evolution of species and the evolution of language are identical in form, although the fundamental mechanism of the former is absent in the latter”. What does he mean?

3.2 The general problem of sound change

If we look at the general patterns of sound change that can be observed for the languages of the world, we can distinguish two basic conditions of sound change, *phonetic conditions* and *systemic conditions*. Phonetic conditions can be further subdivided into *articulatory* and *acoustic conditions*. When trying to explain why certain sound changes can be observed more frequently across different languages of the world, many linguists tend to explain this by invoking phonetic factors. If the sound *p*, for example, turns into an *f*, this is not necessarily surprising given the strong similarity of the sounds. But similarity can be measured in two ways: one can compare the similarity with respect to the production of a sound by a speaker, and with respect to the perception of the sound by a listener. While production of sounds

²Anttila calls this *Sturtevant's paradox*, namely that *regular* sound change produces *irregularity* in language systems, while *irregular* analogy produces *regularity* in language systems.

³See Winter (2014) for a detailed discussion of robustness.

2 Language Change

is traditionally seen as the more important factor contributing to sound change (Hock 1991: 11), there are clear examples for sound change due to misperception and re-interpretation by the listeners (Ohala 1989: 182). Some authors go as far as to claim that production-driven changes reflect regular *internal* language change (which happens gradually during acquisition, or – depending on the theory – also in later stages Bybee 2002), while perception-based changes rather reflect change happening in second language acquisition and language contact (Mowrey and Pagliuca 1995: 48).

While the interaction of production and perception has been discussed in some detail in the linguistic literature, the influence of systemic factors has so far only rarely been regarded. What I mean by this factor is the old structural idea that a language can be seen as a system, and that certain changes in the system may be explained exclusively as resulting from systemic constellations. As a straightforward example, consider the difference in *design space* for the production of consonants, vowels, and tones. In order to maintain pronunciability and comprehensibility, it is useful for the sound system of a given language, to fill in those spots in the design space that are maximally different from each other. The larger the design space and the smaller the inventory, the easier it is to guarantee its functionality. Since design spaces for vowels and tones are much smaller than for consonants, however, these sub-systems are more easily disturbed, which could be used to explain the presence of chain shifts of vowels, or *flip-flop* in tone systems (Wang 1967: 102). Systemic considerations play an increasingly important role in evolutionary theory, and, as shown in List et al. (2016), also be used as explanations for phenomena as strange as the phenomenon of Sapir's *drift* (Sapir 1953).

[Q] *There is a lot of discussing in the linguistic literature with respect to the time when sound change occurs: should it occur during the life time of a human being, or should it rather occur only at the time of acquisition? What data would we expect for both scenarios?*

3.3 The historical problem of sound change

The historical problems, i.e., the particular problems of sound changes in particular languages, are usually much better understood than the general or the rational problem, as presented above. As in all cases of historical language comparison, however, typological (general) investigations and particular investigations should ideally guide each other. Unfortunately, general factors are rarely considered when discussing individual proposal for linguistic reconstruction. This was already criticized by Jakobson (1958), who criticized that linguists would rarely consider typological aspects when proposing their reconstructions for unattested languages, but the situation has not changed much in the meantime. The biggest problem in this context seems to be the general lack of cross-linguistic catalogs of attested or proposed sound change processes.

[Q] *Why do linguists often defend to ignore typological evidence in reconstruction?*

4 Towards and improved investigation of sound change

What we need in order to address the rational, the general, and the historical problems of sound change, are, in my opinion, (1) improved models of sound change, specifically models that contrast different depths of analysis and different hypotheses (simple speaker-hearer models, improved models based on articulation and acoustics, exemplar theory, generative grammar accounts, etc.) based on *unified transcription systems*, (2) increased amounts of readily coded data of sound change processes as they have been *inferred* by experts, and (3) improved methods to *analyze* the data, specifically making sure to separating frequency due to inheritance from frequency due to convergence. Linguists nowadays discuss repeatedly the importance of having a catalogue of sound change patterns. While first attempts

to establish such a catalogues have been proposed so far (Kümmel 2008), and other scholars report from their private but so far not open collections of sound change patterns, no unifying attempt has been made so far. What I consider crucial is to work on a database that does not only include the sound change patterns, but also the evidence in form of aligned cognate sets. With the newly published database of *Cross-Linguist Transcription Systems* (CLTS, <https://clts.clld.org>, Anderson et al. 2018), the first step towards a rigorous standardization of transcription systems has already been made.

[Q] Why is it so important to list the evidence along with the patterns?

References

- Anderson, C., T. Tresoldi, T. C. Chacon, A.-M. Fehn, M. Walworth, R. Forkel, and J.-M. List (2018). "A Cross-Linguistic Database of Phonetic Transcription Systems". *Yearbook of the Poznań Linguistic Meeting* 4.1, 21–53.
- Anttila, R. (1976). "The acceptance of sound change by linguistic structure". In: *Recent developments in historical phonology*. Ed. by J. Fisiak. The Hague, Paris, New York: de Gruyter, 43–56.
- Arnaud, A. S., D. Beck, and G. Kondrak (2017). "Identifying cognate sets across dictionaries of related languages". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. (Copenhagen, 09/07–09/11/2017). Association for Computational Linguistics, 2509–2518.
- Běijīng Dàxué 北京大学, ed. (1989). *Hànyǔ fāngyīn zìhuì* 漢語方音字彙 [Chinese dialect character pronunciation list]. 2nd ed. Běijīng 北京: Wénzì Gǎigé 文字改革.
- Blasi, D. E., S. Moran, S. R. Moisiuk, P. Widmer, D. Dediu, and B. Bickel (2019). "Human sound systems are shaped by post-Neolithic changes in bite configuration". *Science* 363.1192, 1–10.
- Bodt, T. A. and J.-M. List (2019). "Testing the predictive strength of the comparative method: An ongoing experiment on unattested words in Western Kho-Bwa languages". *Papers in Historical Phonology* 4.1, 22–44.
- Bybee, J. L. (2002). "Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change". *Language Variation and Change* 14, 261–290.
- Campbell, L. (1999). *Historical linguistics. An introduction*. 2nd ed. Edinburgh: Edinburgh Univ. Press.
- Chen, E. (2019). "Phonological reconstruction of Proto-Kampa consonants". *Berkeley Papers in Formal Linguistics* 2.1, 1–56.
- Chen, M. (1972). "The time dimension. Contribution toward a theory of sound change". *Foundations of Language* 8.4, 457–498. JSTOR: 25000618.
- Coseriu, E. (1973). *Sincronia, diacronia e historia. El problema del cambio lingüístico* [Synchrony, diachrony, and history. The problem of linguistic change]. Madrid: Biblioteca Románica Hispánica.
- Everett, C., D. E. Blasi, and S. G. Roberts (2015). "Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots". *Proceedings of the National Academy of Sciences of the United States of America* 112.5, 1322–1327.
- Fox, A. (1995). *Linguistic reconstruction. An introduction to theory and method*. Oxford: Oxford University Press.
- Grassmann, H. (1863). "Ueber die aspiraten und ihr gleichzeitiges vorhandensein im an- und auslaute der wurzeln". *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 12.2, 81–110.
- Grimm, J. (1822). *Deutsche Grammatik*. 2nd ed. Vol. 1. Göttingen: Dieterichsche Buchhandlung. Google Books: MnsKAAAAIAAJ.
- Hill, N. (2016). "A refutation of Song's (2014) explanation of the 'stop coda problem' in Old Chinese". *International Journal of Chinese Linguistics* 2.2, 270–281.
- 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.
- Hock, H. H. (1991). *Principles of historical linguistics*. 2nd ed. Berlin: Mouton de Gruyter.
- Hockett, C. F. (1965). "Sound change". *Language* 41.2, 185–204. eprint: 411873 (jstor).
- Jakobson, R. (1958). "Typological studies and their contribution to historical comparative linguistics". In: *Proceedings of the Eighth International Congress of Linguistics*. Oslo, 17–35.
- Jarceva, V. N., ed. (1990). *Lingvističeskij énciklopedičeskij slovar* (Linguistical encyclopedical dictionary). Moscow: Sovetskaja Enciklopedija.
- Kiparsky, P. (1988). "Phonological change". In: *Linguistics. The Cambridge survey*. Vol. 1: *Linguistic theory. Foundations*. Ed. by F. J. Newmeyer. Cambridge et al.: Cambridge University Press, 363–415.
- Klimov, G. A. (1990). *Osnovy lingvističeskoj komparativistiki* [Foundations of comparative linguistics]. Moscow: Nauka.
- Kolipakam, V., F. M. Jordan, M. Dunn, S. J. Greenhill, R. Bouckaert, R. D. Gray, and A. Verkerk (2018). "A Bayesian phylogenetic study of the Dravidian language family". *Royal Society Open Science* 5.171504, 1–17.
- Kondrak, G. (2000). "A new algorithm for the alignment of phonetic sequences". In: *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*. (Seattle, 04/29–05/03/2000), 288–295.
- Kümmel, M. J. (2008). *Konsonantenwandel* [Consonant change]. Wiesbaden: Reichert.
- Labov, W. (1994). *Principles of linguistic change*. Vol. 1: *Internal factors*. Malden, Oxford, and West Sussex: Wiley-Blackwell.
- (2001). *Principles of linguistic change*. Vol. 2: *Social factors*. Malden, Oxford, and West Sussex: Wiley-Blackwell.
- Labov, W. (1981). "Resolving the Neogrammarian Controversy". *Language* 57.2, 267–308. JSTOR: 413692.
- List, J.-M. (2014). *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press.
- (2019). "Automatic inference of sound correspondence patterns across multiple languages". *Computational Linguistics* 1.45, 137–161.
- 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.
- 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 (2018a). *LingPy. A Python library for quantitative tasks in historical linguistics*. URL: <http://lingpy.org>.
- List, J.-M., M. Walworth, S. J. Greenhill, T. Tresoldi, and R. Forkel (2018b). "Sequence comparison in computational historical linguistics". *Journal of Language Evolution* 3.2, 130–144.
- Lottner, C. F. (1862). "Ausnahmen der ersten lautverschiebung". *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 11.3/4, 161–205. JSTOR: 40844790.
- Meier-Brügger, M. (2002). *Indogermanische Sprachwissenschaft*. In collab. with M. Fritz and M. Mayrhofer. 8th ed. Berlin and New York: de Gruyter.
- Meillet, A. (1954). *La méthode comparative en linguistique historique* [The comparative method in historical linguistics]. Repr. Paris: Honoré Champion.
- Mowrey, R. and W. Pagliuca (1995). "The reductive character of articulatory evolution". *Rivista di Linguistica* 7, 37–124.
- Ohala, J. J. (1989). "Sound change is drawn from a pool of synchronic variation". In: *Language Change: Contributions to the study of its causes*. Ed. by L. E. Breivik and E. H. Jahr. Berlin: Mouton de Gruyter, 173–198.
- Osthoff, H. and K. Brugmann (1878). *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. Vol. 1. Leipzig: Hirzel.
- Paul, H. (1886). *Principien der Sprachgeschichte*. 2nd ed. Halle: Max Niemeyer. prinziendersp01 paulgoog: ia.
- Rask, R. K. (1818). *Undersøgelse om det gamle Nordiske eller Islandske sprogs oprindelse* [Investigation of the origin of the Old Norse or Icelandic language]. Copenhagen: Gyldendalske Boghandlings Forlag. GoogleBooks: cWgJAAAAQAAJ.
- Sapir, E. (1953). "Language. An introduction to the study of speech".
- Verner, K. A. (1877). "Eine Ausnahme der ersten Lautverschiebung [An exception to the first sound shift]". *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 23.2, 97–130.
- Wang, W. S.-Y. (1967). "Phonological features of tone". *International Journal of American Linguistics* 33.2, 93–105. JSTOR: 1263953.
- Wang, W. S.-Y. (1969). "Competing changes as a cause of residue". *Language* 45.1, 9–25. JSTOR: 411748.
- (2006). *Yǔyán, yǔyīn yǔ jìshù* 語言、語音與技術 [Language, phonology and technology]. Shanghai 上海: Xiānggāng Chéngshì Dàxué.
- Weinreich, U., W. Labov, and M. I. Herzog (1968). "Empirical foundations for a theory of language change". In: *Directions for historical linguistics: A symposium*. Ed. by W. P. Lehmann and Y. Malkiel. Austin: University of Texas Press, 95–189 (–195?).
- Winter, B. (2014). "Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality". *BioEssays* 36, 960–967.

Cross-Linguistic Studies on Semantic Change

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we look at the phenomenon of semantic change and will look at recent attempts to investigate it from a rigorously data-driven, cross-linguistic perspective.

1 Semantics and semantic change

It is well known and not surprising for practitioners of historical linguistics that semantics and semantic change are topics that are very difficult to handle systematically. The reason for this lies in what Sperber (1923: 1) calls the *psychological factors* of meaning, which are much more difficult to grasp and describe than it is to give logical definitions of certain concepts.

Apart from the general question where to allocate semantic change (in the domain of the lexicon or the domain of pragmatics, or as a transition between the two, see (Traugott 2012)), the reason for the problems one faces when dealing with semantic change can be found in the structural differences between sign form and sign meaning and the resulting processes by which both entities change. While the formal part of the linguistic sign is characterized by its sequential structure and sound change is characterized by the *alternation* of segments, the meaning part is better described as some kind of *conceptual network*, and semantic change is not based on an alternation but on the *accumulation* and *reduction* of potential referents,¹ for example by a reorganization of the sign's *reference potential* (List 2014: 36). Although change in meaning is traditionally considered to be notoriously irregular and unpredictable, with scholars emphasizing that “there is [...] little in semantic change which bears any relationship to regularity in phonological change” (Fox 1995: 111), it is also obvious that a large number of observed pathways of semantic change can be observed to occur independently in many different language families of the world. In some sense, we face the same problems we also found for the handling of regular sound change patterns. If we want to study pathways of semantic change cross-linguistically, we will need to find a way to make our data comparable. That this can be cumbersome and difficult could be observed for the Catalogue of Semantic Shifts (Zalizniak 2018, Zalizniak et al. 2012), which originally presented a larger collection of observed semantic change processes, but ultimately has problems to provide a rigorous specification of the different meanings that were tracked.²

[Q] How can we imagine this process of accumulation and reduction to take place, and what is meant by “reference potential”?

2 Fixing meanings: Concepticon

In 1950, Morris Swadesh (1909–1967) proposed the idea that certain parts of the lexicon of human languages are universal, stable over time, and rather resistant to borrowing. As a result, he claimed that this part of the lexicon, which was later called *basic vocabulary*, would be very useful to address the

¹This can already be found in the work of Herman Paul (1846–1921), who emphasizes that there is always an “extension or restriction of the extent of the meaning” and that “only the succession of extension and restriction allows the emergence of a new, from the original one completely different meaning” (Paul 1880 [1886]: 66, my translation).

²To my knowledge, the authors are currently working on a new version that will hopefully cope with the problems of the older version and also provide an increase in data (see <http://datsemshift.ru>).

problem of subgrouping in historical linguistics (Swadesh 1950: 157). He illustrated this by proposing a first list of basic concepts, which was, in fact, nothing else than a collection of concept labels, as shown below:

I, thou, he, we, ye, one, two, three, four, five, six, seven, eight, nine, ten, hundred, all, animal, ashes, back, bad, bark, belly, big, [...] this, tongue, tooth, tree, warm, water, what, where, white, who, wife, wind, woman, year, yellow. (ibid.: 161)

In the following years, Swadesh refined his original concept lists of basic vocabulary items, thereby reducing the original test list of 215 items first to 200 (Swadesh 1952) and then to 100 items (Swadesh 1955). Scholars working on different language families and different datasets provided further modifications, be it that the concepts which Swadesh had proposed were lacking proper translational equivalents in the languages they were working on, or that they turned out to be not as stable and universal as Swadesh had claimed (Alpher and Nash 1999, Matisoff 1978). Up to today, hundreds of different concept lists have been compiled for various purposes.

[Q] For what other purposes might scholars propose concept lists?

2.1 Concept lists

Concept lists are collections of concepts which scholars decided to compile at some point. In an ideal concept list, concepts would be described by a *concept label* (*elicitation gloss*) and a short *definition*. Most published concept lists, however, only contain a concept label. On the other hand, certain concept lists have been further expanded by adding structure, such as rankings, divisions, or relations. Concept lists are compiled for a variety of different purposes. The purpose for which a given concept list was originally defined has an immediate influence on its structure. Given the multitude of use cases in both synchronic and diachronic linguistics, it is difficult to give an exhaustive and unique classification scheme for all concept lists which have been compiled in the past. We find lists produced for historical language comparison (Swadesh 1952), subdivided lists of stable and less stable concepts (see Yakhontov's list mentioned in Starostin 1991), lists of the "most stable" concepts across all times and cultures (), classical questionnaires for linguistic field work (BDS), ranked lists (Starostin 2007), and many concept lists used in psycholinguistics, e.g., to study language acquisition (Ferguson 1964), to conduct naming tests (Ardila 2007), or to study specific semantic domains (Snoek 2013).³

[Q] What is meant by "naming tests" in this context?

2.2 Linking concept lists

While all the concept lists which have been published so far constitute language resources with rich and valuable information, we lack guidelines, standards, best practices, and models to handle their interoperability. Language diversity is often addressed with region- or language-specific questionnaires. This makes it difficult to integrate and compare these resources. The Concepticon (<https://concepticon.clld.org>, List et al. 2016) is an attempt to overcome these difficulties by linking the many different concept lists which are used in the linguistic literature. In order to do so, we offer open, linked, and shared data in collaborative architectures, and by now quite advanced workflows for curating and testing the data we have assembled so far. In the Concepticon project, all entries from different concept lists are partitioned into sets of labels referring to the same concept – so called *concept sets*. Each concept

³See List (2018) for details on the history of concept list compilation.

2 Language Change

set is given a unique identifier (`Concepticon ID`), a unique label (`Concepticon Gloss`), a human-readable definition (`Concepticon Definition`), a rough semantic field, and a short description regarding its ontological category. Based on the availability of resources, we further provide metadata for concept sets (e.g. by including links to the Princeton WordNet University 2010).

[Q] *Why could one not instead just start from Princeton WordNet as the source of definitions and senses? Why does the Concepticon need its own range of concept glosses?*

2.3 Examples

As a simple example for typical problems involving the linking of concept lists, consider the concepts given in the table below. Here, the four lists apparently intend to denote the same concept 'dull'. From the Chinese terms used in the lists by Ben Hamed and Wang (2006) and Chén (1996), however, we can clearly see that the intended meaning is not 'dull' in the sense of 'being blunt (of a knife)', but 'stupid'. Given that both authors originally wanted to render Swadesh's original concept lists in their research, this shows that we are dealing with a translation error here which may well result from the fact that in many concept lists, only 'dull' is used as a concept label, without further specification.

Compiler	Label	Concepticon
Blust (2008)	dull, blunt	DULL
Chén (1996)	呆, 笨 / dull	STUPID
Comrie & Smith (1977)	dull	DULL
Wang (2006)	笨 (不聪明) / dull	STUPID
Swadesh 1952	dull (knife)	DULL

Table 1: Erroneous translations in concept lists

[Q] *What other errors in translations can be possible, when considering Swadesh's original list of 200 concepts?*

3 Cross-Linguistic Data Formats

Linguistics is beyond doubt a data-driven discipline, and most of our daily linguistic work is based on evaluating, creating, and analysing different kinds of data. If one wants to investigate grammatical phenomena, one will need grammatical data, normally example sentences drawn from some kind of corpus. If one want to compare typological aspects of different phenomena, one will again need some kind of corpus in which one can find contrastive examples, or one will have to build this corpus oneself. Even if one simply wants to learn a language which one do not know before, one needs data, as one will need some grammatical descriptions with tables, example sentences, as well as a good dictionary which helps us how to translate words from the foreign language into our own mother tongue.

[Q] *In what subfield of linguistics can data-free research be carried out?*

3.1 Data problems

The problem of data in linguistics is that it is all too often not FAIR in the sense of Wilkinson et al. (2016): **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable.

It is still very difficult to find particular datasets, since linguistic journals often do not have a policy on supplementary data and may lack resources for hosting data on their servers. It is also often difficult to access data, and many papers which are based on original data are still being published without the data and having to request the data from the authors is sometimes a more serious obstacle than it should be. Due to idiosyncratic formats, linguistic datasets also often lack interoperability and are therefore not reusable. (Forkel et al. 2018: 2)

While it was less common to share one's data, or to even compile data directly, in the research of the nineties, and it was beyond doubt even difficult to find a good repository to share one's data up to the end of the first decade of the second millennium, it is disappointing to see to which degree modern linguistic research still fails to be based on FAIR data. While it is clear, that data sharing may be difficult for ethical reasons, there are still many people who think they own their data. While nobody should be required to share their data before a publication, it is clear, however, that a publication that does not offer the data is irreproducible and therefore scientifically questionable (see Berez-Kroeker et al. 2018 for the distinction between reproducible and replicable research).

[Q] *The idea of reproducible research is nice, but how can one avoid to be scooped by colleagues who just grab the data and write papers on them?*

3.2 Data standards

The Cross-Linguistic Data Formats initiative (CLDF, <https://cldf.clld.org>, Forkel et al. 2018) comes along with: (a) standardization efforts, (b) software APIs which help to test and use the data, and (c) working examples for best practice. (a) points to linguistic meta-data-bases like Glottolog (<https://glottolog.org>, Hammarström et al. 2018), Concepticon (List et al. 2016), and the Cross-Linguistic Transcription System initiative (CLTS, <https://clts.clld.org>, Anderson et al. 2018). These databases help scholars to make explicit what data (what *languages*, what *concepts*, what *sounds* they are working with, and additionally aid them in merging different datasets into larger data collections. They aim, in brief, at increasing the *comparability* of linguistic data. (b) points to software (currently written in Python and R), which helps users to test how well their data conforms to the standards established by the CLDF initiative. The software contributes to the *transparency* of the data, as it requires data to be presented in both machine- and human-readable formats. (c) points to existing datasets which have been created by different scholars and try to illustrate how the standards can be used and implemented. These working examples (see, e.g., Sagart et al. 2019) increase both the *availability* of data, they also make them more *findable*, as they are shared on public repositories, with the necessary metadata that makes it easy to search for data in CLDF format, as well as contributing to *transparency* and *comparability*. At the moment, we are trying to lift the CLDF initiative to the next level, by working on new workflows that help for a more efficient creation and curation of cross-linguistic data. A first example for these efforts is the CLICS² database (List et al. 2018), which we will discuss in the next section.

[Q] *What other possibilities apart from sharing examples of best practice and providing standards would we have to encourage and propagate data sharing in linguistics?*

4 Cross-linguistic approaches to semantic change

We have repeatedly seen and discussed how notoriously difficult it is to study semantic change systematically, given that, once it comes to “meaning, one has as a guide only a certain probability based on common sense, on the personal evaluation of the linguist, and on the parallels that he can cite” (Wilkins 1996: 264). Interestingly, however, the often-invoked differences between semantic change and sound

2 Language Change

change become much less striking when we stop to think about sound change as something ultimately *regular*. In the last session, we have discussed the regularity of sound change a lot, and one of the important aspects was that the apparent regularity is nothing else than a change on a higher level, not at the level of the word alone, a change of the phoneme system, as emphasized early by Bloomfield (1933 [1973]: 351). If we look at the *substance* of sound change, at concrete patterns, and the incredible number of different sound segments which scholars propose to have found in certain languages (Anderson et al. 2018), however, sound change does not seem much more chaotic than semantic change. On the contrary: if it is possible to establish a first reference catalogue of phonetic transcriptions, and if we trust that the initial work done in the Concepticon project has been done thoroughly enough, and if we further keep in mind that diachronic patterns often can also be observed synchronically, we may be able to work on feasible solutions to at least approximately reconstruct basic semantic structure from cross-linguistic data.

[Q] How does semantic change surface in synchronic linguistic data?

4.1 Polysemy, homophony, and colexification

Polysemy and homophony are two seemingly contrary concepts in linguistics. However, in the end they describe both the same phenomenon, namely that a word form in a given language can have multiple meanings. François (2008) therefore suggests to replace the two interpretative terms by the descriptive term colexification. Colexification in this context only means that an individual language “is said to colexify two functionally distinct senses if, and only if, it can associate them with the same lexical form” (ibid.: 171).

[Q] How can the distinction between interpretative and descriptive terminology be understood?

4.2 Colexification networks

If one has enough data, it is considerably easy to construct *concept networks* from cross-linguistic colexifications (Cysouw 2010). The starting point are semantically aligned word lists for a large amount of different languages from different language families. By counting, in how many languages, or in how many language families a certain colexification recurs, we can further *weight* the edges of the network, as shown in Figure 1.

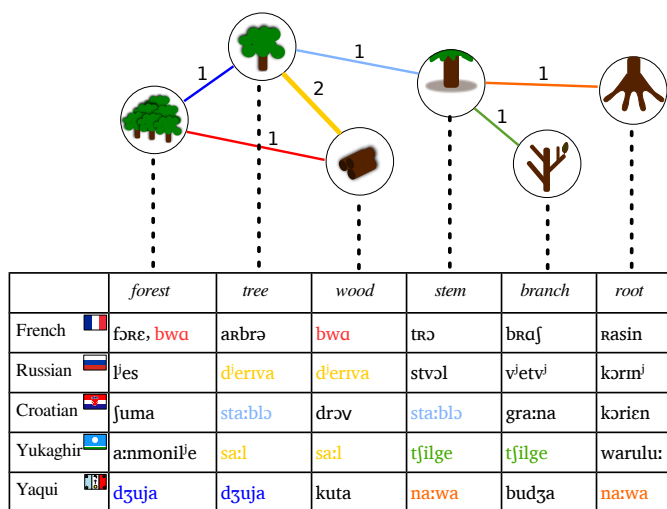
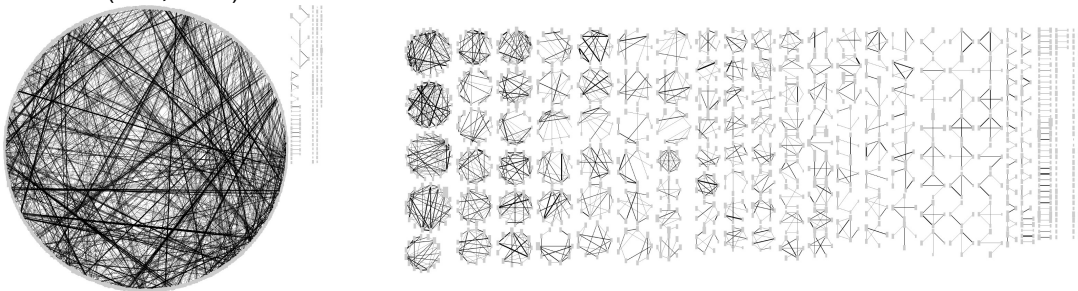


Figure 1: Reconstructing colexification networks from multi-lingual wordlists.

[Q] Is there any straightforward way to derive directed graphs from weighted, undirected colexification networks?

4.3 Analyzing colexification networks

Taking a colexification network alone does not necessarily help us in answering questions regarding semantic change or human cognition. This is due to the increasing complexity of colexification networks, the more concepts and languages we add. The graphic below, for example, shows a network which has been constructed from an analysis of 195 languages covering 44 language families (List et al. 2013). What we need is a network analysis which uses specific algorithms to analyse the structure of the network more properly. In concrete, analyses for *community detection* can help us to partition the networks into groups which correspond to important *semantic fields*. The term *community* was first coined in social network analysis, where it was used to identify communities of people in social networks. In a broader sense, a community refers to “groups of vertices within which the connections are dense but between which they are sparser” (Newman 2004: 4). In List et al. (2013), we used the algorithm by Girvan and Newman (2002) to analyse the network on the left. The result is given in the graphic on the right, where the originally almost completely connected network has been partitioned into 337 communities, with 104 being relatively big (5 and more nodes, covering a rather large parts of the 1289 concepts in our original database (879, 68%).



(a) complete networks

(b) analysed network

Figure 3: Comparing clustered and unclustered colexification networks.

[Q] Below a community from the network is shown, in which meanings which center around “tree” and “wood” have been grouped together. What can we learn from the network? What can’t we learn?

4.4 Database of Cross-Linguistic Colexifications

CLICS² (<https://clics.clld.org>, List et al. 2018) is an online database of colexifications in currently 1220 language varieties of the world. CLICS² superseded the original Database of Cross-Linguistic Classifications, which established a computer-assisted framework for the interactive representation of cross-linguistic colexification patterns (Mayer et al. 2014). While the original CLICS database was low in terms of cross-linguistic coverage and difficult to maintain, the strict adherence to the format specifications based on the CLDF initiative made it possible to grow the data drastically, from originally 221 language varieties in the original version up to 1220 varieties in the current version.⁴

4.5 Data curation and aggregation in CLICS²

The major advancement of CLICS² was a new framework for data curation and aggregation, entirely built on the CLDF strategies. Essentially, this workflow consists of four major stages, which can be carried

⁴We are currently preparing an update that will further increase the coverage to more than 2000 language varieties.

2 Language Change

out independently from each other. These stages include the *mapping of concepts* to Concepticon, the *referencing of sources* in the original data, the *linking of languages* to Glottolog, and the *cleaning of lexical entries* using a dedicated suite of Python scripts. Once data are prepared in this form and rendered in PDF, aggregating data from different sources into a larger database is extremely straightforward. Since the investigation of colexification patterns furthermore not requires to compare word forms *across* languages, but only *inside*, no further normalization (e.g., of the transcriptions) is needed.

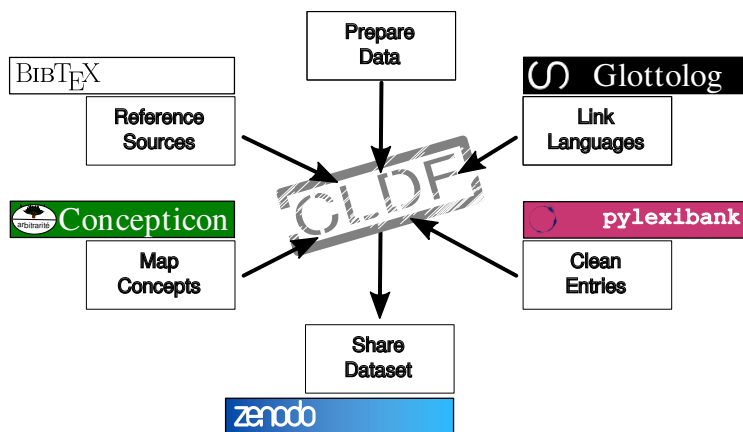


Figure 4: Workflow for data aggregation and curation in CLICS².

[Q] What pitfalls should one avoid when trying to clean lexical entries?

4.6 Examples

The visualization framework used in CLICS is based on an interactive, force-directed, graph layout, written in JavaScript. The basic idea behind this visualization is to allow users to inspect both all the data underlying a given colexification (ideally up to allowing to trace the original datasets, the word forms, and the original elicitation glosses), while at the same time offering a bird's eye view on the global distribution of a given colexification pattern. This is illustrate in the screenshot in Figure 2, where the cluster around words for “tree” and “wood” is shown.

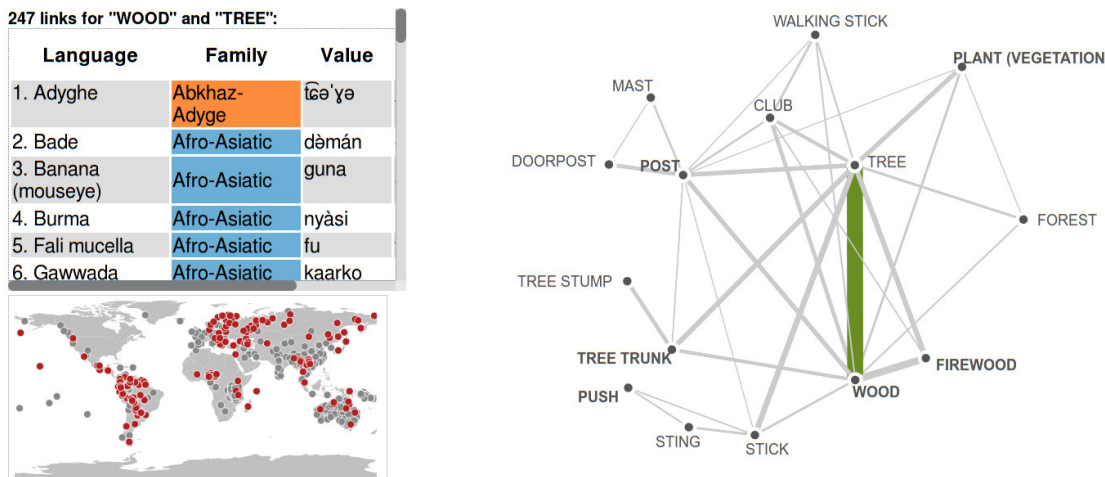


Figure 2: Screenshot from the CLICS² database (see infomap_2_WOOD).

[Q] What exactly does this visualization tell us?

5 Beyond colexification networks

In contrast to the problem of sound change, the identification, the inference of cross-linguistically recurring polysemies can be rather straightforwardly done, by avoiding any distinction between polysemy and homophony in a first place, and then searching for those patterns which recur often enough in big colexification networks. Colexification networks as proposed in the CLICS² database, however, do not solve all problems. First of all, they are a convenient way to present the data to linguists who are interested in the investigation of polysemy patterns due to their individual research. The colexification data as it was assembled with help of our improved CLDF data curation workflows, however, offer much more potential for future investigations. This is shown, for example, by Gast and Koptjevskaja-Tamm (2018) who study areal aspects of polysemy patterns, as well as by (Georgakopoulos and Polis 2018), who present new ideas to add a diachronic dimension. Additionally, there is a lot of potential for studies that use the colexification data in order to check linguistic, cognitive, and psychological theories and hypotheses.

[Q] What theories could, for example, be tested, with help of polysemy patterns?

References

- Allen, B. (2007). *Bai Dialect Survey*. Dallas: SIL International. PDF: <http://www.sil.org/silesr/2007/silesr2007-012.pdf>.
- Alpher, B. and D. Nash (1999). "Lexical replacement and cognate equilibrium in Australia". *Australian Journal of Linguistics: Journal of the Australian Linguistic Society* 19.1, 5–56.
- Anderson, C., T. Tresoldi, T. C. Chacon, A.-M. Fehn, M. Walworth, R. Forkel, and J.-M. List (2018). "A Cross-Linguistic Database of Phonetic Transcription Systems". *Yearbook of the Poznań Linguistic Meeting* 4.1, 21–53.
- Ardila, A. (2007). "Toward the development of a cross-linguistic naming test". *Archives of Clinical Neuropsychology* 22.3. Special Issue: Cultural Diversity, 297–307.
- Ben Hamed, M. and F. Wang (2006). "Stuck in the forest: Trees, networks and Chinese dialects". *Diachronica* 23, 29–60.
- Berez-Kroeker, A. L. et al. (2018). "Reproducible research in linguistics: A position statement on data citation and attribution in our field". *Linguistics* 56.1, 1–18.
- Bloomfield, L. (1973). *Language*. London: Allen & Unwin.
- Chén Bǎoyà 陈保亚 (1996). *Lùn yǔyán jiēchù yǔ yǔyán liánméng 论语言接触与语言联盟 [Language contact and language unions]*. Běijīng 北京: Yǔwén 语文.
- Cysouw, M. (2010). "Semantic maps as metrics on meaning". *Linguistic Discovery* 8.1, 70–95.
- Ferguson, C. A. (1964). "Baby talk in six languages". *American Anthropologist* 66.6, 103–114.
- Forkel, R., J.-M. List, S. J. Greenhill, C. Rzymiski, S. Bank, M. Cysouw, H. Hammarström, M. Haspelmath, G. A. Kaiping, and R. D. Gray (2018). "Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics". *Scientific Data* 5.180205, 1–10.
- Fox, A. (1995). *Linguistic reconstruction. An introduction to theory and method*. Oxford: Oxford University Press.
- François, A. (2008). "Semantic maps and the typology of colexification: intertwining polysemous networks across languages". In: *From polysemy to semantic change*. Ed. by M. Vanhove. Amsterdam: Benjamins, 163–215.
- Gast, V. and M. Koptjevskaja-Tamm (2018). "The areal factor in lexical typology. Some evidence from lexical databases". In: *Aspects of linguistic variation*. Ed. by D. Olmen, T. Mortelmans, and F. Brisard. Berlin and New York: de Gruyter, 43–81.
- Georgakopoulos, T. and S. Polis (2018). "The semantic map model: State of the art and future avenues for linguistic research". *Language and Linguistics Compass* 12.2. e12270 LNCO-0727.R1, e12270-n/a.
- Girvan, M. and M. E. Newman (2002). "Community structure in social and biological networks". *Proceedings of the National Academy of Sciences of the United States of America* 99.12, 7821–7826.
- Hammarström, H., R. Forkel, and M. Haspelmath (2018). *Glottolog*. Version 3.3. URL: <http://glottolog.org>.
- List, J.-M. (2014). *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press.
- (2018). "Towards a history of concept list compilation in historical linguistics". *History and Philosophy of the Language Sciences* 5.10, 1–14.
- List, J.-M., A. Terhalle, and M. Urban (2013). "Using network approaches to enhance the analysis of cross-linguistic polysemies". In: *Proceedings of the 10th International Conference on Computational Semantics – Short Papers. IWCS 2013* (Potsdam, 03/19–03/22/2013). Association for Computational Linguistics. Stroudsburg, 347–353.
- List, J.-M., M. Cysouw, and R. Forkel (2016). "Concepticon. A resource for the linking of concept lists". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation. "LREC 2016"* (Portorož, 05/23–05/28/2016). Ed. by N. C. C. Chair, K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis. European Language Resources Association (ELRA), 2393–2400.
- List, J.-M., S. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2018). "CLICS². An improved database of cross-linguistic colexifications assembling lexical data with help of cross-linguistic data formats". *Linguistic Typology* 22.2, 277–306.
- Matisoff, J. A. (1978). *Variational semantics in Tibeto-Burman. The organic approach to linguistic comparison*. Institute for the Study of Human Issues.
- Mayer, T., J.-M. List, A. Terhalle, and M. Urban (2014). "An interactive visualization of cross-linguistic colexification patterns". In: *Visualization as added value in the development, use and evaluation of Linguistic Resources. Workshop organized as part of the International Conference on Language Resources and Evaluation, 1–8*.
- Newman, M. E. J. (2004). "Analysis of weighted networks". *Physical Review E* 70.5, 056131.
- Paul, H. (1886). *Principien der Sprachgeschichte*. 2nd ed. Halle: Max Niemeyer. prinziendersp01paulgoog: i.a.
- Sagart, L., G. Jacques, Y. Lai, R. Ryder, V. Thouzeau, S. J. Greenhill, and J.-M. List (2019). "Dated language phylogenies shed light on the ancestry of Sino-Tibetan". *Proceedings of the National Academy of Science of the United States of America* 116 (21), 10317–10322.
- Snoek, C. (2013). "Using semantically restricted word-lists to investigate relationships among Athapaskan languages". In: *Approaches to Measuring Linguistic Differences*. Ed. by L. Borin and A. Saxena. Berlin: Mouton de Gruyter, 231–248.
- Sperber, H. (1923). *Einführung in die Bedeutungslehre*. Bonn and Leipzig: Kurt Schroeder.
- Starostin, S. A. (2007). "Computer-based simulation of the glottochronological process (Letter to M. Gell-Mann)". In: S. A. Starostin: *Trudy po yazykoznaniiyu [S. A. Starostin: Works in Linguistics]*. LRC Publishing House, 854–861.
- Starostin, S. A. (1991). *Altajskaja problema i proischozhenije japonskogo jazyka [The Altaic problem and the origin of the Japanese language]*. Moscow: Nauka.
- Swadesh, M. (1950). "Salish internal relationships". *International Journal of American Linguistics* 16.4, 157–167. JSTOR: 1262898.
- (1952). "Lexico-statistic dating of prehistoric ethnic contacts. With special reference to North American Indians and Eskimos". *Proceedings of the American Philosophical Society* 96.4, 452–463. JSTOR: 3143802.
- (1955). "Towards greater accuracy in lexicostatistic dating". *International Journal of American Linguistics* 21.2, 121–137. JSTOR: 1263939.
- Traugott, E. C. (2012). "Pragmatics and language change". In: 249–565.
- University, P. (2010). *WordNet. A lexical database for English*. Online Resource. Princeton.
- Wilkins, D. P. (1996). "Natural tendencies of semantic change and the search for cognates". In: *The comparative method reviewed. Regularity and irregularity in language change. The comparative method reviewed. Regularity and irregularity in language change*. Ed. by M. Durie. With an intro. by M. D. Ross and M. Durie. New York: Oxford University Press, 264–304.
- Wilkinson, M. D. et al. (2016). "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific Data* 3.160018, 1–9.
- Zalizniak, A. A. (2018). "The Catalogue of Semantic Shifts: 20 years later". *Russian Journal of Linguistics* 22.4, 770–787.
- Zalizniak, A. A., M. Bulakh, D. Ganenkov, I. Gruntov, T. Maisak, and M. Russo (2012). "The catalogue of semantic shifts as a database for lexical semantic typology". *Linguistics* 50.3, 633–669.

Empirical Approaches to Studying Language Contact

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we look at the phenomenon of language contact and will look at recent computational attempts to investigate it from a rigorously data-driven, cross-linguistic perspective.¹

1 Similarities

No matter whether one is interested in inherited or borrowed traits, without resorting to some notion of similarity across languages, it is not possible to study historical language relations. Depending on what traits (comparative concepts, in the sense of Haspelmath 2010) we inspect, languages can resemble in various ways. They can share similar words, but also similar structures. While some similarities may give us concrete hints regarding shared histories, many of the similarities we can observe are coincidental or based on general (“universal”) tendencies in the languages of the world. More systematically, we can distinguish similarities that are: (1) *coincidental* (simply due to chance), (2) *natural* (being grounded in human cognition), (3) *genealogical* (due to common inheritance), and (4) *contact-induced* (due to lateral transfer). As an example for the first type, consider Modern Greek θεός [θεός] ‘god’ and Spanish *dios* [dios] ‘god’. Although both words look and sound similar, this is a coincidence, as we see from their oldest ancestors, Old Latin *deivos* and Mycenaean Greek *thehós* (Meier-Brügger 2002: 57f). As an example for the second type, consider Chinese *māmā* [ma⁵⁵ma⁰] 媽媽 ‘mother’ vs. German *Mama* [mama] ‘mother’. Both words are similar, but only because they reflect general principles of early language acquisition (Jakobson 1960). An example for genealogical similarity are German *Zahn* [tsa:n] and English *tooth* [tu:θ], both going back to Proto-Germanic **tanθ-*. Contact-induced similarity is reflected in English *mountain* [mauntɪn] and French *montagne* [mõtɑ̃ʒ], with the former borrowed from the latter. We can display those similarities in some kind of decision tree as shown in Figure 1.

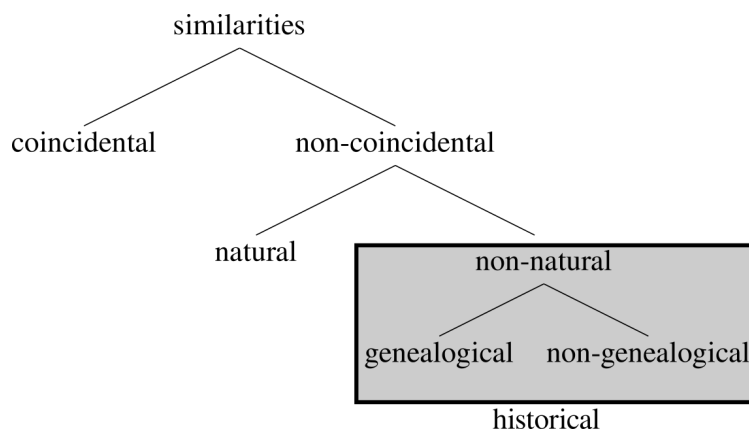


Figure 1: Four reasons for similarities among languages.

[Q] Figure 1 shows a decision tree of the four basic reasons for similarities that can be observed for languages. Why are the last two types on the right labelled “historical” in this figure?

¹This lecture is in part based on List (forthcoming).

2 Contact-induced similarities

The ways in which languages can influence each other vary greatly. While *lexical borrowing* is probably the most frequent way in which language contact surfaces, it is by no means the only possibility, and even the ways in which lexical borrowing can take place, vary greatly, specifically if one keeps in mind that the linguistic sign in one language does not need to be transferred in its entirety into another language, with both its original form and its original meaning. In addition to these “practical” aspects of contact-induced processes of language change, we also have to ask ourselves the question to which degree the type of language contact in a given situation might lead to different processes. Thus, Ross (2013) points out that contact-induced change in the case of bilingual speakers might be fundamentally different in the outcome than contact-induced change as it can be observed for second language learners. In the case of bilinguals, furthermore, Ross (ibid.) assumes that deeper levels of interference, like grammatical calquing and syntactic restructuring, only occur among preadolescents, while classical lexical calquing (i.e., *loan translation*) can also occur among adults. In general, one can say that, while many aspects of language contact have been studied so far, most studies have done so on the basis of particular languages, and no attempts have been made to unify the inference and the analysis of the available data supporting language contact.

[Q] (Weinreich 1953 [1974]) distinguishes between direct borrowing, loan transfer, and hybrid transfer, which can be defined as the expansion of the denotation range of a linguistic sign in the donor language which is phonetically similar to the form of the sign in the recipient language. Can you find examples for this process?

3 Classical approaches to studying language contact

While historical linguistics has developed sophisticated techniques to prove that language similarities are genealogical, the techniques for identifying contact-induced similarities are less homogeneous, involving detailed sifting of multiple pieces which are only in combination convincing. In this regard, techniques for contact detection are not much different from other, more specific, types of linguistic reconstruction, such as the “philological reconstruction” of ancient pronunciations (Jarceva 1990, Sturtevant 1920), the reconstruction of detailed etymologies (Malkiel 1954), or the reconstruction of syntax (Willis 2011). Despite the difficulty in determining exact workflows, we can identify a couple of proxies that scholars use to assess whether a given trait has been borrowed or not.

[Q] What is the specific problem when dealing with multiple pieces of evidence in the historical sciences?

3.1 Direct evidence

The most straightforward way to study language contact is by means of *direct evidence*. The fact that Guǎngzhōu Chinese [t^hai³³ iœŋ²¹] 太陽 ‘sun’ is a recent borrowing from Mandarin Chinese, for example, is easy to prove when comparing modern sources of the dialect with older ones. While sources from the 1960s (CIHUI) list only the form [jit²²t^hœu²¹₃₅] 熱頭, more recent vocabulary collections list exclusively the former form (Liú Lǐlǐ 刘俐李 et al. 2007). If languages are well-documented across time, we can often directly see when a word enters their lexicon. If there is no direct evidence, scholars need to resort to indirect techniques to prove that traits arose from contact. In contrast to general language change, contact-induced change does not proceed in a largely regular manner, but can be seen as a disruptive and chaotic event that may occur but might as well not occur during language history.

[Q] Although direct evidence seems to be the safest guess we have regarding the investigation of language contact, what specific shortcomings can we still encounter here?

2 Language Change

3.2 Conflicts with genealogical explanations

One important class of hints for language contact are conflicts with genealogical explanations. A first type of conflicts is represented by similarities shared among unrelated or distantly related languages. Since these conflicts arise from the supposed phylogeny of the languages under consideration, we can speak of *phylogeny-related arguments* for interference. A second conflict involves the traits themselves, most prominently observed in the case of irregular sound correspondence patterns. We can call these cases *trait-related arguments* for contact. A third type of argument can be derived from distributional properties of shared traits. We can call these *distribution-based arguments* for contact.

[Q] Compare borrowings, such as German *Job* from Modern English, English *mountain* (from Old French), and German *Damm* 'dam' from Low German. What type of conflicts to these represent, respectively?

3.3 Areal proximity

Given that language contact requires the direct contact of speakers of different languages, it is self-evident that areal proximity, including proximity by means of travel routes (or in modern times, by means of communication), is a necessary argument when proposing contact relations between different varieties. Given the potential complexity of these routes, however, and the limited access to sufficient data, most of the research devoted to language contact is either based on the investigation of concrete scenarios (one language whose speakers are known to be bilinguals, for example), or areal proximity is taking as a starting point. Although we know that areal proximity in a literal sense bears certain shortcomings (see, e.g., the analysis by Brockmann and Helbing 2013 on the spread of diseases), it tends to be a good starting point when working with languages whose history is still largely unknown.

[Q] There is one peculiar process of "borrowing" sometimes pointed out in the literature, whose counterpart in biology would be the interbreeding of revived dinosaurs from Jurassic Park with modern crocodiles or other species. What process could this be?

3.4 Borrowability

Since direct evidence confirms that linguistic interference does not act to the same degree on all levels of linguistic organisation, the notion of *borrowability* also plays an important role. Although scholars tend to have different opinions about the concept, most would probably agree with the borrowability scale proposed by (Aikhenvald 2007: 5), which ranges from "inflectional morphology" and "core vocabulary", representing aspects resistant to borrowing, up to "discourse structure" and the "structure of idioms", representing aspects easy to borrow. How core vocabulary can be defined, and how the borrowability of individual concepts can be determined and ranked, however, has been subject to controversial debates (Lee and Sagart 2008, Starostin 1995, Tadmor 2009, Zenner et al. 2014).

[Q] Scholars have been trying for a long time to infer universal borrowing rates. What complicates this search, if one considers closely what "borrowability" actually implies?

4 Computational approaches to studying language contact

Despite the large number of quantitative applications during the last two decades, computational approaches to infer contact situations are still in their infancy. As of now, none of the few approaches proposed so far can compete with the classical methods. The reasons for this are twofold. First, given

the multiple types of evidence employed by the classical approaches, the formalization of the problem of borrowing detection is difficult. Second, given the limited number and suitability of datasets annotated for different types of linguistic interference, scholars have a hard time in developing algorithms, since they lack data for testing and training.

[Q] A third reason for the poor performance of computational methods for the detection of language contact is the “practice” of dealing with borrowings in the field of classical historical and areal linguistics. What is so peculiar about this practice?

4.1 Phylogeny-based approaches to borrowing detection

The basic idea behind all phylogeny-based approaches to borrowing detection is that truly cognate traits should evolve without conflict along the true phylogeny of a given language family. If traits are in conflict with the phylogeny, this is assumed to be a direct hint that these traits were borrowed. Consequently, this also means that the traits which were assumed to be cognate were wrongly annotated when creating the dataset. As an example, consider the scenario for the evolution of words meaning ‘human being’ in Romance, Germanic, and Celtic languages in Figure 2. While we find reflexes of Latin *persona* ‘mask’ in Italian and French (and also Spanish, but not in this particular dataset used for this example), we also find the word *person* in English. By inferring how the words most probably evolved along the given phylogeny, we can see a conflict involving the reflexes of Latin *persona*, as they evolve two times on the tree, one time in Romance, and one time in English. This conflict of the evolution of one character in the phylogeny can be interpreted as resulting from a borrowing event, and we know, of course, that this is true for the case of English *person*. Quite a few approaches that have been published so far make use of this technique and idea (Cathcart et al. 2018, List et al. 2014, Nakhleh et al. 2005), but specifically for the investigation of borrowings in lexical data, the methods were shown to often overestimate the amount of borrowings (Jäger 2018).

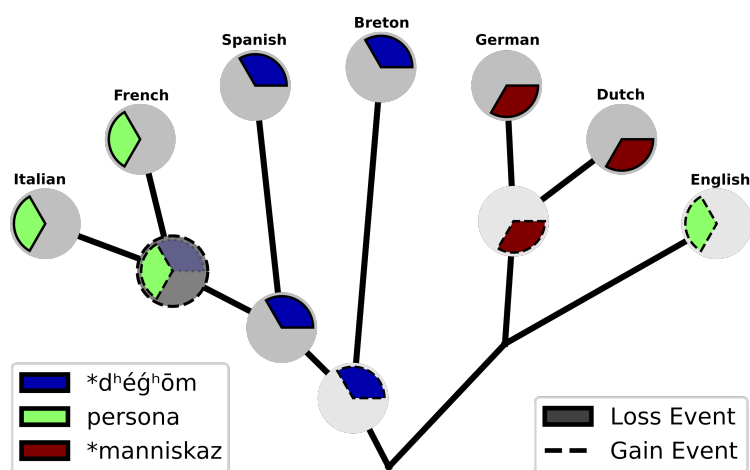


Figure 2: Phylogeny-related conflicts as a hint on borrowings.

[Q] What is the central problem of phylogeny-based arguments for borrowing, not only in quantitative, but also in qualitative approaches?

4.2 Sequence-based approaches to borrowing detection

We have seen before, that regular sound correspondences are usually seen as some kind of proof that words in different languages have been inherited. We have also seen that sophisticated techniques for

2 Language Change

automated word comparison are now available in very stable software packages and can be readily used (List 2017). The most sophisticated techniques in this context measure not only the *surface similarity* of words in different languages, but also the degree to which these similarities are *regular*. For the purpose of identifying borrowings, however, methods that measure only the surface similarity of words have proven more useful, given that – in contrast to regularly inherited words – lexical borrowings show a high degree of surface similarity with the words from which they were copied into the recipient language. When comparing word similarities across unrelated languages, as first proposed by Ark et al. (2007), surface similarities alone can serve as a proxy for borrowing detection. As shown in follow-up studies (Mennecier et al. 2016, Zhang et al. 2019), the cut-off point, or threshold, by which words are automatically judged to be similar or not is crucial for the success of sequence-based approaches to contact inference. In order to determine these thresholds, annotated data is needed, in which linguists have marked which words they consider as obvious borrowings. Zhang et al. (2019) show that the rather simple, historically informed Sound-Class-Based Alignment (SCA) approach (List 2012) largely outperforms earlier approaches, such as the modified edit distance algorithm by (Heeringa 2004), or the rather sophisticated PMI-based scoring system (Wieling et al. 2012).

[Q] *Is there any way to employ the same principle of surface similarity when studying borrowings in related languages?*

4.3 Borrowability-accounts

The idea that lexical concepts could be ranked by the expected borrowability of their counterparts in human languages was most prominently proposed by Swadesh (Swadesh 1952, Swadesh 1955), but even in the work of Antoine Meillet (1866–1936) we can find statements emphasizing that certain concepts tend to be more stable and less prone to borrowing (Meillet 1965). The idea, that concepts can be ranked by their relative borrowability, however, does not provide a concrete method to determine borrowings. While borrowability is regularly employed in classical approaches to studying language contact, an automated account requires a formalized procedure. The first to define such a procedure was Sergey Yakhontov (1926–2018), who proposed to divide a concept list into a stable and a less stable part. Whenever the proportion of related words between two or more languages would be higher in the stable compared to the unstable sublist, he would take this as evidence for deeper genetic relationship. If the proportion showed the opposite behavior, with few words in the stable and many related words in the unstable part, this was taken as evidence for contact.² Interestingly, the idea itself was later re-invented independently by scholars from different backgrounds. Thus, Chén (1996), but used different sublists to resolve questions of language contact in South East Asia. Chén’s principle was then also used to study the affiliation of Bai (Wang 2006), a question that is still unresolved up to today (Lee and Sagart 2008). (McMahon et al. 2005) reinvented Yakhontov’s sublist principle a third time, but while Yakhontov and Chén had divided one list into two, McMahon et al. derived two very small lists from a big one, a stable list, labelled as “hihi”, and an unstable list, labelled as “lolo”. By computing Neighbor-Nets from the lexical distances derived from the sublists, they tried to identify borrowings comparing the networks.

[Q] *What were the criteria of the scholars for the division or ranking of their specific sublists?*

²Although Yakhontov never published any study about this idea, his principle was employed by many colleagues, in whose work, especially that of Sergei Starostin (1953–2005), we find the procedure described in due detail (Starostin 1991).

5 Towards a computer-assisted framework for areal linguistics

Despite projects devoted to a cross-linguistic investigation of language contact in the past (Haspelmath and Tadmor 2009) there is still a large gap between what linguists have claimed to happen in particular languages in comparison with g-linguistics accounts on language contact phenomena. Given how often projects like the *World Loanword Database* (WOLD, <https://wold.clld.org>, *ibid.*) are quoted in the literature, it seems that there is a general interest among linguists to look at contact phenomena from a general linguistics perspective as opposed to the descriptions of certain phenomena described for particular languages. The limitations of the WOLD, on the other hand,³ show, that it is time to re-think to which degree the investigation of language contact, even if it would only address the “low” level of lexical borrowing, could be carried out in a more rigorous framework that makes use of advanced techniques for inference, modeling, and analysis. In the following, initial ideas for such a *computer-assisted* approach to studying language contact, will be presented.

[Q] *What is the inference procedure underlying the borrowing judgments in the WOLD?*

5.1 LIFTing data

The most crucial advantage of the new cross-linguistic era of historical linguistics and linguistic typology is that it is much easier nowadays to assemble large amounts of raw data for different languages of the world. The disadvantage is the poor state (in terms of comparability) of most of these resources. To address the problem of comparability, one needs to *lift* the data to another level. The term *lift* can be taken literally here, but it can also be used as a mnemotechnic device for the most important aspects of this procedure. In this sense, *lift* refers to (1) the *linking of data* to the major reference catalogs (Concepticon, List et al. 2016, and Glottolog Hammarström et al. 2018), (2) the *identification of comparable subsets of the data* (as, for example, described in List et al. 2018), (3) the *fixing* of individual errors, and (4) the *transformation* of regular differences in the transcription systems. Table 1, below, gives an example of datasets we have lifted during the last two years for the CALC project to investigate language contact phenomena in South-East Asia. In this collection, all transcriptions are unified and have been transformed according to the prescriptions of the CLTS initiative (CLTS, <https://clts.clld.org>, Anderson et al. 2018).

Dataset	Concepts	Languages	Words	Conceptlist
Sociolinguistic Research on Monpa (Abraham et al. 2005)	305	30	8213	1
Bai Dialect Survey (BDS)	499	9	4546	1
Chinese Dialect Vocabularies (CIHUI)	738	18	18069	1
Lexical Cognates in Western Kho-Bwa (Bodt and List 2019)	530	8	3958	1
Sui Dialect Research (Castro and Pan 2015)	508	16	9693	1
Yi Varieties in Heqing (Castro et al. 2010)	529	6	3101	1
Zhuang Dialects in Hongshui He (Castro and Hansen 2010)	488	20	11186	1
Miao and Yao Language (Chén 2012)	794	25	21573	1
Wordlists in Selected Languages of Nepal (Hale 1973)	679	13	11041	1
Phonological Database of Chinese Dialects (YINKU)	180	40	10178	1
Collection of Basic Words in Chinese Dialects (Liú Lǐ 刘俐李 et al. 2007)	201	19	4302	1
Naga Languages of North-East India (Marrison 1967)	646	40	27441	1
rGyalrongic Languages Database (Nagano and Prins 2013)	871	10	10685	0
Notes on the Southern Chin Languages (So-Hartmann 1988)	280	8	2171	1
Tibeto-Burman Phonology and Lexicon (Sūn 1991)	905	51	50434	1
TOTAL	2061	313	196591	

Table 1: Data, lifted in the CALC project.

³Here, I mean specifically the small amount of languages used in the sample, and the from today’s perspective rather low level of annotation, which makes it difficult to check the inferences, as criteria for labeling a given word as a borrowing are given only in prose.

2 Language Change

[Q] Given that there are a total of 2061 different concepts in the database, more than any single dataset contains, how can one identify a balanced subset of those data?

5.2 New approaches to phonetic distance

In order to infer borrowings on a large scale, it is important to design methods that work with a high precision, even if this means that their recall is low. The major argument would be that if we have enough data, we would find interesting patterns that can be further analysed, but if we inflate our data with false positives, this would be deleterious. Since borrowings are reflected differently in terms of phonetic similarity between source and target words than cognates, specific algorithms will be needed, which should – ideally – make use of phonetic feature information. Given that we have already lifted our data to the level of a unified transcription system, we can here profit from the feature system underlying the CLTS initiative and create a new algorithm that yields fine-grained pronunciation distances based on articulatory features. In this way, one could then further enhance the procedure for borrowing detection for unrelated languages, as it was outlined in (Menecier et al. 2016).

[Q] What should one keep in mind when designing a feature-based algorithm that essentially returns phonetic similarity or distance scores?

5.3 Initial ideas on stratification

In situations in which languages exhibit more than just sporadic language contact, we can often find that certain parts of a recipient language and its descendants show *layers* or *strata* of words that were introduced in a certain period of contact. This is well-known and well-observed, for example, for the Hmong Mien languages (Ratliff 2010), which have borrowed a lot of words from Chinese, but also share a certain number of words with neighboring Tai-Kadai languages. In a similar way, it is assumed that the Bai varieties, a group of Tibeto-Burman languages spoken in Yunnan, have been heavily influenced by Chinese (Lee and Sagart 2008, Wang 2006). In order to find out, which words belong to a given *layer* or *stratum* of borrowings, linguists manually identify certain sound correspondence patterns, and essentially partition the shared words in groups. The problem of using sound correspondences is that sound change may have masked the layers, so they can no longer be detected. The other problem is that linguists usually do this on the basis of pairwise language comparisons, while there are usually many more languages that exhibit a certain amount of contact. A first idea for the automated assessment of contact strata could therefore consist in a graph-based clustering analysis which essentially clusters previously identified shared (borrowed) words into groups, based on the languages in which they occur. Together with a strict annotation of all identified layers, this would reflect a much more complete picture of all evidence available.

[Q] How could linguists improve their explicitness when it comes to the inference of strata?

5.4 New analyses

Last not least, we will essentially need new analyses for any inferences we can make with respect to language contact. Here, I think specifically of the very interesting theories of interference, as the one mentioned before by Ross (2013), that the processes of language contact change in dependence of the age of the speakers. In a similar way, what we should look into, is the question of *which* words are most frequently borrowed, specifically also to see if we find that the concepts are similar from a global perspective, or not. In fact, the idea of a global *borrowability scale*, as pursued by some researchers, may well

be wrong, since we know very well that contact happens on very specific sociolinguistic settings. While we observe striking parallels for the patterns of denotations, as we have seen in the CLICS² database (List et al. 2018), it may well be that there are no universal tendencies when looking at borrowings.

[Q] What other interesting questions could be asked, if a first set of globally inferred data on borrowings was available?

References

- Abraham, B., K. Sako, E. Kinny, and I. Zeliang (2005). *Sociolinguistic research among selected groups in Western Arunachal Pradesh: Highlighting Monpa*. Dallas: SIL International.
- Aikhenvald, A. Y. (2007). "Grammars in contact. A cross-linguistic perspective". In: *Grammars in contact*. Ed. by A. Y. Aikhenvald and R. M. W. Dixon. Oxford: Oxford University Press, 1-66.
- Allen, B. (2007). *Bai Dialect Survey*. Dallas: SIL International. PDF: <http://www.sil.org/silesr/2007/silesr2007-012.pdf>.
- Anderson, C., T. Tresoldi, T. C. Chacon, A.-M. Fehn, M. Walworth, R. Forkel, and J.-M. List (2018). "A Cross-Linguistic Database of Phonetic Transcription Systems". *Yearbook of the Poznań Linguistic Meeting* 4.1, 21-53.
- Ark, R. van der, P. Menecier, J. Nerbonne, and F. Manni (2007). "Preliminary identification of language groups and loan words in Central Asia". In: *Proceedings of the RANLP Workshop on Acquisition and Management of Multilingual Lexicons*. (Borovets, 09/03/2007), 13-20.
- Běijīng Dàxué 北京大学, ed. (1964). *Hànyǔ fāngyán cihui* 汉语方言词汇 [Chinese dialect vocabularies]. Běijīng 北京: Wénzì Gǎigé 文字改革.
- Boot, T. A. and J.-M. List (2019). "Testing the predictive strength of the comparative method: An ongoing experiment on unattested words in Western Kho-Bwa languages". *Papers in Historical Phonology* 4.1, 22-44.
- Brockmann, D. and D. Helbing (2013). "The Hidden Geometry of Complex, Network-Driven Contagion Phenomena". *Science* 342.12, 1337-1342.
- Castro, A. and B. Hansen (2010). *Hongshui He Zhuang dialect intelligibility survey*. Dallas: SIL International.
- Castro, A. and X. Pan, eds. (2015). *Sui dialect research*. Guizhou: SIL International.
- Castro, A., B. Crook, and R. Fleming (2010). *A sociolinguistic survey of Kua-nsi and related Yi varieties in Heqing county, Yunnan province, China*. Heqing: SIL International.
- Cathcart, C., G. Carling, F. Larson, R. Johansson, and E. Round (2018). "Areal pressure in grammatical evolution. An Indo-European case study". *Diachronica* 35.1, 1-34.
- Chén Bǎoyà 陈保亚 (1996). *Lùn yǔyán jiēchù yǔ yǔyán liánméng* 论语言接触与语言联盟 [Language contact and language unions]. Běijīng 北京: Yǔwén 语文.
- Chén Qiguāng 陈其光 (2012). *Miàoyáo yǔwén* 妙药语文 [Miao and Yao language]. Běijīng: Zhōngyāng Mínzú Dàxué 中央民族大学 [Central Institute of Minorities].
- Hale, A. (1973). *Clause, sentence, and discourse patterns in selected languages of Nepal IV: word lists*. Kathmandu: Summer Institute of Linguistics and Tribhuvan University Press. STEDT: <https://stedt.berkeley.edu/stedt/cgi/rootcanal.pl/source/AH-CSDPN>.
- Hammarström, H., R. Forkel, and M. Haspelmath (2018). *Glottolog*. Version 3.3. URL: <http://glottolog.org>.
- Haspelmath, M. (2010). "Comparative concepts and descriptive categories". *Language* 86.3, 663-687.
- Haspelmath, M. and U. Tadmor, eds. (2009). Berlin and New York: de Gruyter.
- Heeringa, W. J. (2004). "Measuring dialect pronunciation differences using Levenshtein distance". PhD thesis. Groningen: Rijksuniversiteit Groningen.
- Hóu Jīngyí 侯精一, ed. (2004). *Xiàndài Hànyǔ fāngyán yīnkù* 现代汉语方言音库 [Phonological database of Chinese dialects].
- Jakobson, R. (1960). "Why 'Mama' and 'Papa'?" In: *Perspectives in psychological theory: Essays in honor of Heinz Werner*. Ed. by B. Kaplan and S. Wapner. New York: International Universities Press, 124-134.
- Jarceva, V. N., ed. (1990). *Lingvističeskij énciklopedičeskij slovar* (Linguistical encyclopedical dictionary). Moscow: Sovetskaja Enciklopedija.
- Jäger, G. (2018). "Global-scale phylogenetic linguistic inference from lexical resources". *Scientific Data* 5, 180189, 1-16.
- Lee, Y.-J. and L. Sagart (2008). "No limits to borrowing: The case of Bai and Chinese". *Diachronica* 25.3, 357-385.
- List, J.-M. (2012). "SCA: Phonetic alignment based on sound classes". In: *New directions in logic, language, and computation*. Ed. by M. Slavkovik and D. Lassiter. Berlin and Heidelberg: Springer, 32-51.
- (2017). "A web-based interactive tool for creating, inspecting, editing, and publishing etymological datasets". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. System Demonstrations*. Valencia: Association for Computational Linguistics, 9-12.
- (forthcoming). "Automated methods for the investigation of language contact situations". *Language and Linguistics Compass* ????, 1-19.
- List, J.-M., S. Nelson-Sathi, H. Geisler, and W. Martin (2014). "Networks of lexical borrowing and lateral gene transfer in language and genome evolution". *Bioessays* 36.2, 141-150.
- List, J.-M., M. Cysouw, and R. Forkel (2016). "Concepticon. A resource for the linking of concept lists". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation. "LREC 2016"* (Portorož, 05/23-05/28/2016). Ed. by N. C. C. Chair, K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Odiijk, and S. Piperidis. European Language Resources Association (ELRA), 2393-2400.
- List, J.-M., S. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2018). "CLICS². An improved database of cross-linguistic colexifications assembling lexical data with help of cross-linguistic data formats". *Linguistic Typology* 22.2, 277-306.
- Liú Lǐ 刘俐李, Wáng Hóngzhōng 王洪钟, and Bǎi Yīng 柏莹 (2007). *Xiàndài Hànyǔ fāngyán héxīncí, tèzhēng cǐjī* 现代汉语方言核心词·特征词集 [Collection of basic vocabulary words and characteristic dialect words in modern Chinese dialects]. Nánjīng 南京: Fēnghuāng 凤凰.
- Malkiel, Y. (1954). "Etymology and the Structure of Word Families". *Word* 10.2-3, 265-274.
- Marrison, G. E. (1967). "The classification of the Naga languages of North-East India". PhD thesis. London: School of Oriental and African Sciences.
- McMahon, A., P. Heggarty, R. McMahon, and N. Slaska (2005). "Swadesh sublists and the benefits of borrowing: An Andean case study". *Transactions of the Philological Society* 103, 147-170.
- Meier-Brügger, M. (2002). *Indogermanische Sprachwissenschaft*. In collab. with M. Fritz and M. Mayrhofer. 8th ed. Berlin and New York: de Gruyter.
- Meillet, A. (1965). *Linguistique historique et linguistique générale*. Paris: Libr. Champion.
- Menecier, P., J. Nerbonne, E. Heyer, and F. Manni (2016). "A Central Asian language survey". *Language Dynamics and Change* 6.1, 57-98.
- Nagano, Y. and M. Prins (2013). *rGyalrong Languages Database*. Osaka: National Museum of Ethnology.
- Nakhleh, L., D. Ringe, and T. Warnow (2005). "Perfect Phylogenetic Networks: A new methodology for reconstructing the evolutionary history of natural languages". *Language* 81.2, 382-420. JSTOR: 4489897.
- Ratiff, M. (2010). *Hmong-Mien language history*. Canberra: Pacific Linguistics.
- Ross, M. D. (2013). "Diagnosing contact processes from their outcomes. The importance of life stages". *Journal of Language Contact* 6.1, 5-47.
- So-Hartmann, H. (1988). "Notes on the Southern Chin languages". *Linguistics of the Tibet-Burman Area* 11.2, 98-119.
- Starostin, S. A. (1991). *Altajskaja problema i proischozhenije japonskogo jazyka* [The Altaic problem and the origin of the Japanese language]. Moscow: Nauka.
- (1995). "Old Chinese vocabulary: A historical perspective". In: *The ancestry of the Chinese language*. Ed. by W. S.-Y. Wang. Berkeley: University of California Press, 225-251.
- Sturtevant, E. H. (1920). *The pronunciation of Greek and Latin*. Chicago: University of Chicago Press. Internet Archive: [pronunciationgr00unkngoog](http://www.archive.org/details/pronunciationgr00unkngoog).
- Swadesh, M. (1952). "Lexico-statistic dating of prehistoric ethnic contacts. With special reference to North American Indians and Eskimos". *Proceedings of the American Philosophical Society* 96.4, 452-463. JSTOR: 3143802.
- (1955). "Towards greater accuracy in lexicostatistic dating". *International Journal of American Linguistics* 21.2, 121-137. JSTOR: 1263939.
- Sūn Hóngkāi 孙宏开, ed. (1991). *Zàngmǐnyǔ yǔyīn hé cihui* 藏缅语语音和词汇 [Tibeto-Burman phonology and lexicon]. Zhōngguó Shèhuì Kēxué 中国社会科学 [Chinese Social Sciences Press].
- Tadmor, U. (2009). "Loanwords in the world's languages. Findings and results". In: *Loanwords in the world's languages. A comparative handbook*. Ed. by M. Haspelmath and U. Tadmor. Berlin and New York: de Gruyter, 55-75.
- Wang, F. (2006). *Comparison of languages in contact. The distillation method and the case of Bai*. Taipei: Institute of Linguistics Academia Sinica.
- Weinreich, U. (1974). *Languages in contact. With a preface by André Martinet*. 8th ed. The Hague and Paris: Mouton.
- Wieling, M., E. Margaretha, and J. Nerbonne (2012). "Inducing a measure of phonetic similarity from pronunciation variation". *Journal of Phonetics* 40.2, 307-314.
- Willis, D. (2011). "Reconstructing last week's weather: Syntactic reconstruction and Brythonic free relatives". *Journal of Linguistics* 47.2, 407-446.
- Zenner, E., D. Speelman, and D. Geeraerts (2014). "Core vocabulary, borrowability and entrenchment". *Diachronica* 31.1, 74-105.
- Zhang, M., S. Yan, W. Pan, and L. Jin (2019). "Phylogenetic evidence for Sino-Tibetan origin in northern China in the late Neolithic". *Nature* 569, 112-115.

3 Beyond Language Change

The three sections deal with various aspects of historical linguistics that may be of interest for those who want to study language change from a pragmatic perspective, discussing questions of language modeling, the evolution of speech acts, and the evolution of poetry.

Modeling Language Change

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we try to investigate how linguists have tried to model different aspects of language change in the past. Here, we concentrate on external language history, on lexical change, and sound change.

1 Trees, webs, and networks

When discussing how languages change, it is useful to follow the broad distinction, proposed by Gabelentz (2016) to distinguish between *external* and *internal* language history (*innere und äußere Sprachgeschichte*). While internal language history refers to different stages of the same language, external language history refers to the *phylogeny* of a language family, i.e. to the processes by which one language diversifies into multiple descendant languages. In order to discuss the details of modeling language change, we will start from external language history and focus on the history of the classical family tree model.

[Q] In biology, the term *ontogeny* is often contrasted with *phylogeny*, as it concentrates on the life stages of a given organism (birth, adolescence, death). Would it make sense to use this term in relation to internal language history?

1.1 Dendrophily and August Schleicher

Scholars like Jacob Grimm had a rather fuzzy understanding of the historical relatedness of languages, and many scholars kept thinking that contemporary languages could be directly “derived” from each other. This changed in the mid of the 19th century, when scholars started to take the idea that languages seem to evolve in tree-like patterns more seriously. While this idea had been around for some time before the advent of “modern” historical linguistics (List et al. 2016), it was not until scholars like August Schleicher (1821-1868) started to propagate the idea not only in words, but also in illustrations (Schleicher 1853, Schleicher 1861), that the family tree model of language history was accepted as something useful to discuss in historical linguistics.

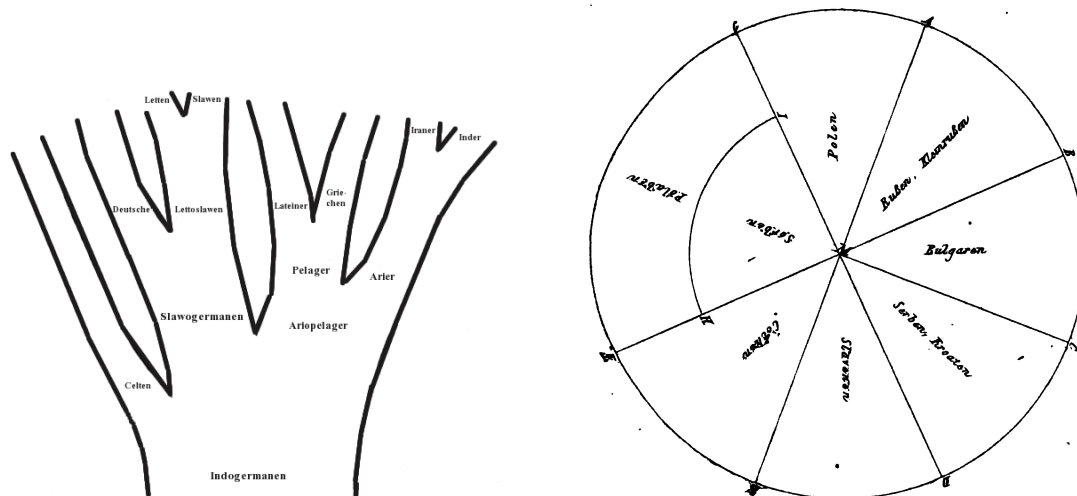


Figure 1: Schleicher’s early tree from 1853, and an attempt to visualize the wave theory by Schmidt (1875).

[Q] *If you compare Schleicher’s early tree drawing from 1853 with modern phylogenetic trees, they will look quite different, in terms of abstraction. What could this reflect about the thoughts of the authors?*

1.2 Dendrophily and Johannes Schmidt

While Schleicher probably euphorically assumed that the family tree model would settle some heated debates in the young discipline of historical linguistics, the phase of “dendrophily” did not exist for long. In 1872, Johannes Schmidt expressed serious doubts in many aspects of historical linguistics, calling not only the idea of a reconstructed proto-language as a “scientific fiction” (Schmidt 1872: 31), but also suggested to replace the idea of the family tree “by the image of a wave that spreads out from the center in concentric circles becoming weaker and weaker the farther they get away from the center” (ibid.: 27). Since then, the theory of wave-like spread of languages, has been quoted a lot in the literature in historical linguistics and is often mentioned as a well-acknowledged opponent of the model of tree-like evolution.

[Q] *The attempt to visualize his “wave theory” by Schmidt (1875) can be seen as symptomatic for further visualizations of the wave theory published afterwards (Geisler and List 2013). What is the specific problem with this visualization?*

1.3 Arachniophily and Hugo Schuchardt

While the strong divide between trees and waves, or dendrophilists and dendrophilists still characterizes the field of historical linguistics, we find already rather early quite modern accounts on the modeling of external language history. Among these is Hugo Schuchardt (1842-1927), who was not content with the simplifications enforced upon language history by the wave model, emphasizing, that the diversification process by which languages split, was characterized by the diffusion of traits among closely related varieties. From this, he concluded: “We connect the branches and twigs of the tree with countless horizontal lines and it ceases to be a tree” (Schuchardt 1870 [1900]). What is important about this idea is not necessarily the fact that the tree Schuchardt describes is no longer a tree, but rather more importantly that the tree is – and this was not mentioned – a very specific *network*, in which processes of inheritance are modeled as a tree, while processes of transfer are handled by adding horizontal lines.

[Q] *If we accept that a phylogenetic network is essentially a tree to which specific horizontal edges were added, can we model all aspects of external language history, or will there remain problematic cases?*

1.4 Save the trees

While biologists by and large accept the usefulness of the tree model along with its limitations, while they specifically also work on extended network-based models in which different processes (usually broken down to vertical vs. lateral relations) of inheritance and transfer can be handled, linguists have remained remarkably stubborn with respect to the wave theory, which has so far not been well explained, and for which no real model exists. The reason seems to be that it was never quite clear, where the historical dimension of the wave model can be found (Jacques and List 2019), while it is clear that any model of language evolution should have a way to model time. Thus, it seems that it may be more useful to not see the wave model as an opponent of the tree model, but rather treat it as a theory of the *actuation of change*. When seen in this way, however, one should stop crediting Schmidt for the theory, as Schmidt is explicitly

3 Beyond Language Change

trying to replace the family tree. In a recent article, Jacques and List (2019) defend the usefulness of the tree model for addressing explicit problems in language change, especially emphasizing how trees can help to identify directional preferences of processes without having any prior evidence with respect to directional tendencies. In biology, the problem of reducing language evolution to a single tree is often also circumvented by using large *forests* of equally powerful but slightly different trees to test, for example, how some features supposedly evolved along a given set of trees (Jäger and List 2018).

[Q] Jacques and List (2019) also discuss the problem of incomplete lineage sorting, which essentially refers to situations in which two extant species seem similar, by sharing a certain amount of material exclusively, because the shared traits were carried along in form of allele variations in the ancestral populations (Rogers and Gibbs 2014: 351). Can a similar problem be found in historical linguistics?

2 Modeling lexical change

While we find many interesting challenges when trying to model external language history, the general picture becomes even more complicated when try to model internal language history. While many linguists probably see the greatest challenge in questions of grammaticalization, it is enough to look into those aspects of language that have been rather thoroughly investigated to find enough challenges to start with. One such aspect is *lexical change*. In a broad sense, lexical change refers to the way in which the lexicon of a human language evolves. In a narrower sense, which we will maintain here, it concentrates on the processes that affect the linguistic signs of a language during its history.

[Q] What are the major processes that can affect a linguistic sign?

2.1 Three dimensions of the linguistic sign

When concentrating on the words and how they are affected during language history, we need to identify the major processes that constitute the changes that affect them. Following Gévaudan (2007: 15-17), we can distinguish three different dimensions along which words can change, namely, the *semantic dimension* (a given word can change its meaning), the *morphological dimension* (new words are formed from old words by combining existing words or deriving new words with help of affixes), and the *stratic dimension* (languages may acquire words from their neighbors and thus contain strata of contact).

[Q] In the second session, we have discussed quickly a slightly extended model of the linguistic sign. To which degree does this model remind of the dimensions of lexical variation by Gévaudan?

2.2 Lexical change and sound change

The focus on three dimensions along which a word can change deliberately excludes sound change. Excluding sound change is justified by the fact that, in the majority of cases, the process proceeds independently from semantic change, morphological change, and borrowing, while the latter three processes often are, of course, cases where sound change may trigger the other three processes – for example, in cases where sound change leads to homophonous words in a language that express contrary meanings, which is usually resolved by using another word form for one of the concepts. An example for this process can be found in Chinese, where *shǒu* (in modern pronunciation) came to mean both “head” and “hand” (spelled as 首 and 手). Nowadays, *shǒu* remains only in expressions like *shǒudū* 首都 “capital”, while *tóu* 头 is the regular word for “head”. interact. Since the number of these processes where we have sufficient evidence to infer that sound change triggered other changes is rather small, we will do better to ignore it when trying to design initial models of lexical change.

[Q] People keep repeating that models do not necessarily need to be realistic. But if they are not realistic, what can we in the end gain from them?

2.3 Lexical replacement

Important work on lexical change goes back at least to the 1950s, when Morris Swadesh (1909–1967) proposed his theory of *lexicostatistics* and *glottochronology* (Lees 1953, Swadesh 1952). What was important in this context was not the idea that one could compute the divergence time of languages, but the *data model* which Swadesh introduced. This data model is represented by a word-list in which a particular list of concepts is translated into a particular range of languages. While former work on semantic change had been mostly *onomasiological* – form-based, taking the word as the basic unit and asking how it would change its meaning over time – the new model used concepts as a *comparandum*, investigating how word forms *replaced each other* in expressing specific contexts over time. This *onomasiological* or concept-based perspective has the great advantage of drastically facilitating the sampling of language data from different languages. Swadesh’s *concept-slot model* can be seen as some kind of a chest of drawers, in which each drawer represents a specific concept and the content of a drawer represents the words one can use to express that given concept. In such a model, lexical change proceeds by *replacement*: a word within a given concept drawer can be kicked out of the drawer in order to make place for another word. Unfortunately, we do not find many attempts to test the characteristics of this model in simulation studies. The only one known to me is a posthumously published letter from Sergey Starostin (1953-2005) to Murray Gell-Mann (Starostin 2007), in which he describes an attempt to account for his theory that a word’s replacement range increases with the word’s age (“Comparative-historical linguistics and lexicostatistics”) in a computer simulation.

[Q] How can one explain what Starostin calls the “aging of words”, i.e., the fact that the longer a word is part of a language, the more likely it is to be replaced?

2.4 Gain and loss

An alternative to Swadesh’s concept-based model of lexical replacement is to treat a language as a bag of words in which – over time – certain words are added, and certain words are deleted. This model is very popular in evolutionary biology, where *gene families* correspond to the words in our bag of words, and evolution is modeled as a process of gene family gain or gene family loss (Cohen et al. 2008). The model is very easy to be applied to linguistics, where the gene family has a counterpart in the etymological *root* or the *word family*. Biologists have described the stochastic characteristics of different gain-loss models, and software packages that help to employ the models for inference of phylogenies are also available (Ronquist and Huelsenbeck 2003). While gain-loss models are frequently used by linguists to infer phylogenies (Gray and Jordan 2000, Sagart et al. 2019), they are less frequently used for plain simulation studies. Here, the only attempts that I know of are one study by Greenhill et al. (2009), where the authors used the TraitLab software (Nicholls et al. 2013) to simulate language change along with horizontal transfer events, and a study by Murawaki (2015), in which (if I understand the study correctly) a gain-loss model is used to model language contact.

[Q] What are the advantages and disadvantages of the gain-loss model in comparison with the concept slot model of lexical change?

2.5 Modeling lexical change with semantic shift

For the moment, no attempt to model morphological change as part of a model for lexical change is known to me (at least not from the perspective of g-linguistics). The problem of the gain-loss and the concept-slot models to account for semantic change, however, can be overcome by turning to *bipartite graph models of lexical change* (see Newman 2010: 32f for details on bipartite graphs). In such a model, the lexicon of a human language is represented by a bipartite graph consisting of *concepts* as one type of node and *word forms* (or forms) as another type of node. The association strength of a given word node and a given concept node (or its “reference potential”, see List 2014: 21f), i.e. the likelihood of a word being used by a speaker to denote a given concept, can be modeled with help of *weighted edges*. This model naturally accounts for *synonymy* (if a meaning can be expressed by multiple words) and *polysemy* (if a word can express multiple meanings). Lexical change in such a model would consist of the *re-arrangement* of the *weights* in the network. Word loss and word gain would occur if a new word node is introduced into the network or an existing node gets dissociated from all of the concepts. We can find this idea of bipartite modeling of a language’s lexicon in the early linguistic work of Sankoff (1969: 28-53), as reflected in the Figure 2 below.

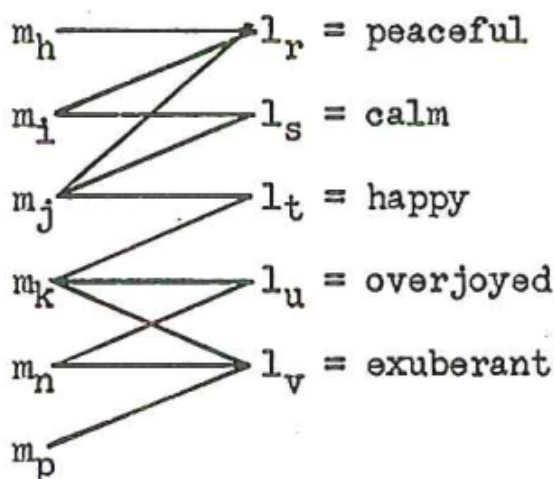


Figure 2: Bipartite graph model by Sankoff (1969: 36).

[Q] What are the advantages and what are the disadvantages of these models in comparison to the concept slot and the gain-loss models? And how would one model lateral transfer (borrowing)?

3 Modeling sound change

When discussing the empirical study of sound change in the third session, we have already covered many of the most interesting aspects of the phenomenon. We have, however, not directly discussed what the consequences of the different theories would be when trying to *model* the process. In the following, we will quickly try to look more closely at the explicit aspects of sound change modeling. When discussing sound change, we need to distinguish *mechanisms*, *types*, and *patterns*. Mechanisms refer to how the process “proceeds”, the types refer to the concrete manifestations of the process (like a certain, concrete change), and patterns reflect the systematic perspective of changes (i.e., their impact on the sound system of a given language, see List 2014).

[Q] +++

3.1 Sound change mechanisms

The question regarding the mechanism is important, since it refers to the dispute whether sound change is happening simultaneously for the whole lexicon of a given language, i.e., whether it reflects a change in the inventory of sounds, or whether it jumps from word to word, as the defenders of lexical diffusion propose, whom I mentioned before (cf. Wang 1969). While nobody would probably nowadays deny that sound change can proceed as a regular process (Labov 1981), it is less clear to which degree the idea of lexical diffusion can be confirmed.¹ What is interesting about the debate about diffusion are the more general implications of the concepts of gradualness and abruptness of change. Since classical sound change would assume gradualness of phonetic change that captures the whole lexicon (and therefore could be said to appear *abrupt* with respect to the lexicon, scholars have stated that the only “grammatical innovation” that could be affected by this change would be the phonetic rules “that assign realizations to phonological categories” (Bermúdez-Otero 2007: 503). Lexical diffusion, on the contrary, would affect the lexical representation of individual words. If we assume, similar to the thoughts by (Ross 2013) on the age during which certain kinds of linguistic interference happen, that humans may adapt new realizations of particular lexical items during their whole lifetime, while they are more likely to acquire phonetic rules in childhood, that we would assume that the two mechanisms present two distinct processes which may even leave an impact in the linguistic data we observe (Sankoff 2018).

[Q] Scholars like Bermúdez-Otero (2007) and Bybee (2002) emphasize a distinction between gradualness (as opposed to abruptness) of change with respect to sounds and with respect to the lexicon. Does this distinction seem reasonable in the light of the discussions at hand?

3.2 Sound change types

When trying to model sound change with respect to its typology, one needs to investigate both the concrete evidence that we could theoretically draw from the languages in the world, and the models that would explain these (e.g. the assumption that sound change induced by speakers or pronunciation differs from sound change induced by listeners or perception). We have also already seen that some of these theories exist (Hock 1991, Ohala 1989), but we face the problem that there is not nearly enough data to support any of the ideas that have been proposed so far in the literature. Initial work on creating a general typology of sound change (similar to a typology of cross-linguistic polysemies, as we have proposed with List et al. 2019) has been carried out (Kümmel 2008), but the major work of finding a way to compare the major tendencies of sound change processes across a large sample of the world’s languages, i.e., the *typology of sound change*, has not been carried out so far. The reason why we are missing this typology is that we are missing clear-cut machine-readable accounts on annotated, aligned data where scholars provide their proto-forms for the reconstructed languages along with their proposed sound laws in a system that can in fact be tested and run (to allow to estimate also the exceptions or where those systems fail).

[Q] Can you think of any examples for sound change that is rather induced by perception than by production?

3.3 Sound change patterns

Even more difficult than modeling the types of sound change, i.e., the tendencies of which one would assume that they derive from the characteristics of speech sounds alone, is the modeling of what is called *patterns of sound change* in this context. Since sound change tendencies are not only initiated by

¹Technically, the theory is dangerous, since it allows a high degree of freedom in the analysis, which can have a deleterious impact on the inference of cognates (Hill 2016). But this does not mean, of course, that the process itself does not exist.

3 Beyond Language Change

the general properties of speech sounds, but also by the linguistic systems in which these speech sounds are employed, it may not be enough to look only at the tendencies of sound change, while ignoring the actual sound systems which assemble particular sounds. *Modeling sound change* thus implies not only to investigate the major system-independent tendencies of sound change and how they come about, but it requires also to investigate to which degree certain tendencies may shift due to specific constellations of *systemic pressure*. An example for an influential attempt to model this are Vennemann's *preference laws of syllable structure* (Vennemann 1988). According to this model, sound change may at times result from specific violations of universal syllabic constraints. In some sense, we can see Vennemann's attempt as a specific *model* for sound change. The problem is, however, that the model was never empirically tested, but rather merely illustrated with help of examples.

[Q] While scholars occasionally mention the importance of the systemic perspective in sound change, there have been no real attempts to separate the two aspects in a concrete reconstruction of a particular language. But how could the two aspects be separated anyway?

3.4 Simulation studies

We can find quite a few published papers devoted to the simulation of certain aspects of sound change, but so far, we do not (at least to my current knowledge) find any comprehensive account that would try to feed some 1,000 words to a computer and see how this “language” develops – which sound laws can be observed to occur, and how they change the shape of the given language. What we find, instead, are a couple of very interesting accounts that try to deal with certain aspects of sound change. (Winter and Wedel 2016) and Wedel for example test agent-based exemplar models, in order to see how systems maintain contrast despite variation in the realization. Hamann (2014: 259f) gives a short overview of other recent articles. Au (2008) presents simulation studies that aim to test to which degree lexical diffusion and “regular” sound change interact in language evolution. Dediu and Moisik (2019) investigate, with the help of different models, to which degree vocal tract anatomy of speakers may have an impact on the actuation of sound change. Stevens et al. (2019) present an agent-based simulation to investigate the change of s to ʃ as it is observed in some languages. This summary of literature is very eclectic, especially because I have only just started to read more about the different proposals out there. What is important for the problem of sound change simulation is that, to my knowledge, there is no approach yet ready to run the full simulation of a given lexicon for a given language, as stated above. Instead, the studies reported so far have a much more fine-grained focus, specifically concentrating on the dynamics of speaker interaction.

[Q] If – what Dediu and Moisik find – turns out to be true, this would have specific consequences on the investigation of tendencies of sound change. Why?

4 General thoughts on modeling

This overview has only touched a very small part of the topics that all those who want to model language change in the one or the other way have to deal with. The major message that I wanted to transport with this course is that we need a more conscious discussions about the modeling tasks in historical linguistics, but maybe also about linguistics in general. While linguists often use the term model, and talk about different approaches in the modeling of a particular problem, the discussions often miss the practical, empirical aspect of language. As a result, the multitude of the different models that have been proposed in the literature so far, were just built in the heads of the people who proposed them along with their students. Any *evaluation* was only carried out on examples that were eclectically selected in order to

illustrate the models or to challenge them. In a paper that deals with gain-loss models of language history and how binary character data can be used to infer phylogenetic trees with such a model, Atkinson and Gray (2006) discuss how “realistic” models in science have to be in the end.

When biologists model evolution, they lie: they lie about the independence of character state changes across sites; they lie about the homogeneity of substitution mechanisms; and they lie about the importance of selection pressure on substitution rates. But these are lies that lead us to the truth. (ibid.)

In historical linguistics (and probably also in linguistics in general), we often wish that the models we develop would directly reflect the complex reality of language. What we often misunderstand, however, is that models that reflect reality in its entirety will have to be so complex that we would not be able to learn anything from them.

[Q] In physics, scholars at times accuse the string theory of being “too powerful”, as it could explain almost everything. What exactly is it that they worry about?

References

- Atkinson, Q. D. and R. D. Gray (2006). “How old is the Indo-European language family? Illumination or more moths to the flame?” In: *Phylogenetic methods and the prehistory of languages*. Ed. by P. Forster and C. Renfrew. Cambridge, Oxford, and Oakville: McDonald Institute for Archaeological Research, 91–109.
- Au, C.-P. (2008). “Acquisition and evolution of phonological systems”. Taipei: Academia Sinica.
- Bermúdez-Otero, R. (2007). “Diachronic phonology”. In: *The Cambridge handbook of phonology*. Ed. by P. de Lacy. New York: Cambridge University Press, 497–517.
- Bybee, J. L. (2002). “Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change”. *Language Variation and Change* 14, 261–290.
- Cohen, O., N. D. Rubinstein, A. Stern, U. Gophna, and T. Pupko (2008). “A likelihood framework to analyse phyletic patterns”. *Philosophical Transactions of the Royal Society B*.
- Dediu, D. and S. Moisik (2019). “Pushes and pulls from below: Anatomical variation, articulation and sound change”. *Glossa* 4.1, 1–33.
- Gabelentz, G. von der (2016). *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Ed. by M. Ringmacher and J. McElvenny. Berlin: Language Science Press.
- Geisler, H. and J.-M. List (2013). “Do languages grow on trees? The tree metaphor in the history of linguistics”. In: *Classification and evolution in biology, linguistics and the history of science. Concepts – methods – visualization*. Ed. by H. Fangerau, H. Geisler, T. Halling, and W. Martin. Stuttgart: Franz Steiner Verlag, 111–124.
- Gray, R. D. and F. M. Jordan (2000). “Language trees support the express-train sequences of Austronesian expansion”. *Nature* 405, 1052–1055.
- Greenhill, S. J., T. E. Currie, and R. D. Gray (2009). “Does horizontal transmission invalidate cultural phylogenies?” *Proc. Biol. Sci.* 276.1665, 2299–2306.
- Gévaudan, P. (2007). *Typologie des lexikalischen Wandels. Bedeutungswandel, Wortbildung und Entlehnung am Beispiel der romanischen Sprachen*. Stauffenburg-Linguistik 45. Tübingen: Stauffenburg.
- Hamann, S. (2014). “Phonological changes”. In: *Routledge Handbook of Historical Linguistics*. Ed. by C. Bownen and B. Evans. London and New York: Routledge.
- Hill, N. (2016). “A refutation of Song’s (2014) explanation of the ‘stop coda problem’ in Old Chinese”. *International Journal of Chinese Linguistics* 2.2, 270–281.
- Hock, H. H. (1991). *Principles of historical linguistics*. 2nd ed. Berlin: Mouton de Gruyter.
- Jacques, G. and J.-M. List (2019). “Save the trees: Why we need tree models in linguistic reconstruction (and when we should apply them)”. *Journal of Historical Linguistics* 9.1, 128–166.
- 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.
- Kümmel, M. J. (2008). *Konsonantenwandel* [Consonant change]. Wiesbaden: Reichert.
- Labov, W. (1981). “Resolving the Neogrammarian Controversy”. *Language* 57.2, 267–308. JSTOR: 413692.
- Lees, R. B. (1953). “The basis of glottochronology”. *Language* 29.2, 113–127. JSTOR: 410164.
- List, J.-M. (2014). *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press.
- 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.
- List, J.-M., S. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2019). *CLICS. Database of Cross-Linguistic Colexifications. Version 2.1.0*. URL: <https://clics.cild.org>.
- Murawaki, Y. (2015). “Spatial structure of evolutionary models of dialects in Contact”. *PLOS ONE* 10.7, e0134335.
- Newman, M. E. J. (2010). *Networks. An Introduction*. Oxford: Oxford University Press.
- Nicholls, G. K., R. J. Ryder, and D. Welch (2013). *TraitLab: A MatLab package for fitting and simulating binary tree-like data*.
- Ohala, J. J. (1989). “Sound change is drawn from a pool of synchronic variation”. In: *Language Change: Contributions to the study of its causes*. Ed. by L. E. Breivik and E. H. Jahr. Berlin: Mouton de Gruyter, 173–198.
- Rogers, J. and R. A. Gibbs (2014). “Comparative primate genomics: emerging patterns of genome content and dynamics”. *Nat. Rev. Genet.* 15.5, 347–359.
- Ronquist, F. and J. P. Huelsenbeck (2003). “MrBayes 3: Bayesian phylogenetic inference under mixed models”. *Bioinformatics* 19.12, 1572–1574.
- Ross, M. D. (2013). “Diagnosing contact processes from their outcomes. The importance of life stages”. *Journal of Language Contact* 6.1, 5–47.
- Sagart, L., G. Jacques, Y. Lai, R. Ryder, V. Thouzeau, S. J. Greenhill, and J.-M. List (2019). “Dated language phylogenies shed light on the ancestry of Sino-Tibetan”. *Proceedings of the National Academy of Science of the United States of America* 116 (21), 10317–10322.
- Sankoff, D. (1969). “Historical linguistics as stochastic process”. Dissertation. Montreal: McGill University.
- Sankoff, G. (2018). “Language change across the lifespan”. *Annual Review of Linguistics* 4.1, 297–316.
- Schleicher, A. (1853). “Die ersten Spaltungen des indogermanischen Urvolkes The first splits of the Indo-European people”. *Allgemeine Monatsschrift für Wissenschaft und Literatur* 3, 786–787.
- (1861). *Compendium der vergleichenden Grammatik der indogermanischen Sprache*. Vol. 1: *Kurzer Abriss einer Lautlehre der indogermanischen Ursprache*. Weimar: Böhlau.
- Schmidt, J. (1872). *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen* On the genetic relations among the Indo-European languages. Weimar: Hermann Böhlau.
- (1875). *Zur Geschichte des indogermanischen Vocalismus. Zweite Abteilung*. Weimar: Hermann Böhlau.
- Schuchardt, H. (1900). *Über die Klassifikation der romanischen Mundarten. Probe-Vorlesung, gehalten zu Leipzig am 30. April 1870* On the classification of Romance dialects. Test lecture given in Leipzig on April 30, 1870. Graz. URL: <http://schuchardt.uni-graz.at/cgi-bin/print.cgi?action=show&type=pdf&id=724>.
- Starostin, S. A. (2007). “Computer-based simulation of the glottochronological process (Letter to M. Gell-Mann)”. In: S. A. Starostin: *Trudy po jazykoznaniju* [S. A. Starostin: *Works in Linguistics*]. LRC Publishing House, 854–861.
- Starostin, S. A. (1989). “Sравnitel’no-istoričeskoe jazykoznanie i leksikostatistika [Comparative-historical linguistics and lexicostatistics]”. In: *Lingvističeskaja rekonstrukcija i drevnejšaja istorija Vostoka* [Linguistic reconstruction and the oldest history of the East]. Vol. 1: *Materialy k diskussijam na konferencii* [Materials for the discussion on the conference]. Ed. by S. V. Kullanda, J. D. Longinov, A. J. Militarev, E. J. Nosenko, and V. A. Shnirelman. Moscow: Institut Vostokovedenija, 3–39; English translation: – (2000). “Comparative-historical linguistics and lexicostatistics”. In: *Time depth in historical linguistics*. Trans. from the Russian by I. Peiros. Vol. 1. Papers in the prehistory of languages. Cambridge: McDonald Institute for Archaeological Research, 223–265.
- Stevens, M., J. Harrington, and F. Schiel (2019). “Associating the origin and spread of sound change using agent-based modelling applied to /s/- retraction in English”. *Glossa* 4.1, 1–30.
- Swadesh, M. (1952). “Lexico-statistic dating of prehistoric ethnic contacts. With special reference to North American Indians and Eskimos”. *Proceedings of the American Philosophical Society* 96.4, 452–463. JSTOR: 3143802.
- Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change*. Berlin, New York, and Amsterdam: Mouton de Gruyter.
- Wang, W. S.-Y. (1969). “Competing changes as a cause of residue”. *Language* 45.1, 9–25. JSTOR: 411748.
- Winter, B. and A. Wedel (2016). “The co-evolution of speech and the lexicon: Interaction of functional pressures, redundancy, and category variation”. *Topics in Cognitive Science* 8, 503–513.

Evolution of Speech Acts

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we discuss some historical aspects on the evolution of speech acts, and then present some ideas and tips with respect to the creation and curation of databases in diversity linguistics.

1 Initial considerations on speech acts

Many ideas that we consider as important breakthroughs in modern linguistics can often be found – at least in some initial form – when looking at the literature of the older scholars of the 19th century. Reading Paul (1880 [1886]), for example, is a pleasure for all those who are interested in general aspects of historical linguistics and language change. Even the work of Grimm (1822) can be full of surprises, but specifically the work of Georg von der Gabelentz (1840-1893) often seems extremely modern, both with respect to his grammar of Chinese (Gabelentz 1881 [1953]) or his book on general linguistics, called *Die Sprachwissenschaft* (Gabelentz 2016). When looking at the history of speech act theory, it is therefore also not surprising that we find first ideas on acting with language already in Gabelentz, confirming that “pragmatic ideas, descriptions and claims have been in the wind for a long time and the pragmatic shift did not come out of nowhere” (Staffeldt 2017: 1). For example, Gabelentz (2016: 108) makes a distinction between what he calls *logical modalities* with *psychological modalities*, which he describes as “the relationship of the speaker to the things being said, whether he informs, asks, exclaims, orders or pleads [...]” (translation by Staffeldt 2017: 2.2). Figure 1 shows a summary image of the different communicative forms of speech identified by Gabelentz.

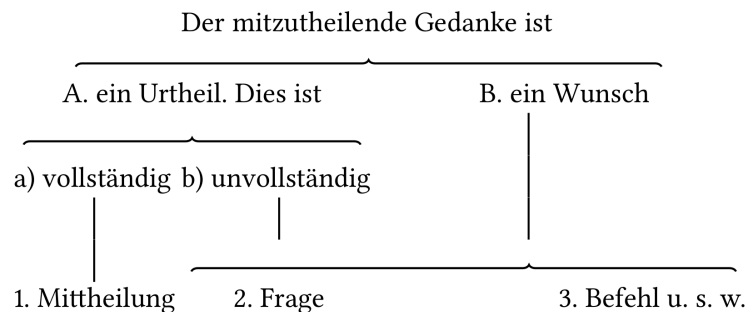


Figure 1: Communicative forms of speech in Gabelentz (2016: 336), discussed in Staffeldt (2017)

[Q] Is it useful to look always for predecessors in the history of linguistics?

1.1 Illocutionary acts

If we look at illocutionary acts as they were first postulated by Austin (1962) and then categorized by Searle (1975) and Searle (1976), we find five basic types of illocutionary acts, namely

- (A) assertatives (make sure the speaker is telling the truth)
- (B) directives (make sure the hearer does as one wants)
- (C) commissives (make sure the speaker commits to an action)

- (D) expressives (make sure the speaker expresses attitudes and emotions)
- (E) declarations (modify the reality with your act)

While the classification of illocutionary acts, or their distinction from other types of speech acts has been discussed in the literature for a long time, it is surprising that we do not often find examples where scholars discuss the concrete distribution of speech acts in the languages of the world. It seems specifically important to make sure which of the examples Searle provides necessarily recur across the world's languages, and which may not.

[Q] *If you take "thanking" as an expressive, what can you tell about its cultural importance and its global distribution?*

1.2 Evolution of illocutionary acts

Not only the synchronic distribution, also the question of how speech acts actually evolve, how they originate, how they change, and under which circumstances, has not often been in the focus of linguistic interest. While it may not be important for general speech act theory that some of the concepts of certain acts being vividly discussed by scholars are in fact about actions that may well not recur across all cultures (like thanking, for example), it is important for the field of linguistics in general, and for diversity linguistics in particular, since here, specifically when working on general linguistics, it is of great interest to see to which degree certain forms of speech acts are universal, and which are not.

[Q] *Which of Searle's five classes seems to be the culturally most diverse?*

2 Cross-linguistic perspectives on illocutionary acts

Before we start discussing and illustrating how a data-driven analysis of the cross-linguistic aspects of illocutionary acts could be best carried out, we should discuss a couple of classical examples that show where historical linguistics or typologists have tried to investigate pragmatic phenomena (or specifically speech acts) cross-linguistically.

[Q] *Do you know any diachronic studies on speech acts?*

2.1 Directives

If we take directives in their simplest form, as commands, we find that many of the world's languages use similar techniques to express turn a phrase into a command, specifically those addressing the direct counterpart of an utterance (i.e., second person). While the design space is considerably large here, ranging from particles via specific lexical items up to the use of bare verbal stems (Aikhenvald 2010: 18), it seems possible to find even some tendencies, in so far as "Synthetic languages tend to mark imperatives with inflectional means. And isolating and highly analytic languages will employ particles (short independent function words) as command markers" (ibid.). Given the importance of the imperative mood in many language's grammar, it is not surprising that linguistic research has been quite thorough in this regard, and that many aspects of both the evolution and the distribution of imperative constructions across the world's languages are relatively good understood. Thus, we often find similar strategies to avoid the imperative (e.g., by asking a question instead), as illustrated in depth by Aikhenvald (ibid.: 288), as well as we know that those constructions that are used to express the imperative may also

3 Beyond Language Change

quite often give rise to new functions (e.g., from imperative to conditional, as we can observe in Russian and German). Apart from the basic sources from which imperative constructions arise, scholars have also investigated basic reasons *why* this should happen. If we take it as a primary rule of evolution that synchronic variation is the pool of future changes, the obvious pool for the change of imperative constructions are those constructions which people use in order to disguise them. With time, these strategies can become the dominant ones, and lead to a shift (Aikhenvald 2010: 342).

[Q] Givón (2005: 172) draws a continuum between prototypical imperatives (*Pass the salt!*) and prototypical interrogatives (*Was there any salt here?*). Can we use a similar way of reasoning to draw a continuum between imperatives and conditionals?

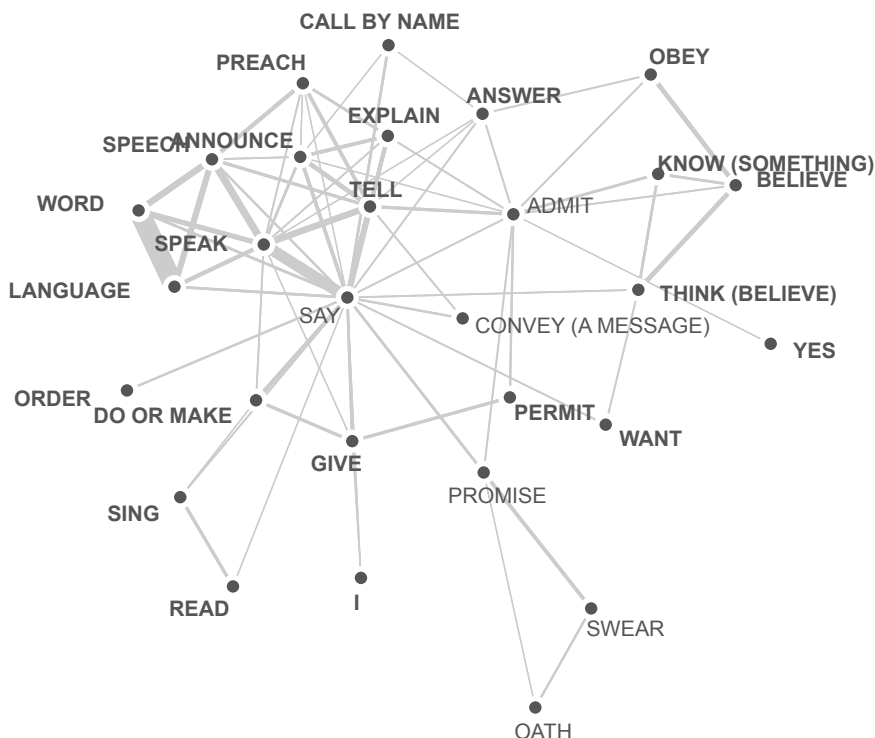


Figure 2: Subgraph for “PROMISE” in CLICS (List et al. 2019).

2.2 Commissives

If we look at typical expressions that speakers use to confirm something, and at the words with which they colexify in the languages of the world, one can use this as a starting point for an investigation into the evolution of the constructions that languages use to express commissives. A very simple and quick example can be drawn from the CLICS database (ibid.), where we can simply search for colexifications involving “YES”, we find that the concept “YES” is mostly colexified with concepts like “GOOD”, “CORRECT (RIGHT)”, “TRUE”, and “CERTAIN”. A second cluster involving “YES” links it to “AND” and “IF”. A spurious link can be found for “NO”, and another link, which may again be spurious, links to “ADMIT”. Given that most of these concepts have counterparts that we may also use in our languages to express the concept of “YES”, it is not surprising to also find this reflected in the languages of the world. Again, the fact that synchronic variation is high when it comes to express the meaning “YES”, we will assume to find quite a few transitions in language history.

[Q] When inspecting the graph in Figure 2 for “PROMISE”, what is the likely direction of the links in which the concept is involved?

2.3 Expressives

Unfortunately, not all concepts can be found in the CLICS database. For this reason, the case of expressives, especially ways to express “THANK (SOMEBODY)” in the languages of the world, cannot be directly carried out in a cross-linguistic way. What is also interesting with respect to the concept of thanking is that this expressive act is obviously not universal, as many linguists have repeatedly reported that certain communities would not only not have any counterparts for our “thank you”, but that the concept was also difficult to understand. What thanking prototypically does (in my opinion) is to fill some perceived gap that results from having received something by another person. In order to deal with this, there are different possibilities for the receivers. They can try to compensate, by giving the person who gave them something or helped them something in return. But most of the time, it is not possible to do so immediately, which is why it may seem even more important to find a strategy to let the person know that the original act has not been forgotten. This is expressed in expressions like English *Thanks!* and German *Danke!*, which originally go back to Proto-Germanic **þankjan* “to think”, of which they are derived (Pfeifer 1993: s.v. “danke”). In other languages, the feeling of gratitude is expressed by turning to higher forces, like in Russian, where *spasíbo* originally meant something like “may got repay you”, an expression, which independently evolved in some parts of Southern Germany, where people express their gratitude by saying “vergelt’s Gott”. Portuguese *obrigado* points to the feeling of an obligation, and Spanish *gracias*, going back to Latin *gratia*, which has been derived from *grātus* “pleasing”, seems to be used to express the joy about having received something. In all these cases, the languages have “normalized” or “ritualized” the situation in which one person receives something from another person and wants to express gratitude, but the ways in which this has been done differ.

[Q] *If some languages don't have words to express gratitude directly, is it possible that the speakers also do not perceive the feeling of gratitude?*

3 How to study evolution cross-linguistically?

While we have by now quite a lot of different resources for linguistic research, including databases of world-wide languages samples, which are getting larger and larger, there are still many questions that have not been thoroughly investigated. I would suspect that speech acts and how they evolve is one example, but it may well also simply be that I missed the relevant literature, when I searched for it, as it may happen that outsiders overlook the core work that has been done, just because they do not use the correct vocabulary. In any way, it seems to be useful to reflect a bit about the exact procedure by which data can be assembled in historical linguistics and linguistic typology. Based on these procedures, we can then discuss some general recommendations for data handling.

[Q] *What problem would you like to address by assembling a global, cross-linguistic database?*

3.1 Cross-linguistic studies from scratch

If one realizes that a certain topic can only be investigated by building a database from scratch, it is important to devote a good amount of time to the planning of the database, and also to include an initial test phase. A cross-linguistic study from scratch does not necessarily involve actual field work. Bigger linguistic database like WALS (Dryer and Haspelmath 2013) have been built with help of linguistic consultants who searched existing grammars of the world’s languages for the relevant information. In cases like WOLD (Haspelmath 2010), this seems to have been a bit different, since the contributors provided actual word lists of the languages, but the sample was also much smaller. The biggest problem when

3 Beyond Language Change

building datasets from scratch is that scholars tend to through out information they deem irrelevant at the time of collection. The ASJP project (Wichmann et al. 2016), for example, planned to provide standardized 40-concept wordlists for as many languages of the world as possible. As a result, scholars only submitted 40 words or sometimes even less, although they occasionally had a lot more. Since the project uses a simplified alphabet for phonetic transcription, scholars who had good-quality original transcriptions could not provide the original transcriptions with the simplified forms, so although the information was available, it has now been lost. To avoid these problems, it is quite important to think not only of what one would personally like to do with the data one assembles, but also what other people might want to do with it.

[Q] How can one make sure that a project remains feasible, i.e., that one can really finish to achieve a sample of sufficient size?

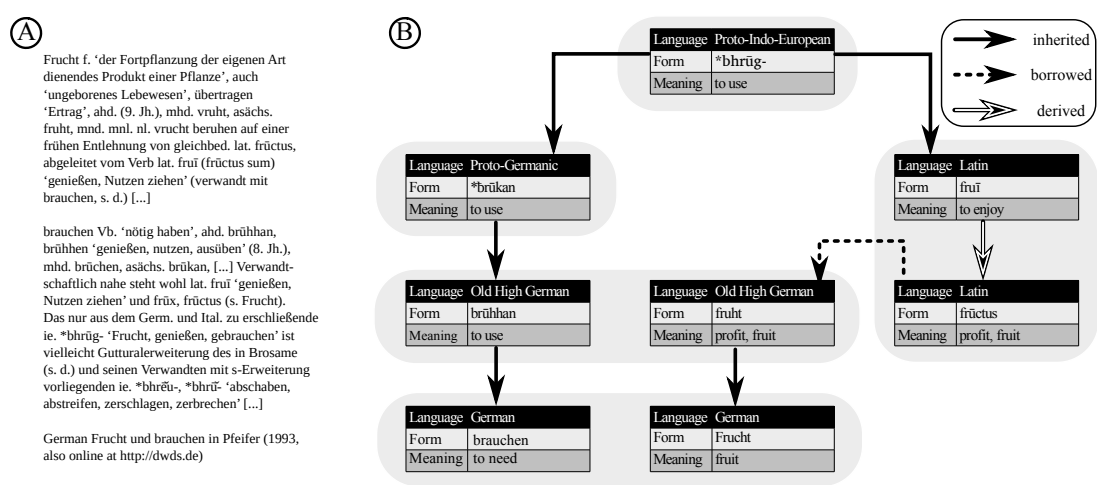


Figure 3: Example for an etymological dictionary entry (left) and a udata-base-like representation (right).

3.2 Aggregation studies

Aggregation studies are the counterpart of studies from scratch. While from-scratch studies have a direct target and try to achieve this by collecting the data for this very purpose, aggregation studies are more based on scholars checking to see what is there and how it can be best combined. An example for such a study is the assembly of the new CLICS² database (List et al. 2019), which was assembled by merging 15 different datasets which at times also considerably differ in quality. While aggregation also requires scholars to invest time, specifically by making the data comparable (linking to Glottolog, mapping to Concepticon), the data can often be assembled much more rapidly than with help of from-scratch studies. The disadvantage of aggregation studies is that they may not cover all points needed for an analysis, that they are substantially skewed (in terms of languages and concepts in the sample), and it may be difficult to enlarge a database further, if one relies on publicly available studies. Their advantage is that they are much less biased methodologically: for CLICS, there was never a pre-selection of which concepts to include, since the goal of the project is to provide as much data as possible for the languages of the world.

[Q] What are the risks and the advantages of the two strategies for data assembly?

4 Basic techniques for data management

Data play an increasingly important role in comparative linguistics. That this is the case should in fact be obvious from the sheer fact that comparative linguistics involves the comparison of languages. Without data, without a thorough comparison of as many commonalities as one can find between two or more languages, it is difficult up to impossible to arrive at convincing conclusions. Despite the great importance of data for the discipline, however, we find only a few articles in the history of linguistics that explicitly deal with data and data collection. Exceptions are Swadesh's "A punchcard system of cognate huntin" in which the scholar describes how he assembles data for lexicostatistic comparison (Swadesh 1963), or the instructions which Gabelentz (2016) gives in quite a few points on what he calls "Collectaneen zum Sprachvergleiche" which he uses both for comparative linguistic work as well as for work on a particular language. But the big textbooks on historical linguistics, which are often cited, do not discuss data at all, we find nothing in Lehmann (1967), nothing in Fox (1995), in Croft (1990), or in Campbell (2013).¹

[Q] *What is the reason for the missing awareness of data in linguistics?*

4.1 Types of data in historical linguistics

Linguists who work in historical linguistics and linguistic typology are much more hungry for data than many of their colleagues from theoretical linguistics. The reason for this is surely that we cannot use our intuition to produce new sentences which we could check for their syntactic properties when dealing with languages from the past. Furthermore, to understand the past, we need to know about current distribution, both to learn from them and to compare them against our findings. For this reason, historical linguistics and linguistic typology have until now assembled massive amountso of linguistic data in different forms, ranging from wordlists, grammars, dictionaries, or field work notes. In order to work with these data, linguists sieve through these dictionaries, wordlists, and field work notes in the search for cognates, for examples that prove their theory, or for so far unrecognized other types of similarities between specific languages. As specific data format in historical linguistics, with a long tradition, are etymological dictionaries, in which the reconstructed proto-forms for a language are arranged in dictionary form, and the reflexes in the descendant languages are listed, along with additional information in prose (see e.g. PFEIFER for an example of such a dictionary for German).

[Q] *In some sense, we could call etymological dictionaries a database, since they assemble all kind of knowledge, albeit in form of prose, along with references to past analyses. However, there is a certain type of data, that may come even closer to an etymological dictionary. Which could this be?*

4.2 Data problems in diversity linguistics

There are many different problems with data in linguistics in general, and in diversity linguistics in particular. We can assign the problems to three different aspects, namely (a) availability of data, (b) the transparency of data, and (c) the comparability of data. Specifically the availability of data is a general problem in diversity linguistics, as it still happens quite a lot that scholars publish papers without also making the data freely available. That this results in research that is not reproducible does not seem to worry some colleagues, but for those who try to raise the scientific standards of our research, it is quite annoying to see how many papers are still delivered and accepted without data and code. We also find a rather high number of grammars where authors do not properly cite where they took the examples from:

¹Campbell lists quite some data in form of tables, but there is no chapter in the book, to my knowledge, that would recommend how data management should be carried out.

3 Beyond Language Change

It is disappointing that so many among the authors of newly commissioned articles did not cite their data; this failing is particularly perplexing in the case of those authors who benefited from the generosity of agencies that explicitly require archiving in public repositories. The move toward open data is still in its early days. (Hill 2017: 306)

A further problem in our research is the lack of transparency when it comes to data management. There are multiple ways in which scholars present their findings on cognacy, but the majority makes it rather hard to understand what is going on, specifically because scholars often do not use alignments (List et al. 2018) to point to the sound correspondences they identified. This all yields to a situation where scholars have accumulated large amounts of data and analyses for the languages of the world, but where the majority of these data is still largely incomparable. It is clear that the comparison across languages is a problem of itself (Haspelmath 2010), and there may be many obstacles, but if we want to achieve more in diversity linguistics, we need to work on drastically increasing the comparability of our data.

[Q] Scholars often repeat the importance of accessibility, transparency, and comparability are crucial for scientific research and the scientific method. But why is this after all the case?

4.3 Recommendations for data organization

At times we may think it would be important to use specific software to curate our data. This assumption is erroneous, since it is not the software, but rather the abstract *schema* which we use to curate our data, which counts. Using software can help in preparation, especially when dealing with annotation. Much more important, however, is the *model* or the *schema* in which one curates the data. The Cross-Linguistic Data Formats initiative (CLDF, <https://cldf.clld.org>, Forkel et al. 2018) offers recommendations and rules for data organization in historical linguistics and linguistic typology. Figure 1 provides two general recommendations which can be seen as crucial for any data enterprise.

<p>(a) One Value per Cell</p> <p>Many datasets that have been published in the past place multiple values in the same cell of their data. This is most frequently the case with elicitation meanings for which multiple translations could be found. Since scholars are rarely explicit about the separators or the techniques by which they handle these problems, many different ways to address multiple translations per meaning have been used in the past, ranging from additional columns up to secondary characters indicating multiple values in a cell (commas, slashes, pipes), and datasets may even mix the different techniques. To avoid these problems, CLDF specifies to use long tables throughout all applications.</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td colspan="4">NEITHER:</td></tr> <tr><td>Meaning</td><td>English</td><td>German</td><td>Dutch</td></tr> <tr><td>bark</td><td>bark</td><td>Rinde, Borke</td><td>bast</td></tr> <tr><td colspan="4">NOR:</td></tr> <tr><td>Meaning</td><td>English</td><td>German</td><td>Dutch</td></tr> <tr><td>bark</td><td>bark</td><td>Rinde</td><td>bast</td></tr> <tr><td>bark</td><td>*</td><td>Borke</td><td>---</td></tr> <tr><td colspan="4">BUT:</td></tr> <tr><td>ID</td><td>Meaning</td><td>Language</td><td>Form</td></tr> <tr><td>1</td><td>bark</td><td>English</td><td>bark</td></tr> <tr><td>2</td><td>bark</td><td>German</td><td>Rinde</td></tr> <tr><td>3</td><td>bark</td><td>German</td><td>Borke</td></tr> <tr><td>4</td><td>bark</td><td>Dutch</td><td>bast</td></tr> </table>	NEITHER:				Meaning	English	German	Dutch	bark	bark	Rinde, Borke	bast	NOR:				Meaning	English	German	Dutch	bark	bark	Rinde	bast	bark	*	Borke	---	BUT:				ID	Meaning	Language	Form	1	bark	English	bark	2	bark	German	Rinde	3	bark	German	Borke	4	bark	Dutch	bast																		
NEITHER:																																																																							
Meaning	English	German	Dutch																																																																				
bark	bark	Rinde, Borke	bast																																																																				
NOR:																																																																							
Meaning	English	German	Dutch																																																																				
bark	bark	Rinde	bast																																																																				
bark	*	Borke	---																																																																				
BUT:																																																																							
ID	Meaning	Language	Form																																																																				
1	bark	English	bark																																																																				
2	bark	German	Rinde																																																																				
3	bark	German	Borke																																																																				
4	bark	Dutch	bast																																																																				
<p>(b) Anticipate the Need of Multiple Tables</p> <p>When a certain complexity of analysis is reached, multiple tables become inevitable in linguistic datasets. Unfortunately, the need of multiple tables is often not readily anticipated, and datasets do not transparently state how to link across tables. Especially formats for cognate coding show great variation in this regard, ranging from multiple sheets in spreadsheet software that were manually created up to customized formats in which additional information is encoded in form of markup, such as colored cells or text in italic or bold font. All these attempts are very error prone and lead to data-loss, especially if only certain parts of the data are shared. To avoid these problems, CLDF specifies to turn to multiple tables whenever this is needed, but to make it explicit in the metadata, how tables should be linked.</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td colspan="4">NEITHER:</td><td style="text-align: right;">--SHEET-B</td></tr> <tr><td>Meaning</td><td>English</td><td>German</td><td>Dutch</td><td></td></tr> <tr><td>bark</td><td>A</td><td>B, A</td><td>C</td><td></td></tr> <tr><td colspan="4">NOR:</td><td style="text-align: right;">--SHEET-A</td></tr> <tr><td>Meaning</td><td>English</td><td>German</td><td>Dutch</td><td></td></tr> <tr><td>bark</td><td>bark</td><td>Rind, Borke</td><td>bast</td><td></td></tr> <tr><td colspan="4">BUT:</td><td></td></tr> <tr><td colspan="2"></td><td style="text-align: center;">--TABLE-A</td><td colspan="2" style="text-align: center;">--TABLE-B</td></tr> <tr><td>ID</td><td>Meaning</td><td>Language</td><td>Form</td><td>ID</td><td>Cognacy</td></tr> <tr><td>1</td><td>bark</td><td>English</td><td>bark</td><td>1</td><td>bark-A</td></tr> <tr><td>2</td><td>bark</td><td>German</td><td>Rinde</td><td>2</td><td>bark-B</td></tr> <tr><td>3</td><td>bark</td><td>German</td><td>Borke</td><td>3</td><td>bark-A</td></tr> <tr><td>4</td><td>bark</td><td>Dutch</td><td>bast</td><td>4</td><td>bark-c</td></tr> </table>	NEITHER:				--SHEET-B	Meaning	English	German	Dutch		bark	A	B, A	C		NOR:				--SHEET-A	Meaning	English	German	Dutch		bark	bark	Rind , Borke	bast		BUT:							--TABLE-A	--TABLE-B		ID	Meaning	Language	Form	ID	Cognacy	1	bark	English	bark	1	bark-A	2	bark	German	Rinde	2	bark-B	3	bark	German	Borke	3	bark-A	4	bark	Dutch	bast	4	bark-c
NEITHER:				--SHEET-B																																																																			
Meaning	English	German	Dutch																																																																				
bark	A	B, A	C																																																																				
NOR:				--SHEET-A																																																																			
Meaning	English	German	Dutch																																																																				
bark	bark	Rind , Borke	bast																																																																				
BUT:																																																																							
		--TABLE-A	--TABLE-B																																																																				
ID	Meaning	Language	Form	ID	Cognacy																																																																		
1	bark	English	bark	1	bark-A																																																																		
2	bark	German	Rinde	2	bark-B																																																																		
3	bark	German	Borke	3	bark-A																																																																		
4	bark	Dutch	bast	4	bark-c																																																																		

Figure 4: Basic recommendations for data management.

CLDF comes along with a software package that can be used to check if data conforms to the format standards, with information on the current standards we support (wordlists, grammatical feature datasets), and examples of best practice. Essentially, CLDF can be created in multiple ways (there's no specific software required to produce a CLDF package), but we recommend users to make use of text editors, or spreadsheet editors, for convenience.

[Q] *Why should the second principle be important?*

4.4 The importance of annotation

Annotation is very important for the analysis of languages and texts. The general idea of annotation is to enrich a given resource by adding more information to it than there was before (Milà-Garcia 2018). Our research question will determine what information we prefer to add to our resources. Interlinear-glossed text, for example, offers a meta-language in order to distinguish grammatical from content words. Interlinear-glossed text thus helps linguists to understand the content of a phrase. In a similar way, the annotation of morpheme structures, as proposed in Hill and List (2017) employs the same idea in order to analyze the meanings of morphemes inside different words of the same language. We can distinguish two types of annotation: *inline* and *stand-off* (Eckart 2012). Inline annotation changes the original data directly, by adding tags. Stand-off annotation, on the other hand, does not touch the original data, but instead makes references to them. Most annotation frameworks mix the two types, to get the best of the two worlds: stand-off-annotation is very flexible, while inline-annotation is often much easier to accomplish in practice.

[Q] *Why is inline annotation easier to realize?*

References

- Aikhenvald, A. Y. (2010). *Imperatives and commands*. Oxford: Oxford University Press.
- Austin, J. L. (1962). *How to do things with words. The William James Lectures delivered at Harvard University (1955)*. Oxford: Clarendon Press.
- Campbell, L. (2013). *Historical Linguistics*. 3rd ed. Edinburgh: Edinburgh University Press.
- Croft, W. (1990). *Typology and universals*. Cambridge: Cambridge University Press.
- Dryer, M. S. and M. Haspelmath, eds. (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Eckart, K. (2012). "Resource annotations" . In: ed. by A. . Clarin-D. Berlin: DWDS, 30–42.
- Forkel, R., J.-M. List, S. J. Greenhill, C. Rzymiski, S. Bank, M. Cysouw, H. Hammarström, M. Haspelmath, G. A. Kaiping, and R. D. Gray (2018). "Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics" . *Scientific Data* 5.180205, 1–10.
- Fox, A. (1995). *Linguistic reconstruction. An introduction to theory and method*. Oxford: Oxford University Press.
- Gabelentz, G. von der (2016). *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Ed. by M. Ringmacher and J. McElvenny. Berlin: Language Science Press.
- Gabelentz, G. v. d. (1953). *Chinesische Grammatik. Mit Ausschluss des niederen Stiles und der heutigen Umgangssprache*. Repr. Berlin: Deutscher Verlag der Wissenschaften. [Original edition: Leipzig: Weigel, 1881].
- Givón, T. (2005). *Context as other Minds: The pragmatics of sociality, cognition and communication*. Amsterdam and Philadelphia: John Benjamins.
- Grimm, J. (1822). *Deutsche Grammatik*. 2nd ed. Vol. 1. Göttingen: Dieterichsche Buchhandlung. Google Books: MnsKAAAAIAAJ.
- Haspelmath, M. (2010). "Comparative concepts and descriptive categories" . *Language* 86.3, 663–687.
- Hill, N. W. (2017). "The State of Sino-Tibetan" . Review of Thurgood and Lapolla (2017) *The Sino-Tibetan Languages*. Second Edition. *Archiv Orientalní* 85, 305–315.
- 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.
- Lehmann, W. P., ed. (1967). *A Reader in Nineteenth Century Historical Indo-European Linguistics*. Bloomington: Indiana University Press.
- List, J.-M., M. Walworth, S. J. Greenhill, T. Tresoldi, and R. Forkel (2018). "Sequence comparison in computational historical linguistics" . *Journal of Language Evolution* 3.2, 130–144.
- List, J.-M., S. J. Greenhill, C. Anderson, T. Mayer, T. Tresoldi, and R. Forkel (2019). *CLICS. Database of Cross-Linguistic Colexifications. Version 2.1.0*. url: <https://clics.cldf.org>.
- Milà-Garcia, A. (2018). "Pragmatic annotation for a multi-layered analysis of speech acts: A methodological proposal" . *Corpus Pragmatics* 2.1, 265–287.
- Paul, H. (1886). *Principien der Sprachgeschichte*. 2nd ed. Halle: Max Niemeyer. prinziendersp01paulgoog: ia.
- Pfeifer, W., comp. (1993). *Etymologisches Wörterbuch des Deutschen*. 2nd ed. 2 vols. Berlin: Akademie. URL: <http://www.dwds.de/>.
- Searle, J. R. (1975). "A taxonomy of illocutionary acts" . In: *Language, mind, and knowledge*. Ed. by K. Gunderson. Vol. 7. Minneapolis: University of Minnesota Press, 344–369. URL: <http://hdl.handle.net/11299/185220>.
- (1976). "A classification of illocutionary acts" . *Language in Society* 5.1, 1–23.
- Staffeldt, S. (2017). "Speech act theory and Georg von der Gabelentz" . *History and Philosophy of the Language Sciences* 9.1, 0.
- Swadesh, M. (1963). "A Punchcard System of Cognate Hunting" . *International Journal of American Linguistics* 29.3, 283–288.
- Wichmann, S., E. W. Holman, and C. H. Brown (2016). *The ASJP database*. Jena: Max Planck Institute for the Science of Human History.

Evolution of Poetry

Johann-Mattis List (DLCE, MPI-SHH, Jena)

Abstract

In this session, we discuss poetry and language in the context of evolution, how rhymes can be used as evidence in historical linguistics, and how they can be annotated.

1 Poetry and language

So far, the relation between poetry and language has not been the key interest of research, neither in linguistic typology, nor in historical linguistics, nor in the theoretical frameworks on grammar. This is a pity, since the interaction between the two is so close that one cannot be thought without the other. While poetry needs a language in which it can be realized (and this language may be signed, spoken, and also written, to some degree), language is to a larger degree also influenced by the speakers intention to speak nicely, even if this is at times ignored in linguistics. It therefore seems useful to have a closer look at the importance of poetry for language, both in general, synchronic, and in diachronic terms.

[Q] In which context can one often hear that a certain language variety has been strongly influenced by poetry?

1.1 The sixth function of language

That language serves not only as a tool for communication, but may also simply be used to sound or look nice, has already explicitly stated by Roman Jakobson (1896-1982), who included a specific *poetic function* of language among his six functions of language.

The set (Einstellung) towards the MESSAGE as such, focus on the message for its own sake, is the POETIC function of language. [...] This function, by promoting the palpability of signs, deepens the fundamental dichotomy of signs and objects. (Jakobson 1960: 356)

If the way we express things is not only based on the message we want to transfer, but also on factors which are genuinely related to aesthetical factors, this means that we cannot study the development of languages without taking the poetic function of language into account. While a synchronic analysis of linguistic structures might do without taking the form of the message into account, a historical analysis can surely not do without it, since “poetic factors” may well be the source of certain linguistic developments (be they regular or sporadic).

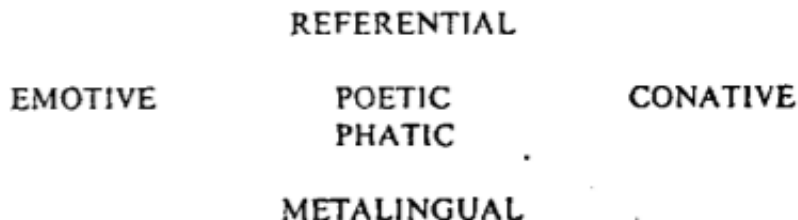


Figure 1: Six functions of language in Jakobson (1960).

[Q] Jakobson (1960) distinguishes six functions of language, the referential function (serves to refer to something), the emotive function (serves to express something about the speaker), the conative function (serves to try to influence the hearer), the phatic function (to make sure communication works fine), the metalinguistic function (to talk about language), and the poetic function (to talk nicely). Figure 1 shows how Jakobson himself arranged the different functions. Why does he place them in this way?

1.2 The evolution of poetic traditions

Poetic traditions evolve in close interaction with three key factors, namely *cognition*, *culture*, and *communication*. These three factors influence each other in different degrees and ultimately yield the patterns in the evolution of particular poetry traditions which we can observe over time. Communicative (or linguistic) traditions are mostly vertically inherited by the generation of parent speakers, but wide-spread transfer of linguistic material through contact is also well-known in language history. Cultural traditions are easily spread also across linguistic boundaries under specific cultural settings. Cognitive factors are supposed to be universal among populations but we do not know to which degree first-language acquisition or external factors (where people live, in which climate, etc.) may influence the way we perceive the world. As an example for a specific tradition of poetry, consider rhyming. Rhyming is not universal among the world's poetic traditions. We may therefore ask, (1) under which circumstances rhyming evolves independent of contact, (2) how easily rhyming traditions are lost, (3) how strongly cultural factors (traditions of archaic rhyming, copying of the masters, etc.) influence them, and (4) to which degree rhymes accepted by a population conform to rules of linguistic (phonetic, phonological) similarity (or dissimilarity).

[Q] Which of the three factors mentioned here seems to be the most important to describe the evolution of poetic traditions?

1.3 Language structure and poetry

Language structure most likely favors specific traditions of poetry, but so far, we do not know, what aspects of linguistic structure is tight to certain types of poetry. So we may again ask ourselves, (1) what kinds of linguistic structure can be shown to influence poetic traditions, (2) how linguistic structure copes with traditions if poetry is superimposed by another culture, and (3) under which circumstances cognitive factors (e.g., comprehensibility or memorializability of poetry) influence the emergence of poetic traditions.

[Q] Here, the questions that are being asked are directed at the restrictions that language may impose on poetry. But what does poetry have to offer to language, or to their users?

2 Chinese rhyme analysis

The analysis of rhyme patterns is one of the core methods for the reconstruction of Old Chinese phonology. It emerged when scholars of the Suí 隋 (581–618) and Táng 唐 (618–907) dynasties realized that old poems, especially those in the Book of Odes (Shījīng 詩經 ca. 1050–600 BCE), were full of inconsistencies regarding the rhyming of words. While the first reaction was to attribute inconsistencies to a different, less strict attitude towards rhyming practiced by the ancestors (as advocated by Lù Dè míng 陸德明, 550–630), or to a habit of the elders to switch the pronunciation in certain words in order to make them rhyme (a practice called *xíéyīn* 諧音 'sound harmonization', Baxter 1992: 153). Later scholars from the Míng 明 (1368–1644) and Qīng 清 dynasties (1644–1911) realized that the inconsistencies in

3 Beyond Language Change

the rhyme patterns reflect the effects of language change (Baxter 1992: 153-157). This is illustrated in Table 1.

Chinese Text	Translation	RW	Patterns	MCH	OCBS-Rhyme
燕燕於飛	The swallows go flying	<i>fēi</i> 飛	A	*pjij	*-ər
下上其音	falling and rising are their voices;	<i>yīn</i> 音	B	*ʎim	*-əm
之子於歸	This young lady goes to her new home,	<i>guī</i> 歸	A	*kjwɨj	*-əj
遠送於南	far I accompany her to the south.	<i>nán</i> 南	B	*nom	*-əm
瞻望弗及	I gaze after her, can no longer see her,	[<i>jí</i> 及]	-	[*gɨp]	[*-əp]
實勞我心	truly it grieves my heart	<i>xīn</i> 心	B	*sim	*-əm

Table 1: Changing rhymes in Chinese poetry.

Assuming that rhyming was originally rather consistent, with rhyme words being mostly identical in the pronunciation of nucleus and coda, the analysis of rhyme words makes it not only possible to establish rhyme categories but also to interpret them further phonetically or phonologically. The classical approach for rhyme analysis, which is called *sīguàn shéngqiān fǎ* 絲貫繩牽法 ‘link-and-bind method’ (Gēng 2004), or *yùnjiǎo xìlián fǎ* 韻腳系聯法 ‘rhyme linking method’ (Lǚ 2009), consists of roughly two steps: In a first step, groups of Old Chinese words, mostly represented by one Chinese character and identified to rhyme with each other in a given text are collected. In a further step, these groups are compared with each other. If identical words are found in different groups, those groups can be combined to form larger groups. This procedure is then repeated until categories of rhymes can be identified that ideally do not show any more transitions among each other. This approach is essentially similar to the ‘linking method’ *xìlián fǎ* 系聯法 see Liú 2006: 56-67), first proposed in Chén Lǐ’s 陳禮 (1818–1882) *Qièyùnkǎo* 切韻考 (1848), by which characters used in *fǎnqiè* 反切 readings in rhyme books are clustered into groups of supposedly common pronunciations for initials and rhymes. In both approaches, similarities in pronunciation are indirectly inferred by spinning a web of direct links between characters.

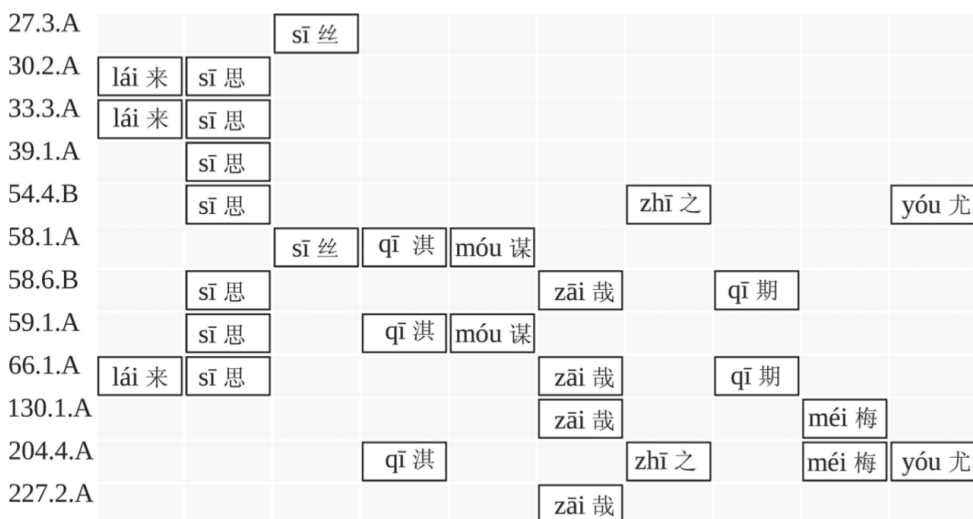


Figure 2: Illustrating the linking method for rhyme analysis.

[Q] Figure 2 illustrates the linking method for the zhī 之 group in the Book of Odes. What is the obvious drawback of this method?

2.1 Network Approach to Rhyme Analysis

The crucial idea of our computer-assisted approach to rhyme analysis is to construct a *network of rhyme patterns* in which nodes represent rhyme words and connections between nodes represent how often those rhymes co-occur in the Book of Odes. The following graphic illustrates this procedure for two stanzas of the Shījīng:

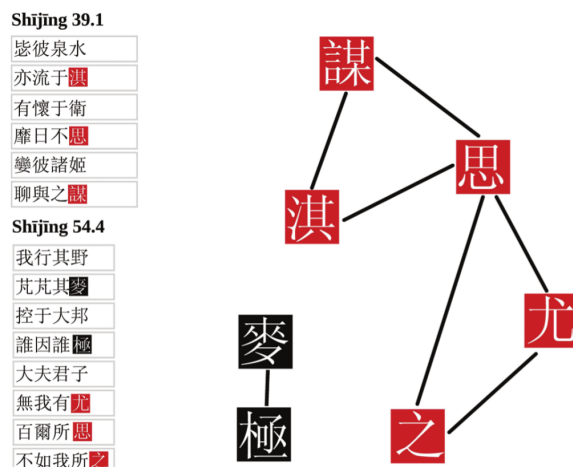


Figure 3: Construction of rhyme networks.

The major advantage of this representation is that we can apply various methods for network analysis to data which was assembled in this form. As a result, we can investigate the rhyme network and test to which degree different reconstruction systems offer a consistent view on Old Chinese rhyming. As a very simple test, we can check whether a given reconstruction system conforms to the principle of *vowel purity* (Ho 2016) which expects words with similar vowels to rhyme more often than words with different vowels. Our test, which is reported in List et al. (2017) could show that most of the Old Chinese reconstruction systems which postulate 6 vowels correspond more closely to vowel purity than other reconstruction systems with more or less vowels. Even by eyeballing the figure above, in which vowel quality is reflected with help of colors following the OC reconstruction system by Baxter and Sagart (2014), one can see that words rhyming with each other tend to have the same vowel.

[Q] If six-vowel reconstruction systems perform better on vowel purity, does this automatically mean that they are better in general?

2.2 The Shījīng Rhyme Browser

In order to make it more convenient for the readers to investigate the data underlying this paper in full detail, an interactive web-based application was created. This freely available Shījīng Browser (<http://digling.org/shijing/>) lists all potential rhyme words in tabular form along with additional information including the *pīnyīn* transliteration, the Middle Chinese reading, the reconstruction by Baxter and Sagart (ibid.), the reading by Pān (2000), the GSR index (Karlgren 1957), and the number of poem, stanza, and section. With help of interactive search fields, the data can quickly be filtered, enabling the users to search for specific poems, for specific characters, or for specific readings. When clicking on the “Poem” field in the application, a window pops up and shows the whole poem, in which all rhyme words are highlighted. In certain cases, where potential alternative rhymes were identified, this is marked in an additional column. In a recently modified version, we contrast rhyme annotations by Wáng (1980 [2006]) with those given in Baxter (1992) (<http://digling.org/shijing/wangli/>, List 2017). The table below gives an example on the organization of the interface.

3 Beyond Language Change

Text	Stanza	MCH	Pān Wúyùn	OCBS	Wáng Lǐ	Starostin	Rhyme	Group
遵彼汝墳，伐其條枚	1.AB	mwoj	muwul	m ^h əj	muəi	məj	A	微
未見君子，惄如調飢	1.CD	tsiX	kril	Cə.kəj]	kiei	krəj	A	脂
遵彼汝墳，伐其條枚	2.AB	sijH	ph-ljuɔds	s-ləp-s	jiet	silhəps	B	質
既見君子，不我遐棄	2.CD	khjiH	khids	[k ^h][t]-s	khiet	khjɔs	B	質
魴魚鱗尾	3.A	mj+jX	muwul?	[m]əj?	miuəi	məj?	C	微
王室如燬	3.B	xjweX	qh ^w ral?	[ŋ](r)aj?	xiuəi	h ^w ej?	C	微
雖則如燬	3.C	xjweX	qh ^w ral?	[ŋ](r)aj?	xiuəi	h ^w ej?	C	微
父母孔迺	3.D	nyeX	mjel?	n[ə][r]?	njiei	n(h)ej?	C	脂

Table 2: Rhyme browser for the book of Odes.

[Q] What could be the problem of comparing rhymes in books other than the Book of Odes?

3 Annotating rhymes in poetry

Having seen that it pays off, in general, to work on computer-assisted approaches in those cases where large amounts of data have to be handled, we might want to step back a bit from the very specific question of Chinese historical phonology and the rhyming practice, and rather ask, similar to what we did in Session 7, what we could do if we had a large database of poetry, and what questions we would like to ask. Once we have determined this sufficiently, we should decide what kind of data we want to have. In fact, most of the questions have already been discussed in the first section of this session. The question that now remains for us is how we can actually handle world-wide data on poetry in such a way that we can address these questions? In the following, we will first look at how it is actually being done, and then develop an alternative framework from the problems we observe in the current practice.

[Q] What specific questions would you like to ask about the evolution and typology of poetry?

3.1 Current annotation practice

When analyzing rhymes in poetry, one of the most crucial questions is what rhymes with what and where it rhymes. We can call such an analysis (which is a true analysis, since we may assume that experts commit errors in their assessment of either what the majority of language users think or what the author intended) a *rhyme judgment analysis*, similar to the term *cognate judgment*, which reflects the identification of potential cognate words by experts or algorithms. The ways in which scholars share their respective rhyme judgments in the literature is very diverse and makes a formal comparison of different rhyme analyses difficult. The problem here lies only to some degree in missing digital versions of important contributions, which would be merely a problem for pure computational approaches. A more significant problem is that many authors report their rhyme judgments in a form that is insufficiently explicit to infer the individual judgments made on individual poems and stanzas. Apart from scholars who presented only the results of their analyses, without providing the evidence, we also often find analyses that are extremely difficult to inspect, due to the way they present their judgments. In this sense, only a small amount of rhyme analyses is truly explicit. Among the few explicit rhyme analyses, we again face the problem that scholars differ widely in the formats they use for annotation, and also in the depth of annotation provided.

We have seen before that one can roughly distinguish between *inline* and *stand-off* annotation (Eckart 2012).¹ As an example illustrating the difference between the two annotation styles, consider the rhyme annotation employed by Baxter (Baxter 1992) as compared to the one by Wáng Lì Wáng 1980 [2006], for poem 109 (second part of stanza 2 in the Book of Odes). While Wáng Lì provides the rhyme judgements inline, Baxter (p. 625) basically uses a stand-off annotation by listing all relevant data in tabular form:

Character	Pinyin	MCH	OCH	Rhyme
哉	zāi	tsoj	*tsi	B
其	jī	ki	*k(r)ji	B
之	zhī	tsyi	*tji	B
之	zhī	tsyi	*tji	B
思	sī	si	*sji	B

彼人是哉(tzə)! 子曰何其(giə)
 心之憂矣,其誰知之(tjiə)?
 其誰知之(tjiə)?
 蓋亦勿思(sio)! (*Shijing*, 109.2)

Figure 4: Rhyme annotation in Baxter (1992) and Wáng (1980).

[Q] In order to test their algorithm on automated rhyme detection, Haider and Kuhn (2018) uses a corpus in which poems are separated into stanzas, and stanzas are separated into lines, and rhyming is annotated by providing an attribute for each stanza, which reflects which line rhyme with which line, similar to the practice in school, using letters of the alphabet. What huge disadvantage has this system?

3.2 Requirements for annotation frameworks

Our basic ideas for a useful annotation framework require: (1) simplicity, (2) exhaustiveness, (3) flexibility. Simplicity means that people should be able to apply our format prescriptions with a minimal amount of work, using standard off-the-shelf tools, like text or spreadsheet editors, rather than complex new tools that would have to be created specifically for rhyme analysis. Exhaustiveness means that we wish to be able to reflect all knowledge that can be formalized in a given rhyme analysis. While we would always allow adding ad-hoc information in note-fields, we want to offer a high degree of granularity in annotations, allowing, for example, the inclusion of phonetic transcriptions and phonetic alignments (List 2014). Flexibility allows for a quick extension of the data when needed, using mechanisms already offered by the framework. In order to achieve all these goals, we draw largely from our experience with the enhanced annotation and computer-assisted manipulation of wordlists in historical linguistics (Hill and List 2017) and their subsequent inclusion into the CLDF specifications.

[Q] The Zen of Python says: “Simple things should be simple, complex things should be possible”. What does this mean exactly?

3.3 Preliminary framework for rhyme annotation

Based on the discussions of the desiderata and past experiments which proved the particular insufficiency of certain annotation forms, the core annotation of a poem or a poem collection, as proposed in (List et al. 2017) now contains the following main components:

- ID: the identifier, which is a numerical ID.
- POEM: a name for the given poem.

¹While inline annotation manipulates the original data directly, for example, by adding tags, stand-off annotation only references the original data, without directly modifying it. Most annotation frameworks, however, typically use a mixture between the two types, although it is clear that stand-off annotation has the advantage of allowing for far more flexibility, especially if adding multiple layers of annotation to a given resource.

3 Beyond Language Change

- STANZA: the stanza of the poem (usually a numeric value, preceded by the name of the poem).
- LINE_IN_SOURCE: the line of the poem as we find it in the source from which the data is taken (especially containing original punctuation etc.).
- LINE: a double-segmented version of the line, in which words are separated with help of + as a separator, and spaces can be used to represent segments of phonetic values (similar to the format adopted by the LingPy software package to represent phonetic sequences and alignments).
- LINE_ORDER: A numerical value that provides the order of the lines of a poem in a given stanza.
- RHYMEIDS: A list of numerical identifiers, indicating which words in a the LINE rhyme by assigning the same ID to different words, using 0 to indicate that a given word does not rhyme.
- ALIGNMENT: A double-segmented version of the line that can, however, store aligned content, differing from the data in LINE, as well. This data comes in handy when trying to check questions of phonetic similarity of rhyme words, or of vowel purity, which would greatly facilitate automatic analyses as the one presented in List et al. (2017).

With these eight columns provided, poems can be annotated in a very straightforward way, regardless of the language in which they were written. One can, of course, add many more columns, depending on specific characteristics of the datasets, but for the general rhyme annotation, we think that these fields will be sufficient for most of the cases; it substantially exceeds rhyme annotation frameworks that have been proposed so far in terms of detail.

[Q] *What is the obvious drawback of this annotation schema?*

3.4 PoePy: Python library for quantitative handling of rhymes

We have developed a software API, called PoePy (<https://github.com/lingpy/poepy>), that allows one to parse, manipulate, and convert files following our new rhyme annotation schema in a convenient way, with help of the Python language. The framework builds heavily on LingPy, a Python library for quantitative tasks in historical linguistics (List et al. 2018), as well as SinoPy, a Python library for specialized tasks in Chinese historical linguistics (List 2018). The GitHub site of our API offers additional information for installing and using our software library. PoePy can read datasets in our general format mentioned above, it can also be used to align rhyme words, provided they are readily assigned to the data, and it can convert the data to different formats, that ease rhyme pattern inspection. Our stanza 2 from Ode 109 of the Shī- jīng, for example, can be rendered directly in the following tabular form, that greatly facilitates seeing the rhyme structure of the poem.

ID	STANZA	LINE	R:467	R:468
1733	109.2	園有棘	kiak	
1734	109.2	其實之食	djiak	
1735	109.2	心之憂矣		
1736	109.2	聊以行國	kuak	
1737	109.2	不我知者		
1738	109.2	謂我士也罔極	qiak	
1739	109.2	彼人是哉		tzə
1740	109.2	子曰何其		giə
1742	109.2	其誰知之		tjiə
1744	109.2	蓋亦勿思		sio

Table 3: Tabular rendering of rhymes in PoePy.

[Q] *How could the display be further enhanced?*

3.5 Examples

As a first example, consider the first stanza of Bob Dylan’s song “I want you” (from the album *Blonde on Blonde*, 1966). Here the rhyme patterns are more complex than in many other poems, but rhyming is in parts also more lax, with more imperfect rhymes, reflecting the typical style of Dylan’s poetry.

ID	ST	LINE	R:1	R:2	R:3
1	1.1	The guilty undertaker <i>sighs</i>	s - ai s		
2	1.1	The lonesome organ grinder <i>cries</i>	k r ai s		
3	1.1	The silver saxophones <i>say</i>	s - æi -		
4	1.1	I should <i>refuse_you</i>		r i f j u: s j u:	
5	1.1	The cracked bells and washed-out <i>horns</i>			h - ɔ r n s
6	1.1	Blow into my face with <i>scorn,</i>			s k ɔ r n -
7	1.1	but it’s not that way, I wasn’t <i>born</i>			b - ɔ r n -
8	1.1	<i>to lose_you</i>		- - - l u: s j u:	

Table 4: Bob Dylan’s “I want you”.

A further example is the song “Te doy una canción” by Silvio Rodríguez (from the album *Mujeres*, 1978), in which none of the three rhyme pairs which we have annotated in stanza 1.2 rhymes perfectly. One might thus assume that rhyming was generally not intended in this song, but we find a very similar pattern in stanza 1.4., and songs in which the words *tú* “you” and *luz* “light” co-occur in potential rhyming position are very frequent in Spanish songs. Our hope is, that with a growing body of datasets in this form, we may learn more about the difference between rhymes which are intended and rhymes which might occur simply by chance.

ID	ST	LINE	R:1	R:2	R:3
7	1.2	Te doy una canción si abro una <i>puerta</i>	puer ta		
8	1.2	Y de las sombras sales <i>tú</i>		tú	
9	1.2	Te doy una canción de <i>madrugada,</i>	madruga da		
10	1.2	Cuando más quiero tu <i>luz</i>		luz	
11	1.2	Te doy una canción cuando apareces			
12	1.2	El misterio del <i>amor</i>			a mor
13	1.2	Y si no apareces, no me importa:			
14	1.2	Yo te doy una <i>canción</i>			can ción

Table 5: “Te doy una canción” by Silvio Rodríguez.

References

- Baxter, W. H. (1992). *A handbook of Old Chinese phonology*. Berlin: de Gruyter.
- Baxter, W. H. and L. Sagart (2014). *Old Chinese. A new reconstruction*. Oxford: Oxford University Press.
- Eckart, K. (2012). “Resource annotations”. In: ed. by A. Clarin-D. Berlin: DWDS, 30–42.
- Gēng Zhènshēng 耿振生 (2004). *20 shiji Hànyǔ yǔyīnxué fāngfǎ lùn* 20 世纪汉语音韵学方法论 [20th century’s methods in traditional Chinese phonology]. Běijīng 北京: Běijīng Dàxué 北京大學.
- Haider, T. and J. Kuhn (2018). “Supervised rhyme detection with Siamese recurrent networks”. In: *Proceedings of Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. (Santa Fe, 08/25/2018), 81–86.
- 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.
- Ho, D.-a. (2016). “Such errors could have been avoided. Review of “Old Chinese: A new reconstruction”, by William H. Baxter and Laurent Sagart”. *Journal of Chinese Linguistics* 44.1, 175–230.
- Jakobson, R. (1960). “Why ‘Mama’ and ‘Papa?’”. In: *Perspectives in psychological theory: Essays in honor of Heinz Werner*. Ed. by B. Kaplan and S. Wapner. New York: International Universities Press, 124–134.
- Karlgren, B. (1957). “Grammata serica recensa”. *Bulletin of the Museum of Far Eastern Antiquities* 29, 1–332.
- List, J.-M. (2014). *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press.
- (2017). *Vertikale und laterale Aspekte der chinesischen Dialektgeschichte* [Vertical and lateral aspects of Chinese dialect history]. Jena: Max Planck Institute for the Science of Human History.
- (2018). *SinoPy: A Python library for quantitative tasks in Chinese historical linguistics*. Version 0.3.1. URL: <https://github.com/lingpy/sinopy>.
- List, J.-M., S. J. Greenhill, and R. D. Gray (2017a). “The potential of automatic word comparison for historical linguistics”. *PLOS ONE* 12.1, 1–18.
- List, J.-M., J. S. Pathmanathan, N. W. Hill, E. Baptiste, and P. Lopez (2017b). “Vowel purity and rhyme evidence in Old Chinese reconstruction”. *Lingua Sinica* 3.1, 1–17.
- List, J.-M., S. Greenhill, T. Tresoldi, and R. Forkel (2018). *LingPy. A Python library for quantitative tasks in historical linguistics*. URL: <http://lingpy.org>.
- Lǚ Xiǎonán 刘晓南 (2006). *Hànyǔ yīnyùn yánjiū jiàochéng* Reader in traditional Chinese phonology [汉语音韵研究教程]. Běijīng 北京: Běijīng Dàxué 北京大學.
- Lǚ Shèngnán 吕胜男 (2009). “A brief study of the methodology of the study of ancient rhyme. And Concurrently on the study of the rhyme of “Jinwen Shangshu” 古韵研究方法论发微. 兼论今文《尚书》用韵研究 [History of ancient Chinese linguistics]”. *Nányáng Shītān Dàxué Bào (Shèhuì Kēxué Bǎn)* 南阳师范学院学报(社会科学版) [Journal of Nanyang Normal University (Social Sciences)] 8.2, 57–61.
- Pān Wùyún 潘悟云 (2000). *Hànyǔ lìshǐ yīnyùnxié* 汉语历史音韵学 [Chinese historical phonology]. Shànghǎi 上海: Shànghǎi Jiàoyù 上海教育.
- Wáng Lǐ 王力 (2006). *Hànyǔ shǐgǎo* 汉语史稿 [History of the Chinese language]. Repr. Běijīng 北京: Zhōnghuá Shūjú 中华书局.