Description, Comparison and Typology of Languages: An Interview with Martin Haspelmath

Martin Haspelmath[1]

ReVEL - Diachronic linguistics and language typology are fields of study that have a long history. How do you see the evolution of these fields and what is their future?

MARTIN HASPELMATH - Actually, serious diachronic linguistics is much older than typology, going back to the early 19th century, with breakthroughs such as Bopp's comparative Indo-European studies, and Rask's and Grimm's comparative Germanic studies. Historical-diachronic linguistics has thus seen exciting discoveries since the 1820s, and it was a fairly mature field by the 1880s. By contrast, a sustained tradition of typological studies has existed only since the 1960s – Greenberg's 1963 paper was a milestone. For serious cross-linguistic studies, one needs to have access to good descriptions of languages from around the world, and few linguists had the means or the interests to pursue this kind of work in the 19th century and in the first half of the 20th century, when national or nationalist ideas were at the forefront of public interest in many countries. After the Second World War, more and more people took a global perspective, and more and more countries in Africa and Asia became independent, full-fledged members of the international community. It therefore seems to me that the interest in worldwide comparison is also a consequence of the decolonization of the 1960s and 1970s: Increasingly, the languages of Asia and Africa were not just of interest for Christian missionaries and a few anthropologists, but were seen as a key part of the cultural wealth of these countries. Westerners increasingly came to accept that their countries were not at the top of some hierarchy, but just some world areas with a specific shared history.

[1] Department of Linguistic and Cultural Evolution - Max Planck Institute for Evolutionary Anthropology

Moreover, I think that as universal values (such as human rights) became more and more prominent, scholars also got more interested in what is universal across languages, in an empirical sense. It is probably not an accident that Greenberg's (1963) foundational work on typology and Chomsky's (1965) early work on universal grammar appeared around the same time. This also had to do with the shift from Europe to the United States after the war: In the 19th century, historical linguistics was typically seen as a way of studying one's nation's past, but North Americans were very diverse in their ethnic origins and did not have a long history in their own countries; by contrast, U.S. linguists had access to a wide range of indigenous languages, as well as diverse immigrant languages. So this favoured a broader, universal perspective on language and languages.

It is impossible to predict the future, but let me point out a few trends of recent years. First, diachronic linguistics has become more computational and quantitative, and it has increasingly fine-grained findings, but there do not seem to have been any major breakthroughs in understanding. There is now a general recognition that language change can only be understood in the context of social variation within a community, but there is an extremely wide variety of factors that have been associated with language change. *De facto*, we do not understand language change very well, and it is perhaps best seen as largely random, like other kinds of social and cultural changes. Second, linguistic typology has increasingly focused on diversity, and some of the hopes of the 1970s and 1980s have not been fulfilled: Chomskyan linguists have given up the hope that there are a few dozen parameters that explain syntactic variation across languages, and Greenbergian typologists have increasingly focused on large-scale geographical patterns that sometimes obscure universal trends. So typology has not had big breakthroughs either, at least of the sort that the optimism of earlier decades may have suggested. Our research is getting more wide-ranging and more sophisticated, but it's very difficult to achieve greater depth of understanding.

ReVEL - From your perspective, what are some of the prominent past and recent advancements by diachronic linguistics or language typology for the understanding of grammar and human language?

MARTIN HASPELMATH - After the somewhat pessimistic remarks that I just made, let me focus on two aspects of comparative grammar research where I see true

progress: Word order universals and universals of asymmetric coding. Greenberg's word order universals (first highlighted in his 1963 paper) inspired many linguists in the 1970s and 1980s to seek grand explanatory proposals, but parameter-setting models and simple head-dependent approaches did not work well. In my view, the breakthrough came with Dryer's (1992) paper and Hawkins's (1994) book, where it was proposed that the kinds of word order patterns that are found most frequently are also those that exhibit the greatest efficiency of processing. When a verb-object language also has prepositions and postposed possessors, this means that dependency length is minimized and constituents can be recognized more easily. The recent paper by Futrell, Levy and Gibson (2020) adduces a lot of evidence for this view.

The topic of asymmetric coding is less well known as a general topic of typology, and many linguists treat it under the heading of "markedness": Plural is considered "marked" in contrast to singular, future tense is "marked" in contrast to present tense, indefinite subjects are "marked" in contrast to "unmarked" definite subjects, and so on. But the "markedness" idea has never been made precise, and it has turned out that an explanation in terms of efficient coding is possible here as well: The "unmarked" cases are invariably the more frequent and thus more predictable ones, so it makes good functional sense to code them by zero or by a shorter form. I have summarized some of the evidence for this in my 2021a paper.

So I think that many regularities of order and coding can be explained by communicative efficiency, and I see this as true progress over the last few decades. We also know that these functional explanations presuppose a view of languages as culturally evolving systems, analogous to biological systems with their functional adaptations. But cultural evolution and adaptation is not as well understood as biological evolution yet, and I see this as a major challenge for the future.

ReVEL - One of your most well-known works is the *World Atlas of Language Structures* (WALS), an essential resource for linguistic research, as it provides data on the geographical distribution of grammatical structures for a large sample of the world's languages. Recently, you have also been involved in Glottobank, as a senior advisor in the Grambank project. Can you tell us a little bit about these projects?

MARTIN HASPELMATH - While the *World Atlas of Language Structures* (2005) was created as part of Bernard Comrie's department at our Max Planck Institute, *Grambank* (2023) is a successor project created under Russell Gray's leadership (Gray basically came to replace Comrie in 2015 as director of a Max Planck department studying worldwide comparative linguistics). While *WALS* was based on datasets that were created by different scholars in different institutions, originally for different research questions, and our task consisted in collecting them and making them accessible in a uniform way, the *Grambank* project is a much more top-down enterprise: It is based on a set of 195 grammatical questions, and the data were supplied by linguists (mostly students or postdocs) who were employed specifically to gather the data. So there is a difference in the methodology, but I see the two projects as complementary. Grambank has a lot more data, but the descriptions of the features are richer and deeper in WALS. Both kinds of approaches will remain important in the future, I think. It will not be easy to get as much funding for an equally ambitious successor project to Grambank, but for the time being, there seems to be some funding to continue enlarging Grambank.

One issue is what happens to such online databases when the funding ends (e.g. when a leading researcher retires), or when current web environment changes. Many valuable web resources have simply disappeared over the last two decades. Fortunately, our department (led by Robert Forkel) has created a data standard for such cross-linguistic data (called CLDF: Cross-Linguistic Data Formats), which should make it easier to work with the underlying data. If the web applications disappear at some point, the data should still be available and usable.

ReVEL - Some of your latest studies have been addressing the concept of general linguistics. For example, in the article "General linguistics must be based on universals (or non-conventional aspects of language)", you claim that "one must study universals if one wants to make general claims". What is your conception of general linguistics? How does one study general linguistics nowadays?

MARTIN HASPELMATH - General linguistics simply means the study of Human Language in general, and it has not changed since Saussure's time (whose famous 1916 book was called *Cours de linguistique générale*). But since the 1960s, the term

"theoretical linguistics" has been used more frequently, at least in English-speaking contexts. However, the two are not synonymous: Theoretical linguistics is opposed to applied linguistics (the study of language(s) for practical purposes), and it need not be general – in fact, there is a lot of theoretical particular linguistics. For example, a study of French verbal inflection is a contribution to French theoretical linguistics, but not necessarily to general linguistics. I think that general linguistics needs to be based on broadly comparative studies, because otherwise we cannot distinguish well between what is general across languages and what is particular to an individual language. Why is it necessary to emphasize this? Because many linguists have made far-reaching proposals about universal categories or architectures on the basis of an individual language – in the past, often on the basis of English, but more recently, also on the basis of small minoritized languages (as emphasized, for example, by Andrew Nevins in his 2022 book). I think that these proposals are often speculative and premature.

ReVEL - We will finish this interview by kindly asking you what bibliographic references you recommend for someone if they are interested in studying the fields of language description, diachronic linguistics, and language typology.

MARTIN HASPELMATH - For language description, it is probably best to look at specific high-quality descriptions of particular languages, such as those that have appeared in Language Science Press's book series *Comprehensive Grammar Library*. But overview books such as Aikhenvald (2015) may be useful as well. And for greater depth, handbooks such as Shopen (ed.) 2007 remain essential.

For language change, I don't have any particular recommendations for beginners – there are many textbooks, and the most ambitious book (Croft's 2000 "Explaining language change") is not easy to read. However, for philosophically minded readers, I can recommend the old (1994) book "On language change" by Rudi Keller, which is written in a very accessible style. It provides an interesting historical perspective and contains some key insights about language change as cultural evolution.

Finally, for linguistic typology, Moravcsik (2013) and (for more advanced readers) Song (2018) are the best introductory texts. But I also recommend browsing

the maps of *WALS* and *Grambank*, because they give one very good impressions of worldwide cross-linguistic variation.

References

AIKHENVALD, Alexandra Y. *The art of grammar:* A practical guide. Oxford: Oxford University Press, 2015.

CHOMSKY, Noam A. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press, 1965.

CROFT, William. *Explaining language change:* An evolutionary approach. Harlow: Longman, 2000.

DRYER, Matthew S. The Greenbergian word order correlations. *Language*, v. 68, n. 1, p. 81–138, 1992.

FUTRELL, Richard; LEVY, Roger P.; GIBSON, Edward. Dependency locality as an explanatory principle for word order. *Language,* v. 96, n. 2, p. 371–412, 2020. Disponível em: https://doi.org/10.1353/lan.2020.0024.

GREENBERG, Joseph H. Some universals of grammar with particular reference to the order of meaningful elements. *In:* GREENBERG, Joseph H. (ed.). *Universals of language*. Cambridge, MA: MIT Press, 1963. p. 73–113.

HASPELMATH, Martin. General linguistics must be based on universals (or nonconventional aspects of language). *Theoretical Linguistics,* v. 47, n. 1–2, p. 1–31, 2021a. Disponível em: https://doi.org/10.1515/tl-2021-2002.

HASPELMATH, Martin. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics,* v. 57, n. 3, p. 605–633, 2021b. Disponível em: https://doi.org/10.1017/S0022226720000535.

HAWKINS, John A. *A performance theory of order and constituency*. Cambridge: Cambridge University Press, 1994.

KELLER, Rudi. *On language change:* The invisible hand in language. London: Routledge, 1994.

MORAVCSIK, Edith A. *Introducing language typology*. Cambridge: Cambridge University Press, 2013.

SHOPEN, Timothy (ed.). *Language typology and syntactic description* (3 volumes). Cambridge: Cambridge University Press, 2007.

SONG, Jae Jung. *Linguistic typology*. Oxford: Oxford University Press, 2018.