

Martin Haspelmath

# The challenge of making language description and comparison mutually beneficial

DOI 10.1515/lingty-2016-0008

Received March 20, 2016

**Abstract:** I argue that the distinction between comparative concepts and descriptive categories helps language describers and typologists to benefit from each other because describers are free to set up their own categories, typologists are free to define their own concepts, comparison need not involve complete systems, and interlinear translation can be either based on comparative concepts or descriptive categories. A similar distinction also exists in other disciplines that deal with cultural concepts.

**Keywords:** categories, comparison, methodology

There is one type of critical comment on the proposal of conceptually separating descriptive categories for analysis of particular languages and comparative concepts for language typology (Haspelmath 2010) that I hear particularly often: the observation that it was precisely the integration of comparative and descriptive endeavors that has led to great progress in both fields over the last few decades. As Volker Gast noted: “Paradoxically, typology has been rather successful” (Lingtyp discussion, 19 January 2016, see Supplementary Online Materials). It seems that the publication of the Shopen volumes (Shopen (ed.) 1985, 2007) was the most important milestone in this development, and this kind of typological overview work does not make an effort to emphasize that it is talking about comparative concepts. On the contrary, it typically presents the distinctions it draws as potentially relevant for any fieldworker. So why am I saying that there is a need for a somewhat novel conceptual distinction?

I would say that the two enterprises of description and comparison need to be separated conceptually, but integrated in practice. Of course, comparative linguists should take a very close look at the languages they compare, and describers should continue to pay attention to what is happening in typology and be inspired by it, or even do it themselves (perhaps on a smaller scale, looking at various types found in their world region). Description and

---

**Martin Haspelmath**, Max-Planck-Institut für Menschheitsgeschichte, Kahlaische Strasse 10, 07745 Jena, Germany, E-mail: haspelmath@shh.mpg.de

comparison should by all means be mutually beneficial, and I would argue that the conceptual separation, rather than being an impediment, actually contributes to this goal in a number of ways.

First, the realization that typologists' comparative concepts are not discoveries but instruments for discoveries (unlike chemical elements, which are true discoveries about what can exist in nature) is liberating for language describers, because they can set up completely new categories, with completely new labels, without having to worry whether what they are positing has been found before in other languages. Of course, often we find phenomena that are similar to previous findings by our colleagues, but there are also genuine novelties (such as verbs for 'and', Brown & Dryer 2008). And categories that look familiar may show some strange features. Thus, when we read about a Jussive mood or a Perrelative case in some language, the transparent labels are helpful for getting an initial idea about the category, but the capitalization reminds us that the nature of these language-particular categories is not exhausted by the labels. And when a label is used with several different senses (e.g., "conative"; Vincent 2013), we need not be worried. This is merely a practical problem (alleviated by the use of capital letters for language-particular categories), not a theoretical problem.

Second, the realization that comparative concepts are prerequisites for comparative results, not research results themselves, relieves the typologist of the burden of having to hit upon the "right" concepts. In the generative approach, typological comparison is necessarily based on the universal categories of UG, so a generative linguist works with three unknown elements simultaneously: the unknown universal categories, the unknown language-specific analyses, and the unknown cross-linguistic generalizations. This means that a lot of intuition-based assumptions (sometimes about all three) have to be made, and the empirical research results cannot be as objective as they are if the comparative concepts are rigorously defined. Zwicky (1985) described how generative analyses are more like doctors' diagnoses, and in Haspelmath (2015) I point out that definition-based studies are more objective than necessarily subjective diagnoses. It is true that intuition also plays a role in the choice of comparative concepts (Lazard 2005), but not in their definition and in the testing of hypotheses based on them.

Third, if we understand that our category-based comparative concepts (such as "ergative case" or "serial verb construction", cf. Haspelmath (2016) for the latter) are not different in principle from our grammar-external comparative concepts (such as nonverbal stimuli in fieldwork experiments, or text passages in parallel texts, e.g., Wälchli & Cysouw (2012)), it will be easier to accept that language comparison is necessarily partial. Just as nonverbal stimuli and parallel text passages can never contain all possible concepts or

situations, it will never be possible to compare languages with respect to all their features, and many features that are interesting to compare in some languages are necessarily uninteresting in many others. By contrast, description must aspire to be complete. Thus, the scopes of comparison and description overlap only partially, and no effort needs to be made to make them completely congruent. In the Lingtyp discussion, William Croft proposed a comparative definition of relative clauses that “limit[s] the predicate head of the modifying proposition to action concepts” (18 January 2016). This yields a semantic definition that does not include English attributive adjectives, but that would also exclude English relative clauses like *the woman who knows the answer*, thus introducing another incongruence. For typological purposes, this does not matter, because English relative clauses with non-adjectival non-action predicates (like *know*) have no special typological properties, so nothing is lost. More generally, Nichols & Bickel (2005) advocate “exemplar-based” definitions of comparative concepts, explicitly acknowledging that typology can often make do with a small characteristic property, without requiring the complete facts.

Fourth, the distinction between descriptive categories and comparative concepts makes us realize that helping readers of example sentences by providing interlinear glosses can be done in two different ways: by providing renderings that approximate the language-particular classes, or by providing renderings that view the example from a comparative perspective. What is called “Augmented” or “Aorist” from a language-specific perspective (and glossed AUG and AOR in a grammar) may well be called “plural” or “past tense” (glossed PL and PST) in a typological context that glosses over the subtle distinctions between these notions and employs comparative concepts. Both of these glossings are correct, and one can choose between them (or even mix them) opportunistically. Interlinear glosses are not abbreviations of deep analyses, but reading aids to the reader.

Language describers are sometimes frustrated with typologists because these may appear to use strange concepts (such as “subject” for Chinese word order, where it has no place, cf. LaPolla & Poa (2006)), but if it is recognized that the typologists’ concepts are separate from the descriptive categories, this frustration should disappear. Conversely, typologists are often frustrated with describers because their grammars do not answer the questions of the typologists in a neat way and use strange terminology. But grammar writers often have good reasons for choosing special terms and for describing a phenomenon in an unexpected place (e.g., describing attributive property words in the chapter on relative clauses, because they do not behave much differently). Of course, if grammar writers also have typologist readers in mind and additionally include

references in appropriate places, this is even better, but it will never be possible to anticipate all questions that a typologist might have.

The distinction between comparative concepts and descriptive categories may be less prominent in other disciplines, but at least in structuralist anthropology, the idea that each society needs to be described in its own terms has been important since Boas's times, and Pike's distinction between emic (culture-specific) and etic (outsider) concepts is apparently better known in anthropology than in linguistics. Clearly, a special set of comparative concepts must be used in comparative anthropology (e.g., in studies of political organization such as Currie et al. (2010), where bands, tribes, chiefdoms, and states are distinguished, based on much earlier ethnographic work). Why does the distinction seem controversial in our field? Perhaps the grammatical concepts of linguistics have had more influence on typologists simply because the grammatical systems are so compelling (thus, nobody could doubt the wisdom of describing Russian by means of language-specific categories such as Dative case or Perfective aspect). Grammatical terminology has long been carried over from one language to the next (Latin to German, English to Japanese, and so on), with smaller or greater adjustments, so it has long seemed to linguists that they are using the same categories that are used in description also for comparison. But upon further reflection, when taking the perspective of a larger number of languages, it becomes clear quickly that language-particular notions do not work for comparison. Greenberg (1963) knew well that "we are essentially using semantic notions" (rather than language-particular syntactic categories) when comparing the orders of major clause constituents, and likewise Swadesh's list of 100 meanings could not possibly refer to meanings that are equally relevant for comparison (Swadesh's goal) and description.

**Acknowledgements:** I am grateful to Susanne Maria Michaelis and Edith Moravcsik for discussion and for comments on this paper.

## References

- Brown, Lea & Matthew S. Dryer. 2008. The verbs for 'and' in Walman, a Torricelli language of Papua New Guinea. *Language* 84. 528–565.
- Currie, Thomas E., Simon J. Greenhill, Russell D. Gray, Toshikazu Hasegawa & Ruth Mace. 2010. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* 467 (7317). 801–804.

- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86. 663–687.
- Haspelmath, Martin. 2015. Defining vs. diagnosing linguistics categories: A case study of clitic phenomena. In Joanna Błaszczak, Dorota Klimek-Jankowska & Krzysztof Migdalski (eds.), *How categorical are categories? New approaches to the old questions of noun, verb and adjective*, 273–303. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics* 17. 291–319.
- LaPolla, Randy J. & Dory Poa. 2006. On describing word order. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 269–295. Berlin: Mouton de Gruyter.
- Lazard, Gilbert. 2005. What are we typologists doing? In Zygmunt Frajzyngier, Adam Hodges & David S. Rood (eds.), *Linguistic diversity and language theories*, 1–23. Amsterdam: Benjamins.
- Nichols, Johanna & Balthasar Bickel. 2005. Locus of marking: Whole-language typology. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press. <http://wals.info/chapter/25>
- Shopen, Timothy (ed.). 1985. *Language typology and syntactic description*. 3 vols. Cambridge: Cambridge University Press.
- Shopen, Timothy (ed.). 2007. *Language typology and syntactic description*. 2nd edn. 3 vols. Cambridge: Cambridge University Press.
- Vincent, Nigel. 2013. Conative. *Linguistic Typology* 17. 269–289.
- Wälchli, Bernhard & Michael Cysouw. 2012. Lexical typology through similarity semantics: Toward a semantic map of motion verbs. *Linguistics* 50. 671–710.
- Zwicky, Arnold M. 1985. Clitics and particles. *Language* 61. 283–305.