

## The interplay between comparative concepts and descriptive categories (Reply to Newmeyer)

MARTIN HASPELMATH

In his reply to my paper 'Comparative concepts and descriptive categories in cross-linguistic studies' (Haspelmath 2010, henceforth H10), Newmeyer (henceforth N) defends the widespread old idea that there are cross-linguistic categories and that the same kinds of concepts are and should be used both in cross-linguistic studies and in language-particular description. However, his critique does not amount to a full-blown endorsement of the categorial universalist position, which in its prototypical generative form assumes a single set of categories or features for both purposes. For instance, when N says that 'it is only by means of working out the interplay between the language-particular and the language-independent that we can hope to understand either' (p. 2-3), he implies that the two are not identical, and that the relationship between them is non-trivial, much as I have emphasized.

The bulk of N's reply is devoted to listing various problems in H10's examples that he sees as undermining its general claims, especially with respect to the notions of subject, thematic roles, adjectives, clause, and word. In this brief reply, I will try to show that none of them are real, and that a closer consideration of these cases actually strengthens my position.

N first observes that despite my criticism of the equation of language-particular 'subject' categories across languages, I do use 'subject' as a comparative concept in some of the generalizations and definitions (in particular, in (3) and (13)). I did not give the relevant definition in H10 for lack of space, hoping that it could be inferred, but here it is: 'the agent of a simple transitive clause' (the same definition is assumed by Greenberg 1963). This is of course a very different concept from 'Subject in Tagalog' (defined in terms of case-marking by *ang*, very often different from the agent NP) and from 'Subject in English' (perhaps defined as in N's (6), comprising expletive *there*, which surely is not an agent). Thus, this underscores rather than undermines my central point. N's own concrete proposal for defining 'subject' ('[an entity that manifests] more of the core properties of subjects than of constructs that contrast with subjects', p. 12) is also a possible comparative concept, provided that the 'core properties' can be defined in a universally applicable way. However, as I noted in H10 (§7.3) in the discussion of Keenan's (1976) definition of 'subject', it is unclear how one can draw up such a list of core properties on a principled basis. And most crucially, it is unlikely that such a definition will be found useful in language-particular studies. Descriptive categories are normally defined in terms of specific conditions, not in vague quantitative terms ('more properties than').

N also criticizes me for using semantic roles such as 'recipient' (in 2) and 'agent' (in 16) (even though he uses the latter concept himself in (6)). But he overlooks that the problems with such semantic roles have never arisen in a typological context. Nobody has ever seen problems with defining the comparative concept 'ergative' in terms of 'agent'. It is only when one tries to describe particular languages with a single universal set of semantic roles that one runs into problems. But we can describe *I play the sonata* in purely grammatical terms, without reference to general semantic roles.<sup>1</sup> Generative

---

<sup>1</sup> The price is that it needs to be stipulated that in English, *\*The sonata plays me* is not possible, but at least it works, and every grammar needs a massive amount of stipulation anyway.

linguists try to use as many cross-linguistic categories in the description of individual languages as possible, and this often leads to insurmountable problems. Typology is happy to limit its generalizations to clear cases of agents, patients and recipients. Admittedly, we cannot formulate universals over sentences like *I play a sonata*, but typology does not strive for exhaustiveness of coverage. Comparative linguists can thus simply leave them aside, whereas descriptivists HAVE TO find a category for them, and this is difficult if one only wants to use cross-linguistic categories.

Let us look at adjectives next. N is of course right that 'it makes all the difference in the world for generalizations about Adjective-Noun order' how 'adjective' is defined (p. 7). Greenberg and Dryer chose them in a way that abstracts from their verb-like vs. non-verb-like behavior, but it is of course conceivable that more interesting generalizations about adjective ordering could be found with a different definition (see Dryer 1988:197-198). Note that N has not shown that 'a purely semantic definition is inadequate' (p. 7), merely that it is not the only possibility (of course not). N suggests that a comparative definition of 'adjective' that is more similar to language-particular definitions is superior, but his own proposal does not work: Whether property-concept words are 'a separate word class' or not cannot be objectively determined. There are many languages where property-concept words are described as a subclass of verbs by some linguists, but as a separate word class by other linguists. But there is no generally accepted set criteria that could distinguish between subclasses and separate classes. The two descriptions are thus notational variants, and 'separate word class' status cannot be used in a definition of a comparative concept 'adjective'. One could, of course, employ more specific criteria, e.g. the need for special relative-clause marking in attributive use, the need for a special copula in predicative use, etc. But all such comparative definitions of 'adjective' would be different from language-particular descriptive categories, underscoring again my central point.

Another concept that I left undefined for reasons of space is 'clause'. This might be defined as 'an expression that contains one predicate and potentially at least some of its arguments and that can be independently negated'. The main issue in defining 'clause' cross-linguistically is how to treat 'complex predicates' consisting of two verbs such as serial verb constructions and structures with certain auxiliary-like or light verbs ('clause union', 'restructuring'). It seems to me that from a cross-linguistic point of view, the possibility of negation corresponds best to our intuition about what should count as a clause. In particular languages, there may of course be quite different criteria for defining something like the clause, but these are not helpful at the comparative level. N claims that Givón's (1980) gradient or prototype-based generalizations could not be captured in my system, but the opposite is true. In order to express universal claims in an explicit and readily testable way, one needs discrete comparative concepts of the sort that I exemplify at length in H10. Appeal to gradience or prototypicality often has a useful heuristic role, and Givón's work has been highly stimulating, but in the end, the big picture has to be dissolved into fine-grained discrete comparative concepts.

The problems with defining the 'word' notion that N mentions are well-known, and I have highlighted them recently in my own work (Haspelmath 2011). My tentative definition of 'word' in H10 is an example of a conceivable comparative concept that I would not endorse myself, so N's specific point is well-taken. But my conclusion from these problems is different from N's: I think that all comparative concepts that incorporate the notion 'word' (or

'morphological') are suspect, and should probably be replaced.<sup>2</sup> More generally, I do not claim that all the comparative concepts used in H10 are unproblematic or optimal. What I do claim is that they must be distinguished from descriptive categories, and that the generalizations based on them are only as good as the definitions.

In addition to the problems he sees with my examples, N makes one important general point: He observes that in generative linguistics, categories 'do not admit to definition outside of the formal system in which they partake' (p. 9). This attitude accounts for general lack of concern for defining categories in generative work. For example, I noted in Haspelmath (2011) that linguistics textbooks in the earlier part of the 20th century were often at pains to provide a careful definition of the 'word', whereas since the 1970s, mainstream textbooks have invariably ignored the issue, simply assuming that a 'word' notion has a place in the theory. Similarly, syntactic categories such as 'adjective' or 'anaphor' are not defined in generative work, and instead specific analyses are proposed that make use of these notions, as part of a 'web of interconnecting assumptions that constitute theories'. This is a possible approach for language-particular analyses (i.e. theories of the mental grammar of speakers of a language), at least if the analyses are complete (which they rarely are). But for cross-linguistic studies, it inevitably leads to circularity. Di Sciullo & Williams (1987) are free to use an undefined 'word' concept for their analysis/theory of English, but when they go on to make claims about languages in general (e.g. that in general, 'words' and 'phrases' are different kinds of objects), these claims are untestable (at least in practice), because there are no criteria for telling whether an element in another language is a 'word' in the same sense as 'word' in the grammar of English. Similarly, Chomsky's (1981: ch. 3) binding theory makes good sense as a theory of some English anaphoric expressions ('pronouns' such as *she*, and 'anaphors' such as *himself*), but since the notions of 'pronoun' and 'anaphor' are not defined in general terms, it is impossible to test the claim that this binding theory is not just true for English, but for languages in general (see Haspelmath 2008, n. 5).

N says that I 'steadfastly reject' the idea that we have to 'work out the interplay between the language-particular and the language-independent' (p. 2-3), but I just see this interplay differently. In my view, language description and typology can in principle be done independently of each other, but in practice, they are of course profiting enormously from each other. As Dryer (2006:210) notes, 'typological work had a tremendous impact on descriptive work' over the last few decades, and typology is relying on better and better descriptions becoming available. Thus, prospects for a fruitful interplay in the future are excellent, and neither descriptivists nor typologists have to content themselves with a secondary role in this common enterprise of understanding human languages.

## References

- Chomsky, Noam A. 1981. *Lectures on government and binding*. Dordrecht: Foris.  
Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.  
Dryer, Matthew S. 1988. Object-verb order and adjective-noun order: dispelling a myth. *Lingua* 74(2-3). 185–217.

---

<sup>2</sup> As a result, in my own typological work on 'cases/adpositions', I prefer the more general term 'flag', which is neutral with respect to the affix/word distinction (e.g. Haspelmath 2005).

- Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and basic linguistic theory. In Felix K. Ameka, Alan Dench, & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 207–234. Berlin: Mouton de Gruyter.
- Givón, T. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4. 333-378.
- 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. 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3(1). 1–21.
- Haspelmath, Martin. 2008. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40-63.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(4).
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. To appear in *Folia Linguistica*.
- Keenan, Edward L. 1976. Towards a universal definition of 'subject'. In Charles N. Li (ed.), *Subject and topic*, vol. 305, 303-333. New York: Academic Press.