Yury Lander* and Peter Arkadiev
On the right of being a comparative concept

DOI 10.1515/lingty-2016-0014
Received March 22, 2016; revised July 2, 2016

Abstract: We provide a critical review of the distinction between “comparative concepts” and “descriptive categories”, showing that in current typological practice the former are usually dependent on the latter and are often vague, being organized around prototypes rather than having sharp boundaries. We also propose a classification of comparative concepts, arguing that their definitions can be based on similarities between languages or on differences between languages or can also be “blind” to language-particular facts. We conclude that, first, comparative concepts and descriptive categories are ontologically not as distinct as some typologists would like to have it, and, second, that attempts at a “non-aprioristic” approach to linguistic description and language typology are more of an illusion than reality or even a desideratum.

Keywords: categories, comparison, methodology, relative clause

1 Introduction

Haspelmath (2010) argued for the importance of distinguishing between descriptive categories, i.e., categories which are relevant for grammars of specific languages and are defined on language-internal basis, and comparative concepts, which he described as a kind of abstract categories “invented” by typologists in order to compare languages with each other. The validity of this distinction is beyond doubt. Indeed, there is no need to assume that all properties of language-particular categories are identical to the properties of more abstract patterns studied by typologists: for example, it is by no means the case that the English Saxon Genitive construction always has the properties which are usually associated with the typological concept of adnominal possessive construction, or that the Russian...
Perfective ideally fits the crosslinguistic concept of perfective aspect as formulated, e.g., by Dahl (1985: 69–84).

The nature of comparative concepts is somewhat controversial, though. Are they really only parts of conceptual systems which exist independently of descriptive categories? (see van der Auwera & Sahoo (2015) for strong arguments against such a view.) Should they be defined as precisely as possible, like measurements in quantitative studies? These issues are addressed in the first part of this paper (Sections 2 and 3). In the second part (Section 4), we propose a simple typology of comparative concepts based on their relations with the actual linguistic data. Section 5 summarizes our conclusions.

For the sake of exposition, we will use a limited number of categories (primarily non-restrictive relative clauses, but not only them), but this is due to the reasons of space. We hope to present more elaborated case studies in the future.

2 Comparative concepts need not be independent of descriptive categories

Haspelmath (2010: 680) claims that “[c]omparative concepts are motivated and defined in a way that is quite independent of linguistic categories (though of course not independent of the facts of languages)”. It seems to us that the parenthesized part of this statement is sometimes underestimated and even contradicts the idea of independent motivation. Most comparative concepts are by no means arbitrary. They are usually intended to provide a perspective on the phenomena that are already known and therefore owe their existence to concrete data. If comparative concepts are not felt to be relevant for the grammars of different languages, they are usually not viable. For example, Raible (2001: 7) recalls Dante Alighieri, “who, at the beginning of the 14th century, in his *de vulgari eloquentia*, classified the Romance languages according to their expressions for *YES* into *(h)oc, oïl, and si(c) languages*, yet the comparative concept “the phonetic content of *YES*” does not seem to be relevant for linguistic typology and hence is not taken seriously anymore.¹

As an example of a workable concept, take non-restrictive relative clause constructions, where a relative clause “makes a comment about an NP or other constituent, without delimiting its reference” (Andrews 2007: 207).

¹ Cf., however, the accepted names of Bosnian-Croatian-Serbian dialects named after the form of the interrogative pronoun ‘what’: Štokavian, Kajkavian, and Čakavian (we thank Michael Betsch for reminding us). It is certainly not a coincidence that frequent and salient words such as ‘yes’ or ‘what’ may serve as bases for language names.
Such constructions are contrasted with restrictive relative clause constructions, where the relative clause restricts the reference of the noun phrase – as in (1b) vs. (1a) (both examples from Quirk et al. 1985: 1258):

(1) a. Then he met Mary, who invited him to the party.
   b. That’s the girl (that) he met at the party.

The concept of non-restrictive relative clause is defined primarily on a semantic basis, yet it has formal manifestations. In particular, there are languages where there are marking contrasts between restrictive and non-restrictive relative clauses. For instance, in Hausa non-restrictives differ from restrictives in the tone pattern of the relative pronoun and the range of inflectional person-aspect forms (Jaggar 1998, 2001: 536–538). In English relative constructions with *that* or lacking a complementizer, illustrated in (1b), are not normally used non-restrictively (Quirk et al. 1985: 1258). Also, independently of formal marking, non-restrictive relative clauses are usually more “external” within noun phrases than restrictive clauses, as seen in (2) (from Hawkins 2004: 240):

(2) a. Students that major in mathematics, who must of course work hard, ...
   b. *Students, who must of course work hard, that major in mathematics ...

It is unlikely that the concept of non-restrictive relative clause would have been used actively if it had no formal manifestation in grammars.\(^\text{2}\)

Note, however, that for some concepts, the feeling of relevance for grammars may be due to particular theoretical frameworks (e.g., the notions “control” and “raising”, cf. Davies & Dubinsky 2004, Serdobolskaya 2009). In fact, even the classification of Romance languages on the basis of the expressions for ‘yes’ made sense for its author because of his “desire to vindicate the claims of the *lingua di sì*, especially against the exaggerated and monopolistic pretensions of the *lingua d’oco*” (Ewert 1940: 356), and it may well be that some of the modern comparative concepts may serve similar purposes. Even so, we believe that the main

\(^{2}\) It is sometimes thought that the universal usefulness of such concepts as non-restrictive relative clause is based on their semantic definition. This matter does not seem that obvious to us (see also Stassen (2010) for some discussion). In particular, it is possible to find comparative concepts which have no semantic basis and yet are crosslinguistically relevant. For example, Nichols (1988, 1992) argued that adnominal inalienable possessive constructions do not have strict semantic content and are just archaic possessive constructions which remained intact in contexts requiring the expression of the possessor. Yet for Nichols inalienable possessives definitely serve as a comparative concept.
requirement for comparative concepts is that they should be useful for cross-
linguistic generalizations, theory-dependent or other.

We conclude that comparative concepts should be allowed to be based on
descriptive categories and the latter should be allowed to be thought of as mani-
festations of comparative concepts. (More on this point below, as well as in van der
Auwera & Sahoo (2015)). Any comparison requires that we find the objects com-
pared similar to some extent, belonging to the same kind – and hence abstract
away from some of their specific properties. For sure, this leads to some “desensi-
tization”, but that is a sacrifice which is a prerequisite for comparison.

3 Comparative concepts may be vague

If we allow some comparative concepts to reflect language-particular descriptive
categories, we may want these concepts to also reflect the fact that descriptive
categories are often vague. Natural languages display much variation, so that not
only speech communities but also individual speakers themselves may be inconsis-
tent in using linguistic patterns. However, this variation usually does not arise
in the most frequent contexts, and the fact that the probability of variation is larger
in certain contexts of use is just as typologically relevant as the fact that something
is consistent throughout all or the majority of languages. The main problem with
non-vague comparative concepts is that they are not able to capture the “consis-
tent inconsistency” of some contexts and the “consistent consistency” of others.

Returning to non-restrictive relatives, we observe that some kinds of non-
restrictive modification are more easily covered by constructions which are typi-
cally used restrictively than others. For example, at first glance non-restrictive
relatives do not seem to be allowed in West Circassian (also known as Adyghe), a
polysynthetic language of the Northwest Caucasian family. Here it is strictly
impossible to form a relative clause construction with a personal pronoun as the
head. When eliciting examples with proper names as heads, Circassian speakers
emphasize that there are other discourse referents with the same name and hence
betray the restrictiveness of such constructions. Nonetheless, our corpus of West
Circassian texts contains occasional occurrences of relative clauses used non-
restrictively, as in (3), where the internally-headed relative clause just adds
information on the epic hero Sosruko’s horse named Tkhozhiye:

3 West Circassian is not unique in this respect: a list of languages lacking non-restrictive
relatives can be found in Cinque (2013: 202).
They saw Sosruko's Tkhozhiye, (riding) on whom he was raising the dust.

Further, there exist contexts where non-restrictive relatives are not clearly differentiated from restrictive ones. For example, Wierzbicka (2002: 291), while arguing against a strict contrast between restrictive and non-restrictive clauses, states that "in the case of indefinite heads the distinction is hard to sustain at all". It seems, therefore, that non-restrictives can be contrasted with restrictive modification to different degrees. For example, non-restrictives headed by personal pronouns are more strongly contrasted than non-restrictives headed by proper nouns. The idea of comparative concepts should not turn away from this. For the general concept of non-restrictive relative clauses, it is equally important that there are no problems with treating relatives modifying indefinites as non-restrictives and that there is variation in treating certain relatives modifying indefinites as non-restrictives and that there is variation in treating personal pronouns as non-restrictives. The farther from this prototype a relative is, the less it is expected to behave as non-restrictive. It is worth noting that all prototype-based comparative concepts presuppose similar prototype-based effects in unrelated languages.
Variation is related to instability and hence to diachronic processes which determine language-specific descriptive categories. Vague prototype-based comparative concepts presumably can manage with this language-specific nature. Thus, we suspect that languages may provide different grammaticalized cut-off points between restrictive and non-restrictive modification, but this will not pose any problems if restrictive and non-restrictive modification are vague comparative concepts. The same perspective may apply to some other contrasts, even the most general ones such as the contrast between morphology and syntax and the related “indeterminacy” of the notion of word (cf. Dixon & Aikhenvald 2003; Haspelmath 2011). The contrast between words and complex phrases may certainly differ in various languages and may even be fuzzy in one and the same language, but this does not invalidate comparative concepts like “word” and “phrase” (pace Haspelmath 2011). We have no doubts that “word” in West Circassian, which may incorporate much information on oblique arguments, as in (4a), is not the same as “word” in English. Yet to the extent that there is robust evidence that Circassian languages make a contrast between syntax and morphology – for example, their speakers may introduce oblique arguments not only morphologically but also by means of postpositions, as in (4b) – we may identify the contrast between words and complex phrases (and hence the contrast between morphology and syntax) in West Circassian with a similar contrast in English and other languages.

(4) a. a-š’ hantχwaps θ-f-a-be-hazərə-ʁ 
that-obl millet.soup 3sg.10-ben-3pl.erg-caus-ready-pst

b. a-š’ paje hantχwaps a-be-hazərə-ʁ 
that-obl for millet.soup 3pl.erg-caus-ready-pst

‘They made millet soup for him.’

There are also vague comparative concepts such as “modality” or “subject” for which defining any non-arbitrary prototype may be impossible. At the same time, such concepts still presuppose some internal structuring: there are different, albeit interrelated kinds of modality (van der Auwera & Plungian 1998) and there are different, albeit interrelated kinds of subjects (see, for instance, the attempt to distinguish between subject and pivot in Dixon (1994) and between most prominent argument and pivot in Falk (2006)). We conclude that all vague comparative concepts should be able to impose some kind of ordering on the language-particular descriptive categories and on the instantiations of these categories. Non-vague concepts may

After Keenan (1976), subject is sometimes cited as a prototype-based concept: see van der Auwera & Gast (2010), but see Lander & Tyshkevich (2015) for arguments against this view.
also be dealt with in this vein, since they treat all instantiations as manifesting the concept equally and this is also a kind of ordering.

4 Towards a typology of comparative concepts

The preceding sections have outlined our position on what comparative concepts can be. In this section, we present a sketch typology of comparative concepts. In particular, we suggest that there are three different kinds of comparative concepts.

Type 1 comparative concepts are those that are based on similarities between (some) languages. All of the concepts discussed above belong to this type, since they are based on observations holding for several languages: many languages display non-random effects related to the contrasts between morphology and syntax, restrictive and non-restrictive relative clauses, and inalienable and alienable adnominal possessive constructions. It is for (at least some of) these comparative concepts that we have argued for the usefulness of a vague model. The reason for this lies in the very nature of such comparative concepts: being based on similarities between languages, they often incorporate properties that matching categories in individual languages may possess, but don’t have to. For example, the comparative concept of “perfect”, as defined by Dahl (1985: 129–139), refers to a number of distinct functions such as “resultative”, “experiential”, or “hot news”, each of which is subject to crosslinguistic variation.

It should be noted that some of these concepts represent epiphenomena of other categories. It is probable that the concept of non-restrictive relative clause, which we used above for exemplification, is exactly of this kind, since non-restrictives can often be thought of simply as a kind of relatives which should be placed farther from the prototype of restrictive clauses. Unlike the latter, the alleged prototype of non-restrictives presumably need not favour grammaticalization of dedicated marking.

Type 2 comparative concepts are quite special, because they are based not on similarities, but rather on differences between languages. Typical representatives of this type are elementary meanings or functions, such as those constituting the basis of many semantic maps; cf. the following formulation from Haspelmath (2003: 217; emphasis ours): “A function is put on the map if there is at least one pair of languages that differ with respect to this function”. It is precisely this type of comparative concept which underlies both the whole semantic map approach and such prominent notions as “cross-linguistic gramm type” (Bybee & Dahl 1989), see also Evans (2010) on the application of “etic
grids” in semantic typology,\(^7\) as well as on the necessity to consider (more) language-specific “emic” characterizations of certain constellations of functions. As has been argued by Dahl (1985), Bybee & Dahl (1989), and much subsequent work, a lot of categories allow for valid typological generalizations in terms of the aforementioned crosslinguistic gram types, and at least some of them tend to have typical formal manifestation in different languages (e.g., the language-particular categories corresponding to the above-mentioned crosslinguistic gram type “perfect” are overwhelmingly expressed by analytic means in particular languages).

Importantly, type 2 comparative concepts are not expected to be vague. Indeed, each elementary function can be simply defined as an entry in a questionnaire (e.g., in the classic Tense-Aspect Questionnaire of Dahl (1985)), and hence the concept is both universally applicable (to the extent that a particular questionnaire sentence can be translated into a given language – something usually assumed without further qualification) and rigidly defined (again, to the extent that a linguist devising the questionnaire makes sure that its entries are not semantically vague, let alone ambiguous). It is the clusters of such elementary functions, themselves comparative concepts of type 1, that are necessarily fuzzy.

Finally, type 3 comparative concepts pretend not to be based on descriptive categories at all. Actually, many comparative concepts pass into this kind when they do not clearly apply to all languages, but yet are claimed to be universal. An important example of this kind is the concept of basic word order, which is usually “identified with the order that occurs in stylistically neutral, independent, indicative clauses with full noun phrase (NP) participants, where the subject is definite, agentive, and human, the object is a definite semantic patient, and the verb represents an action, not a state or an event” (Siewierska 1994). As argued by Brody (1984), Mithun (1987), Siewierska (1994), and La Polla & Poa (2006), in some languages this concept does not appear as a relevant descriptive category.

Such comparative concepts may look somewhat shady, but they can reveal generalizations about the organization of languages. For example, the fact that actors tend to precede undergoers in transitive clauses is certainly remarkable even if we discuss it with respect to topic-oriented languages such as Chinese or ergative languages where undergoers may have some properties that are usually associated with subjects.

\(^7\) Probably the first time that the method of “etic grids” has explicitly been applied to cross-linguistic comparison was by the eminent Russian lexicologist Nikita Tolstoj (1968) in his work on the lexical typology of Slavic languages and semantic reconstruction.
Yet type 3 comparative concepts should be approached with care. Consider the concept of adjective/noun word order (Dryer 2013). When taken without any language-particular background, its unmarked adjective–noun order makes Hausa (Chadic; Nigeria) a rare language, because its basically right-branching order seems to be violated by the prenominal position of the adjective. On closer inspection, however, it turns out that Hausa displays the phenomenon called “dependency-marking reversal in noun-attributive constructions” by Malchukov (2000) or “possessive-like attributive constructions” by Ross (1998), where a word described as an adjective on semantic grounds appears as the head of the nominal phrase and the noun which it modifies semantically is expressed as its possessor. See the possessive construction in (5a) and the construction with an adjective in (5b):

(5) a. kudi-n Audù
money-of Audu
‘Audu’s money’ (Jaggar 2001: 332)
b. kanānā-n yārā
small-of children
‘small children’ (Jaggar 2001: 344)

Not taking language-particular facts into account in this and similar cases may lead to doubtful generalizations concerning word order correlations and branching principles, which is a consequence of the use of a type 3 comparative concept.

5 Discussion

In the preceding sections we have argued that, despite Haspelmath’s contention that comparative concepts have to be independent of categories used for language description, the current practice of typologists shows that many comparative concepts which have proven to be useful (i.e., yield interesting and valid crosslinguistic generalizations) are in fact grounded in language-particular categories, do not always have universal applicability, and are often vague categories with a prototype core and fuzzy boundaries. Further, we have identified at least three types of comparative concept: concepts based on crosslinguistic similarity, concepts based on crosslinguistic differences, and concepts based on abstraction from both similarities and differences. We claim that all three types of comparative concept are legitimate, but we also emphasize that they have different ontological status and should not be conflated with each other.
Another point we would like to make concerns the so-called “non-aprioristic approach” to both language description and linguistic typology, which is advocated by prominent typologists like Croft (2001), Dryer (2006), and Haspelmath (2009, 2014). Though we agree that it is methodologically wrong to assume either that categories from one language can be successfully used in the description of another language, or that all languages have the same categories instantiating some universally given categories, we doubt whether a fully “non-aprioristic” and “framework-free” approach will be productive or indeed possible both in language description and linguistic typology. When a linguist approaches a language in order to write a grammatical description of it or sets out for a crosslinguistic investigation of a particular phenomenon, s/he by necessity entertains some implicit or explicit expectations about what languages (of the world or of a particular area or genetic group) may look like, what kinds of categories and structures are (more or less often) found in languages, and how different elements interact in linguistic systems. These expectations may come from different sources, such as the previous acquaintance with other languages, information provided by linguistic typology, as well as particular theoretical frameworks, and they will necessarily determine the range of questions the linguist will ask and, at least partly, the categorization and description of the linguistic data or the typological analysis s/he will produce.

Our point with respect to this is twofold, and, in our view, fairly obvious. First, since no (good) linguist is a “tabula rasa” void of any views about language, one should not pretend to be completely free of any assumptions, but rather explicate them and make sure that different assumptions do not contradict each other. That is why, we believe, there are no reasons to give up theory-dependent comparative concepts if the relevant assumptions are clearly formulated. Second, despite claims that language description does not and should not “depend on the theoretical results of typologists” (Haspelmath 2010: 683), we argue that the more a descriptive linguist knows about similarities and differences between languages discovered in typological and other comparative studies, the better the description s/he is likely to produce, simply because being informed about other languages and typological variation (and co-variation) in general enables one to ask more sophisticated research questions and see systematic patterns in the data which otherwise might appear just chaotic. As van der Auwera & Sahoo (2015: 139) put it, there is nothing wrong with trying out a comparative concept supported by 25 languages on the 26th language. On the contrary, it is highly recommended. The 26th language should most certainly be described in its own terms, but that does not mean that one should start from ‘categorial’ or ‘conceptual’ scratch each time one sets out to describe a new language.
Connected to this is the question of logical and epistemological links between comparative concepts and descriptive categories. If, following Haspelmath, it is methodologically flawed to attempt to “discover” the manifestations in individual languages of alleged universal categories (which Croft, Dryer, Haspelmath, and many other typologists claim simply not to exist), and if descriptive categories of particular languages only match (or don’t match) the definitions of universal comparative concepts, but do not “manifest” or “instantiate” them (see Haspelmath 2010: 680), then what is the relation between typological generalizations formulated in terms of comparative concepts and the structures of particular languages? If word order typology is based on semantic roles of “agent” and “patient” rather than on grammatical relations of “subject” and “object”, can we expect this typology to be informative of those languages whose syntax is not sensitive to these particular semantic notions? If the answer to this question is in the negative (as Matthew Dryer acknowledges in the Lingtyp discussion, 19 January 2016; see Supplementary Online Materials), then how can we interpret word order correlations among different constructions (e.g., between the order of verb and object and the order of adposition and its dependent noun phrase)? Indeed, if our typology aims at some explanatory and predictive value, it should allow the linguist to draw inferences of the following kind: “If (i) in the majority of the languages where the nominal object precedes the verb the noun phrase precedes the adposition, and (ii) in language L the nominal object precedes the verb, then (iii) it should be expected that language L has postpositions”. However, such an obvious implication would be fundamentally misguided because its second premise and the conclusion appeal to language-particular descriptive categories, while the first premise is formulated in terms of a universal comparative concept which is not necessarily compatible with the categories of language L. In order for this kind of logic – which appears to us to belong to the normal everyday toolkit of descriptive and theoretical linguists – to be coherent and legitimate, the relation between comparative concepts and descriptive categories should be such that they are not just two entirely disjoint and independent sets of notions, but rather (i) comparative concepts should have a firm basis in language-particular categories, and (ii) descriptive categories should be seen as meaningfully corresponding to comparative concepts of typology, with the two ultimately being definable by properties and features of the same nature.8

8 See van der Auwera & Sahoo’s (2015: 139) strong statement that “descriptive categories and comparative concepts are the same sort of entities [...] they are both sets of properties and in one way or another, properties can be shared across languages and in that case they define the comparative concept”.

Acknowledgements: We thank Michael Betsch, Anna Dybo, Maxim Fedotov, Dmitry Gerasimov, Martin Haspelmath, Maria Kholodilova, Tetiana Liubchenko, Tanja Mortelmans, Sorin Paliga, Benjamin Saade, Hedwig Skirgård, Nathan W. Hill, and Fernando Zúñiga for useful comments on an earlier version of this paper, and Frans Plank for encouragement to write it. None of these colleagues is responsible for the views expressed here. This material is partly based upon work supported by the Russian Science Foundation, grant No. 14-18-03270.

Abbreviations: 3 = 3rd person; ABS = absolutive; BEN = benefactive; CAUS = causative; DIR = directive; DYN = dynamicity; ERG = ergative; IO = indirect object; LOC = locative applicative; OBL = oblique; PL = plural; POSS = possessive; PST = past; REL = relative; SG = singular.

References


Brody, Jill. 1984. Some problems with the concept of basic word order. Linguistics 22. 711–736.


van der Auwera, Johan & Kalyanamalini Sahoo. 2015. On comparative concepts and descriptive categories, such as they are. Acta Linguistica Hafniensia 47. 136–173.
