Comparative concepts and descriptive categories in crosslinguistic studies

Martin Haspelmath

DISCUSSION NOTES

Comparative concepts and descriptive
categories in crosslinguistic studies

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology

In this discussion note, I argue that we need to distinguish carefully between descriptive categories, that is, categories of particular languages, and comparative concepts, which are used for crosslinguistic comparison and are specifically created by typologists for the purposes of comparison. Descriptive formal categories cannot be equated across languages because the criteria for category assignment are different from language to language. This old structuralist insight (called CATEGORIAL PARTICULARISM) has recently been emphasized again by several linguists, but the idea that linguists need to identify 'crosslinguistic categories' before they can compare languages is still widespread, especially (but not only) in generative linguistics. Instead, what we have to do (and normally do in practice) is to create comparative concepts that allow us to identify comparable phenomena across languages and to formulate crosslinguistic generalizations. Comparative concepts have to be universally applicable, so they can only be based on other universally applicable concepts: conceptual-semantic concepts, general formal concepts, and other comparative concepts. Comparative concepts are not always purely semantically based concepts, but outside of phonology they usually contain a semantic component. The fact that typologists compare languages in terms of a separate set of concepts that is not taxonomically superordinate to descriptive linguistic categories means that typology and language-particular analysis are more independent of each other than is often thought.*

Keywords: language typology, comparative concept, descriptive category, crosslinguistic category

1. INTRODUCTION: HOW TO COMPARE LANGUAGES. The purpose of this discussion note is to argue that crosslinguistic comparison should be based on comparative concepts created by the typologist, rather than on crosslinguistic categories that are instantiated in different languages, and to show how comparative concepts differ from language-specific descriptive categories. Although in practice typologists generally work with such special comparative concepts, this distinction between comparative concepts and descriptive categories has not been articulated clearly before.

More commonly, linguists tend to assume that there is a substantial set of universally available CROSSLINGUISTIC CATEGORIES (such as adjective, passive voice, accusative case, future tense, second person, subject, affix, clitic, phrase, WH-movement) from which languages may make a selection (Newmeyer 2007), and which are used both for description/analysis and for comparison. Typological research would then simply consist in identifying adjectives, passives, and so on in each language that has the category, and examining the ways in which the properties of the categories vary across languages. In this approach to crosslinguistic comparison, which we can call CATEGORIAL UNIVER-SALISM, it is one of the main tasks of comparative linguists (i.e. typologists) to determine what these crosslinguistic categories are. For linguists working on individual

664 LANGUAGE, VOLUME 86, NUMBER 3 (2010)

languages, this means that language-particular categories can be expected to be drawn from a relatively small set, that they can be equated with categories of other languages, and that they can be identified with (or instantiate) the crosslinguistic categories. Thus, comparative linguistics is an important necessary prerequisite for analyses of particular languages. Categorial universalism has been uniformly adopted in generative typology since its beginnings, and it appears to be implicitly assumed by many other linguists as well (e.g. Payne 1997, Corbett 2000, Van Valin 2005, Dixon 2010). It is sometimes even assumed that particular categories are universal in the sense not only of being universally available, but also of being universally instantiated (e.g. nouns and verbs in Baker 2003, adjectives in Dixon 2004).

The present discussion note starts out from an alternative view of the tasks of comparative linguistics. Following recent work on the foundations of grammatical typology by various authors (Dryer 1997, Croft 2001, Lazard 2006, Haspelmath 2007, Cristofaro 2009), I assume that grammatical categories are not crosslinguistic entities (either universally available or universally instantiated). Each language has its own categories, and to describe a language, a linguist must create a set of DESCRIPTIVE CATEGORIES for it, and speakers must create mental categories during language acquisition. These categories are often similar across languages, but the similarities and differences between languages cannot be captured by equating categories across languages. It was one of the major insights of structuralist linguistics of the twentieth century (especially the first half) that languages are best described in their own terms (e.g. Boas 1911), rather than in terms of a set of preestablished categories that are assumed to be universal, although in fact they are merely taken from an influential grammatical tradition (e.g. Latin grammar, or English grammar, or generative grammar, or 'basic linguistic theory'). This alternative, nonaprioristic approach to categories can be called CATEGORIAL PARTICULAR-ISM. In this approach, language-particular analyses can be carried out independently of comparative linguistics.[1]

Categorial particularism appears to make crosslinguistic comparison more challenging, but I argue that there exists a coherent and viable methodology for typological research that is compatible with it, which has in fact been employed by most researchers in the Greenbergian tradition (e.g. Greenberg 1963, Mallinson & Blake 1981, Comrie 1989, Dryer 1992, Croft 2003, Haspelmath et al. 2005, Song 2011). This is the use of COMPARATIVE CONCEPTS, that is, concepts specifically designed for the purpose of comparison that are independent of descriptive categories. However, linguists have often been unclear about the way that the apparent paradox of comparability of incommensurable systems can be resolved. It is my goal here to explicate this approach and defend it against challenges from a categorial-universalist perspective such as Newmeyer 2007.

I first give an overview of the crucial notion of comparative concept and characterize descriptive categories, showing that they must be different in different languages, and then argue against the use of crosslinguistic categories. The heart of the paper is §5, where I give concrete examples of well-known grammatical comparative concepts and show that the corresponding language-particular categories are crucially different. Following this, I address the terminological issues arising from this distinction and review earlier approaches to grammatical comparison, showing that very few earlier authors

---

[1] I use the terms 'comparative linguistics' and 'typology' interchangeably in this discussion note. 'Typology' is often associated with specifically nongenerative approaches, so I generally prefer the broader (and more transparent) but longer term 'comparative linguistics'.

have made this important distinction explicit, even though in practice many linguists distinguish the two notions implicitly. I then ask how comparative concepts are chosen, concluding that no general answer can be given because multiple perspectives of comparison can be adopted simultaneously without contradiction. Finally, §9 emphasizes that comparative concepts are not simply generalizations over linguistic categories, and that typology cannot be based on the comparison of categories in the sense of structurally coherent units of languages.

2. TYPOLOGISTS USE COMPARATIVE CONCEPTS. Typologists have often observed that crosslinguistic comparison of morphosyntactic patterns cannot be based on formal patterns (because these are too diverse), but has to be based on universal conceptual-semantic concepts (e.g. Stassen 1985:14, 2011, Croft 1990:11–12, 1995:88, 2003:13–14, Heger 1990/91, Givón 2001:20–23, Song 2001:10–12, Haspelmath 2007). As Newmeyer (2007:136) rightly emphasizes, however, 'typological generalizations need to make reference to the specific form in which these universal concepts are realized as well' (see also Rijkhoff 2009, 2010). Typologists make generalizations about phenomena such as case affixes, gender, adpositions, passive constructions, and relative clauses, and none of these can be defined in purely conceptual-semantic terms.

Thus, I claim that what crosslinguistic grammatical research is based on in general is comparative concepts. Comparative concepts are concepts created by comparative linguists for the specific purpose of crosslinguistic comparison. Unlike descriptive categories, they are not part of particular language systems and are not needed by descriptive linguists or by speakers. They are not psychologically real, and they cannot be right or wrong. They can only be more or less well suited to the task of permitting crosslinguistic comparison. They are often labeled in the same way as descriptive categories, but they stand in a many-to-many relationship with them (§9). Comparative concepts are universally applicable, and they are defined on the basis of other universally applicable concepts: universal conceptual-semantic concepts, general formal concepts, and other comparative concepts. Comparative concepts have often been used implicitly in the typological literature, but there has not been any detailed and explicit discussion of the difference between comparative concepts and language-particular descriptive categories.

Comparative concepts are needed for stating empirically testable universal claims. Consider the generalization in 1.

(1) In all languages with a dative and an accusative case, the dative case marker is at least as long as the accusative case marker.

To test this claim, we need to identify dative and accusative case markers across languages, and we need to measure phonological length in a crosslinguistically meaningful way. It is relatively easy to see that dative cases (or dative-like cases) cannot be equated across languages; that is, there is no crosslinguistic dative category. The Russian Dative, the Korean Dative, and the Turkish Dative are similar enough to be called by the same name, but there are numerous differences between them and they cannot be simply equated with each other. Clearly, their nature is not captured satisfactorily by saying that they are instantiations of a crosslinguistic category 'dative'.[2]

However, 'dative case' can be defined as a comparative concept, as in 2.

---

[2] Grammatical labels with an initial capital are used for language-particular descriptive categories, and lower-case is used for comparative concepts; see §6.

(2) A dative case is a morphological marker that has among its functions the coding of the recipient argument of a physical transfer verb (such as 'give', 'lend', 'sell', 'hand'), when this is coded differently from the theme argument.

This definition is based on the conceptual-semantic concepts 'recipient' and 'physical transfer verb', as well as the comparative concepts 'morphological' and 'argument'. The comparative concept 'morphological', in turn, presupposes the comparative concept 'word', perhaps definable as a 'segment string that cannot be interrupted by a free form without changing the meaning', and the comparative concept 'argument' can be defined as 'referential phrase that can fill a verb's semantic valency position'. These definitions are still not composed of primitive concepts, but they do not (at least not evidently) introduce language-particular concepts, so they are (or can be made, with suitable refinement) crosslinguistically applicable. Thus, given the definition of the comparative concept 'dative' in 2, we can go on to test the generalization in 1.

This is done by matching the phenomena of languages with the comparative concepts. The Russian Dative matches the 'dative case' concept in 2, and so does the Finnish Allative, but the Nivkh Dative-Accusative case, for instance, does not match it, despite its name, as it is used to express the causee in causative constructions, but not the recipient of a physical transfer verb. Note that we cannot say that the Russian Dative and the Finnish Allative 'instantiate' the 'dative case' concept, because these categories have many more properties than are contained in the definition in 2. This is the crucial difference between comparative concepts as proposed here and the crosslinguistic categories that I reject. On the categorial-universalist view, categories in particular languages instantiate crosslinguistic categories, and they can thus be equated across languages. On the view presented here, the fact that two language-particular categories both match a comparative concept just means that they are similar in the relevant respect, but not that they are 'the same' in any sense (see §9 for further discussion).

I use the term COMPARATIVE CONCEPT (rather than COMPARATIVE CATEGORY) in order to emphasize that the comparative concepts are typologists' constructs, not part of the structure of languages. Since Aristotle, we have known that speakers need categories (semantic categories and formal-grammatical categories) to be able to use language, and linguists need (corresponding) categories to describe individual languages. Comparative concepts, by contrast, are a sort of metacategory that is irrelevant to language learning or language description/linguistic analysis. Linguists tend to use the same grammatical terms for descriptive categories and comparative concepts, but these two uses of the terms refer to different kinds of entities.

3. DESCRIPTIVE LINGUISTS USE LANGUAGE-PARTICULAR DESCRIPTIVE CATEGORIES. A large part of linguists' activity is descriptive in the sense that it focuses on the analysis of a particular language, whether one's main concern is theoretical or applied.[3] To describe a language, one needs categories, because it is not possible to list all the acceptable sentences of a language. Out of the countless possibilities, linguists typically value those descriptions most highly that 'do justice' to the language in an optimal way. Well into the nineteenth century, it was common to find descriptions of caseless languages in

---

[3] I follow Lyons (1981:34–35) in contrasting DESCRIPTIVE LINGUISTICS (the study of particular languages) with GENERAL LINGUISTICS (the study of language in general). The THEORETICAL–APPLIED distinction is orthogonal to the descriptive–general distinction. I only address theoretical concerns here—that is, by 'descriptive' I mean 'theoretical-descriptive'. (Instead of 'description', many linguists use the equivalent term 'analysis'.)

terms of the Latin six-case model (nominative: *John*, accusative: *John*, dative: *to John*, genitive: *of John*, ablative: *from John*, vocative: *o John*). Descriptions of this sort worked, in that they allowed the users of the description to distinguish between acceptable and unacceptable expressions in the language, but they did it in a way that has been unacceptable to linguists since the advent of structuralism at the latest. At least as theoretical linguists, we are not satisfied with getting the facts right, but we also want to gain insights into the inner workings of the language, into the way its various elements cohere. This means that English should be described nonaprioristically, with descriptive categories designed for English, not with categories designed for Latin, Arabic, or any other language. As the American structuralists recognized, this emancipation from the Latin (or Standard Average European) model must be extended to all other languages. According to Boas (1911:81), the descriptive categories chosen by the authors of the *Handbook of American Indian languages* 'depend entirely on the inner form of each language'. See Gil 2001 for a recent compelling formulation of the argument against apriorism in linguistic analysis.

A very different view has come to prevail in the second half of the twentieth century, however, according to which there exists a set of crosslinguistic categories (some of them perhaps even universal) from which languages may choose. According to this view, English happens not to have a dative case, but the Korean dative case instantiates the same crosslinguistic category as the Icelandic dative case (see Maling 2001); adjectives in Tariana are adjectives in the same sense as adjectives in Russian (see Aikhenvald 2004); or subjects in Tagalog are subjects in the same sense as in English (see Kroeger 1993). On this view, the task of descriptive work is to identify the phenomena of a language with these preestablished crosslinguistic categories. These categories are often assumed to be part of the innate UNIVERSAL GRAMMAR (substantive universals, Chomsky 1965:§1.5), but sometimes linguists simply assume their crosslinguistic validity without any cognitive commitment.[4]

More recently, a number of linguists have argued against the view that grammatical categories are crosslinguistic, and have essentially returned to the structuralist position that each language has its own categories and that there are no crosslinguistic categories (Lazard 1992, 2005, 2006, Dryer 1997, Culicover 1999:Ch. 2, Croft 2000, 2001, Haspelmath 2007, 2010, Mielke 2008, Cristofaro 2009).

The argument for categorial particularism is simple: the criteria used for identifying categories such as cases, word classes, and grammatical relations are themselves language-particular. Subjects in English are identified by a set of criteria that is only partially comparable to the set of criteria that might be used to identify subjects in Tagalog (cf. Schachter 1976). Adjectives in Mandarin Chinese are identified by a set of criteria that is quite different from the criteria applied in Russian (cf. McCawley 1992). Many criteria are completely lacking in many of the languages. Since there is no principled way of deciding which criteria are relevant, linguists that assume crosslinguistic categories are often (indeed, usually) in disagreement over category assignments: Is Thai *dii* a verb or an adjective (cf. Prasithrathsint 2000)? Are the French subject pronominals *je*, *tu*, *il*

---

[4] A referee claims that on a noninnatist view, it is incoherent to assume that children learning different languages on the basis of different data learn the same categories. While I certainly agree that it is more natural to assume that if categories are not given in advance, they cannot end up the same, I could still imagine that learners of different languages somehow converge on the same categories (e.g. because the categories are so evidently useful). Thus, I treat generative categorial universalism together with nongenerative categorial universalism (e.g. Van Valin 2005, Dixon 2010) here, even though their views of course differ in important ways.

(etc.) pronouns or agreement markers (cf. De Cat 2005)? Is the English genitive marker *'s* an enclitic word or a phrasal suffix (cf. Zwicky 1987)? Such category-assignment controversies cannot be resolved, because the answer one gives depends on one's choice of criteria, and this choice is 'opportunistic' (Croft 2001:30, 41, 2009). Linguists inevitably choose the criteria in such a way that they obtain the result that fits their general perspective best. But since perspectives differ, different linguists arrive at different categories, and it is impossible to tell which category assignment is correct. The solution is to accept that categories are language-particular, and to describe languages in their own terms.

Language-particular descriptive categories are required in phonetics and phonology as well. Port and Leary (2005) argue against the assumption of 'a fixed, universally available alphabet-like phonetic inventory', and Ladd (2011) likewise notes that phonetic research of the last decades is not compatible with the 'universal categorization assumption', the idea that there is a closed universal inventory of possible segment types. The distinctive features of modern phonology were originally conceived of as language-particular categories (Trubetzkoy 1939), and it was only later in work by Roman Jakobson, Morris Halle, and others that they were reconceptualized as crosslinguistic categories. More recently, a number of phonologists have advocated a return to the Trubetzkoyan view. In Mohanan and colleagues' (2010:§1) approach, 'UG does not provide an inventory of specific distinctive features … Inventories of features, contrastive segments and feature-constraints emerge in an individual grammar in the course of language development through exposure to data'. Similarly, Mielke's (2008) 'emergent feature theory' argues against innate, crosslinguistic features and in favor of language-particular features. Mohanan and colleagues (2010:§6) recognize that this has consequences for crosslinguistic comparison, which cannot be based on a universal list of predefined features: 'What is needed is a cross-linguistically valid currency of distinctive features: such a currency can obtain without reference to a set of features stipulated in UG'. In other words, phonological typology needs comparative phonological concepts, and it is such comparative concepts that good phonological typology (e.g. Maddieson 1984) is based on, often implicitly. Maddieson (1984), for instance, does not discuss the theoretical basis for comparison explicitly, but he essentially compares segments across languages in terms of the IPA symbols. As D. R. Ladd (p.c.) observes, the International Phonetic Alphabet is useful as a set of widely understood comparative concepts for crosslinguistic comparison, but not as 'a set of symbols for representing all the possible sounds of the world's languages' (International Phonetic Association 1999:159).

The distinction between descriptive categories and comparative concepts is even more familiar from lexical comparison. While simplistic approaches such as Swadesh-list-based comparison make the simplifying assumption of a one-to-one correspondence between lexical meanings and words, and thus between words across languages, reality is more complex: words in one language are often in semantic many-to-many relationships with words in another language. Any systematic lexical comparison of languages needs to work with a somewhat arbitrarily chosen set of standardized lexical meanings (e.g. the World Loanword Database, Haspelmath & Tadmor 2009a,b). These lexical meanings are comparative concepts, and the meanings of individual languages are language-particular semantic categories. Neither can be reduced to the other.

4. CROSSLINGUISTIC CATEGORIES CANNOT BE THE BASIS FOR COMPARATIVE LINGUISTICS. The approach to language comparison in terms of crosslinguistic categories has not been successful. If there were such a limited set of universally available categories from

which languages may choose, we would expect that grammatical research from various perspectives would gradually converge on these categories. Linguists differ widely in the kinds of categories they assume, however, and a common experience of typologists is that each new language presents them with something that they have never seen before. While there is a core set of phenomena that tend to be quite similar across languages (noun, verb, demonstrative, personal pronoun, subordinate clause, question-word fronting, etc.),[5] there is a large periphery of construction types that are not readily identifiable across languages. And where there is broad agreement about certain categories and construction types, it is sometimes hard to avoid the impression that the agreement is based more on grammatical traditions (influential frameworks, or influential languages such as English or Dyirbal) than on real convergence of analyses. As Bach (2004:56–57) notes, early generative grammar 'took over with no substantial justification the categories of traditional grammar ... The initial empirical base was English and as this base was broadened to include more and more different languages, these categories were naturally taken over for the "new" languages'.

The Boasian position that each language has its own categories avoids the problems of category-assignment controversies and bias by traditions. Its most glaring potential problem is that it makes languages seem incommensurable, rendering language comparison impossible. And indeed, as Greenberg (1974:42) noted, American structuralists 'largely ignored typology'. However, I am arguing here that language comparison should not be (and in the best cases is not) based on structural linguistic categories at all, but on comparative concepts. Thus, there is no conflict or tension between a Boasian and a Greenbergian approach.

It should be noted that the arguments against crosslinguistic categories apply with equal force if one adopts an innatist stance. A referee thinks that generative linguists 'can claim that linguistic categories are innate and such a claim seems impervious to most arguments against crosslinguistic categories'. But since generative linguists do not have direct access to mental grammars either, they have to justify their category assignments in the same way as linguists who do not make assumptions about innateness. One may of course assume that children are born with a subject category, but if linguists cannot identify subjects crosslinguistically for lack of consistently applicable criteria, then speakers can hardly learn subjects either. Culicover (1999:Ch. 2), coming from a generative background, arrives at a particularist conclusion without even considering linguistic diversity in detail, solely on the basis of learnability considerations. He observes that syntactic categories must be taken to have subcategories, comprising items with specific properties not shared by all members of the category (e.g. pronouns as a subcategory of nouns). Sometimes these subcategories are small, consisting only of a single member (e.g. postposed degree adverbs, of which English has only one: *enough*). Culicover notes that if such small and idiosyncratic categories are also assumed to be drawn from a universal set, this entails 'that all conceivable sets of elements that could share any formal property at all are universally available syntactic categories, including every individual lexical item' (1999:40). He concludes that this renders the categorial-universalist position vacuous, and that it must be abandoned.

---

[5] As Dryer (1997:123) emphasizes, there is no need to appeal to crosslinguistic categories to explain these similarities, because functional explanations are generally available for them. For instance, the fact that nouns and verbs contrast grammatically in most languages is evidently due to the fact that all human beings tend to use time-stable thing concepts as referents and ephemeral process concepts as predicates (Croft 1991:Ch. 2).

5. EXAMPLES OF COMPARATIVE CONCEPTS. Comparative concepts are necessary for the formulation of crosslinguistic generalizations (or 'language universals'). If we were not interested in such generalizations, we could do without comparative concepts, and of course many descriptive linguists, psycholinguists, and sociolinguists do not use comparative grammatical concepts in their work at all. To give the reader a better sense of what I am talking about, I give seven concrete examples of comparative grammatical concepts from morphosyntax and how they are used in crosslinguistic studies. In each of the following subsections, I first cite a proposed generalization, and then give a definition for one of the key concepts in the generalization. As in our first example in §2, the definitions appeal only to conceptual-semantic concepts, to general (not linguistics-specific) formal concepts (such as 'precede', 'identical', 'overt', which I assume to be unproblematic), and to other comparative concepts. The definitions of comparative concepts must not contain language-particular components. Space limitations here do not allow me to define all of the concepts used in the definitions and to justify all aspects of the definitions and the universality of the concepts used in them. A complete account will of course have to be fully explicit, but here the point is primarily to show that comparative concepts are quite different from descriptive categories used for the analysis of particular languages.[6]

5.1. ADJECTIVE. Comparative part-of-speech concepts such as 'adjective' are necessary for stating the well-known Greenbergian generalizations (see 3). Greenberg and other word-order typologists have worked with a definition of 'adjective' along the lines of 4.

    (3) GENERALIZATION: If a language has dominant SOV word order and the genitive follows the governing noun, then the adjective likewise follows the noun. (Greenberg 1963, universal 5)[7]

    (4) DEFINITION: An adjective is a lexeme that denotes a descriptive property and that can be used to narrow the reference of a noun.

This definition makes use of the comparative concept 'lexeme' and the conceptual-semantic concepts 'property' and 'narrow the reference' (the latter is necessary to exclude words like *size* and *beauty*, which denote properties, but are not normally used to modify a noun).

Importantly, it is irrelevant for this definition whether a language has a separate word class that would be called 'Adjective' (i.e. a descriptive category), or whether it uses its 'Noun' or 'Verb' categories to attribute properties to nominal referents. In the latter case, a Verb (= descriptive category) may be an adjective (= comparative concept) for the purposes of crosslinguistic comparison (cf. Dryer 2005a:354). Since descriptive categories and comparative concepts are different kinds of entities, there is no contradiction here. (There is a terminological problem, however, which I address in §6 below.)

5.2. FUTURE TENSE. Tense-aspect concepts have also been widely compared across languages by typologists. Let us consider 'future tense'.

---

[6] All seven examples concern concepts that include a meaning element in their definition. However, morphosyntactic typology also widely uses comparative concepts in whose definitions meaning plays no role: word, affix, phrase, clitic, and so on. Everything I say about comparative concepts applies to them as well (as it does to phonological comparative concepts), but I do not discuss them here because the illustration would be even more complex, and I am less certain of their actual usefulness for typology (cf. Haspelmath 2011 for serious doubts concerning the notion 'word').

[7] This generalization has been called into question by Dryer (1988, 1992, 2005a). This is irrelevant here, because my purpose is merely to explain how comparative concepts function in typology. If Greenberg and Dryer had not assumed this sense of the term 'adjective' (as a comparative concept), neither the formulation of the generalization nor its refutation would have been possible.

(5) GENERALIZATION: In all languages, markers of future tense are less bound than markers of present tense or past tense, or equally bound, but never more so. (Ultan 1978:91)

(6) DEFINITION: A future tense is a grammatical marker associated with the verb that has future time reference as one prominent meaning.

This definition makes use of the conceptual-semantic concept 'future time reference' and the comparative concepts 'verb' (defined in a manner analogous to the definition of 'adjective' that we just saw) and 'grammatical marker' (to delimit future tenses from temporal adverbs like *tomorrow*). The vague qualifications 'associated with' and 'prominent' will have to be made more precise in one way or another to make the generalization truly testable.

What is important here is that the generalization in 5 cannot be formulated in purely conceptual-semantic terms, because it is not supposed to make claims about temporal adverbs or adverbial clauses. Moreover, future tenses are normally not synonymous across languages. For instance, the Spanish Future tense (e.g. *vendrá* 'will come') is also used to express probability, but not habituality (Butt & Benjamin 2000:213–16), while the Lezgian Future tense (e.g. *qwe-da* 'will come') is also used to express habituality, but not probability (Haspelmath 1993:141–42). The generalization in 5 is intended to cover both forms, but clearly they cannot be 'the same category' in any sense. They are different descriptive categories, but they both match the comparative concept in 6. Note that 6 does not require that future time reference is the primary sense of the form in question, and a form such as the Korean -*keyss*-form, which is called 'Volitional Mood' (as a descriptive category) by Chang (1996:128–31), can also express future tense and thus falls under 6 and 5, even though volition and probability are its primary senses.

5.3. QUESTION-WORD MOVEMENT. Comparative concepts are not restricted to concepts that are the counterparts of (morpho)syntactic categories. Comparative linguists compare entire constructions across languages using comparative concepts, for example 'question-word movement' (often called WH-movement even in languages where question words do not begin with WH-).

(7) GENERALIZATION: Question-word movement is always to the left. (Hawkins 1999:273, Dryer 2005b)

(8) DEFINITION: Question-word movement is a syntactic pattern in which a question word (or question-word phrase) occurs in a special position in which its non-question-word counterpart would not normally occur.

This definition makes use of the general formal concept '(special) position', and the comparative concepts 'syntactic pattern', 'question word', and 'question-word phrase' (the latter two are defined immediately below in §5.4). Note that it contains no conceptual-semantic concept.

Importantly, no claim is made that question-word movement is a crosslinguistically uniform process. We know that it is not: in some languages, both content interrogatives and relative clauses undergo a similar movement process (e.g. in English, where both have been unified under the language-particular label WH-movement), in others only content interrogatives do, and in yet others a good case can be made that there is a general focus fronting rule that applies to question-word phrases and to focused non-question-word phrases alike. Thus, the fronting rules in these languages are all different descriptive categories,[8] but they all fall under the comparative concept in 8.

---

[8] It is true that abstract analyses are possible in which they are all the same process (e.g. Move α, Chomsky 1981:Ch. 6), but such abstract analyses are very hard to test. Here I focus on crosslinguistic generalizations that are readily testable.

5.4. QUESTION WORD. For a full understanding of 'question-word movement', we still need to define 'question word' (often called WH-word) as a comparative concept.[9]

> (9) DEFINITION: A question word is a word that can be used as a question pronoun (or adverb), that is, to represent the questioned content in a content question.

As in the case of the future tense, the definition does not assume that question words are restricted to the use as question pronouns, and this use does not even have to be their primary use. In many languages, question words are also used as indefinite pronouns (Haspelmath 1997:§7.3, Bhat 2000) or as relative pronouns, and the question use may not even be more prominent than other uses. Such pronouns have therefore sometimes been called 'ignoratives' (Wierzbicka 1980:Ch. 8) or 'epistememes' (Mushin 1995). But the only meaning that is universally available to what we call question words is the question-pronoun use, so this is the decisive criterion in the definition of the comparative concept in 9. Again, we can say without contradiction that in a certain language an Epistememe (= descriptive category) is a question word (= comparative concept).

5.5. RELATIVE CLAUSE. Another comparative-concept type that has figured prominently in crosslinguistic studies is 'relative clause'.

> (10) GENERALIZATION: If the relative clause precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun, or both. (Greenberg 1963, universal 24)
>
> (11) DEFINITION: A relative clause is a clause that is used to narrow the reference of a referential phrase and in which the referent of the phrase plays a semantic role.

This definition is very similar to the definition in Keenan and Comrie's (1977) influential article on the typology of relative clauses. But while Keenan and Comrie consider it 'an essentially semantically based definition' (1977:63), I emphasize here that it is a comparative concept that contains not only conceptual-semantic components such as 'narrow the reference' and 'semantic role', but also the formal concept 'clause' (itself a comparative concept). This concept is necessary to delimit relative clauses from attributive adjectives in languages like English, and from possessors, which are semantically like relative clauses, but are not clauses.

As in the earlier examples, a construction does not have to be a relative clause in a particular language to qualify as a relative clause for the purposes of crosslinguistic studies. Thus, Comrie and Horie (1995) and Comrie (1998), following Matsumoto (1988, 1997), argue that Japanese clauses such as 12a instantiate the same construction as in 12b and 12c.

> (12) a. gakusei ga katta hon
>   [student NOM bought] book
>   'the book that the student bought'
>   b. gakusei ga hon o   katta   zizitu
>   [student NOM book ACC bought] fact
>   'the fact that the student bought the book'
>   c. dareka   ga doa o tataku oto
>   [someone NOM door ACC knock] sound
>   'the sound of someone knocking at the door'

---

[9] 'Question-word phrase' is simply defined as 'a phrase that contains a question word'.

Matsumoto's and Comrie's 'Noun-Modifying Construction' is thus a descriptive category within Japanese. Japanese has no category that closely corresponds to the descriptive category of Relative Clauses in English, but for crosslinguistic studies of relative clauses, the construction in 12a can be taken as a relative clause (in the comparative sense; cf. Comrie & Kuteva 2005).

5.6. REFLEXIVE PRONOUN. Reflexive pronouns are preferred in direct-object function (e.g. *She saw herself*) and are used less commonly in adnominal possessive function (cf. **She saw herself's son*). This can be expressed as a universal generalization.

> (13) GENERALIZATION: If a language uses a reflexive pronoun for an adnominal possessor that is coreferential with the subject, then it also uses a reflexive pronoun for a subject-coreferential direct object. (Haspelmath 2008b)
> (14) DEFINITION: A reflexive pronoun is a specialized anaphoric expression with a coreferential reading that can only be used in the same clause as the antecedent.

The definition of 'reflexive pronoun' thus has both a semantic ('anaphoric') and a formal ('in the same clause') component. The definition mentions only part of the usual properties of reflexive pronouns and is silent about the c-command requirement for the antecedent, subject orientation, and the distinction between local and long-distance reflexives. Few languages will have a descriptive category that corresponds exactly to 14, but for some typological purposes this definition is sufficient.

5.7. ERGATIVE CASE.

> (15) GENERALIZATION: In all languages with an ergative case, it has at least some overt allomorphs. (cf. Dixon 1979:§2.3)
> (16) DEFINITION: An ergative case is a morphological marker that has among its functions the coding of the agent of typical transitive clauses, when this is coded differently from the single argument of intransitive clauses.

In most languages with cases that match the definition in 16, they are called ergative, but it is not possible to equate them across languages, because they may have quite a few additional functions as well (instrumental, locative, possessive, general oblique; cf. Dixon 1994:57). Within each language, there may well be reasons to regard the case not as polysemous, but as having a general meaning that happens to comprise the meaning of the transitive agent. In some languages, the corresponding descriptive category is traditionally known by a different name (Narrative case in Georgian, Relative case in Eskimo). Note also that the generalization in 15 and the definition in 16 are neutral with respect to the competing analyses of split ergative systems: Dixon's, according to which languages like Dyirbal have an ergative-absolutive case system coexisting with a nominative-absolutive case system (in the 1st and 2nd person pronouns), and Goddard's (1982), according to which such languages have a tripartite ('ergative-accusative-nominative') system. On the latter analysis, 'Ergative' has a different meaning as a descriptive category, but the comparative concept and the generalization are not affected.

5.8. SUMMARY. Comparative concepts are concepts created by comparative linguists for the purpose of formulating readily testable crosslinguistic generalizations (such as those in 3, 5, 7, etc.). They are potentially applicable to any human language. Their definitions contain other universally applicable concepts of three kinds: universal conceptual-semantic concepts, general formal concepts (such as 'precede', 'overt'), and other (more primitive) comparative concepts.

Comparative concepts allow linguists to identify comparable grammatical phenomena in different languages, but by identifying a phenomenon in a particular language as

a match of a comparative concept, nothing is claimed about the way in which that phenomenon should be analyzed within the language (what kind of descriptive category should be used for it). Comparative concepts and descriptive categories are quite different kinds of entities that should not be confused (for more on the differences, see §9).

6. TERMINOLOGY: WHAT TO CALL COMPARATIVE CONCEPTS AND DESCRIPTIVE CATEGORIES. If there were a limited set of crosslinguistic categories from which languages choose, the terminological aspect of crosslinguistic research would be easy: we would simply have to settle on a unique term for each crosslinguistic category or feature. Since the crosslinguistic categories would be used by descriptivists, generalists, and typologists alike, these could all use the same terms.

But with the distinction between comparative concepts and descriptive categories that is proposed here as an alternative, terminology becomes a serious issue, because the entities that need to be named are multiplied: comparative concepts are linguist-specific (in the sense that every linguist is free to define her or his own concepts), and descriptive categories are language-particular.

Sometimes descriptive categories have been identified by arbitrary numbers ('case 3'), and another possibility is to simply name a category after its exponent (e.g. the *-ing*-form in English, Quirk et al. 1985:92). But generally linguists prefer transparent names that make it easier to remember them, and language-particular labels that are very similar to the corresponding comparative concepts clearly facilitate communication among specialists of different languages. Since the 1980s, another convention for descriptive categories has gained some ground, especially among typologists who want to contrast descriptive categories with comparative concepts. Following Comrie (1976:10), Bybee (1985:141), and Dahl (1985:34), grammatical labels with an initial capital refer to language-particular descriptive categories (e.g. 'the Russian Perfective aspect', 'the Spanish Imperfect tense'), while ordinary lower-case spelling is used for comparative concepts. This makes sense because descriptive categories are akin to proper names in that they refer to unique entities (one can never say that a category is 'an Imperfect', just as one can never say that a city is 'a Warsaw'). Croft (2001:50) proposes to extend this convention to syntactic categories (e.g. 'the Kutenai Verb') and constructions (e.g. 'the Tagalog Actor Focus', 'the English Relative Clause'), which are just as language-particular as the tense and aspect forms discussed by Comrie, Bybee, and Dahl. While the practice of capitalizing descriptive categories is not very widespread yet, it has occasionally been adopted in reference grammars (e.g. Haspelmath 1993, Maslova 2003). As indicated in n. 2, this practice has also been adopted in this discussion note.

No specific terminological convention is proposed for (linguist-specific) comparative concepts here. It might in general be good to adopt more semantic terms (e.g. possessor instead of genitive, irrealis instead of subjunctive, coreferential instead of reflexive), but this is not a general solution, because as we saw, comparative concepts are by no means always or even typically purely semantic concepts. Typologists normally use the same kinds of terms that are used in grammars of particular languages, and although this can give rise to misunderstandings, it seems unavoidable because we want the terminology to be transparent.

7. EARLIER APPROACHES TO COMPARISON OF GRAMMARS. A referee questioned whether the ideas presented in this discussion note were new, so it is useful to briefly recapitulate the history of comparative grammatical research from the point of view of comparative concepts. In this section, I show that with very few exceptions, the approach

described here has not been articulated before, even though it has often been practiced, especially since Greenberg 1963.

7.1. NINETEENTH-CENTURY LINGUISTICS. The idea that grammatical categories exist independently of particular languages and that different languages may have (some of) the same categories has a long history, going back at least to the seventeenth-to-eighteenth-century approach known as *grammaire générale*. The earliest large-scale typological studies, such as von Humboldt 1830 on the dual, von der Gabelentz 1861 on the passive, and de la Grasserie 1896 on the article, do not show any awareness of the distinction between language-particular categories and comparative concepts. Just as language descriptions at the time tended to work with the categories inherited from Latin and Greek grammar, typology also worked with these categories. That each language has its own categories was an insight of the early twentieth century, both among Boasian descriptivists and Saussurean generalists. As we saw in §4, however, the structuralists of the first half of the twentieth century were not much concerned with crosslinguistic studies (with some exceptions such as Hjelmslev and Trubetzkoy).

7.2. GENERATIVE LINGUISTICS. It was only in the 1960s that many general linguists became seriously interested in typological comparison again, especially under the influence of Greenberg (1963) and Chomsky (1965). Greenberg and his students worked with comparative concepts as conceived of here, but they did not worry much about their relation to the descriptive categories of individual languages. Chomsky 1965, a work that became even more influential in the field, explicitly espoused categorial universalism: categories (or their more elementary components, features) are universal in the sense that they are part of the innate universal grammar (UG). In the Chomskyan approach, individual languages need not instantiate all categories of UG, but every category of a particular language is an instantiation of a category of UG. Of course, over the decades the categories that are actually assumed to be universal have varied (and as in generative phonology, the morphosyntactic categories have often been decomposed into more abstract features), but the basic assumptions have remained remarkably constant.

The generative universalist position on categories, which has been dominant since the 1960s, has a very simple view of categories: there is only one set of them, and they are used by descriptivists, generalists, and typologists alike. Sometimes a distinction is made between 'pretheoretical' or 'descriptive' concepts on the one hand and 'theoretical' concepts on the other, but only the 'theoretical' concepts are taken seriously. These are used both for language-particular analysis and for language comparison. These crosslinguistic categories thus conflate descriptive categories and comparative concepts, as understood here.[10]

In practice, however, this creates considerable difficulties, because there is no agreement on what the crosslinguistic categories are, and on how individual languages should be analyzed in their terms. (Recall that I observed earlier, in §3, that category-assignment controversies cannot be resolved because there is no principled basis for choosing the relevant criteria.) Yet in the generative approach, analyses of individual languages in terms of crosslinguistic categories are a prerequisite for doing typology:

---

[10] This generative view has been adopted without further discussion by the knowledge-engineering proposals known as 'General ontology for linguistic description' (Farrar & Langendoen 2003, http://linguistics-ontology.org). The conclusions of this discussion note would argue for a reassessment of these proposals. (See also Good 2004, who clearly recognizes that in addition to a general ontology, descriptive grammars make crucial use of a 'local ontology', that is, language-particular descriptive categories.)

> Assigning category membership is often no easy task … Is Inflection the head of the category Sentence, thus transforming the latter into a[n] Inflection Phrase (IP)? … Is every Noun Phrase dominated by a Determiner Phrase (DP)? … There are no settled answers to these questions. Given the fact that we are unsure precisely what the inventory of categories for any language is, it is clearly premature to make sweeping claims … (Newmeyer 1998:338)

But as long as we know what the facts are and how the comparative concepts should be applied, we can leave questions such as these aside and do typology anyway. And for a Boasian/Saussurean typologist or grammar writer, questions such as these would not even make sense, because categories like IP and DP are not assumed to be universal to begin with.

If the principles-and-parameters program were leading to an increasingly clear view of what the universal categories and parameters are, the universalist approach would gain plausibility, despite arguments such as Dryer's and Croft's that crosslinguistic formal categories cannot be identified in principle. But as both critics and advocates of the Chomskyan approach have noted (e.g. Newmeyer 2004, 2005:§3.2, Tomasello 2005:§3, Baker 2008:352), the principles-and-parameters program has not led to an increasingly clear picture of the substantive and formal universals of universal grammar (see also Haspelmath 2008a).

It seems to me that the conflation of descriptive categories and comparative concepts in presumed 'crosslinguistic categories' has made it very difficult for generative linguists to engage in broad crosslinguistic studies. On the generative view, language-particular analysis is not possible without a clear picture of UG, and language comparison is not possible without a thorough analysis of the languages that are compared. But our picture of UG has remained murky, and all our language-particular analyses are very preliminary. As a result, virtually all major large-scale crosslinguistic studies of the last two decades (e.g. almost all the chapters of *The world atlas of language structures*, Haspelmath et al. 2005) have been carried out in the context of the Greenbergian approach, using comparative concepts. Generative linguists have hardly contributed to broadly comparative studies, limiting themselves mostly to 'one language at a time' comparisons (Croft 2009). And those large-scale crosslinguistic studies that were carried out by generativists (e.g. Cinque 1999, Julien 2002, Gianollo et al. 2008) have in practice used comparative concepts, rather than crosslinguistic categories whose instantiations in individual languages are well motivated. To the extent that the large-scale crosslinguistic work has been fruitful, it argues for the approach that is advocated here.

7.3. OTHER TYPOLOGICAL APPROACHES. Outside of generative linguistics, there is usually no explicit claim that categories can be equated across languages, that is, that the same categories can be used for both typology and language-particular description/analysis. But typologists often ignore the distinction between language-particular descriptive categories and crosslinguistically applicable comparative concepts. In textbooks like Payne 1997, Whaley 1997, and Tallerman 1998, in prominent studies of particular phenomena such as Diessel 1999 and Corbett 2000, and in entire approaches such as the St. Petersburg school of typology (Nedjalkov & Litvinov 1995), role-and-reference grammar (Van Valin 2005), and functional discourse grammar (Hengeveld & Mackenzie 2008), the issue is simply left aside.

Other typologists have emphasized the fact that crosslinguistic work must be based on conceptual-semantic, or more generally functional comparative concepts (Stassen 1985:14, Croft 1990:11–12, Heger 1990/91, Givón 2001:20–23, Haspelmath 2007). This view recognizes that language-particular categories can be highly diverse, but it locates typological diversity exclusively in the simple mapping of form and function. It would thus not be capable of delving deeper into the typology of ergative constructions,

relative clauses, question-word movement constructions, and so forth, that is, phenom-
ena that cannot be defined in exclusively semantic or functional terms (cf. §3, and
Rijkhoff 2009, 2010).

But there is also a group of nongenerative typologists that have explicitly claimed
that the categories used for description and comparison are the same. Most clearly,
Lehmann (1989:142) formulates the following principle:

> Describe your language in such a way that the maxim of your description could serve, at the same time,
> as the principle of general comparative grammar—and, thus, as the maxim of description of any other
> language.

According to Lehmann, 'the description of a specific language is a concretization of
general comparative grammar' (1989:142), so in these formulations, his 'general com-
parative grammar' seems to come close to Chomskyan universal grammar. Similarly,
Dahl (1985:31) claims that 'the overwhelming majority of all categories found in the
[tense-mood-aspect] systems of the world's languages are chosen from a restricted set
of category types', and in this result he sees a 'refutation of the relativist view that
"every language must be described in its own terms" ' (Dahl 1985:32; see also Bybee &
Dahl 1989:52–53).

Yet another approach to the comparison of grammatical patterns involves crosslin-
guistic prototypes, as in Keenan's (1976) article on the notion of subject. Keenan was
specifically concerned with providing a crosslinguistically applicable definition of
'subject' for the purpose of formulating testable universal generalizations, and he listed
over thirty subject properties, proposing that the NP with the greatest number of subject
properties is the subject (of a semantically basic sentence). What Keenan's approach
shares with my proposals is the concern with rigorous crosslinguistically applicable
definitions of comparative concepts, but he says nothing about language-particular
analysis, apparently implying that his subject concept should also be used for analyzing
individual languages (cf. also Comrie 1988, Givón 1995:Ch. 6). But as Dryer (1997:§7)
notes, grammatical relations in individual languages are usually quite discrete, and the
apparent need for fuzzy prototype concepts arises only when one wants to compare lan-
guages. In definitions of comparative concepts, Keenan-style multifactor lists are in
principle acceptable, but Keenan seems to have based his list on a few languages he
happened to know well, not on any empirically observed clustering of properties.[11] In
general, comparative concepts based on fewer factors seem to have greater chances of
leading to deeper insights.

Another multifactor approach to the definition of comparative concepts is Corbett's
'canonical typology', where one 'define[s] the canonical or best [= clearest, indis-
putable] instance, through a set of converging criteria, and use[s] this point in theoreti-
cal space to locate the various occurring types' (Corbett 2007:8). Like Keenan, Corbett
makes it quite clear that his canonical concepts are specifically designed for typological
comparison and are quite different from language-particular categories. Canonical def-
initions, however, are designed primarily to help us understand how different compara-
tive concepts used by different linguists relate to each other (cf. Corbett 2005:31: 'The

---

[11] By contrast, Dahl's (1985) and Bybee and colleagues' (1994) category types (or 'gram types') are de-
rived from empirical crosslinguistic observations (like Lazard's 1992 'focal notions'). And these authors do
not claim that language-particular categories are nothing more than instantiations of the crosslinguistic types.
Bybee and colleagues (1994:148–49) are quite explicit about this, citing an analogy from phonology: 'to say
that a language has a voiceless dental stop does not by any means exhaust the descriptive possibilities or cap-
ture the native speaker's command of such an element. From language to language we find important differ-
ences'. Thus, in practice this approach is quite similar to mine, and Dahl (1985:34) even makes a notational
distinction between category types and language-specific categories.

canonical approach allows us to clarify some of the conceptual problems and misunderstandings that characterize the problem of agreement'). They do not directly help us draw the line between different phenomena to test universal claims, but when we need to draw a line (as in typological databases such as the Surrey Morphology Group's online databases), users 'can be aware how the data relate to their own conceptions and analyses of the area' (Corbett 2005:21). The canonical approach is thus a kind of metatypology that is fully compatible with everything I say here, but the crucial distinction between descriptive categories and comparative concepts is not made explicit in Corbett's work.

A clear distinction between descriptive categories and comparative concepts along the lines advocated here is drawn by Huddleston and Pullum (2002:31–33). They note that grammatical labels such as 'noun', 'preterite', and 'imperative' are often used simultaneously at the 'language-particular level' and at the 'general level', and they insist that 'the levels need to be distinguished, and approached differently'. Language-particular categories are defined formally, whereas general-level concepts may contain semantic elements (cf. Lyons 1966 for similar ideas that influenced Huddleston). These authors do not propose the distinction as a solution to the problem of crosslinguistic comparison (which they are not concerned with), however, and they do not consider the possibility of general-level concepts without direct counterparts at the language-particular level or vice versa. Lazard's (2005, 2006) approach is even closer to mine in that he not only focuses on the differences between similar descriptive categories across languages ('we cannot use linguistic categories [as a standard for typological comparison] because they are language-specific', Lazard 2005:7), but also regards comparative concepts as a different sort of thing. He proposes that as a point of departure for their research, typologists should use 'arbitrary conceptual frameworks', clearly and explicitly defined concepts that are ultimately based on linguists' intuitions (e.g. 'major biactant construction' in the definition of transitivity, Lazard 2002). These are more or less what I have called comparative concepts, the important point being that the comparative concepts are different from language-specific categories, and that they are 'tools for research, not hypotheses susceptible of being verified or falsified' (Lazard 2005:8). A recent paper that follows Lazard and similarly argues for a distinction between 'emic' descriptive categories and 'etic' comparative concepts is Reesink 2008.

In my own research, I came to feel the need for a clear distinction between comparative concepts and descriptive categories when I got involved in the *Atlas of Pidgin and Creole language structures* (Michaelis et al. 2012). This project is similar to the *World atlas of language structures*, but the data on the seventy-odd languages do not come from published sources, but are being provided by language experts on the basis of a questionnaire. The editors had to make it clear repeatedly that the comparative concepts used in the definitions of the features should not be equated with the language-particular descriptive categories that the contributors were used to working with. Thus, it is when language experts and typologists work together directly that the need to distinguish the two kinds of grammatical notions becomes clearest.

8. CHOOSING THE BEST COMPARATIVE CONCEPTS. As I said earlier, comparative concepts are defined by typologists in such a way that they allow them to capture interesting generalizations, or simply salient different language types. Thus, comparative concepts cannot be right or wrong, at least not in the same sense that innate crosslinguistic categories can be right or wrong. They can only be more or less productive, in that they allow the formulation of more or less interesting subdivisions and generalizations.

For example, Mel'čuk (1988) defines 'ergative case' in such a way that the Dyirbal case for transitive subjects is not an ergative, but an instrumental (because coding the transitive agent is not its main function). Mel'čuk's point is to show that case expression, clausal case-marking patterns, and overall syntactic organization do not always coincide, and his concepts are thoroughly defined and motivated. However, this definition does not allow him to formulate the generalization in 15. Thus, Dixon's definition and Mel'čuk's definition are both well motivated while both have their limitations, and they do not contradict each other at all. It would make no sense to claim that one or the other is the 'right' or even the 'best' definition of the term. (In practice, of course, some definitions have more success than others because they turn out to be more useful for more linguists.)

Another example is tense. A possible generalization about tense is 17 (cf. Tonhauser 2008:340, who comes close to making the claim that it is true). In order to test it, we need a crosslinguistically applicable definition of 'tense (marker)'. Two such definitions are given in 18a,b.

(17) GENERALIZATION: Tense markers only occur on verbs; they do not occur on nouns in any language.

(18) a. DEFINITION: A tense marker is a marker that affects the temporal interpretation of its host.
   b. DEFINITION: A tense marker is a marker that occurs as part of a grammatical paradigm, whose occurrence does not depend on the meaning of its host, that encodes a temporal relation between the host time and utterance time (deictic tense) or another contextually given time (relative tense), that does not encode a state change, and that expresses a temporal meaning that may be anaphorically resolved in discourse.

The two definitions in 18a and 18b are two different comparative concepts proposed by different linguists. The definition in 18a is the tense concept used by Nordlinger and Sadler (2004, 2008), and 18b is the tense concept used by Tonhauser (2007, 2008: 337–38). Since they work with different comparative concepts, it is not surprising that they come to different conclusions: while Nordlinger and Sadler show that tense on nouns (in sense 18a) is attested widely, for example, in Guaraní, and hence 17 is wrong, Tonhauser shows that tense on nouns (in sense 18b) does not occur in Guaraní, so that 17 may well be true. Thus, the apparent controversy about the truth of 17 is in fact a controversy about the usefulness of different definitions of the term 'tense marker'. But Nordlinger and Sadler and Tonhauser fail to fully understand the source of their disagreements because both parties seem to assume that 'tense' is a crosslinguistic category. According to Nordlinger and Sadler, '[t]he difference between the two positions amounts to whether or not it is appropriate to consider these nominal temporal markers to be instances of the morphosyntactic category of tense' (2008:328). They cannot give a justification for their view of how 'tense' should be defined, and more generally, they ask, but cannot answer, the question: '[W]hen confronted with unfamiliar or previously undescribed linguistic phenomena, how do we know when to establish a new category to account for it [as done by Tonhauser, who regards the Guaraní markers as being "in a category of their own"], and when to redefine an existing one [as done by Nordlinger & Sadler]?' (2008:329).

Given categorial universalism, this question is inevitable but unanswerable, whereas on categorial particularism, it does not arise: when confronted with a previously undescribed phenomenon, we create a descriptive category for it, and if we want to compare

it with other phenomena, we adjust our comparative concepts in such a way as to arrive at the most insightful comparisons and the most interesting generalizations.

Tonhauser's (2008:334–37) defense of her definition amounts to the observation that her narrower definition allows more interesting crosslinguistic generalizations to be stated, and certainly it would be nice if we could maintain the strong claim in 17 rather than having to abandon it. By contrast, Nordlinger and Sadler's broader definition allowed them to cast their net wide and bring together a fair number of phenomena that had previously not come to linguists' attention. Thus, both comparative concepts have proved productive, and on the present approach, the apparent conflict between the authors simply dissolves.

A referee suggests that a laissez-faire approach to comparative concepts is misguided, and that 'ultimately we want the best set' of comparative concepts. But since the choice of comparative concepts depends on one's goals, and comparative linguists will always have a multiplicity of goals, there will never be a single ('standard') list of such concepts.

9. COMPARATIVE CONCEPTS ARE NOT LINGUISTIC CATEGORIES. I make a terminological distinction between comparative CONCEPTS and descriptive CATEGORIES in order to emphasize that there is no taxonomic relationship between them. Thus, it is not simply the case that the Eskimo Relative case is a manifestation of the same crosslinguistic category 'ergative' (cf. §5.7 above). It cannot even be said to instantiate a higher-level 'category type' (cf. Dahl 1985:31), in the sense of a more abstract category of language structure that abstracts away from language-particular idiosyncrasies. Comparative 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).

A referee who professes to be an 'ardent categorial particularist' nevertheless says that she or he 'expects it to be the very job of typologists and theorists to make generalizations about the very things that show up as language-specific categories of finite verbs, relative clauses, and so forth, in different languages'. But comparative concepts often cross-cut linguistic categories. The Eskimo Relative case is used both for agents of transitive clauses and for adnominal possessors, and since there are other structural similarities between clauses and noun phrases, specialists do not regard this as a coincidence (cf. Woodbury 1985). A typologist who sees an 'ergative case' here sees something quite different from the language expert. Similarly, property words and action words both belong to the same word class in Lao (Enfield 2004). A typologist who sees 'adjectives' in the property words takes a perspective that is not the language expert's perspective.

It is thus not helpful to think of the Eskimo Relative case as 'a kind of ergative case', or of the Japanese Noun-Modifying Construction (§5.5) as 'a kind of relative clause', or of English WH-movement (§5.3) as 'a kind of question-word movement'. These language-particular categories have properties that are not only much more specific than the related comparative concept, but are also additionally used in ways that are quite incompatible with the definition of the comparative concept. The Eskimo Relative case would also be 'a kind of genitive case', English WH-movement would also be 'a kind of relative-pronoun movement', and so on. Every linguistic category would instantiate a multiplicity of more abstract 'comparative categories', and a 'comparative category' could be instantiated multiple times in a single language. With such rampant many-to-many relationships, a taxonomic conceptualization, while logically possible, only obscures matters. Comparative concepts are simply a different kind of general notion from linguistic categories.

In practice, typologists do not generalize over the categories of languages, but over properties of languages that they identify regardless of the categories that speakers seem to have internalized and that structural analysis reveals. One occasionally hears complaints about this practice from structurally minded linguists, exemplified by Matthews (1997:199):

> One cannot just look casually at English and French and say that, because, for example, *je l'aime* translates *I love him*, or *à moi* translates *to me*, both languages 'have' a distinction between subjective and objective. For a careful study of either system might establish that these elements do not stand in a bilateral opposition.

Matthews simply presupposes that 'the basis for comparison lies in the initial structural analysis of each particular system' (1997:199), but this is not how typology works. In fact, it cannot work this way because systems are language-particular and do not lend themselves to comparison in an obvious way. Comparative linguists are willing, and indeed forced, to disregard the structural coherence of language-particular systems, and just focus on the properties that they see through the lens of their comparative concepts.

If one assumes that the explanation of language structure must come from the way language-particular categories and systems are arranged (whether the categories and architectures are innate or not), then it would indeed be very surprising if a typological approach that is oblivious to these matters could succeed. But on a functionalist approach (cf. Dryer 1997), there is no problem, because universal properties of languages are expected to mirror universal functional pressures, not universal categories. We know independently that linguistic categories stand in a very indirect relationship to functional motivations.

The independence of comparative concepts from structural systems is not unique to linguistics, but is characteristic of all cross-group comparison in the social sciences. When we compare wedding dresses across cultures (Foster & Johnson 2003), for example, we will find some widely attested patterns that will be related to the universal functional pressures of marriage, but not necessarily to the ('emic') culture-specific conceptualization that can be attributed to the mental representation of culture members and/or that would be found by structural anthropological analysis. Similarly, in comparative law studies, 'comparative concepts enable us to compare legal rules belonging to different legal systems. Examples of comparative concepts are "adoption" and "unjust enrichment", which refer to legal rules within different systems' (van Baer 1998).

Thus, saying that typologists do not generalize over linguistic categories implies that typology and language-particular analysis are more distinct enterprises than many linguists seem to think (cf. also Haspelmath 2004). Both need each other's research results, but both work with theoretical tools that the other can largely disregard.

10. CONCLUSION. In this discussion note I have shown how crosslinguistic comparison is possible if one adopts the position of categorial particularism, that is, that grammatical categories cannot be equated across languages. Each language has its own categories because the criteria by which the categories are defined (or recognized by learners) are themselves language-particular. A language-particular category set up by a linguist to account for observed speaker behavior is called a descriptive category. Comparative linguists create comparative concepts against which the descriptive categories of particular languages can be matched. These comparative concepts must be universally applicable; that is, they must be based exclusively on more primitive universally applicable concepts: universal conceptual-semantic concepts, general formal concepts, and other comparative concepts (or on extralinguistic situations). This approach has been widely practiced by comparative linguists working in the Greenbergian tradition,

even though not all of them have explicitly adopted categorial particularism and the distinction between descriptive categories and comparative concepts has generally been implicit.

This is very different from the categorial universalism that has been very widely adopted since the 1960s. In the generative approach, it is assumed that the categories (or features) in different languages can be equated, and linguists work with a set of crosslinguistic categories. This would seem to make language comparison more straightforward, but in fact the opposite is the case. On the categorial-universalist view, language-particular analysis is not possible without a clear picture of UG, and language comparison is not possible without a thorough analysis of the languages that are compared. Thus, linguists who work with crosslinguistic categories rarely study larger language samples, and the great majority of testable empirical universals (like those listed in §5, and like the over 2,000 universals in *The universals archive*[12]) have been discovered by linguists who work with comparative concepts as described here.

On the approach defended here, analysts of particular languages do not depend on the theoretical results of typologists, and typologists do not directly compare the theoretical results (i.e. the categories and rules of languages) obtained by language analysts. The analysis of particular languages and the comparison of languages are thus independent of each other as theoretical enterprises. But as the practice of fruitful interaction among typologists and descriptive linguists shows, in practical terms these two kinds of approaches are just part of a larger, coherent endeavor, that of documenting and understanding linguistic diversity.

REFERENCES

AIKHENVALD, ALEXANDRA Y. 2004. The adjective class in Tariana. *Adjective classes: A crosslinguistic typology*, ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald, 97–124. Oxford: Oxford University Press.

BACH, EMMON. 2004. Linguistic universals and particulars. *Linguistics today—Facing a greater challenge*, ed. by Piet van Sterkenburg, 47–60. Amsterdam: John Benjamins.

BAKER, MARK C. 2003. *Lexical categories*. Cambridge: Cambridge University Press.

BAKER, MARK C. 2008. The macroparameter in a microparametric world. In Biberauer, 351–73.

BHAT, D. N. S. 2000. The indefinite-interrogative puzzle. *Linguistic Typology* 4.365–400.

BIBERAUER, THERESA (ed.) 2008. *The limits of syntactic variation*. Amsterdam: John Benjamins.

BOAS, FRANZ. 1911. Introduction. *Handbook of American Indian languages*, vol. 1. (Bureau of American Ethnology, bulletin 40.) Washington, DC: Government Printing Office.

BUTT, JOHN, and CARMEN BENJAMIN. 2000. *A new reference grammar of Modern Spanish*. 3rd edn. London: Arnold.

BYBEE, JOAN L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

BYBEE, JOAN L., and ÖSTEN DAHL. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13.51–103.

BYBEE, JOAN L.; WILLIAM PAGLIUCA; and REVERE D. PERKINS. 1994. *The evolution of grammar*. Chicago: University of Chicago Press.

CHANG, SUK-JIN. 1996. *Korean*. Amsterdam: John Benjamins.

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

CHOMSKY, NOAM. 1981. *Lectures on government and binding*. Dordrecht: Foris.

CINQUE, GUGLIELMO. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

COMRIE, BERNARD. 1976. *Aspect*. Cambridge: Cambridge University Press.

[12] http://typo.uni-konstanz.de/archive

COMRIE, BERNARD. 1988. Passive and voice. *Passive and voice*, ed. by Masayoshi Shibatani, 9–23. Amsterdam: John Benjamins.

COMRIE, BERNARD. 1989. *Language universals and linguistic typology*. 2nd edn. Oxford: Blackwell.

COMRIE, BERNARD. 1998. Rethinking the typology of relative clauses. *Language Design* 1.59–86.

COMRIE, BERNARD, and KAORU HORIE. 1995. Complement clauses versus relative clauses: Some Khmer evidence. *Discourse grammar and typology: Papers in honor of John W. M. Verhaar* (Studies in language companion series 27), ed. by Werner Abraham, T. Givón, and Sandra A. Thompson, 65–75. Amsterdam: John Benjamins.

COMRIE, BERNARD, and TANIA KUTEVA. 2005. Relativization strategies. In Haspelmath et al., 494–501.

CORBETT, GREVILLE G. 2000. *Number*. Cambridge: Cambridge University Press.

CORBETT, GREVILLE G. 2005. The canonical approach in typology. *Linguistic diversity and language theories* (Studies in language companion series 72), ed. by Zygmunt Frajzyngier, Adam Hodges, and David S. Rood, 25–49. Amsterdam: John Benjamins.

CORBETT, GREVILLE G. 2007. Canonical typology, suppletion, and possible words. *Language* 83.8–42.

CRISTOFARO, SONIA. 2009. Grammatical categories and relations: Universality vs. language-specificity and construction-specificity. *Language and Linguistics Compass* 3.441–79.

CROFT, WILLIAM. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

CROFT, WILLIAM. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

CROFT, WILLIAM. 1995. Modern syntactic typology. *Approaches to language typology*, ed. by Masayoshi Shibatani and Theodora Bynon, 85–144. Oxford: Clarendon.

CROFT, WILLIAM. 2000. Parts of speech as language universals and as language-particular categories. *Approaches to the typology of word classes*, ed. by Petra M. Vogel and Bernard Comrie, 65–102. Berlin: Mouton de Gruyter.

CROFT, WILLIAM. 2001. *Radical construction grammar*. Oxford: Oxford University Press.

CROFT, WILLIAM. 2003. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press.

CROFT, WILLIAM. 2009. Methods for finding language universals in syntax. *Universals of language today*, ed. by Sergio Scalise, Elisabetta Magni, and Antonietta Bisetto, 145–64. Dordrecht: Springer.

CULICOVER, PETER W. 1999. *Syntactic nuts*. Oxford: Oxford University Press.

DAHL, ÖSTEN. 1985. *Tense and aspect systems*. Oxford: Blackwell.

DE CAT, CÉCILE. 2005. French subject clitics are not agreement markers. *Lingua* 115. 1195–219.

DE LA GRASSERIE, RAOUL. 1896. De l'article: Morphologie et syntaxe. *Mémoires de la Société de Linguistique de Paris* 9.285–322, 381–94.

DIESSEL, HOLGER. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: John Benjamins.

DIXON, R. M. W. 1979. Ergativity. *Language* 55.59–138.

DIXON, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

DIXON, R. M. W. 2004. Adjective classes in typological perspective. *Adjective classes: A cross-linguistic typology*, ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald, 30–83. Oxford: Oxford University Press.

DIXON, R. M. W. 2010. *Basic linguistic theory*, vols. 1 and 2. Oxford: Oxford University Press.

DRYER, MATTHEW S. 1988. Object-verb order and adjective-noun order: Dispelling a myth. *Lingua* 74.77–109.

DRYER, MATTHEW S. 1992. The Greenbergian word order correlations. *Language* 68.81–138.

DRYER, MATTHEW S. 1997. Are grammatical relations universal? *Essays on language function and language type*, ed. by Joan L. Bybee, John Haiman, and Sandra A. Thompson, 115–43. Amsterdam: John Benjamins.

DRYER, MATTHEW S. 2005a. Order of adjective and noun. In Haspelmath et al., 354–57.

DRYER, MATTHEW S. 2005b. Position of interrogative phrases in content questions. In Haspelmath et al., 378–81.

Enfield, N. J. 2004. Adjectives in Lao. *Adjective classes: A cross-linguistic typology*, ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald, 323–47. Oxford: Oxford University Press.

Farrar, Scott, and D. Terence Langendoen. 2003. A linguistic ontology for the semantic web. *GLOT International* 7.97–100.

Foster, Helen, and Donald Clay Johnson (eds.) 2003. *Wedding dress across cultures*. Oxford: Berg.

Gianollo, Chiara; Cristina Guardiano; and Giuseppe Longobardi. 2008. Three fundamental issues in parametric linguistics. In Biberauer, 109–42.

Gil, David. 2001. Escaping Eurocentrism: Fieldwork as a process of unlearning. *Linguistic fieldwork*, ed. by Paul Newman and Martha Ratliff, 102–32. Cambridge: Cambridge University Press.

Givón, Talmy. 1995. *Functionalism and grammar*. Amsterdam: John Benjamins.

Givón, Talmy. 2001. *Syntax: An introduction*. 2 vols. Amsterdam: John Benjamins.

Goddard, Cliff. 1982. Case systems and case markings in Australian languages: A new interpretation. *Australian Journal of Linguistics* 2.167–96.

Good, Jeff. 2004. The descriptive grammar as a (meta)database. *Proceedings of E-MELD 2004.* Online: http://emeld.org/workshop/2004/jcgood-paper.html.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of grammar*, ed. by Joseph H. Greenberg, 73–113. Cambridge, MA: MIT Press.

Greenberg, Joseph H. 1974. *Language typology: A historical and analytic overview*. The Hague: Mouton.

Haspelmath, Martin. 1993. *A grammar of Lezgian*. (Mouton grammar library 9.) Berlin: Mouton de Gruyter.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2004. Does linguistic explanation presuppose linguistic description? *Studies in Language* 28.554–79.

Haspelmath, Martin. 2007. Pre-established categories don't exist—Consequences for language description and typology. *Linguistic Typology* 11.119–32.

Haspelmath, Martin. 2008a. Parametric versus functional explanations of syntactic universals. In Biberauer, 75–107.

Haspelmath, Martin. 2008b. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6.40–63.

Haspelmath, Martin. 2010. Framework-free grammatical theory. *The Oxford handbook of grammatical analysis*, ed. by Bernd Heine and Heiko Narrog, 341–65. Oxford: Oxford University Press.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, to appear.

Haspelmath, Martin; Matthew S. Dryer; David Gil; and Bernard Comrie (eds.) 2005. *The world atlas of language structures*. Oxford: Oxford University Press.

Haspelmath, Martin, and Uri Tadmor (eds.) 2009a. *Loanwords in the world's languages: A comparative handbook*. Berlin: Mouton de Gruyter.

Haspelmath, Martin, and Uri Tadmor (eds.) 2009b. *World loanword database*. Munich: Max Planck Digital Library. Online: http://wold.livingsources.org/.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75.244–85.

Heger, Klaus. 1990/91. Noeme als Tertia Comparationis im Sprachvergleich. *Romanistisches Jahrbuch* 49/50.6–30.

Hengeveld, Kees, and J. Lachlan Mackenzie. 2008. *Functional discourse grammar: A typologically-based theory of language structure*. Oxford: Oxford University Press.

Huddleston, Rodney D., and Geoffrey K. Pullum (eds.) 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

Julien, Marit. 2002. *Syntactic heads and word formation*. New York: Oxford University Press.

Keenan, Edward L. 1976. Towards a universal definition of 'subject'. *Subject and topic*, ed. by Charles N. Li, 303–33. New York: Academic Press.

KEENAN, EDWARD L., and BERNARD COMRIE. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.63–99.

KROEGER, PAUL R. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.

LADD, D. R. 2011. Phonetics in phonology. *Handbook of phonological theory*, 2nd edn., ed. by John Goldsmith, Jason Riggle, and Alan Yu. Oxford: Wiley-Blackwell, to appear.

LAZARD, GILBERT. 1992. Y a-t-il des catégories interlangagières? *Texte, Sätze, Wörter und Moneme: Festschrift für Klaus Heger*, ed. by Susanne Anschütz, 427–34. Heidelberg: Heidelberger Orientverlag. [Reprinted in *Études de linguistique générale*, by Gilbert Lazard. Leuven: Peeters, 2001.]

LAZARD, GILBERT. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36.141–90.

LAZARD, GILBERT. 2005. What are we typologists doing? *Linguistic diversity and language theories* (Studies in language companion series 72), ed. by Zygmunt Frajzyngier, Adam Hodges, and David S. Rood, 1–23. Amsterdam: John Benjamins.

LAZARD, GILBERT. 2006. *La quête des invariants interlangues: La linguistique est-elle une science?* Paris: Champion.

LEHMANN, CHRISTIAN. 1989. Language description and general comparative grammar. *Reference grammars and modern linguistic theory*, ed. by Gottfried Graustein and Gerhard Leitner, 133–62. Tübingen: Niemeyer.

LYONS, JOHN. 1966. Towards a 'notional' theory of the 'parts of speech'. *Journal of Linguistics* 2.209–36.

LYONS, JOHN. 1981. *Language and linguistics: An introduction*. Cambridge: Cambridge University Press.

MADDIESON, IAN. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

MALING, JOAN. 2001. Dative: The heterogeneity of the mapping among morphological case, grammatical functions, and thematic roles. *Lingua* 111.419–64.

MALLINSON, GRAHAM, and BARRY BLAKE. 1981. *Language typology*. Amsterdam: North Holland.

MASLOVA, ELENA. 2003. *A grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter.

MATSUMOTO, YOSHIKO. 1988. Semantics and pragmatics of noun-modifying constructions in Japanese. *Berkeley Linguistics Society* 14.166–75.

MATSUMOTO, YOSHIKO. 1997. *Noun-modifying constructions in Japanese: A frame-semantic approach*. Amsterdam: John Benjamins.

MATTHEWS, PETER H. 1997. Structural linguistics in the 1990s. *Lingua* 100.193–203.

MCCAWLEY, JAMES D. 1992. Justifying part-of-speech distinctions in Mandarin Chinese. *Journal of Chinese Linguistics* 20.211–46.

MEL'ČUK, IGOR A. 1988. Toward a definition of the concept 'ergative construction'. *Dependency syntax: Theory and practice*, ed. by Igor A. Mel'čuk, 250–63. Albany: SUNY Press.

MICHAELIS, SUSANNE; PHILIPPE MAURER; MARTIN HASPELMATH; and MAGNUS HUBER (eds.) 2012. *Atlas of Pidgin and Creole language structures*. Oxford: Oxford University Press, to appear.

MIELKE, JEFF. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

MOHANAN, K. P.; DIANA ARCHANGELI; and DOUGLAS PULLEYBLANK. 2010. The emergence of optimality theory. *Reality exploration and discovery: Pattern interaction in language and life*, ed. by Linda Uyechi and Lian Hee Wee, 143–58. Stanford, CA: CSLI Publications.

MUSHIN, ILANA. 1995. Epistememes in Australian languages. *Australian Journal of Linguistics* 15.1–31.

NEDJALKOV, VLADIMIR P., and VIKTOR P. LITVINOV. 1995. The St. Petersburg/Leningrad typology group. *Approaches to language typology*, ed. by Masayoshi Shibatani and Theodora Bynon, 215–71. Oxford: Clarendon.

NEWMEYER, FREDERICK J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.

NEWMEYER, FREDERICK J. 2004. Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4.181–234.

NEWMEYER, FREDERICK J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.

NEWMEYER, FREDERICK J. 2007. Linguistic typology requires crosslinguistic formal categories. *Linguistic Typology* 11.133–57.

NORDLINGER, RACHEL, and LOUISA SADLER. 2004. Nominal tense in crosslinguistic perspective. *Language* 80.776–806.

NORDLINGER, RACHEL, and LOUISA SADLER. 2008. When is a temporal marker not a tense? Reply to Tonhauser 2007. *Language* 84.325–31.

PAYNE, THOMAS E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

PORT, ROBERT F., and ADAM P. LEARY. 2005. Against formal phonology. *Language* 81. 927–64.

PRASITHRATHSINT, AMARA. 2000. Adjectives as verbs in Thai. *Linguistic Typology* 4.251–71.

QUIRK, RANDOLPH; SIDNEY GREENBAUM; GEOFFREY LEECH; and JAN SVARTVIK. 1985. *A comprehensive grammar of the English language*. London: Longman.

REESINK, GER P. 2008. Lexicon and syntax from an emic viewpoint. *Studies in Language* 32.866–93.

RIJKHOFF, JAN. 2009. On the (un)suitability of semantic categories. *Linguistic Typology* 13.95–104.

RIJKHOFF, JAN. 2010. Formal, semantic and functional categories in grammatical theory and linguistic typology. Aarhus: University of Aarhus, MS.

SCHACHTER, PAUL. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above. *Subject and topic*, ed. by Charles N. Li, 491–518. New York: Academic Press.

SONG, JAE JUNG. 2001. *Linguistic typology: Morphology and syntax*. Harlow: Pearson Education.

SONG, JAE JUNG (ed.) 2011. *The Oxford handbook of language typology*. Oxford: Oxford University Press, to appear.

STASSEN, LEON. 1985. *Comparison and universal grammar*. Oxford: Blackwell.

STASSEN, LEON. 2011. The problem of cross-linguistic identification. In Song 2011, to appear.

TALLERMAN, MAGGIE. 1998. *Understanding syntax*. London: Arnold.

TOMASELLO, MICHAEL. 2005. Beyond formalities: The case of language acquisition. *The Linguistic Review* 22.183–97.

TONHAUSER, JUDITH. 2007. Nominal tense? The meaning of Guaraní nominal temporal markers. *Language* 83.831–69.

TONHAUSER, JUDITH. 2008. Defining crosslinguistic categories: The case of nominal tense (Reply to Nordlinger and Sadler). *Language* 84.332–42.

TRUBETZKOY, NIKOLAI. 1939. *Grundzüge der Phonologie*. (Travaux du Cercle Linguistique de Prague 7.) Prag: Harrassowitz.

ULTAN, RUSSELL. 1978. The nature of future tenses. *Universals of human language*, vol. 3, ed. by Joseph H. Greenberg, 83–124. Stanford, CA: Stanford University Press.

VAN BAER, CONRAD J. P. 1998. The applicability of comparative concepts. *Electronic Journal of Comparative Law* 2.2.

VAN VALIN, ROBERT D., JR. 2005. *Exploring the syntax-semantic interface*. Cambridge: Cambridge University Press.

VON DER GABELENTZ, HANS CONON. 1861. Über das Passivum. *Abhandlungen der Königlichen Sächsischen Akademie der Wissenschaften, Philologisch-historische Klasse* 3.449–546.

VON HUMBOLDT, WILHELM. 1830. Über den Dualis. *Abhandlungen der historisch philologischen Klasse der Königlichen Akademie der Wissenschaften zu Berlin* 1827.161–87.

WHALEY, LINDSAY J. 1997. *Introduction to typology: The unity and diversity of language*. Thousand Oaks: Sage.

WIERZBICKA, ANNA. 1980. *Lingua mentalis: The semantics of natural language*. Sydney: Academic Press.

WOODBURY, ANTHONY C. 1985. Noun phrase, nominal sentence, and clause in Central Alaskan Yupik Eskimo. *Grammar inside and outside the clause*, ed. by Johanna Nichols and Anthony C. Woodbury, 61–88. Cambridge: Cambridge University Press.

ZWICKY, ARNOLD. 1987. Suppressing the Z's. *Journal of Linguistics* 23.133–48.

Max Planck Institute for Evolutionary Anthropology
Deutscher Platz 6
D-04105 Leipzig, Germany
[haspelmath@eva.mpg.de]