

Towards standardization of morphosyntactic terminology for general linguistics

MARTIN HASPELMATH
MPI-EVA Leipzig and Leipzig University

pp. 35-57 in:

Alfieri, Luca & Arcodia, Giorgio Francesco & Ramat, Paolo (eds.). *Linguistic categories, language description and linguistic typology* (Typological Studies in Language 132). *tsl.132* (Typological Studies in Language 132). Amsterdam: Benjamins. (<https://benjamins.com/catalog/tsl.132>)

Abstract: This paper proposes that just like phonologists, linguists working on morphosyntax should have a core set of standard terms that are understood in exactly the same way across the discipline. Most of these terms are traditional terms that are given a standard retro-definition, because linguists already behave as if these terms had the same meaning for everyone. The definitions are definitions of general concepts (i.e. comparative concepts, applicable to all languages in exactly the same way), but they are expected to be highly similar to language-particular categories with the same labels. If linguists were close to finding out the true natural-kind categories of Human Language that all grammars consist of, there would be no need for definitions, but since this seems to be a remote goal, research on general linguistics must rely on uniformly defined general terms.

1. Terminological consistency and standardization

Standardization has proved highly beneficial in many domains of technology and trade, and it is impossible to imagine today's world without worldwide standards for basic units of time and space (the International System of Units, the Coordinated Universal Time, and others).

In addition to such practical standards that ensure technological interoperability, there are also standard nomenclatures in many fields of science that affect how scientists talk about their subject domain. Biologists have had generally accepted conventions for naming species since the 18th century, and chemists started to organize terminological standards in the 1860s. Linguists have had a standard for representing general phonetic categories of segments since the 1890s: what is now known as the International Phonetic Alphabet (IPA).

In this paper, I would like to make the case for a morphosyntactic counterpart of the IPA: a standard set of morphosyntactic terms for general linguistics. If such a standard were possible, its advantages should be evident to every linguist who has had more than a few years of experience in the field. Many terms in linguistics are used in a variety of ways that are often confusing, and unless one is a specialist in a particular area, one may be unaware of these ambiguities. As a result, automatic literature search is often problematic or impossible, and unfortunately, there is also quite a bit of incomprehension and talking past each other.

A widespread attitude in the field seems to be that the difficulties of our subject matter – the enormous complexities of the many different language systems that linguists are grappling with – make it impossible to have a standard terminology, at least

at the present stage of our knowledge. It is this attitude that I would like to challenge in the present paper. One of the reasons for being more optimistic is that I have observed not only big difficulties, but also quite a bit of sloppy terminological use by linguists. Grammatical terms often change their meanings through a novel use of an existing term that is primarily motivated by the desire to avoid coining a new term. For example, the term *oblique* used to refer to all non-nominative cases (there was the nominative case and the oblique cases). Since the 1970s, however, it has come to be used for all cases apart from the nominative and the accusative, as well as (more generally) for all cases and adpositions that are used for arguments other than the transitive subject and direct object (e.g. Nichols 1984). There was no particular reason for this change, and it would have been easy to coin a new term instead. The confusion generated by the semantic shift of the term *oblique* was thus unrelated to any particular difficulties, and entirely due to the attitude that it does not matter much if an older term is used with a new meaning. And indeed, within a given narrow context, it is often fairly clear what a term means. But from the broader perspective of the entire field of the language sciences, the current level of terminological unclarity and ambiguity is undesirable.

Thus, while there are no doubt many difficulties in the field of general morphosyntax, this need not make it impossible to have a standard terminology. Other fields have difficulties as well (including the phoneticians, who have the IPA), but many fields of science are at least making an effort to have terminological standards. I thus want to argue that the field of general linguistics should have at least a limited set of standard terms (say, a few dozen, comparable to the 107 IPA letters).

It is important to note that my proposal in this paper concerns exclusively terminology for general linguistics, and that I say nothing about language-particular terms here. There are many phenomena that can be talked about only at the level of a particular language, e.g. the French Passé Surcomposé (as in *elle l'a eu vue* 'she had seen her'), or the Genitive Absolute in Ancient Greek, or even the "f-word" in English. There are no counterparts of these phenomena in most other languages, so they are not directly relevant to general linguistics. Terminology of the language-particular kind can perhaps be standardized as well,¹ but in this paper, I will confine myself to terms of general linguistics – in other words, to terms for comparative concepts.

The notion of COMPARATIVE CONCEPTS in grammar has been fairly widely adopted since I first coined the term in 2010 (see Haspelmath 2010; 2018a; Brown & Chumakina 2013; Croft 2016; Dryer 2016), but it seems that many linguists are still not fully clear about it. I will therefore briefly explain the distinction between comparative concepts and language-particular categories in the next section, and contrast both with the notion of (innate) natural-kind categories.

2. Comparative concepts, language-particular categories, and natural kinds

Confusingly, grammatical terms such as *complementizer* or *consonant* are used in three different senses in the literature:

¹ In view of the proliferation of grammatical terminology in different (West) German school textbooks in the 1970s, politicians entrusted some linguistics professors with the task of setting up a standard set of terms for German, English, French and Latin grammar for use in secondary schools. This work has been ongoing for quite a while (see Hennig 2012), showing that grammatical terminology can also be relevant to applied concerns such as teaching of grammar and languages in schools.

- as (category-like) comparative concepts
- as language-particular categories
- as natural-kind categories

A category-like comparative concept is a term that can be applied to any language and that is identified in all languages by the same criteria.² For example, a consonant can be defined as a sound segment that is articulated with at least partial closure of the vocal tract, and a complementizer can be said to be a marker that indicates that the clause in which it occurs is a complement clause. Comparative concepts of this kind are needed for universal claims, e.g. about the position of the complementizer within a complement clause (Dryer 2009). The comparative definition of a complementizer merely says that it is a marker (not that it is a word), because many languages have elements that seem to be part of the verbal morphology but function just like the English word *that* or the Italian word *che*. And in fact, English has a marker like this as well: the suffix *-ing*. This is used as a complementizer in cases like *She considered leav-ing*. Thus, from a comparative perspective, English has not only clause-initial complementizer words, but also an affixed complementizer that is postposed to the verb.

Moving now on to language-particular terms, we can talk about English Complementizers and Latin Consonants (with capitalization of unique entities), but these categories are defined differently from the comparative concepts. In English grammar, a Complementizer is generally said to be a word, so that *-ing* is not considered as a Complementizer. And what counts as a word (or rather, an English Word) is determined by English-specific criteria. Likewise, what counts as a Latin Consonant is determined by Latin-specific criteria. Thus, the semivowels [j] and [w] are usually treated as Consonants for the purposes of (Latin-specific) phonotactics, even though by the phonetic criterion, they would not be consonants because there is no closure of the vocal tract.

Language-particular terms are sometimes written with capitalization, in order to distinguish them clearly from comparative terms (Comrie 1976; Haspelmath 2010: §5), but most of the time, linguists rely on context to make it clear whether they are talking about concepts of general linguistics or about language-particular categories.

Both comparative concepts and language-particular categories must be defined in a precise way. It is sometimes thought that a prototype definition is sufficient for comparative purposes. For example, Dingemanse (2019: 20) says that “typological definitions generally aim to capture the centre of gravity of a phenomenon rather than providing a list of necessary and sufficient properties”, but this is not true. If there were no list of necessary and sufficient properties, it could not be clear whether a phenomenon should be grouped under a comparative concept or not, and it would not be possible to make a cross-linguistic database that records the properties of a sample of languages in terms of the comparative concept.

However, comparative concepts need not be comprehensive, because language comparison is always partial (only language description must ultimately be complete). Languages have many structures that can be compared with similar structures in other languages, but they often have completely unique phenomena that are not amenable to

² A comparative concept need not be category-like. A nonverbal stimulus or a text passage in a parallel text can also be a comparative concept (such concepts are called *etic comparative concepts*; Haspelmath 2018a: 87-88). Another type of comparative concept is the standard lexical meaning, as in the *Concepticon* (List et al. 2019). Such comparative concepts are never confused with descriptive categories, probably because of their specific names.

comparison (e.g. German Weak vs. Strong adjectives, or the Latin Attributive Gerundive construction). Thus, it is possible to have definitions of category-like comparative concepts that focus on a shared core, where the corresponding descriptive categories must be extensionally broader (see the discussion in §5 below). This may give the wrong impression that the comparative concept is vague or covers only a “prototype” (like the definition of *noun* in (5) below).

In addition to comparative concepts and language-particular categories, many linguists also work with natural-kind categories. These are different from language-particular categories in that they occur not only in a single language, but are assumed to be potentially applicable to any language. This is so because they are thought to be innately given (as part of a genetically determined grammar blueprint called “universal grammar”), in advance of language learning (as “pre-established categories”, cf. Haspelmath 2007). Thus, it is often claimed that there is a natural-kind category complementizer (often written COMP or simply C), and that different languages instantiate it in different ways (often by zero, and sometimes even by movement of a verb). Similarly, consonants (often written “[+cons]”) can be thought of as an innate natural-kind category of phonology, instantiated in different ways (often by zero, as in “CV phonology”, Clements & Keyser 1983, or by non-movement in sign languages, Brentari 2002).

Natural-kind categories are very different from comparative concepts and language-particular categories, because they are not instruments for research, but are hypothetical results of research. They have the status of chemical elements in 18th century chemistry, before scientists had figured out what underlies the diversity of chemical compounds (cf. Baker (2001) on the similarities between natural-kind based generative typology and chemistry). Many linguists are skeptical of the natural-kind approach, and even those who assume the basic correctness of the approach admit that the goal of discovering the true natural kinds of the grammar blueprint is still distant. Thus, it would be premature to talk about standardization of natural-kind categories, because we know too little about them, if they exist at all.³

In some other fields, the standard terminology concerns natural-kind categories, e.g. in chemistry (whose elements and compounds are natural kinds) and in biology (whose species are also often considered as natural kinds). And still other fields, such as economics or political science, have no natural-kind categories at all, as far as I am aware. Categories such as income, tax, state, and government are social categories, and nobody would suggest that they are innate properties of the human mind.⁴ Thus, there are three types of disciplines:

- (i) social sciences, which have culture-specific social categories, but no natural-kind categories
- (ii) natural sciences, which have natural-kind categories but no social categories

³ It seems to me that those linguists who think that standardization of grammatical terminology is too difficult often have natural-kind categories in mind, and if so, I fully agree with them. Many linguists do not distinguish clearly between comparative concepts and natural-kind categories, and this may account for the skeptical attitude toward standardization of grammatical terms.

⁴ This might be different with kinship categories like ‘mother’ or ‘brother’, which might conceivably be innate categories, as they seem to be shared with other mammals that have kinship organization, but nothing like humans’ complex culture.

(iii) sciences at the intersection of social and natural sciences such as linguistics, which have both social categories (such as English Complementizer and Latin Consonant) and natural-kind categories (though the latter are controversial in linguistics)

But importantly, ALL sciences have (observer-made) comparative concepts in addition to the independently existing categories (social or natural) that they encounter in the world.

In the natural and social sciences, comparative concepts are not so easily confused with the independently existing categories, so the distinction is not highlighted frequently. But some well-known nomenclature systems in the natural sciences are systems of conventional comparative concepts, e.g. the Yerkes spectral classification of stars in astronomy,⁵ or the International Cloud Atlas Classification of clouds in meteorology.⁶ Likewise, the social sciences have some well-known nomenclature systems of conventional comparative concepts, e.g. the Hornbostel-Sachs system of musical instrument classification in the comparative anthropology of music,⁷ or the Human Development Index in comparative development economics. These systems of comparative concepts exist alongside culture-specific categories, and there is no danger of confusing them.

3. Examples of possible standard definitions of well-known terms

Before continuing the discussion of issues arising in standardization of terms, let us now look at a few concrete examples of terms and their definitions that I think might be suitable for standardization. Naturally, many of these are taken from my earlier or current work.

(1) sentence⁸

A sentence is a maximal clause, i.e. a clause that is not part of another clause.

(2) clause⁹

A clause is a combination of a predicate (full verb or nonverbal predicate) and its arguments plus modifiers.

(3) morph⁹

A morph is a minimal form, i.e. a form that does not consist of other forms.

(4) root¹⁰

A root is a morph that denotes a thing, an action or a property.

⁵ https://en.wikipedia.org/wiki/Stellar_classification#Yerkes_spectral_classification

⁶ https://en.wikipedia.org/wiki/International_Cloud_Atlas

⁷ <https://en.wikipedia.org/wiki/Hornbostel%E2%80%93Sachs>

⁸ <https://dlc.hypotheses.org/1725>

⁹ Haspelmath (2020)

¹⁰ Haspelmath (2012)

- (5) noun¹¹
A noun is a morph (or a root) that denotes a thing, i.e. an object or a person.
- (6) affix¹²
An affix is a non-promiscuous bound form that is not a root.
- (7) bound form¹³
A bound form is a form that cannot occur in isolation.
- (8) marker¹⁴
A marker is a bound form that is not a root.
- (9) A-argument¹⁵
The A-argument is the argument of a two-participant clause that is coded like the ‘breaker’ or ‘killer’ argument of ‘break/kill’, if the other argument (the P-argument) is coded like the ‘broken thing’/‘killed animal’.
- (10) subject¹⁶
The subject of a clause is its A-argument or its S-argument.
- (11) transitive clause¹²
A transitive clause is a clause that has an A-argument and a P-argument.
- (12) ergative construction¹⁷
An ergative construction is a construction with a transitive verb in which the P-argument is coded like the intransitive S-argument, and the A-argument is coded differently.
- (13) argument coding
Argument coding is the marking of an argument’s semantic or syntactic role by means of a flag or a person index.
- (14) passive construction
A passive construction is a construction (i) which shares the verb root with the transitive construction, (ii) whose S-argument corresponds to the transitive P-argument, and (iii) which requires oblique flagging of the

¹¹ A colleague has suggested that noun may be better defined as ‘a root that can be case-marked’, but what does it mean to be case-marked? As seen in (16) and (15) below, ‘case-marker’ can be defined with reference to ‘flag’, which is defined with reference to ‘nominal’. Maybe the latter term can be defined without reference to ‘noun’, but I leave this open here.

¹² Haspelmath (2018b)

¹³ Haspelmath (2013: 2012-213); <https://dlc.hypotheses.org/1779>

¹⁴ A reviewer notes that Pollard & Sag (1994:44-45) provide an exclusively semantic definition of *marker* (“a word whose semantic content is purely logical in nature”), which is perhaps more in line with linguists’ intuitions. But I do not think that a free form would be regarded as a marker, and the “non-root” part of my definition amounts to much the same (see (4)), except that it is easier to apply than a vague concept like “purely logical”.

¹⁵ Lazard (2002)

¹⁶ Haspelmath (2011)

¹⁷ Comrie (1978)

argument corresponding to the transitive A-argument, if it can be expressed at all.

(15) flag¹⁸

A flag is a bound morph that occurs with a nominal and that expresses its semantic role.

(16) case-marker

A case-marker is a flag that is an affix.

(17) serial verb construction¹⁹

A serial verb construction is a monoclausal construction consisting of multiple independent verbs with no element linking them and with no predicate–argument relation between the verbs.

(18) reflexive construction²⁰

A reflexive construction is a grammatical construction

(i) that can only be used when two participants of a clause are coreferential

(ii) and that contains a special form (a reflexivizer) that signals this coreference.

(19) gender system²¹

A gender system is a nomifier system with up to 20 nomifier classes (= gender classes) whose nomifiers are not restricted to occurring on numerals and possibly other adnominal modifiers, or restricted to occurring on possessors.

(20) nomifier system

A nomifier system is a paradigm of grammatical markers which occur on noun-associated forms and each of which expresses (partly reflects, or partly contributes) a broad property of the selecting noun other than person and number.

Several of these definitions will look strange to experienced readers, and many linguists will find it easy to raise objections. But my proposal here is not that these definitions (which happen to be the ones that I use in my own work) should become standard. And it seems that those colleagues who criticize the definitions generally find it less easy to come up with more appropriate definitions, so I have decided to list them anyway here. The main purpose of this article is to make a general case for standard definitions of terms, and by giving some definitions of some basic terms, I provide a proof of concept

¹⁸ Haspelmath (2005); (2019)

¹⁹ Haspelmath (2016)

²⁰ Haspelmath (2021)

²¹ Haspelmath (2018c). I have been asked whether 20 classes is not an arbitrary limit, and indeed it is, but this number is sometimes mentioned in the literature (e.g. Corbett 2007: 242: “the number of genders is not limited to two, nor to three: four is common and twenty is possible”). This arbitrary definition is given here in order to show that some traditional terms cannot be retro-defined in a way that appears natural. But whether a clearly defined concept is “natural” (and what this might mean) is not a question that I address in this paper.

how this can be done concretely, even in difficult cases. (More specific terms such as “free relative clause” or “subsecutive adjective” are presumably easier to define than the basic terms.)

4. Principles for standard morphosyntactic terms

After having seen some concrete examples, let us consider a number of general principles for choosing terms and definitions.

First of all, we would like to have standard definitions of well-known grammatical terms, so many of the definitions in §3 concern well-known terms. This is completely analogous to the IPA, which provides standard definitions of the well-known letters of the Latin alphabet. I call such definitions *retro-definitions*, because they assign a precise meaning to an existing term that does not have a widely recognized precise meaning yet. (Some widely known established terms do have a precise meaning, e.g. *interrogative pronoun*, or *concessive clause*; these terms do not need to be retro-defined, because their definition is not in question.)

But this does not imply that widely needed comparative concepts cannot be described by completely novel terms. Just as the IPA includes many novel letters, we need many additional comparative concepts for morphosyntax, and some of the terms in §3 are quite novel (*morph*, *bound form*, *marker*, and *ergative construction* were added in the 20th century, and *flag* and *nomifier* are even more recent additions). Of course, the list is completely open-ended, and any general linguist who feels the need to use a new concept should feel free to coin a new term. Once a new term has been picked up by a certain number of other linguists, it could be added to the list of standard terms. (This is different from the IPA: phoneticians do not seem to think that the list of possible IPA symbols is completely open-ended.)²² I do not expect the list of category-like comparative concepts ever to be complete, because the range of morphosyntactic constructions that might be compared across languages is very large and open-ended.

Each of the definitions in §3 presupposes a number of other terms, which must either be defined in turn (as comparative concepts), or must be assumed to be generally understood in the same way by everyone. Very basic linguistic concepts such as ‘form’, ‘action’, and ‘semantic role’ can (or must) be left undefined (as primitives), and in addition the definitions may of course contain general nontechnical concepts such as ‘not’, ‘part of’, or ‘require’. But many technical terms will need to be defined in turn, as in the case of *oblique* in the definition of *passive construction*, or *associated form* in the definition of *nomifier system*.²³ Not all of the sample definitions in §3 are thus complete in the sense of being fully comprehensible. Some of them contain terms that are not widely known yet (e.g. *A-argument*, *flag*, *nomifier system*), which I decided to include here in order to show how a number of more familiar terms (*subject*, *case-marker*, *gender system*) are defined.

Ultimately, all standard terms must be defined in such a way that their definitions only include primitive concepts or other well-defined concepts. This is not an easy task,

²² However, the actual descriptions of segment inventories in the world’s languages contain an extremely wide variety of segment types. The Phoible database (Moran & McCloy 2019) includes over 3,000 different segmental comparative concepts, and for some of them, having a non-compound designation by means of a novel letter might well be useful.

²³ Similarly, Corbett (2007: 242) notes that since *gender* is defined in terms of ‘agreement’, “the definition of agreement itself becomes important”.

of course, so I do not foresee it to be finished within a few years, regardless of whether such proposals will be widely accepted or not.

The examples in §3 show that fairly straightforward definitions are possible for frequently used terms, e.g. for *sentence*, *morph*, *root* and *marker*. Such terms are not usually defined by linguists, and are typically learned by ostension, like everyday words. For example, Booij's (2005) morphology textbook does not provide a usable definition of *root*: On p. 29, we read that "Stems can be either simplex or complex. If they are simplex, they are called roots". This would seem to exclude roots which have no inflection and are therefore not stems (like English *gold* or *solid*). For the stereotypical Indo-European language, this may not be a big problem (because most verbs, nouns and adjectives show inflection), but it does not work for languages in general, because many languages have nouns that cannot be inflected, but we would still call them roots. Moreover, Booij provides no definition of *stem*. There are similar problems with the terms *clause* and *sentence*, which are rarely defined in a way that corresponds to the actual usage of the terms. For example, a well-known online glossary defines a sentence as "a grammatical unit that is composed of one or more clauses",²⁴ but this is not the way the term is used, because a clause may of course contain another clause (e.g. a relative clause, or a complement clause), and not all such clauses would be called sentences. It is clear that linguists have simply not invested a lot of energy into providing definitions, and that in many cases it is not difficult to improve on the current situation.

But other frequently used terms are harder to define in such a way that their definition broadly corresponds to their current use. For example, *subject* can be defined only through the terms *A-argument* and *S-argument*, which are not easy to define (see (9), and Haspelmath 2011a). And the term *gender system* is particularly difficult to define. Numeral classifiers are very similar to gender markers, and it seems that the main reason they are never included in discussions of gender systems is that they are stereotypically characteristic of East Asian languages, while gender systems are stereotypically characteristic of European (and African) languages.²⁵ Thus, numeral classifiers have to be specifically excluded, as is done in (19), and a new term (*nomifier*, short for 'nominal classification marker') needs to be introduced as a general term that has numeral classifiers and gender markers as subtypes.

Thus, standardization of grammatical terms has at least two aspects: Retro-definitions of existing widely-used terms (such as *affix*, *sentence*, *subject*) and creation of new terms (such as *A-argument*, *nomifier*) when needed in order to provide retro-definitions.²⁶

What are general principles for retro-definitions? A first principle is that an established term should not be defined in such a way that its definition is at variance with traditional use. It should cover the core of the phenomenon designated by the term (as generally understood), it should cover at least 80% of the cases where the term has been applied, and it should not include too many cases which would not be included traditionally. There are some well-known cases where the meaning of existing terms has

²⁴ <https://glossary.sil.org/term/sentence>

²⁵ Numeral classifiers are very similar to gender markers, cf. Japanese *niwatori san-ba* [chicken three-NUMCL.BA] 'three chickens', Italian *casa nuov-a* [house new-GND.A].

²⁶ But of course, one may also create completely novel terms that may become standard once they prove to be useful and are picked up by many linguists.

been changed by prominent linguists (thus leading to much confusion),²⁷ and this experience should not be repeated. It will often be impossible to find a definition that covers 100% of the traditional usage, because this usage is frequently somewhat inconsistent, but 80% accuracy should be enough to justify continuing the term.

If a traditional term is used so inconsistently that it is not possible to define it in such a way that the definition covers most of its uses, the term should be abandoned. Examples of such terms whose traditional use is not sufficiently coherent are *inflection*, *(non)finite*, and *clitic*; I do not know how to define them in such a way that their definitions would correspond very largely to traditional usage (see Cristofaro (2007) on *finite*, and Haspelmath (2015) on *clitic*). And even though many people still use *word* in a technical sense, it is not clear either how to define it objectively (Haspelmath 2011b).

The definition of a comparative term should be as simple as possible, even if this means that not all cases that are traditionally subsumed under the term are included. For example, if marker is defined as ‘a bound form that is not a root’ (see (8) above), this is a simple definition, but it does not fully capture the intuition that a marker expresses a grammatical meaning. Bound forms such as *however* or *basically* are not roots (according to the definition in (4)), but they would not be considered as typical markers. Thus, the criterion of having a simple definition may conflict with the criterion of having a good match with traditional usage, and one needs to strike a balance between the two criteria.

These principles will not cover all cases, and there will always be a certain amount of arbitrariness in definitions of comparative concepts. This is as it should be, because comparative concepts are not discoveries, but instruments for research (like units of measurement in physics). If there were no arbitrariness, no standardization would be required. The social implications of this will briefly be discussed in the final section.

Another important aspect of retro-definitions is that they may have a restricted extension in comparison to descriptive categories with the same name, as will be discussed next.

5. Shared-core definitions of comparative concepts

The definition of category-like terms need not correspond very closely to language-particular categories, but may be semantically simpler and merely correspond to the shared core of features of different language-particular categories. This is because it is often clearly meaningful to compare languages with respect to salient “core” concepts, regardless of how exactly the language-particular categories are delimited. For example, it is clearly very useful to compare languages with respect to how they express things (physical objects and persons), e.g. how they form plurals and how they express actions involving persons and things (‘the girl took the pens’). We can thus define the term *noun* as a comparative concept as in (5) (‘a root that denotes a thing’). Of course, in most or all languages, the class called “noun” goes beyond this core set of possible denotations (e.g. Russian *vojna* ‘war’, *svoboda* ‘freedom’). Thus, the class of Russian Nouns cannot be defined semantically – this is known to all linguists from their syntax textbooks. But language-particular categories are defined by language-particular criteria (e.g. by being combinable with articles, or showing number distinctions, or inflecting for case), and

²⁷ E.g. *ergative* for ‘unaccusative’ (e.g. Grewendorf 1989), *government* in Chomsky’s (1981) sense (contrasting with the earlier established sense, cf. Lehmann 1983; Kibort 2010), or *anaphor* in Chomsky’s (1981) sense (contrasting with the sense in computational linguistics)

these cannot be applied to all languages. The reason why we call different classes in different languages “nouns” is that they all include roots denoting objects and roots denoting persons, so this is the definition of the comparative concept. The fact that the categories called “noun” in different languages usually include more elements is not relevant to the definition of *noun* as a comparative concept, because the meanings of these other elements play no role in mapping the language-particular classes to the comparative concept.

In a very similar way, Nikolaeva & Spencer (2013: 219-220) mention the example of the term *adjective*, which they propose to define in terms of the concept of ‘gradable property’. This decision leaves aside adjectives like ‘dead’ or ‘blue’ (Nikolaeva & Spencer also mention more exotic cases like ‘alleged’), but again, everyone seems to agree that gradable-property adjectives do indeed represent the shared core of the various categories in different languages that we call adjectives. If a language had a class of words that does not include any gradable property concepts, it would not be called “adjective”.

Another example comes from the domain of gender. Languages with gender classes often have a feminine class, which can be defined semantically as a comparative concept:

“For some values, cross-linguistic comparison is straightforward: feminine gender is the value which includes nouns denoting females, and the interesting typological considerations are what other nouns may be included in this gender value... We need to define the core meanings and functions: we call a gender value the feminine if it includes nouns denoting females, whether or not it also includes diminutives.” (Corbett 2009: 137)

The German Feminine class and the Arabic Feminine class are quite different in their extension (both contain many inanimates), but they share the common core of female animates, which makes it meaningful to compare them.

Similarly, the English preposition *to* and the Russian Dative suffixes *-u/-e/-am* share the recipient meaning (cf. Haspelmath 2010: 666) and can thus be said to match the comparative concept ‘dative’, even though they are otherwise quite different (the English preposition also marks spatial goals, and the Russian Dative case suffixes are also governed by some prepositions).²⁸

Finally, the term *subject* as a comparative concept is defined in (10) in terms of the A-argument of two-argument clauses expressing a physical effect and patientive single-argument clauses, as discussed in detail in Haspelmath (2011a). By contrast, language-particular descriptions must include all the argument of all the verbs, including atypical two-argument verbs (such as ‘to look’ or ‘to like’), and there is much less cross-linguistic uniformity with these other verbs. The literature is full of discussions of how to use the term “subject” with these more heterogeneous verbs, but there is no doubt about physical-effect verbs and patientive single-argument verbs.²⁹

²⁸ Corbett (2009: 137) mentions this example, too: “Similarly we call a case value the dative if used for recipients, whether or not it can also be governed by prepositions.” (However, Corbett is only concerned with “case features”, in which he does not include prepositions.)

²⁹ The need for shared-core comparative concepts is not often mentioned in the literature, but Croft (2016: 378-379) notes something very similar:

“In general, extensionally “large” semantic categories that are given monosemous definitions ... do in fact often fail as comparative concepts. For example, property concepts taken as a broad category (stative, unary valency, gradable, inherent) do not serve well as a comparative concept for

It may well be that Dingemanse's formulation in terms of a "centre of gravity" (cited in §2 above) refers to what I call "shared core" here. However, what I mean here is not a "core" of a phenomenon that obviates the need for a definition in terms of necessary and sufficient conditions. The definitions must be precise (rather than prototype-based, see §6), but the mapping of a language-particular category onto the precisely defined comparative concept is cross-linguistically somewhat variable.

The fact that comparative concepts often refer to a shared core and are extensionally smaller compared to language-particular categories means that not all parts of every language enter the relevant comparison. Inanimate Feminines in German are not part of comparisons of feminine gender classes across languages, and arguments of experiential verbs (which are not typical transitive verbs) are not part of comparisons of subjects across languages. This is not a problem, but it must be kept in mind if one evaluates claims based on such comparative concepts. In order to compare arguments of experiential verbs (such as 'to like'), one needs more fine-grained comparative concepts (cf. the microroles of Hartmann et al. 2014).

6. Stereotypes and prototypes

The literature on grammatical patterns worldwide is full of stereotypes that are widely known, e.g.

- Latin has free word order, but English has rigid word order
- Italian is a pro-drop language, but English cannot drop its personal pronouns
- Turkish is an agglutinating language, and Chinese is an isolating language
- English is poor in inflection, but richer in derivation
- some North American languages have noun incorporation
- Latin makes much use of nonfinite clauses
- German has many compounds
- the Romance languages have clitic pronouns
- English has gender only in personal pronouns

Linguists almost never ask whether these stereotypes are true, but if they were true, this would be very interesting. We see outside of linguistics that many stereotypes are not true (e.g. most dogs are not called *Fido*), but some stereotypes are true (*Smith* is indeed the most frequent surname in the United Kingdom).³⁰ Thus, I find it important to know whether the above stereotypes are true (they may or may not be), and in order to assess them, we need objective definitions of these terms.

This should go without saying, but many linguists seem to treat the terms involved in such stereotypes as somehow having an independent existence, regardless of their definition. For example, at the end of their book on clitics, Spencer & Luís (2012: 321) admit that they have not been able to come up with a definition of the term *clitic* that encompasses all and only those phenomena that they discuss in their book. Still, they do

understanding "adjectives", because semantic subclasses of property concepts – age, dimension, color, value, etc. – exhibit different grammatical behavior in one and the same language ... The usual solution to this problem is to use finer-grained categories, such as the property subclasses, or for core participant roles, the division into A, S, and P."

³⁰ https://en.wikipedia.org/wiki/List_of_most_common_surnames_in_Europe

not conclude from this that the phenomena they discussed may not be coherent (but united merely by the fact that some people use the same label *clitic* for these phenomena). Similarly, Reuland's (2018) overview paper of "reflexives and reflexivity" contains no definition of "reflexive", and the author is apparently not interested in providing one (because he takes his goal as discovering the relevant aspects of the innate grammar blueprint, not as comparing languages systematically). And even though my (2011b) paper on the definition of "word" has been widely cited, many authors continue to use the term *word*, even in technical contexts, as if the term had a clear meaning (perhaps hoping that such a meaning will be provided by someone eventually). Another example of this nonchalant attitude is Massam's (2017) overview paper on noun incorporation. Massam writes: "There is a lot of disagreement about exactly what constitutes noun incorporation ... the field is rich with proposals and counter-proposals as to its true nature" (2017: §7), but she provides no definition. She seems to presuppose that noun incorporation has an independent existence, even if we have not found its definition yet. And Lieber & Stekauer (2009: 14) conclude their introduction to a handbook on compounding by saying that "there are (almost) no reliable criteria for distinguishing compounds from phrases or from other sorts of derived words" – but for some reason, they still say that "it's worth looking further", as if finding a definition were a research result (rather than an indispensable methodological prerequisite).

Thus, many linguists have been unable to provide clear definitions of terms, but have nevertheless been unwilling to abandon the traditional terms. While many have been content to simply ignore these problems (hoping that they are not too serious, and/or that someone else will eventually solve them), others have noticed the problem and have reacted by invoking "prototypes". The idea that linguistic categories are (sometimes) based on cognitive prototypes was made famous by Lakoff (1987) and was explored in more detail by Taylor (1989). This work was based on mental representations of particular knowledge systems, including grammatical knowledge. But the cognitive notion of a prototype category cannot simply be transferred to cross-linguistic categories, which are not represented in any individual speaker's mind. There was some discussion of cross-linguistic prototypes in the 1980s and 1990s (e.g. Newmeyer 1998: Chapter 4), but this idea has not been pursued systematically in recent decades, as far as I am aware. So I do not see any basis for a systematic "prototype view" of cross-linguistic categories. Cross-linguistic phenomena often seem to cluster in certain ways, but the extent to which these clusters are real or based on our stereotypes can be assessed only if we have precise ways of measuring differences between languages.

Thus, I see the standardization of well-known terms as one way to help the discipline move beyond the traditional vague stereotypes.

7. Standard comparative terms and language-particular description

Linguists encounter grammatical terminology most often in the context of particular languages, rather than in a comparative context, as most linguists study particular languages most of the time. Thus, the kind of standardization discussed here might seem not to affect most activities of linguists.

But appearances are deceptive. Even when a linguist talks about a particular language, they typically want their findings to be relevant to other linguists studying similar phenomena in other languages. For example, García García (2019) studies causative verbs in Old English (e.g. *þwīnan* 'dwindle', *þwēnan* 'cause to dwindle'), and puts these verbs explicitly in a typological context. Likewise, Nordlinger (2014) studies

serial verbs in Wambaya (an Australian language), and she puts these constructions in a typological context. The idea that the study of particular languages has relevance for general linguistics has become almost universal, perhaps because there are nowadays many linguists who study languages outside of an applied context (where languages are studied to facilitate language learning, or understanding of important literary text). And when the research question is theoretical rather than applied, then it is attractive to link one's language-particular insights to larger generalizations.³¹

Moreover, even though each language has its own categories (Haspelmath 2007), there are many similarities between the categories of different languages, and we do not want completely different terminologies for different languages. Thus, we call German verb forms like *sind* ('are') Third Person forms, even though they are also used with the polite address form *Sie* ('you') (e.g. *Sind Sie fertig?* 'Are you ready?'). From a language-particular point of view, these verb forms are different from third person forms in, say, French, but it would not serve transparency to give them any other name. Similarly, the Russian Imperative has a use in conditional clauses (E.g. *bud' ja na vašem meste* [be.IMPV I on your place] 'if I were in your place'), so it is different from, say, the Latin Imperative, but it would be strange to give it any other name.

So from the point of view of terminological transparency, it is best if language-particular categories are given names that correspond closely to comparative-concept names. It is therefore also from a purely descriptive, language-particular perspective that standard terminology is relevant; but because of the fundamental difference between comparative concepts and descriptive categories (§2), the standard terms as discussed here are not crucial for language-particular description.

Finally, it should be noted that in practice, IPA characters are used somewhat differently from morphosyntactic comparative concepts as envisaged here, because they are often used for language-particular notation. In fact, linguistics students typically learn IPA characters as a method of transcribing the pronunciation of words in particular languages, and it is only later that they may be confronted with IPA for the purposes of cross-linguistic comparison. But it is a misunderstanding to think of IPA characters as a list of all possible sounds – rather, IPA characters are a list of well-known comparative concepts for comparing sound inventories (Ladd 2011), which happen to work quite well for the practical purposes of transcribing pronunciations. But it should be kept in mind that the IPA, too, arose in a comparative context: English-language teachers in France and French-language teachers in Britain got together in order to create a tool facilitating language learning – which usefully starts with comparing one's original language to the language to be learned. For the purposes of scientifically describing a language (in terms of its own categories), one needs language-particular categories in phonology as well, just as in morphosyntax. The IPA offers a convenient (widely understood) set of symbols, but these are not actually crucial to language-particular description (which must be based on contrasts and phonetic characterizations).

In this connection, a question that sometimes arises concerns the status of gloss abbreviations in interlinear text. There is a standard set of abbreviations in the appendix of the Leipzig Glossing Rules,³² and this has become very popular (e.g. ACC for accusative, GEN for genitive, PL for plural, and so on). But some linguists seem to take it

³¹ The comparative perspective is so deeply engrained in current linguistics that we hardly notice its presence in everyday expressions such as "Turkish has a passive construction", or "Vietnamese has serial verb constructions". Every time when we say that "a language has a category X", we make reference to a category-like comparative concept.

³² <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>

as the main aspect of the rules (while in fact, the gloss abbreviations are merely the appendix), and I fear that a few even think that these categories are intended to be universal categories. However, what is standardized here is merely the relation between the abbreviations and the terms (e.g. the use of INS rather than INSTR for the term *instrumental*). The gloss abbreviations say nothing about the meanings of the terms themselves. And since interlinear glossing is typically done from a language-particular point of view, they are normally interpreted as representing language-particular categories. Thus, in the above gloss [be.IMPV] (for Russian *bud'*), the abbreviation IMPV stands for “(Russian) Imperative”, not for a comparative concept. However, in context where we compare languages (e.g. in a typological study), it makes good sense to provide “comparative glosses”, rather than language-particular glosses. So in a context where word order in conditional clauses in different languages is discussed, it is probably best to gloss the Russian form *bud'* as [be.COND], because the Imperative form serves to indicate a conditional clause in this context. The fact that the actual form is called *Imperative* and is also used for imperative clauses is irrelevant here.

Thus, the technical terms of phonetics/phonology and morphosyntax play a dual role, which helps us understand that some linguists feel that the distinction between comparative concepts and language-particular descriptive categories is somehow problematic or difficult to draw. But it is not a problematic conceptual distinction, and all linguists agree that we should use the same technical terms in both roles (rather than, say, having two completely distinct sets of terms).

8. Concluding remarks

In this paper, I have argued that there should be some standard terminology for morphosyntax, just as there are standards of terminology or notation in other disciplines, including subdisciplines of linguistics such as phonetics. I have given some concrete examples of possible standard definitions of well-known (and novel) terms, and I have addressed some issues and briefly compared the standardization approach with alternative ideas making use of prototypes (or actually stereotypes).

Perhaps the most pressing question that readers have at this point is the question of implementation: What needs to happen so that linguists actually adopt a standard? But this question is outside the scope of the present paper. Adopting a standard is a collective action problem, and there are many different proposals for how to solve such problems. The task that I have set for myself in this paper is merely to address the issue of standardization from a conceptual point of view. The principles discussed in §4 should help to make the proposals acceptable to a maximal set of linguists, but there will always be some arbitrariness in any standard.

And as noted in §2, many linguists seem to conceive of their grammatical terms as universal mental entities (natural-kind categories of the innate grammar blueprint), though this is not often stated clearly. In order for the standardization programme to make progress, it is crucial to make a clear distinction between hypothetical universal innate categories (as true research results) and universally applicable comparative concepts (as historically arbitrary research instruments which are subject to standardization).

Acknowledgements

I thank the editors for inviting me to submit this paper, and two reviewers (including Eitan Grossman) for helpful comments. I also thank the numerous colleagues who discussed the issues raised in the paper in an Academia.edu session. And I am particularly grateful to Erich Round and Greville Corbett for extensive discussions of some potential problems of my approach, as well as Christian Lehmann, Edith Moravcsik and Andrew Spencer for further discussion.

References

- Baker, Mark C. 2001. *The atoms of language*. New York: Basic Books.
- Booij, Geert E. 2005. *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press.
- Brentari, Diane. 2002. Modality differences in sign language phonology and morphophonemics. In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, 35–64. Cambridge: Cambridge University Press.
- Brown, Dunstan & Marina Chumakina. 2013. What there might be and what there is: An introduction to Canonical Typology. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 1–19. Oxford: Oxford University Press.
- Chomsky, Noam A. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Clements, George N & Samuel Jay Keyser. 1985. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin: University of Texas Press.
- Corbett, Greville G. 2007. Gender and noun classes. In Timothy Shopen (ed.), *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*, 241–279. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2009. Universals and features. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 129–143. Dordrecht: Springer.
- Cristofaro, Sonia. 2007. Deconstructing categories: Finiteness in a functional-typological perspective. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 91–114. Oxford: Oxford University Press.
- Croft, William. 2016. Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology* 20(2). 377–393. doi:10.1515/lingty-2016-0012.
- Dingemanse, Mark. 2019. ‘Ideophone’ as a comparative concept. In Kimi Akita & Prashant Pardeshi (eds.), *Ideophones, mimetics, and expressives* (Iconicity in Language and Literature), 13–33. Amsterdam: Benjamins. <https://doi.org/10.1075/ill.16.02din>.
- Dryer, Matthew S. 2009. The branching direction theory of word order correlations revisited. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 185–207. Dordrecht: Springer.
- Dryer, Matthew S. 2016. Crosslinguistic categories, comparative concepts, and the Walman diminutive. *Linguistic Typology* 20(2). 305–331. doi:10.1515/lingty-2016-0009.
- García García, Luisa. 2019. The basic valency orientation of Old English and the causative ja-formation: a synchronic and diachronic approach. *English Language & Linguistics* (XX). 1–25. doi:10.1017/S1360674318000345.
- Grewendorf, Günther. 1989. *Ergativity in German* (Studies in Generative Grammar 35). Dordrecht: Foris.
- Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3(1). 1–21.

- Haspelmath, Martin. 2007. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11(1). 119–132.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. doi:10.1353/lan.2010.0021.
- Haspelmath, Martin. 2011a. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.
- Haspelmath, Martin. 2011b. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.
- Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything: in honor of Edward Keenan* (UCLA Working Papers in Linguistics 17), 109–130. Los Angeles: UCLA.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntax of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2015. Defining vs. diagnosing linguistic categories: A case study of clitic phenomena. In Joanna Błaszczak, Dorota Klimek-Jankowska & Krzysztof Migdalski (eds.), *How categorical are categories? New approaches to the old questions of noun, verb, and adjective*, 273–304. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics* 17(3). 291–319. doi:<http://doi.org/10.1177/2397002215626895>.
- Haspelmath, Martin. 2018a. How comparative concepts and descriptive linguistic categories are different. In Daniël Van Olmen, Tanja Mortelmans & Frank Brisard (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. <http://doi.org/10.5281/zenodo.570000>.
- Haspelmath, Martin. 2018b. The last word on polysynthesis: A review article. *Linguistic Typology* 22(2). 307–326. doi:10.1515/lingty-2018-0011.
- Haspelmath, Martin. 2018c. Toward a new conceptual framework for comparing gender systems and some so-called classifier systems. Stockholm University. doi:10.5281/zenodo.1230569.
- Haspelmath, Martin. 2019. Indexing and flagging, and head and dependent marking. *Te Reo* 62(1). 93–115. doi:10.17617/2.3168042.
- Haspelmath, Martin. 2020. The morph as a minimal linguistic form. *Morphology* 30. 117–134.
- Haspelmath, Martin. 2021. Comparing reflexive constructions in the world's languages. To appear in: Janic, Katarzyna & Puddu, Nicoletta & Haspelmath, Martin (eds.) *Reflexive constructions in the world's languages*. Berlin: Language Science Press.
- Hennig, Mathilde. 2012. Grammatische Terminologie in der Schule: Einladung zur Diskussion. *Zeitschrift für Germanistische Linguistik* 40(3). 443–450. doi:<http://dx.doi.org/10.1515/zgl-2012-0028>.
- Kibort, Anna. 2010. Towards a typology of grammatical features. In Anna Kibort & Greville G. Corbett (ed.), *Features: Perspectives on a key notion in linguistics*, 64–106. Oxford: Oxford University Press.
- Ladd, D. Robert. 2011. Phonetics in phonology. In John A Goldsmith, Jason Riggle & Alan C. L Yu (eds.), *The handbook of phonological theory*, 348–373. Chichester: Blackwell.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.
- Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190.
- Lehmann, Christian. 1983. Rektion und syntaktische Relationen. *Folia Linguistica* 17(1–4). 339–378.
- Lieber, Rochelle & Pavol Štekauer. 2009. Introduction: Status and definition of compounding. *The Oxford handbook of compounding*, 3–18. Oxford: Oxford University Press. 10.1093/oxfordhb/9780199695720.013.0001.

- List, Johann Mattis, Simon Greenhill, Christoph Rzymiski, Nathanael Schweikhard & Robert Forkel (eds.). 2019. *Concepticon 2.0*. Jena: Max Planck Institute for the Science of Human History. <https://concepticon.clld.org/>.
- Massam, Diane. 2017. Incorporation and pseudo-incorporation in syntax. *Oxford Research Encyclopedia of Linguistics*. doi:10.1093/acrefore/9780199384655.013.190 (20 May, 2018).
- Moran, Steven & Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. <https://phoible.org/>.
- Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.
- Nichols, Johanna. 1984. Direct and oblique objects in Chechen-Ingush and Russian. In Frans Plank (ed.), *Objects*, 183–209. London: Academic Press.
- Nordlinger, Rachel. 2014. Serial verbs in Wambaya. In Rob Pensalfini, Myfany Turpin & Diana Guillemin (eds.), *Language description informed by theory*, vol. 147, 263–282. Amsterdam: Benjamins.
- Reuland, Eric. 2018. Reflexives and reflexivity. *Annual Review of Linguistics* 4(1). 81–107. doi:10.1146/annurev-linguistics-011817-045500.
- Spencer, Andrew & Ana R. Luís. 2012. *Clitics*. Cambridge: Cambridge University Press.
- Taylor, John R. 1989. *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Clarendon.