

How to compare major word-classes across the world's languages

Martin Haspelmath

1 Introduction

In this paper, I argue that major word-classes, such as nouns, verbs and adjectives, cannot be compared across languages by asking questions such as those in (1) and (2). Such questions are routinely asked by linguists (functionalists and generativists alike), but these are the wrong questions (cf. Croft 2000), because they make presuppositions which are not valid.

(1) language-particular questions

- Does language X have a noun-verb distinction?
- Does language X have a verb-adjective distinction?
- Does language X have a noun-adjective distinction?

(2) cross-linguistic questions

- Do all languages have a noun-verb distinction?
- Do all languages have a verb-adjective distinction?
- Do all languages have a noun-adjective distinction?

Questions like these would make sense only if we could define noun, verb and adjective as cross-linguistic categories, but cross-linguistic categories do not exist (Croft 2001, Haspelmath 2007, 2010). Categories represent language-particular generalizations and cannot be carried over from language to another one.

This categorial particularist position has been stated clearly by Keenan & Stabler (2003: 1):

“On our approach different languages do have non-trivially different grammars: their grammatical categories are defined internal to the language and may fail to be comparable to ones used for other languages. Their rules, ways of building complex expressions from simpler ones, may also fail to be isomorphic across languages.”

But still, we want to compare languages and extract generalizations from all the observed diversity. Keenan & Stabler propose an approach that is similar to the line of thinking pursued here in that it identifies linguistic invariants without assuming (as many linguists do) that all languages are built from the same building blocks. I will not compare their approach to mine in detail, leaving that as a task for future research.

Edward Keenan has often asked to what extent natural languages differ from logic, and if so, why (e.g. Keenan 1974). In predicate logic, there is no difference between nouns, verbs and adjectives. A formula such as (3), with a predicate *P* and an argument *n*, would be used to render sentences like ‘Charles is a teacher’, ‘A naked one is playing’, or ‘The netting is plentiful’.

(3) $P(n)$

But intuitively, despite the semantic similarity of such predicational sentences, linguists agree that grammars of natural languages do show differences between different word-classes. But they rarely agree about the precise number of such word-classes and about cross-linguistic trends. Why is this so? In this paper, I argue that these problems are to a large extent due to a simple misunderstanding, namely that grammatical categories are cross-linguistic entities. The fundamental insight is due to Croft (1991), (2000), (2001: ch. 2), but even though Croft's work has been widely cited and read, many linguists have not drawn the necessary consequences from his arguments yet. And the way I formulate the approach is somewhat different from Croft's.

I will end up concluding that word-classes cannot be compared directly across languages because of their language-particular nature, but it is not difficult to find strong cross-linguistic trends in the domain of word-class coding.

2 Terminological preliminaries

Before getting to the heart of the issue in the next sections, let us consider some terminological issues in the present section.

The first thing to observe is that while the three terms *noun*, *verb* and *adjective* are used universally and uniformly among linguists (and have been in universal and uniform use in the current senses for over a century), the general term *word-class* that is used here is not nearly as widespread. Alternative terms with considerable currency are *part of speech* and *lexical category*.

The term *part of speech* goes back to classical antiquity (Greek *ta mérē tou lógou*, Latin *partes orationis*) and reflects a period of linguistics that did not distinguish clearly between speech and language, and between classes and items. In modern parlance, grammarians talk about elements of *sentences* (not of *speech*), and about *classes* of such elements, not about “parts” of a sentence. Furthermore, we nowadays distinguish between immediate constituents (of which a sentence has only a few, often as little as two) and ultimate constituents (words or elementary morphs). Nouns, verbs and adjectives are classes of ultimate constituents, not classes of larger, phrasal constituents. Thus, the term *part of speech* is about as opaque semantically as the term *accusative* (which has nothing to do with accusation), but because of its transparent syntax, its opacity is even more confusing. For these reason, I would not recommend it, but of course it has a venerable tradition behind it. The term is enjoying considerable popularity, especially in typological circles (e.g. Anward et al. 1997, Hengeveld et al. 2004, Hengeveld & van Lier 2010).¹

In generative grammar, the term *syntactic category* has been used for classes of constituents (from the word level to the sentence level) since Chomsky (1957); see Rauh (2010) for a detailed account. However, nouns, verbs and adjectives were not considered an interesting issue in generative linguistics for the first few decades; it was simply assumed by almost everyone that all languages have them. In Croft (1991), the term *syntactic category* is used for nouns, verbs and adjectives in a typological context. Since the 1990s, the terminal-node categories have generally been divided into functional categories and lexical categories, so nouns, verbs and adjectives came to be known as *lexical categories* (e.g. Davis & Matthewson 1999, Baker 2003, Chung 2012). This term

¹ Interestingly, Hengeveld & van Lier (2010) use the compound term *parts of speech class*, rather than the old term *part of speech* itself.

is more in line with modern terminology, but in view of the ambiguity of the term *lexical*² and the vagueness of the term *category*, it is not an improvement over the simple term *word-class*.

In this paper, the term *word-class* is used for language-particular categories such as English Nouns or Japanese Verbal Adjectives. (I normally capitalize labels of language-particular categories, cf. Haspelmath 2010: §6).

Note that I will limit myself to major word-classes, ignoring classes such as adpositions, adverbs, particles and interjections.

3 First problem: different criteria in different languages

If one adopts a categorial universalist position, i.e. compares languages starting out with the assumption that they will basically have the same categories, one must be willing to apply different criteria in different languages. For example, to identify “nouns” in Ancient Greek, English and Mandarin Chinese, quite different criteria are commonly applied. This can be seen in (4a-c).

- (4) a. Greek Noun (Dionysius Thrax, *Ars minor*, 2nd c. BCE)
ὄνομα ἐστὶ μέρος λόγου πτωτικόν, σῶμα ἢ πρᾶγμα σημαῖνον
 ‘a Noun is a case-inflected part of speech that denotes a thing or an action’
- b. English Noun (Quirk et al. 1985: 72)
 a Noun is a word that can follow determiners like *the*, *this* and *that*
- c. Mandarin Chinese Noun
 a Noun is a word that can follow a classifier

That the same category should be reflected in different kinds of formal properties in different languages is perhaps not exactly what one would expect, but universalists do not take it as evidence against the assumption that all languages have the same categories. As long as SOME criteria can be found for noun status, universalists are content.

But which properties can be taken as evidence for category assignment? There are no constraints on this – each linguist can make their own choices (this is what Croft 2009 calls “methodological opportunism”). The method is thus subjective and not rigorous (cf. also Post 2008: 377-78). Rigorous comparison requires that languages be compared in terms of concepts that apply in the same way to all languages. We will see some examples of the nonrigorousness of the approach below.

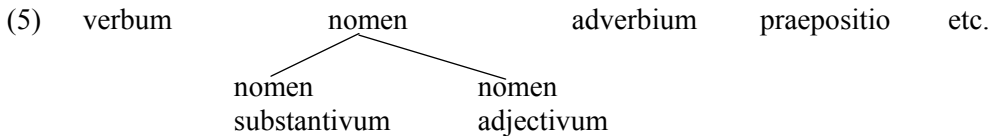
4 Second problem: major classes vs. subclasses

If one starts out by asking whether a language has a distinction between nouns and verbs, or between verbs and adjectives, one presupposes that nouns, verbs, and adjectives can only be major classes. But what if their relationship can equally be described as a

² *Lexical* can mean (1) ‘relating to the lexicon’ (e.g. *lexical rule*, *lexical exception*), and it can mean (2) ‘relating to words’ (e.g. *lexical stress*). To the extent that the lexicon is conceived of as the repertoire of words, these two senses are closely related, but if the lexicon is thought of as a list of all elements that are not fully derivable by rule (as is routinely done by linguists), then *lexical* can also be used for idiomatic phrases or even sentences.

subclass relationship? In actual fact, this is very often the case. If two classes share a property that another class lacks, one can say that the two classes form a macro-class on the basis of this property.

In this way one can, for instance, argue for collapsing nouns and adjectives into a macro-class if they have some similarities. This is in fact the traditional view in Western linguistics, going back all the way to the earliest grammatical descriptions of Greek and Latin. Latin words such as *homo* ‘man’ and *novus* ‘new’ were long regarded as belonging to the category *nomen* (Greek *onoma*, cf. 4a above), which was on a par with *verbum*, *praepositio*, etc. (see 5). This category was later subcategorized into *nomen substantivum* and *nomen adjectivum*, but it was only in 19th century linguistics that substantives and adjectives were regarded as word-classes on a par with verbs and prepositions.³



On the basis of the criteria that were usually considered, this was a very reasonable decision. As Table 1 shows, substantives and adjectives in Latin share many properties that oppose them to verbs (case, intrinsic number, no person, no tense, predicative copula, referential use), so we can say that Latin has two major word classes Nomen and Verbum.

Table 1: Latin *Nomina* and *Verba*

	Nomen		Verbum
	Substantivum	Adjectivum	
inflection	+	+	+
case	+	+	-
intrinsic number	+	+	-
person	-	-	+
tense	-	-	+
copula in predicative use	+	+	-
referential use	+	+	-
attributive use	-	+	-
comparative construction	-	+	-

On the other hand, Substantiva and Adjectiva differ with respect to the last two criteria in Table 1, attributive use and use in comparative constructions (which is possible only for adjectives). The Nomen class thus has two subclasses, Nomen Substantivum and Nomen Adjectivum.

But since Nomina and Verba also share at least one property that opposes them to other words (particles, prepositions, interjections, etc.), namely the possibility to inflect, one might say that the real major classes in Latin are the Flectibilia (inflectible words) and the Nonflectibilia (uninflectible words, which themselves of course fall into several subclasses based on other criteria). This is shown in Table 2.

³ For some reason, substantives came to be called *nouns* in English (where the term *substantive* is rarely used), while in German, the term *nomen* fell out of (common) use.

Table 2: Latin *Flectibilia* and *Nonflectibilia*

	Flectibile		Nonflectibile	
	Nomen		Verbum	
	Substantivum	Adjectivum		
inflection	+	+	+	–
case	+	+	–	
intrinsic number	+	+	–	
person	–	–	+	
tense	–	–	+	
copula in predicative use	+	+	–	
referential use	+	+	–	
attributive use	–	+	–	
comparative construction	–	+	–	

Linguists do not normally set up major word classes that comprise both thing words and action words, but if we go strictly by the criteria of individual languages, there is no justification for this. Limiting major word-classes to noun-like, adjective-like and verb-like classes can be justified only by the assumption that all languages have the categories that we know from school.⁴ Some linguists do this only because they do not think very hard about the differences between languages, but for others, it is an explicit programme. Thus, Chomsky (2001) formulated the principle in (6).

(6) Uniformity Principle

In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.
(Chomsky 2001: 2)

But what would be compelling evidence against the assumption that all languages have nouns, verbs and adjectives as word classes, if just any kind of criterion can be applied (cf. §3)? Would the fact that we have not found a difference between thing words and property words count as compelling evidence? Surely not, because one can always suspect that there is some minute difference which has not come to our attention yet.

And what would be compelling evidence against the assumption that nouns, verbs and adjectives are major word classes in all languages where these can be distinguished in some way? How could one, for instance, argue that Substantiva and Adjectiva are subclasses of Nomina in Latin rather than major word classes? Intuitively, the number and “weight” of the properties which they share should play a role, but there is no rigorous way of making the distinction between major classes that share properties and subclasses. That the distinction is not a substantive one but purely a notational distinction was noted by the first thorough cross-linguistic study of major word-classes, Schachter (1985):

“One might wish to say that in some languages, such as Nootka and Tagalog, nouns and verbs have enough in common grammatically for there to be some question about whether to regard them as two subclasses of a single part of speech rather than two distinct parts of

⁴ Or perhaps by the fact that noun-like, adjective-like and verb-like classes are reasonably large, open classes about which one could write separate sections or even chapters in grammar books. But this is a practical consideration with no significance for the nature of language(s).

speech. Since it seems to be essentially **a matter of terminology**, it need not concern us further.” (Schachter 1985: 13; boldface mine)

Thus, the Uniformity Principle seems to inevitably lead us to the conclusion that all languages have the properties of English school grammar (cf. Haspelmath 2012). Rigorous cross-linguistic research must accept that languages differ considerably, and must try to find similarities, and maybe even universal trends, despite these differences.

5 How to compare languages: not with categories, but with comparative concepts

We thus cannot presuppose that “noun”, “verb” and “adjective” are universally available cross-linguistic categories, because categories of grammar are language-particular. They express language-particular generalizations, or in other words, they are defined with respect to language-particular criteria, and thus they can never be equated across languages. Each language has its own categories (Boas 1911, Lazard 1992, Dryer 1997, Croft 2001, Haspelmath 2007). The questions in (1) and (2) can thus be compared with questions such as those in (7):

- (7) a. What is the line of succession to the German throne?
 b. How many states are there in France?
 c. Who is San Marino’s Minister of Aviation?

These are the wrong questions, because the political organization of countries is different, so categories cannot be equated across countries. Germany has no monarchy, so (7a) can be asked about Belgium, but not about Germany. France is divided into departments (not into states), so (7b) can be asked about the U.S., but not about France, and San Marino has a different set of ministries than other countries. The presuppositions are not fulfilled, so one cannot ask these questions.

In the domain of political organization, countries nowadays often equate their categories to make communication easier. Thus, each country has a “head of state” (whatever the local designation or role in the political system), and for official state visits the heads of state meet and obey certain international rules of protocol. To some extent, these equations then influence the political systems of different countries.

Linguists also often equate categories across languages to make it easier to talk about them, but this has no influence on the languages, of course.⁵ If the extension of terms that are familiar from one language to use in the description of another language is purely for reasons of convenience (it is easier to remember the term “Verb” than the term “category number 3”), then there is no problem. But equating categories across languages in a deeper sense (i.e. for purposes of language comparison) is not possible, because categories are defined by language-particular criteria, as seen in (4a-c). Clearly, then, languages cannot be compared directly on the basis of their grammatical categories. We need a *tertium comparationis* that is not language-particular, but is universally applicable.

To be universally applicable, comparative concepts can be defined on the basis of meaning or sound, but not on the basis of meaning-sound combinations, because these are language-particular.

A very simple definition of comparative concepts for major word-classes is a semantic one:

⁵ However, when bilingual speakers equate the categories of the languages they speak, this may of course lead to mutual accommodation (contact-induced grammatical change).

- (8) a. a noun is a word that denotes a thing or place
 b. a verb is a word that denotes an action or process
 c. an adjective is a word that denotes a property

Typologists have in fact generally worked with this kind of definition, e.g. when determining noun-adjective order or verb-object order in a large number of languages (Greenberg 1963, Dryer 2005a, 2005b).

However, here I would like to propose a more narrowly defined set of concepts, given in (9). Further below (§10) I will explain these notions in more detail.

- (9) a. thing-root: a root that denotes a physical object (animate or inanimate)
 b. action-root: a root that denotes a volitional action
 c. property-root: a root that denotes a property such as age, dimension or value

Unlike the comparative concepts in (8), those in (9) have labels that remind us of their semantic basis. But this is merely a notational difference. It is often more practical to use opaque terms like those in (8). If it is clear that one is talking about comparative concepts, not about language-particular descriptive categories, then using grammar-derived nonsemantic terms such as “noun” does not do much harm.

6 Widely asked wrong questions

As we saw in the introduction, linguists often ask questions such as those in (10) and (11) (repeated from (1)-(2) above):

- (10) language-particular questions
 Does language X have a noun-verb distinction?
 Does language X have a verb-adjective distinction?
 Does language X have a noun-adjective distinction?
- (11) cross-linguistic questions
 Do all languages have a noun-verb distinction?
 Do all languages have a verb-adjective distinction?
 Do all languages have a noun-adjective distinction?

Some linguists have answered no, and others have answered yes. Table 3 gives a summary of recent work that addresses these questions.⁶

⁶ I say “recent work”, because the question of word-class universality is an old one, going back at least to the first half of the 19th century. At the time, it was often said that Indo-European languages are particularly elaborate and developed in their word-class distinctions, and non-Indo-European languages with less developed inflection, or less distinctive inflection for nouns and verbs, were said to lack noun-verb distinctions (e.g. Müller 1861-64).

Table 3: Answers to the questions in (10) and (11) in the literature

		NO	YES
noun-verb distinction	ALL languages	Hengeveld 1992	Baker 2003, Dixon 2010.2
	Iroquoian	Sasse 1988, 1991	Mithun 2000
	Salishan	Kuipers 1968, Kinkade 1983	van Eijk & Hess 1986, Davis & Matthewson 1999, Beck 2012+
	Mundari	Hoffmann 1903, Hengeveld 1992	Evans & Osada 2005
	Tagalog	Gil 2000, Himmelman 2008, Kaufman 2009	Aldridge 2009, Richards 2009
	Polynesian	Mosel & Hovdhaugen 1992	Vonen 2000
noun-adjective distinction	ALL languages	Evans & Levinson 2009	Baker 2003, Dixon 2004
	Quechua	Weber 1989, Hengeveld 1992	Floyd 2011
verb-adjective distinction	ALL languages	Evans & Levinson 2009	Baker 2003, Dixon 2004
	Mandarin Chinese	McCawley 1992	Paul 2005
	Chamorro	Topping 1973	Chung 2012
	Caribbean English	Sebba 1986,	Seuren 1986
	Creole	Winford 1997	

The general trend seems to be that earlier work has tended to deny word-class distinctions, whereas more recent work has tended to (re-)assert word-class distinctions of the familiar type (e.g. Sasse 1988 vs. Mithun 2000 for Iroquoian, Kinkade 1983 vs. Davis & Matthewson 1999 for Salishan, McCawley 1992 vs. Paul 2005 for Mandarin Chinese, etc.). Thus, earlier linguists tended to be lumpers, whereas more recently they have tended to be splitters. It appears that the categorial universalist approach has gained in popularity in recent decades. However, as I will show in the next sections for three examples, the claims that a language is really more like English are often not justified by actual properties of the language, but by a categorial universalist approach. But as we saw earlier, such an approach is incompatible with rigorous cross-linguistic comparison.

7 A wrong question: Are there adjectives in Quechua? (Floyd 2011)

To see what sorts of problems arise if one presupposes the existence of verbs, nouns and adjectives as cross-linguistic categories, we will now look at some concrete examples, beginning with “adjectives” in varieties of Quechua.

Weber (1989: 35-36) adopted a lumping approach and claimed that Huallaga Quechua has no noun-adjective distinction. Both thing-roots and property-roots can be used for reference (as in 12), both thing-roots and property-roots require the copula *ka-* when used predicatively (as in 13), and both thing-roots and property-roots can be used preminally for attributive use (as in 14).

- (12) a. *rumi-ta rika* (reference use)
stone-ACC see.1SG
'I see the stone.'
- b. *hatun-ta rika*
big-ACC see.1SG
'I see the big one.'
- (13) a. *Taqay rumi ka-yka-n.* (predication use)
that stone be-IMPV-3
'That is a stone.'
- b. *Taqay hatun ka-yka-n.*
that big be-IMPV-3
'That is big / a big one.'
- (14) a. *rumi wasi* (attribution use)
stone house
'stone house'
- b. *hatun wasi*
big house
'big house'

This view of Quechua was adopted by a number of other authors, especially by comparative linguists (e.g. Schachter & Shopen 2007, Hengeveld & van Lier 2008). This view is of course not new at all, but corresponds to the centuries-old view that adjectives and substantives are subclasses of the major class *nomen* (see above §4).

Now Floyd (2011) argues that Quechua does have an adjective-noun distinction after all, i.e. the contrast between Weber's and Floyd's view is not unlike the contrast between the earlier Western grammarians and the more modern ones. Floyd points out a number of ways in which property-roots behave differently from thing-roots. For example, adjectives precede nouns, but not vice versa (Floyd 2011: 53):

- (15) a. *Chaypi shuk yurak wasi-ta riku-ku-ni.*
there one white house-ACC see-PROG-1SG
'There I see a white house.'
- b. **Chaypi shuk wasi yurak-ta riku-ku-ni.*
there one house white-ACC see-PROG-1SG

This can be described by saying that there are two word-classes noun and adjective, and that nouns cannot attributively modify adjectives.

But one could alternatively say that there is just a single class of *nomina* (justified by the coding patterns in (12)-(14)), and that different subclasses behave somewhat differently with respect to ordering. We need different subclasses of property-words anyway, because in many languages, they are not fully free in their ordering. For example, in German the ordering in (16a) is perfect, while the ordering in (16b) is very odd (English is of course quite similar).

- (16) German
- | | | |
|----|------------------------------|-----------------------|
| a. | <i>ein schönes großes</i> | ‘a beautiful big one’ |
| b. | ?* <i>ein großes schönes</i> | ‘a big beautiful one’ |

More generally, word-classes almost always have subclasses whose members behave somewhat differently. This is a very important point that is often overlooked.

Floyd (2011) justifies the recognition of adjectives and nouns by the additional phenomena that he takes into account. If one just looks at the paradigm in (12)-(14), one may conclude that thing-words and property-words belong to a single class, but if more facts are considered, then differences emerge. Word-classes must behave in the same way with respect to all “morphosyntactic possibilities”:

“Rather than relying on just one or a few specific features, the basic criterion for establishing a lack of word class distinction that I will respect here is that the morphosyntactic possibilities should be the same for all members of the proposed macro-class exhaustively across the lexicon.” (Floyd 2011: 26)

But if one takes into account ALL features, then one gets many small subclasses. Instead of a “verb” class, one would typically have classes of intransitive verbs, monotransitive verbs, ditransitive verbs, stative verbs, dynamic verbs, and others, or rather (because of cross-classification) stative intransitive verbs, dynamic intransitive verbs, stative monotransitive verbs, dynamic monotransitive verbs, and so on. Instead of a single “noun” class, one would have count nouns, mass nouns, kinship terms, body-part terms, relational nouns, collective nouns, abstract nouns, and so on.

Linguists who prefer to say that adjectives and substantives are different (major) word-classes do not of course deny that there may be some similarities between them that are not shared by verbs or other word-classes. Everyone has to do some lumping, but how far does the lumping go? There does not seem to be an objective way of deciding. Thus, the distinction between the two approaches (ancients vs. moderns, Weber vs. Floyd) boils down to a distinction of terminological preferences.

8 A wrong question: Are there verbs in Tagalog? (Kaufman 2009)

A lack of a noun-verb distinction seems even more radical than a lack of an adjective class, so this issue seems even more important (cf. Evans & Osada 2005). Austronesian languages, and especially Tagalog, have been prominent in these debates, most recently Kaufman (2009) and the commentaries published in the same journal issue (e.g. Aldridge 2009, Richards 2009). Kaufman notes that in Tagalog, action-roots and thing-roots behave alike in reference and predication constructions, as there is no copula in (17b), and the referential use of the action-root does not require more than the nominative marker *ang*.

- | | | | | |
|---------|------------------------------------|-------------------------------------|-------------------|-------------------------------------|
| (17) a. | <i>Nag-íngay</i> | <i>ang</i> | <i>áso.</i> | (action-predicate & thing-referent) |
| | AGENTVOICE-noise | [NOM | dog] | |
| | ‘The dog made noise.’ | | | |
| b. | <i>Áso ang nag-íngay.</i> | (thing-predicate & action-referent) | | |
| | dog | [NOM | AGENTVOICE-noise] | |
| | ‘The one who made noise is a dog.’ | | | |

Moreover, property-roots behave in the same way:

- (18) a. *Ma-bilis ang áso.* (property-predicate & thing-referent)
 STATIVE-quick [NOM dog]
 'The dog is quick.'
- b. *Áso ang ma-bilis.* (thing-predicate & property-referent)
 dog [NOM STATIVE-quick]
 'The quick one is a dog.'

And in attribution, all three root-groupings also behave alike, requiring nothing but the linker morpheme *-ng/na* between the head and the attribute.

- (19) a. *ang áso -ng ma-bilis* (thing-referent & property-attribute)
 NOM dog LK STAT-quick
 'the quick dog'
- b. *ang áso -ng nag-íngay* (thing-referent & action-attribute)
 NOM dog LK ACTORVOICE-noise
 'the dog who made noise'
- c. *ang ma-bilis na nag-ínay* (property-referent & action-attribute)
 NOM quick LK ACTORVOICE-noise
 'the quick one who/which made noise'
- d. *ang nag-íngay na áso* (action-referent & thing-attribute)
 [NOM AGENTVOICE-noise] LK dog
 'the noise-maker who is a dog' (= 19b)

Kaufman concludes that Tagalog has a single macroclass of Nouns. But unsurprisingly, if we adopt Floyd's principle of complete identity of behaviour, then we cannot say that Tagalog has just a single word-class. Most strikingly, action-roots take aspect-modality inflection and voice affixes (e.g. the prefix *nag-*), while thing-roots do not have these possibilities. These morphological differences are very salient, so linguists who have claimed that Tagalog is unlike English with respect to its word-classes have usually said that Tagalog makes no distinction between "syntactic word-classes", only between "morphological word-classes".

But note that aspectual and voice marking is non-uniform across the class of "verbs" in many languages, and that in all languages, verbs have inflectional subclasses. So in the absence of clear criteria that determine what constitutes a major class and what constitutes a subclass, one could maintain that the syntactic uniformity seen in (16)-(18) justifies the postulation of a single major world-class (Verb, or Noun; Kaufman chooses the latter label), with subclasses based on (less important) morphological criteria.

However, there is evidence that syntactically, too, not all roots behave alike: In some contexts a copula seems to be required with thing-roots (Richards 2009: 141), e.g. when the predicate is a complement of a verb of desire:

- (20) a. *Ayo-ko na-ng l-um-angoy.*
 notwant-1SG now-LK swim-ACTORVOICE
 'I don't want to swim anymore.'

- b. *Ayo-ko na-ng maging doktor*
 notwant-1SG now-LK be doctor
 ‘I don’t want to be a doctor anymore.’ (**Ayoko nang doktor.*)

But again, this could be described by setting up different subclasses of the broad (macro-)Noun category, if one decided to give more weight to the criterion of behaviour in ordinary predicative, attributive and referential contexts. Again, there is no objective, rigorous way of deciding.

9 A wrong question: Are there adjectives in Chamorro? (Chung 2012)

According to Topping (1973)’s structuralist (categorical particularist) analysis of Chamorro, this language has two word-classes, Class I (transitive verbs) and Class II (intransitive verbs, nouns, adjectives). Class I is basically identified by the fact that it combines with preposed subject person forms (cf. preposed *hu* in 20a), while Class II words combine with postposed subject person forms (cf. postposed *yu’* in 20b).

- (21) a. *Hu li’i i dangkulu na taotao.* (Class I)
 1SG see the big LK person
 ‘I saw the big person.’ (Chung 2012: 11)
- b. *H<um>ahanao yu’ gi chalan.* (Class II, action-root)
 <AGR>go.PROG 1SG LOC road
 ‘I was going on the road.’ (Chung 2012: 11)

Not only action-roots, but also thing-roots and property-roots combine with postposed subject person forms in this way, so with regard to this criterion, all Class II roots behave alike, justifying Topping’s classification.

- (22) a. *Laña’na puñeta-n taotao hao.* (Class II, thing-root)
 INTJ COMP expletive-LK person 2SG
 ‘My, what a (expletive) person you are.’ (Chung 2012: 11)
- b. *Dangkulu gui’.* (Class II, property-root)
 AGR.big 3SG
 ‘She becomes big.’ (Chung 2012: 11)

Now Chung (2012) claims that on closer inspection, Chamorro has nouns, verbs and adjectives after all. In particular, within Class II, we can distinguish a Noun subclass, because only Nouns can be incorporated, can be prefixed with *mi-* and do not allow subject-predicate agreement. Moreover, we can also distinguish between an Adjective and a Verb subclass, because only Verbs allow a specific external argument. The latter distinguishing criterion is a very subtle one, and it is possible that Topping simply missed it: Chung finds that normally, neither Nouns, nor Adjectives, nor Verbs allow nonspecific (bare indefinite) subjects, so neither ‘A teacher knows us’, ‘A shirt is nice’, nor ‘A teacher is a good person’ is possible in Chamorro. The subject must be specific, as indicated by the specific article *i* (cf. 21a). However, when the predicate is a Noun or an Adjective, this requirement is relaxed: Only the POSSESSOR of the subject has to be specific, so that sentences like (23) are possible, even though the head of the subject noun phrase is a bare indefinite (and lacks the definite article *i*).

- (23) *Bunitu maru' Josephine.*
 AGR.pretty box.kite Josephine
 'Josephine's box kite is pretty.' (Chung 2012: 23)

With verbs, even such "semi-specific" subjects are not possible, the subject has to be fully specific. This feature therefore distinguishes Verbs from Adjectives and Nouns, and Chung takes it as sufficient to claim that Chamorro does have the classical verb-noun-adjective system that is familiar from English.

The ways in which the different kinds of roots differ can this be summarized as in Table 4.

Table 4. Six features of different kinds of roots in Chamorro

features	'see'-type roots	'go'-type roots	'big'-type roots	'person'-type roots
1 passive	+	–	–	–
2 postposed subject person form	–	+	+	+
3 incorporatable	–	–	–	+
4 prefixable with <i>mi-</i>	–	–	–	+
5 subject-predicate agreement	+	+	+	–
6 specific external argument required	+	+	–	–

Clearly, if all these criteria have the same weight, then quite a few different ways of setting up major classes are possible (see Haspelmath 2012), and it is not immediately clear which of the major-class divisions, if any, is better than others.

But Chung does not even ask this question – she primarily asks whether Chamorro can be described/analyzed with the English-type category system that Baker (2003) also argued was universal. Thus, she basically adopts the Uniformity Principle in (6) and asks whether there is sufficient evidence against the hypothesis that Chamorro is like English. A complete lack of formal differentiation between action-roots and property-roots might be reason for the universalist to worry. But Chung has found a piece of evidence for Adjectives (the possibility of "semi-specific subjects), and she takes this as supporting the universalist view.

The problem is not so much that the distinguishing criterion seems particularly far-fetched in this case (much more so than the criteria adduced by Floyd and Richards for the splitting approach in Quechua and Tagalog) – the more general problem is that there is no constraint on what kinds of criteria can be used to set up major categories, and that different criteria are used for different languages. This leads to arbitrary, subjective decisions, and to a nonrigorous methodology.

Instead of asking whether Chamorro can be described with nouns, verbs and adjectives, one might ask whether English, or for that matter all languages, can be described in the Chamorro manner, using Class I (words with objects) and Class II (words without objects). Clearly, the difference between these two types of words is not as salient in all languages as it is in Chamorro (in most languages, pronominal subjects are not coded differently Class I and Class II), but if just any kind of criterion is sufficient to make the distinction, then surely one will find some way of distinguishing between transitive and intransitive words in all languages (if only by the fact that only transitive words can take objects).

10 Comparing languages on the basis of root-groupings

I hope that the above considerations and examples have shown that languages cannot be compared on the basis of language-particular word-classes, because different criteria are used in different languages to establish the word-classes. The cross-linguistic questions in (2) are the wrong questions

Moreover, we have seen that there is no good basis for distinguishing between major classes and subclasses in particular languages. Floyd, Richards and Chung have all made valid observations on Quechua, Tagalog and Chamorro that previous research had overlooked (or at least not highlighted), but this does not invalidate the earlier observations that Quechua, Tagalog and Chamorro are interestingly different from English.

So how can we capture the valid insights of this earlier research, how can we set up comparative concepts that allow us to express the interesting differences between languages in a more general way? Above in §5 I proposed that languages should be compared on the basis of the following semantically based notions:

- (9) a. thing-root: a root that denotes a physical object (animate or inanimate)
- b. action-root: a root that denotes a volitional action
- c. property-root: a root that denotes a property such as age, dimension or value

These comparative concepts appear to have nothing to do with the formal categories of language that linguists pride themselves on being able to discover. However, there are many different kinds of formal categories. Why should some of them be privileged over others? Why should some classes be called *parts of speech*, even though speech has many diverse parts, or *word-classes*,⁷ even though words can be grouped into classes on the basis of many different criteria? Why should thing-roots, action-roots and property-roots have a special status? There are two answers to this. (The second answer will be given in the next section.)

The first answer has to do with the habits of linguists. Even though we rarely admit it, the concepts in (9a-c) are at the basis of what we normally do. If we found strong evidence for grammatical classes of words that have nothing to do with things, actions and properties, we would not call them word-classes. For example, suppose a language has a class of prefixing words such as ‘father’, ‘mouth’, ‘kill’, and ‘eat’, and a class of non-prefixing words such as ‘tree’, ‘knife’, ‘swim’ and ‘sit’, we would not say that these are word-classes, even if the prefixing vs. non-prefixing distinction is important for a number of different regularities in the language. Thus, while saying that “nouns refer to things, verbs to actions and adjectives to properties” may sound naive, from a cross-linguistic perspective, this really is the meaning of these terms.

But why do I suggest that the comparative concepts should be formulated in terms of *roots*, rather than words (as in (8) above)? The problem is that the traditional conception of word-classes is based on the difference between inflected words (word-forms) and lexemes. Word-classes are normally LEXEME CLASSES.⁸

⁷ Note that I hyphenate *word-class* in the more specific sense, in order to distinguish it from ad hoc word classes of other kinds. This is also the reason why I hyphenate the terms *thing-root*, *action-root* and *property-root*.

⁸ It has sometimes been said that word-class is a property of inflected words, so that an inflected form of a verb lexeme could be an adjective (i.e. a participle), or an inflected form of a verb could be a noun (i.e. a verbal noun) (see Haspelmath 1995). Should one say that participles (such as German *helf-end* ‘who helps’) are only adjectives, even though they are forms of the lexeme

But this presupposes that one can make a consistent distinction between inflection and derivation. For example, if the English adverb-forming suffix *-ly* is regarded as an inflectional suffix, then *quickly* is an inflected adverb form of the (adjectival) lexeme *quick*, hence it is an adjective. But if the suffix *-ly* is regarded as a derivational form, then *quickly* is a derived adverb lexeme. It turns out that there is no good general way of distinguishing between the two kinds of processes (Plank 1994, Dixon 2010.1: §5.3), so we cannot make this distinction the basis of our definition. Another serious problem is that there is no good general way of distinguishing inflectional affixes from separate clitic words (Haspelmath 2011).

The solution that I propose here is to consider just the roots in a cross-linguistic context. All languages have a substantial number of thing-roots (e.g. tree, door, child), action-roots (e.g. run, talk, break) and property-roots (e.g. good, old, small). These groupings of roots typically behave similarly (i.e. 'tree' behaves like 'door', 'run' behaves like 'talk', 'good' behaves like 'old', etc.). Thus, we can limit our typological research to roots, and specifically to ways in which languages express the three major root-groupings (thing-roots, property-roots, action-roots).

Of course, languages have many complex expressions ("words") that behave like the roots, and in descriptions of individual languages, we want to describe these, too. So we want to say that both *break* and *enlarge* are Verbs in English, that both *king* and *kingdom* are Nouns, and that both *red* and *reddish* are Adjectives. Likewise, languages have many words (often even roots) that behave like verbs but are not actions (e.g. English *to love*), words that behave like nouns but are not things (e.g. *war*), and words that behave like adjectives but do not denote properties (e.g. *royal*). Thus, if we limit our cross-linguistic comparison to roots, and to roots denoting things, actions and properties, we compare languages only with respect to a part of their vocabulary.

This is a price that we pay for our methodological rigour: The great advantage is that we can readily identify roots in any language (as opposed to "words", which cannot be identified rigorously across languages)⁹, and we can readily identify things, actions and properties¹⁰ (as opposed to "nouns", "verbs" and "adjectives"). But it is easy to see that the phenomena that are still in our purview are at the core of what we are interested in, so while we may lose the fringes, we retain the core. And as a general point, we have to keep in mind that language comparison cannot be all-encompassing anyway: Languages are comparable with respect to many of their features, but we can never draw all features into the comparison. Language comparison is a different enterprise from language description, which must be all-encompassing (all aspects of a language have to be described). Language comparison often works with even smaller core phenomena, e.g. Haspelmath (2005), which considers just the verb 'give', rather than the broader heterogeneous domain of all ditransitive constructions.

helfen 'to help'? Or should one say that they are verbs at the lexeme level, and adjectives at the inflectional level of the word-form? There is no clear answer to this, but the view adopted widely by linguists (often implicitly) has been to say that only the lexeme word-class of a word counts, i.e. inflection does not influence the word-class assignment. (The issue loses much of its relevance in view of what is said further in the text.)

⁹ A further terminological remark: In the present context, *root* can simply be equated with *morph* ('smallest meaningful piece of form'), because linguists will normally call morphs which refer to things, actions or properties *roots* (rather than *affixes*). (There is sometimes some question about how to delimit roots from affixes, cf. Haspelmath 2002: 19-20, and in a cross-linguistic context, it seems best to define roots as 'morphs that denote things, actions, or properties'.)

¹⁰ 'Property' is perhaps not as clear a concept as 'thing' and 'action', so we could limit ourselves to the four core types of properties identified by Dixon (1977): age ('old'), dimension ('small'), value ('good'), colour ('red').

Thus, I propose that we compare languages with respect to their root-groupings, i.e. with respect to the grammatical behaviour of thing-roots, action-roots and property-roots. These could be called “root classes”, but I choose the term *root-grouping* to remind us that these are very special kinds of classes, namely semantic classes of ontological categories of a particular type.

The term *word-class* is retained for language-particular syntactic classes of roots (and other similar elements, often called “words” in language-particular descriptions). But what can be compared is root-groupings, so a typology of word-classes will really be a typology of root-groupings.

As was noted earlier, these comparative concepts are used, for example, in typological work such as Dryer (2005a, 2005b) on adjective-noun order and on verb-object order. Dryer does not quite present it in this way, but in practice, there is no difference. This kind of approach was also adopted by Greenberg (1963), though he was even less explicit about his defining criteria. In addition to allowing us to compare word order across languages, the concepts in (9) can also be used to ask questions about coding, e.g. Which languages use a copula when predicating a thing-root? Are there languages that require a copula with property-roots, but not with thing-roots? This will be briefly discussed in the next section.

11 Comparing languages on the basis of typical associations of root-grouping and propositional-act type

In the last section, we asked why thing-roots, action-roots and property-roots should have special status. The second answer that we can give to this is that in languages in general, they tend to behave in a special way in the three major propositional-act types reference, predication, and attribution (we already saw these in the Tagalog examples in (17)-(19)). In particular,

- when thing-roots are used referentially, they tend to lack special function-indicating coding such as nominalization,
- when action-roots are used as predicates, they tend to lack special function-indicating coding such as copulas, and
- when property-roots are used as attributes, they tend to lack special function-indicating coding such as relative-clause marking or possessive marking.

Thus, the shaded cells in Table 5 show expressions with no extra function-indicating coding, whereas the other cells all have some overt marking (given in boldface in the table; only the elements in the shaded cells lack this special coding). English is quite typical in this regard.

Table 5. Root-groupings and propositional-act types (Croft 1991: 67)

	reference	predication	attribution
thing-roots	WATER	(that) is water	(colour) of water
action-roots	the runn- ing	(it) RUN(-s)	runn- ing (water)
property-roots	the wet- ness	(water) is wet	WET (water)

Thus, the coding of the root-groupings in the different propositional-act types is quite fundamental to the nature of word-classes. It is probably only because of these striking and highly regular coding similarities that the terms “noun”, “verb” and “adjective” have been adopted from Latin into other languages in the Western tradition.

Against this background, we can understand what motivates the lumpers: Weber, Kaufman and Topping noted that in Quechua, Tagalog and Chamorro, the distinctions

made in English in Table 5 are not made in the same way. In Quechua, thing-roots behave in the same way as property roots in attribution, and in Tagalog, even all three root-classes behave in the same in all three propositional-act functions. These are thus salient differences that need to be expressed in some way, because the languages are lumpers not only with respect to English, but also with respect to the cross-linguistic trend.

Since the patterns in Table 5 are very general across languages, one can also set up other kinds of comparative concepts, as in (24).

- (24) a. nouns are roots used for **reference** without special coding (reference-roots)
 b. verbs are roots used for **predication** without special coding (predication-roots)
 c. adjectives are roots used for **attribution** without special coding (attribution-roots)
 d. manner adverbs are roots used for adverbation without special coding (adverbation-roots)¹¹

This is more or less the approach taken by Hengeveld (1992), Hengeveld et al. (2004), Hengeveld & van Lier (2008), van Lier (2009), and Hengeveld & van Lier (2010). This sort of definition of comparative concepts allows Hengeveld and associates to formulate some interesting generalizations, summarized by the parts-of-speech hierarchy in (25):

- (25) parts-of-speech hierarchy:
 predication > reference > attribution > adverbation

The more to the left a propositional act is on the hierarchy, the more likely it is that a language has a specialized word-class for that propositional act. Thus, if a language has a specialized class of nouns (roots used for reference without coding), it also has a specialized class of verbs (roots used for predication without coding), and so on.

12 A brief history of thinking about word-class universality

The study of word-classes has a long history, and in this last section I would like to briefly recall a fascinating account of this history by Bossong (1992), which has not become as widely known as it deserves. Bossong observes that over the centuries, the pendulum has swung back and forth between a particularist and a universalist approach to language diversity. This is summarized in Table 6.

Table 6: The pendulum of particularism and universalism

Antiquity + Renaissance (particularism):

interest only or primarily in language-particular description, no universal claims – but little awareness of categorial differences between languages

Middle Ages + Enlightenment (universalism):

ambitious claims about universal categories of language and thinking – but no interest in differences between languages

19th century (particularism):

ambitious claims about languages differing from each other in their categories, and thus

¹¹ Hengeveld's approach is different from Croft in that it also takes into account manner adverbs. For these, a new kind of function (called „adverbation“ here; this is my term, not Hengeveld's) needs to be set up.

in their thinking – but no attempt at rigorous description (Latin categories often carried over to other languages in practice)

First phase of 20th century (particularism):

rigorous description of languages of different types, with the new insight that each language has its own categories (Franz Boas, Edward Sapir, Ferdinand de Saussure)

Second phase of 20th century (universalism):

ambitious claims about universal categories of language and thinking (Noam Chomsky) – and serious attempts to find the categories of English (or Latin) in all other languages

I would hope that the two approaches will soon be married happily, as in (26):

(26) 21st century (particularism and universalism):

respect for differences between languages, no ethnocentrism, no confusion of universal categories and universal trends

If one recognizes that language description and language comparison are two distinct enterprises, one can show respect for the differences in descriptions, but at the same time bring out the generalizations in comparative work.

13 Conclusion

Languages differ in more ways than we might naively suspect. But they also show many striking similarities that seem to reflect their functional unity. These similarities cannot be captured by setting up a set of universal categories and asserting that languages make all (or many, or some) of these categorial distinctions. Equating categories across languages but using different criteria in different languages is not a rigorous methodology. It leads to arbitrary, subjective decisions and unresolvable debates.

What needs to be done is to compare languages in terms of a special set of comparative concepts. Very promising work of this kind in the domain of word-classes, root-groupings and propositional-act functions has been done by Bill Croft and Kees Hengeveld, but many linguists are still trying to ask questions such as those in (1) and (2) which cannot be answered, because they are the wrong questions.

References

- Aldridge, Edith. 2009. Minimalist questions for the nominalist analysis of Tagalog syntax. *Theoretical Linguistics* 35(1). 51–62.
- Anward, Jan, Edith Moravcsik & Leon Stassen. 1997. Parts of speech: A challenge for typology. *Linguistic Typology* 1(2). 167–184.
- Baker, Mark C. 2003. *Lexical categories: verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Beck, David. 2012+. Uni-directional flexibility and the noun–verb distinction in Lushootseed. In Jan Rijkhoff and Eva van Lier (eds.), *Flexible word classes: A typological study of underspecified parts-of-speech*. Oxford: Oxford University Press. (To appear)

- Boas, Franz. 1911. Introduction. In Franz Boas (ed.), *Handbook of American Indian Languages*, 1-83. Washington, DC: Bureau of American Ethnology.
- Bossong, Georg. 1992. Reflections on the history of the study of universals: the example of the partes orationis. In Michel Kefer & Johan van der Auwera (eds.), *Meaning and grammar: cross-linguistic perspectives*, 3-16. Berlin: Mouton de Gruyter.
- Chomsky, Noam A. 1957. *Syntactic structures*. 's-Gravenhage: Mouton.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Chung, Sandra. 2012. Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 38: 1-56.
- Croft, William. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.
- Croft, William. 2000. Parts of speech as language universals and as language-particular categories. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 65-102. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2009. Methods for finding language universals in syntax. In Sergio Scalise, Elisabetta Magni, & Antonietta Bisetto (eds.), *Universals of language today*, 145-164. Berlin: Springer.
- Davis, Henry & Lisa Matthewson. 1999. On the functional determination of lexical categories. *Revue québécoise de linguistique* 27(2). 29-69.
- Dixon, R.M.W. 1977. Where have all the adjectives gone? *Studies in Language* 1. 19-80.
- Dixon, R.M.W. 2004. Adjective classes in typological perspective. In R.M.W Dixon & Alexandra Y. Aikhenvald (eds.), *Adjective classes: A cross-linguistic typology*, 1-49. Oxford: Oxford University Press.
- Dixon, R.M.W. 2010.1. *Basic linguistic theory*. Vol. 1. Methodology. Oxford: Oxford University Press.
- Dixon, R.M.W. 2010.2. *Basic linguistic theory*. Vol. 2. Grammatical topics. Oxford: Oxford University Press.
- Dryer, Matthew S. 1997. Are grammatical relations universal? In Joan L. Bybee, John Haiman, & Sandra A. Thompson (eds.), *Essays on Language Function and Language Type: Dedicated to T. Givon*, 115-143. Amsterdam: Benjamins.
- Dryer, Matthew S. 2005a. Order of adjective and noun. In Martin Haspelmath, Matthew S Dryer, David Gil, & Bernard Comrie (eds.), *The world atlas of language structures*, 346-349. Oxford University Press.
- Dryer, Matthew S. 2005b. Order of object and verb. In Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.), *The world atlas of language structures*, 338-341. Oxford: Oxford University Press.

- Evans, Nicholas & Stephen C Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(05). 429–448.
- Evans, Nicholas & Toshiki Osada. 2005. Mundari: The myth of a language without word classes. *Linguistic Typology* 9(3). 351–390.
- Floyd, Simon. 2011. Re-discovering the Quechua adjective. *Linguistic Typology* 15. 25–63.
- Gil, David. 2000. Syntactic categories, cross-linguistic variation and universal grammar. In Petra M. Vogel & Bernard Comrie (eds.), *Approches to the typology of word classes*, 173–216. Berlin: Mouton de Gruyter.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin. 1995. Word-class-changing inflection and morphological theory. In: Booij, Geert & van Marle, Jaap (eds) *Yearbook of Morphology 1995*, 43–66. Dordrecht: Kluwer.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.
- Haspelmath, Martin. 2005. Ditransitive constructions: The verb ‘give.’ In Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 426–429. Oxford: Oxford University Press.
- Haspelmath, Martin. 2007. Pre-established categories don’t exist: consequences for language description and typology. *Linguistic Typology* 11. 119 - 132.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(2). 31–80.
- Haspelmath, Martin. 2012. Escaping ethnocentrism in the study of word-class universals. *Theoretical Linguistics* 38: 91-102.
- Hengeveld, Kees & Eva van Lier. 2008. Parts of speech and dependent clauses in Functional Discourse Grammar. *Studies in Language* 32(3). 753–785.
- Hengeveld, Kees & Eva van Lier. 2010. An implicational map of parts of speech. *Linguistic Discovery* 8(1). 129–156.
- Hengeveld, Kees, Jan Rijkhoff & Anna Siewierska. 2004. Parts-of-speech systems and word order. *Journal of Linguistics* 40(3). 527–570.
- Hengeveld, Kees. 1992. *Non-verbal predication: theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In: Austin, Peter K. & Musgrave, Simon (eds.) *Voice and grammatical functions in Austronesian languages*. Stanford: CSLI, 247–293.
- Hoffmann, John. 1903. *Mundari Grammar*. Calcutta: The Secretariat Press.

- Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35(1). 1–49.
- Keenan, E.L. 1974. Logic and language. In *Language as a Human Problem*, ed. M Bloomfield & E. Haugen, 187–196, New York: W.W. Norton and Company.
- Keenan, Edward L. & Stabler, Edward P. 2003. Linguistic invariants and language variation. 12th International Congress on Logic Methodology and Philosophy of Science (LMPS'03), Oviedo, Spain, August 7–13, 2003 (http://www.freewebs.com/philomanian/Sociolinguistics/linguistic_invariants_and_language_variation.pdf)
- Kinkade, M. D. 1983. Salish evidence against the universality of noun and verb. *Lingua* 60(1). 25–39.
- Kuipers, A. H. 1968. The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21.610–26.
- Lazard, Gilbert. 1992. Y at-il des catégories interlangagières? In Susanne Anschutz (ed.), *Texte, Sätze, Wörter und Moneme: Festschrift für Klaus Heger zum 65. Geburtstag*, 427–434. Heidelberg: Heidelberg Orientverlag.
- McCawley, James D. 1992. Justifying part-of-speech assignments in Mandarin Ch. *J. of Chinese Linguistics* 20(2). 211–246.
- Mithun, Marianne. 2000. Noun and verb in Iroquoian languages: Multicategorisation from multiple criteria. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 397–420. Berlin: M. de Gruyter.
- Mosel, Ulrike & Even Hovdhaugen. 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Müller, F. Max. 1861–64. *Lectures on the science of language*. London.
- Paul, Waltraud. 2005. Adjectival modification in Mandarin Chinese and related issues. *Linguistics* 43(4). 757–793.
- Plank, Frans. 1994. Inflection and derivation. In R.E. Asher (ed.), *Encyclopedia of Language and Linguistics*, vol. 3, 1671–1678. Oxford: Pergamon.
- Post, Mark. 2008. Adjectives in Thai: Implications for a functionalist typology of word classes. *Linguistic Typology* 12(3). 339–381.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Rauh, Gisa. 2010. *Syntactic categories: their identification and description in linguistic theories*. Oxford: Oxford University Press.
- Richards, Norvin. 2009. Nouns, verbs, and hidden structure in Tagalog. *Theoretical Linguistics* 35(1). 139–152.
- Sasse, Hans-Jürgen. 1988. Der irokesische Sprachtyp. *Zeitschrift für Sprachwissenschaft*.
- Sasse, Hans-Jürgen. 1991. Predication and sentence constitution in universal perspective. In: Zaefferer, Dietmar (ed.) *Semantic universals and universal semantics*, Berlin: Foris, 75–95.

- Schachter, Paul & Timothy Shopen. 2007. Parts-of-speech systems.. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. I, 1–60. 2nd ed. Cambridge: Cambridge University Press.
- Schachter, Paul. 1985. Parts-of-speech systems. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. I, 3–61. Cambridge: Cambridge University Press.
- Sebba, Mark. 1986. Adjectives and copulas in Sranan Tongo. *Journal of Pidgin and Creole languages* 1(1). 109–121.
- Seuren, Pieter A. M. 1986. Adjectives as adjectives in Sranan: A reply to Sebba. *Journal of Pidgin and Creole languages* 1(1). 123–134.
- Topping, Donald. 1973. *Chamorro reference grammar*. Honolulu: University Press of Hawai'i.
- van Eijk, J. P. & T. Hess. 1986. Noun and verb in Salish. *Lingua* 69(4). 319–331.
- van Lier, Eva H. 2009. *Parts of speech and dependent clauses: A typological study*. Leiden: LOT.
- Vonen, Arnfinn Muruvik. 2000. Polynesian multifunctionality and the ambitions of linguistic description. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 479–487. Berlin: Mouton de Gruyter.
- Weber, David John. 1989. *A Grammar of Huallaga (Huánaco) Quechua*. Berkeley: U. of California Press.
- Winford, Donald. 1997. Property items and predication in Sranan. *Journal of Pidgin and Creole languages* 12(2). 237–301.

Affiliation

Martin Haspelmath
Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany
haspelmath@eva.mpg.de