

The structural uniqueness of languages and the value of comparison for language description

Martin Haspelmath

Max Planck Institute for Evolutionary Anthropology | Leipzig University

This paper shows why it is not a contradiction to say that each language is structurally unique and must be described with its own categories, but language description profits enormously from typological knowledge. It has sometimes been suggested that the Boasian imperative (“each language should be described in its own terms”) leads to un insightful analyses, and that language description should instead be “typologically informed”. But the Boasian imperative is not at all incompatible with an intimate connection between description and comparison: Comparative (or typological) knowledge is highly valuable both for making our descriptions transparent and comprehensible, and for helping describers to ask a wide range of questions that would not have occurred to them otherwise. Since we do not know whether any of the building blocks of languages are innate and universal for this reason, we cannot rely on general frameworks (of the generative type) for our descriptions, but we can use typological questionnaires and other kinds of comparative information as a scaffold. Such scaffolds are not theoretical components of the description, but are important methodological tools.

Keywords: language description, linguistic typology, comparative grammar, methodology of linguistics

1. Languages are comparable but structurally different

Languages are different and similar in confusing ways. This paper aims to clarify a widespread misconception: That the evident similarities between languages, which are now universally recognized and emphasized, **MUST REFLECT SIMILARITIES IN THEIR GRAMMATICAL AND LEXICAL STRUCTURES**. Languages can be similar

even when they are structurally different, and this has consequences for the practice of language description.

Clearly, different languages are commensurable to the extent that translation from one language to another is not only possible, but not even particularly difficult in many cases. So it would be an exaggeration to say that a language represents a conceptual world of its own that outsiders cannot penetrate unless they learn it as thoroughly as the native speakers. This is what makes it possible to know about other cultures, and to compare languages and cultures successfully. The subfield of comparative linguistics (or “linguistic typology”) is thriving, and works such as the *World atlas of language structures* (Haspelmath et al. 2005) are generally regarded as successful and interesting, just as anthropologists have long compared a wide range of cultures from around the world (e.g. Boas 1896; Ember & Ember 1998; Candea 2018; Slingerland et al. 2020).

But on the other hand, translation is not trivial, as languages differ not only in the shapes of their words and morphs, but also in their syntactic and semantic structures.¹ Translation would be trivial if each word had an exact semantic match in the other language, and if languages had fully equivalent grammatical constructions. In many simple cases, this does indeed seem to be the case. For example, there does not seem to be a noticeable problem with translating the English sentence (1) into Mandarin Chinese, Russian and Hebrew, as seen in (2a–c).

(1) *Kim ate three apples yesterday.*

(2) a. Mandarin

老金昨天吃了三个苹果。

Lao-Jīn zuótiān chī-le sān-gè píngguǒ.

Kim yesterday eat-PFV three-CLF apple

b. Russian

Ким вчера съела три яблока.

Kim včera s-”e-l-a tri jablok-a

Kim yesterday PFV-eat-PST-F.SG three.ACC apple-SG.GEN

c. Hebrew

תפוחים שלוש אכל אתמול קים

tapuḥ-im shlosh-a etmol axal Kim

apple-PL three-M yesterday ate.3SG Kim

A superficial look at such simple cases could lead one to conclude that by and large, our languages have equivalent means of expressing our thoughts, and that learning another language boils down to learning the words plus a few unmoti-

1. On the problems of translation, see the recent anthropological perspective in Hanks & Severi (2014).

vated rules such as word order rules (e.g. that the ‘yesterday’ word typically follows the subject in Mandarin, and that it can occur between the verb and the object in Hebrew), plus idiosyncratic rules of inflection (e.g. that the Hebrew plural suffix is *-im* in *tapuh-im* ‘apple’, not *-ot* as in many other words).

However, linguists have long been aware that this view is too simple. Different languages often use different structures or different building blocks to render equivalent ideas. This goes beyond the dropping or adding of obligatory elements, such as dropping the plural marker in Mandarin, adding the gender marker to the verb in Russian and to the numeral in Hebrew, and adding the classifier *-ge* in Mandarin.

Different languages have different tense-aspect systems (as illustrated by the perfective markers in Russian and Mandarin) and different case systems (as illustrated by the surprising genitive case in Russian *jablok-a* ‘apple’), and they use their plural markers differently (as also illustrated by the singular form *jabloka* used after the numeral *tri*). Different languages have different determiner systems (English and Hebrew could easily add a definite article to the object: *the apples*, *ha-tapuhim*, but Russian cannot), different ways in which classificatory information is relevant for modifiers (cf. the Hebrew masculine numeral *shlosh-a*, where Russian lacks a gender contrast with *tri* ‘three’, and where Mandarin requires the classifier), and different ways in which causative clauses are formed: English can say *Lee made Kim eat three apples yesterday*, but the counterparts of this would look rather different in the other languages.

Thus, the morphosyntactic structures of languages do not map onto each other straightforwardly, and the same applies to their phonological systems and their semantic systems. Even with simple verbs like ‘eat’, this can be illustrated by the contrast between English and German. German *essen* is generally a good translation of English *eat*, but it cannot be used for animals (when a dog eats an apple, this is expressed with a different verb, *fressen*). The differences between languages with respect to grammatical meanings are increasingly attracting the attention of semanticists, as illustrated by works like Plungian (2011) and Bochnak & Matthewson (2015).

The structural differences between languages have long been known, and they were particularly emphasized in the first half of the 20th century when linguists got increasingly interested in language structures for their own sakes (not merely as reflecting interesting histories). Lyons (1968: 55) discussed semantic differences and gave the example of kinship terms in Russian and English, noting that where English has the two words *brother-in-law* and *son-in-law*, Russian has the four different words *zjat* ‘(sister’s or daughter’s husband), *šurin* (wife’s brother), *svojak* (wife’s sister’s husband), and *dever*’ (husband’s brother). Lyons noted that the lack of a perfect match does not mean that English *brother-in-law* has several differ-

ent meanings (= is polysemous), or that Russian *zjat'* is polysemous. They simply have different meanings that do not map straightforwardly onto each other:

Every language has its own semantic structure. To the degree that the meanings of one language can be brought into one-to-one correspondence with those of another, we will say that the two languages are semantically isomorphic (have the same semantic structure). The degree of semantic isomorphism between different languages varies considerably. (Lyons 1968: 55)

This view is often associated with “structuralism” (as the mainstream approach of linguists who were working before the 1970s is sometimes called), but it has never been controversial.² It was not emphasized by linguists in the 19th century and has been in the background since the 1970s, but nobody has claimed that there is something wrong with the general view that languages are structurally different. (For fairly clear statements of the structural uniqueness of phonemes, reasserting the insights of Trubetzkoy (1939), see Lass (1984) and Simpson (1999); in this paper, I focus on morphosyntax.)

So this is what I mean by *structural uniqueness* of languages: Even though in many simple cases, languages are quite similar, there are also many other cases where they exhibit phonological, morphosyntactic and semantic structures which are different from some other languages. So there is no claim that everything about a language is different from every other language. Clearly, many meanings map very accurately across very many languages (e.g. those of English *three*, Russian *tri*, Mandarin *sān*, and so on), and many phonemes are very similar across different languages (e.g. Mandarin /a/ and Hebrew /a/; but English only has /æ/ and /ʌ/, and Russian /a/ has a very peculiar stress-conditioned distribution).

So all I mean by structural uniqueness is that languages are not always structurally isomorphic (in Lyons’s sense), and it is of course these *non-isomorphisms* that have led to most of the discussions within comparative linguistics. They present us with a challenge, and as I briefly noted at the very beginning, few linguists would say that the structural differences imply that languages are not similar and cannot be compared. Languages are obviously similar when they are isomorphic, but it is possible to see similarities even in the absence of isomorphisms. This was already illustrated by the Russian kinship terms mentioned above, which are not so completely different from English kinship terms that comparison is impossible. But what are the consequences of this for language description?

In the next section, I will elaborate somewhat on the nature of description of structurally unique systems. We will see that description does not rely on com-

2. See also my blogpost “We are all structuralists” (<https://dlc.hypotheses.org/2356>).

parison from a theoretical perspective, but in subsequent sections (§§ 3–5), I will illustrate how comparative information can be very valuable for our descriptive methods, so that it is important that descriptions should be “typologically informed” (Himmelman 2016: 475).

2. Describing the structures of a language in its own terms

The structures of a language are described by identifying the classes of elements (classes of sounds, classes of forms, classes of concepts) that are necessary for formulating the combinatory rules. As every linguist knows, this is done by determining contrasts within the language-particular system. A set of sounds is grouped under a phoneme (i.e. a structural element) if there are no contrasts internal to the set, and a set of concepts is grouped under a meaning if the only contrast is with other meanings in the system. For example, German /i/ and /ɪ/ form a contrast (e.g. in *Miete* [mi:tə] vs. *Mitte* [mɪtə]), but in Russian, the two sounds [i] and [ɪ] never form a contrast. In English, *brother* and *sister* contrast by sex (while age or profession or hair colour are irrelevant), whereas in Indonesian, *kakak* ‘elder sibling’ and *adik* ‘younger sibling’ contrast by age (while sex or profession or hair colour are irrelevant; this example is from Evans 2011, an excellent overview of cross-linguistic semantic non-isomorphisms).³

Grammatical classes, such as noun and adjective, or dative and genitive, or present tense and future tense, are likewise identified by system-internal distributional criteria. Again, this has never been controversial in linguistics. Many syntax textbooks note that while one learns in primary school that a noun is a word that denotes persons, things and places, this is not how linguists identify nouns (e.g. Tallerman 1998: 30). An English Noun is a word that can occur after the definite article *the* and form a complete nominal with it, and an English Future-tense construction is a construction with one of the auxiliaries *will*, *shall*, or *going to*. Semantic criteria typically give the same results, but when they do not, it is the system-internal criteria that are decisive. For example, English *war* denotes an event rather than a thing, but we can say *the war*, so it is an English Noun. And future time reference can sometimes be achieved with a different construction (e.g. *The restaurant opens at 19h00 tonight*), but this construction is not an English Future-tense construction.

3. As emphasized recently by LaPolla (2020), the idea that “each language manifests a unique set of cognitive categories” goes back at least to Wilhelm von Humboldt (LaPolla 2020: 148) and has a long tradition in comparative linguistics.

If we were content with describing a language in very superficial terms, we could alternatively use a rough notional approach. We might say, for example, that Mandarin *-le* in (2a) is a kind of past time reference marker, that Russian *-a* in *jablok-a* is a kind of plural marker, and that English *three* is an inanimate numeral in *three apples*. These descriptions would not be completely wrong, and one might say that they are useful as first approximations.

But linguists are more ambitious and want to describe each language comprehensively and in a way that does justice to the language, i.e. that does not distort it by applying inappropriate categories (such as animacy classes of numerals in English, where they are unmotivated, because there is no difference between *three dogs* and *three apples*). Especially fieldworkers (and anthropologists more generally) have a long tradition of criticizing descriptions that impose a preconceived viewpoint on a different language and culture. Mithun (2001: 51) highlights the importance of collecting natural texts if we want to “learn about the language in its own terms”, and Epps (2011: 648) notes that fieldworkers should “produce descriptions in formats that will enable and facilitate comparison across languages, but also remain true to the languages themselves, without forcing them into ill-fitting predetermined categories (Gil 2001; Haspelmath 2007, 2010b).” Since Boas (1911), this approach has often been described by the slogan that one should “describe a language in its own terms”, and we can call this the “Boasian imperative”. LaPollda & Poa (2006: 270) put it exactly in these terms:

Each language is a unique set of language-specific conventions, and so each language should be described in its own terms.

Now this may sound as if comparative linguistics could not contribute anything to the task of describing a language. But this is of course the precise opposite of what describers have experienced over the past few decades, during which we saw grammatical descriptions improve markedly, clearly as a result of greater awareness of the broad cross-linguistic patterns. This progress in the quality of grammars was highlighted by Dryer (2006), who credited the three “Shopen volumes” with contributing substantially to this (see Shopen (ed.) 1985, 2007).

Thus, describing a language in its own terms cannot mean that describers should ignore the insights from comparative linguistics and general theoretical ideas (see § 3–4 below). What it means is that the categories chosen for description are not DETERMINED by the typologists or general theorists. If a general typology tells me that most languages have a simple five-vowel system and that contrasts among high vowels (such as /i/ vs. /ɪ/) are unusual, this may lead me to look more carefully at the German vowel system. But in the end, it is language-internal contrasts like *Miete*/*Mitte* which decide on the analysis, rather than general ideas about vowel systems, claims about universal feature systems, and so on.

Likewise, if a typological theory tells me that “prototypical nouns” should denote objects (Croft 1991), it may lead me to check more carefully whether English *war* can occur with a definite article, but once this is clear, it is the English-specific criterion that decides on the analysis of English. And if a general theory of sonority specifies that fricatives should not pattern with liquids but with stops, this does not contradict the observation that in Russian, the voiced labial fricative /v/ patterns as a sonorant with respect to the Russian-specific Regressive Voicing Assimilation rule (so a word like *svoboda* ‘freedom’ is not pronounced [zv-], but [sv-]; Halle 1959: 63). Bickel (2015: § 2) expresses this by saying that our descriptions should not “force languages into Procrustean beds”.

Again, there is no real controversy about all this. Just as language-particular morphosyntactic criteria take precedence over semantic criteria in determining word classes, they also take precedence over typological expectations. However, there are many linguists (influenced by the Chomskyan tradition) who think that the Boasian imperative is not in conflict with a completely different idea, namely that there is an innate grammar blueprint which specifies a smallish number of building blocks (not more than a few hundred innate features and categories) that all languages are made up of. The idea of this approach is that a full analysis of a (part of a) language must be stated in terms of these innate building blocks (also called elements of “universal grammar”; see Baker 2001, who compares these hypothetical building blocks to the elements of chemistry).⁴ But even though many linguists are sympathetic to this general idea and attempt to make use of a set of specific innate building blocks, not much is known for sure about these (Haspelmath 2007). We do not know which tense-aspect categories are innately specified, in a way that would rule out certain kinds of descriptions of the Russian Perfective aspect prefix *s-* and the Chinese Perfective aspect suffix *-le* in (2). We do not know what is innately specified about classifiers and gender markers in such a way that it would constrain our descriptions of the Feminine gender marker *-a* in Russian and the Chinese classifier *-ge*. And the same applies to phoneme systems, word classes, and everything else in grammar (as well as in lexical semantics). So in practice, it is not possible (yet?) to describe languages by means of a universal set of innate building blocks, and much of the research in this tradition is quite speculative.⁵ There is thus no coherent practical alternative to the description of a

4. See also Davis et al. (2014) for a recent statement that explicitly attempts to reconcile the Boasian and the Chomskyan approach (see also my blogpost “Marrying Boas and Chomsky”, <https://dlc.hypotheses.org/775>).

5. Perhaps the closest that linguists have come to proposing a smallish set of innate building blocks that actually work for description is the Natural Semantic Metalanguage proposal by Goddard & Wierzbicka (2002). However, this proposal has not made it into the mainstream, for

language in its own terms, as opposed to preconceived concepts that were developed for some other language.⁶ The Boasian imperative has not lost its relevance in the 21st century (see also Muro (2015) on grammatical “anisomorphism” and the Boasian view; and Epps et al. 2017 for a wider perspective).

Now that we have seen the structural uniqueness of languages and the need for language-particular criteria in identifying the categories of languages, let us get back to the similarities among languages, and how they can help us describe particular languages, even in the absence of a smallish set of innate building blocks. I will argue that comparison is valuable for description in two ways: (i) to help descriptions achieve transparency (§ 3), and (ii) to help describers ask a wide range of potentially relevant questions (§ 4–5).

3. Comparison is valuable for transparency

First and most obviously, knowing about other languages helps authors describe a new language in an accessible way, and it helps readers understand the description.

Each language has its own classes, in the sense that they are defined in ways specific to each language (as we saw in the last section), and one might draw the conclusion from this that each class should get a different idiosyncratic name. This approach was sometimes adopted around the middle of the 20th century (cf. Dryer 1997: 117), but the grammars written according to this method are notoriously hard to understand. At that time, it was also thought that formulaic notations on the model of mathematics or logic might lead to greater scientific rigour of linguistic research results, but by the late 1970s, it had become clear to all language describers that their descriptions should not only be accurate and well-argued, but also comprehensible and transparent.⁷

reasons that I do not understand well. I find it much more practical than most other proposals of innate building blocks.

6. A salient way in which this point is demonstrated is that research papers in the Chomskyan tradition which rely on innate building blocks usually adopt the expository strategy of describing their phenomena twice: first in “neutral” terms, before adding a “theory-oriented analysis” that describes the same facts using the hypothesized innate elements. One may hope that the theory-oriented description will somehow turn out to be the “right” one, but the first (“neutral”) description is the only practical one.

7. Aikhenvald (2015: 6) emphasizes the value of transparency from the perspective of longevity: “A reference grammar should last beyond the life-span of its author. This is the main reason why it should not be cast in any of the time-line formalisms which come and go...”

For example, if there are two constructions with two-place verbs, one of which has no particular marking on the agent nominal and the verb, while the other construction has a marker on the verb and an oblique postposition on the agent nominal, then the latter should be described with the well-known label “passive” (if it is less frequently used than the former), because construction pairs of this type are found in many languages and are often called “active” and “passive”. And if there are two classes of nouns that behave differently in adnominal possessive constructions, and if the first class contains nouns like ‘hand’, ‘eye’, ‘nose’ and ‘hair’, while the second includes nouns like ‘house’, ‘garden’, ‘bicycle’, and ‘money’, then it is not helpful to call them “class I” and “class II”, because these opaque terms are difficult to remember. It is much better to call them “inalienable” and “alienable”, because similar classes exist in many languages, and this distinction is well-known to linguists.

It is worth noting here that this is a purely practical consideration, with no theoretical import. If our theoretical goal is merely to understand the behaviour of the speakers of one language, then idiosyncratic terms are fully adequate to that goal. And if our practical goal is merely to learn one language (e.g. if a Buddhist scholar wants to learn Pali in order to read the texts written in this language), then again there is no problem at all. People who just study a single language do not need transparent terminology from comparative linguistics.

The difficulty with idiosyncratic language-particular terminology is entirely due to the implicitly comparative nature of what many linguists do, especially linguists who describe small languages that have little prestige and no practical applications outside their speech community. Language documenters working in the tradition of Himmelmann (1998) and Gippert et al. (2006) take such an implicitly comparative approach. They are not merely cultural workers who want to preserve particular cultural manifestations for posterity, but as Himmelmann (2016: 474) puts it:

One of the purposes of language documentations is to serve as a primary data pool for typological inquiry.

For such comparative purposes, it would be counterproductive to use idiosyncratic names for the categories of each language. Language describers want to be understood by comparative linguists, and also by linguists working on other languages of the region (given that linguists often come together at regional conferences, e.g. about languages of South America, or languages of the Caucasus, or languages of Southeast Asia). And of course, language describers often engage in comparisons themselves.

So the solution that has generally been adopted, and that has proved entirely adequate, is to use familiar terminology from general linguistics when a language-

particular grammatical meaning, form class, or construction is very similar to meanings/classes/constructions in other languages that are generally known by a well-known term. For example, there is nothing wrong with using the term “feminine” (originally developed for Greek) also for a Russian gender class that is similar to the Greek class, and there is no problem with using the term “perfective” (originally developed for Russian) also for Mandarin *-le*, which shows clear similarities with the corresponding Russian category.

The use of identical labels for similar grammatical phenomena is particularly salient in interlinear glosses, where linguists typically use abbreviated grammatical labels such as SG (singular), PST (past tense), CLF (classifier), as illustrated in (2a–c) above. The authors of the Leipzig Glossing Rules⁸ have proposed to standardize these abbreviations, but crucially, the abbreviations are merely standard representations of labels, and there is no standardization of the meanings of the terms implied in them (see Bond 2010: 251).⁹

But while general terms are convenient for language-particular categories, there is a clear difference between concepts of general linguistics (see § 6) and language-particular classes, so when we want to be precise, we can use Comrie’s (1976) convention of capitalizing language-particular terms (cf. Croft 2001; Haspelmath 2010b: § 5). Above I used the terms “English Noun” and “English Future tense”, with capitalization because these classes are different from the general concepts “noun” and “future tense”, which are defined purely semantically.¹⁰ The context often makes it clear whether we are talking about general concepts or language-particular classes, so this convention need not be applied rigidly. But it is useful to be aware of it because it is easy to be confused in some contexts. Linguists often mix language-particular contexts and comparative contexts, and then having a way to distinguish the two meanings is useful. For example, a linguist may say that “nouns with plural meaning rarely take a plural suffix in Mandarin, but they almost always take a plural suffix in English.” Here it is clear that “noun” is used in the general sense, not in a language-particular sense, even though the term is applied to two different languages. But this simultaneous use of “noun” for two languages means that it must be defined here in a general way (see note 10), not in language-particular ways. These mixed uses of grammatical terms may lead to confusion, which is compounded by the problem that many linguists assume innate building blocks, often implicitly (as noted in the preceding section).

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

9. For a proposal to standardize the meanings of grammatical terms, see Haspelmath (2021a).

10. From a cross-linguistic perspective, a noun is indeed best defined as a root that denotes a thing (rather than an action or a property), as discussed in Haspelmath (2021b).

Now while suitable existing labels are often available, this is not always the case. In such cases, describers should feel free to create a new label. Linguists are often reluctant to coin new terms, but there is no good reason for this. Languages are very diverse, and it cannot be assumed that everything that we find in a language is sufficiently similar to something that was described before. Moreover, there is no clear way of telling whether the similarities are “sufficient”.

It should always be remembered that a label is not more than a practical device, and that language-particular categories are never exhaustively characterized by it. In his survey of the relation between description and typology, Bond (2010: 252) emphasizes this point:

anguishing over finding an appropriate label for a category is in many ways moot. Essentially, this is because assigning a label to a category does not describe it: the variables which are important for cross-linguistic comparison are actually much more fine-grained than category labels conventionally indicate.

And here is another consideration that may be worth pointing out: Since assigning a label is not the same as describing the phenomenon, no good purpose is served if a grammatical description cites typological work when introducing a language-particular phenomenon. For example, when Berghäll (2015: 297) introduces adverbs of Mauwake (a language of Papua New Guinea), she cites Schachter’s (1985) definition of adverbs (“words functioning as modifiers of constituents other than nouns”). But this does not tell us anything specific about Mauwake adverbs, and it is not the grammar author’s task to educate readers who did not know how the term “adverb” is generally used.

Unfortunately, the crucial (and widely recognized) distinction between general (comparative) concepts and unique language-particular categories has sometimes been conflated with the alleged view (not held by anyone, as far as I know) that language description should be completely unrelated to comparison. Lehmann (2018: § 1.1) writes:

Since at least some of the linguistic elements and categories of [a] language are unlike those of any other language, it has often been claimed that every language must be described in its own terms, implying that concepts used in the description of one language are useless in the description of another language. This is a non sequitur.

Lehmann is of course right that similarities between languages are “useful”, but they are useful primarily because we can use the same terms for very similar categories in different languages (for transparency), and not because the use of these terms could replace language-particular analysis. So in this passage, he misrep-

resents the Boasian imperative.¹¹ The Boasian imperative does not say that one needs different terminology for each language, or that describers should not be inspired by what is known about other languages (as we will see in § 4).

4. Comparison is valuable for asking questions

In addition to being valuable for transparent (and easily comparable) descriptions, awareness of the results of comparative linguistics is also valuable for heuristic purposes. As Himmelmann (2016: 475) notes, “typology is particularly useful for language description by making researchers aware of the diversity of possible linguistic structures and by providing examples of phenomena that may be of use in analyzing a given language”. For example, in quite a few languages, person names can take plural marking with an associative interpretation, as in (4).

- (4) Hungarian (Moravcsik 2003: 469)
Péter-ék
 Peter-ASS.PL
 ‘Peter and his family or friends or associates’

English, Spanish, German and Russian do not allow such uses of plural suffixes, so associative plurals have not been widely discussed in the literature, and without a course in comparative grammar (or without reading the three volumes edited by Shopen 2007), prospective language describers might not be aware of this possibility and overlook such a phenomenon in their language. For this reason, the more diverse phenomena from other languages we know, the better we are in a position to ask additional questions about a new language.

A very concrete way in which general comparative considerations have been found useful by many describers is by providing questionnaires and stimulus kits. Bond (2010: 250) writes:

Perhaps the most tangible tools that typologists create for field linguists are questionnaires and stimulus kits. Questionnaires are very useful for linguistic research, but stimulus kits are particularly important for language documentation because they remove some of the language bias that questionnaires introduce.

11. See also Himmelmann (2016: 475), who misrepresents the Boasian imperative when he describes it as a conception “of language description as an enterprise completely separate from typological... inquiry”. Language description and typology are intimately connected in practice, but conceptually distinct in theory (see also my discussion of Himmelmann’s criticism in this blogpost: <https://dlc.hypotheses.org/953>). I should also note here that the present paper was partly inspired by the interesting critical remarks in Himmelmann (2019).

A recent volume of the journal *Language Documentation and Conservation*, Lahaussois & Vuillermet (2019) (eds), brings together a range of papers about questionnaires and similar resources.¹²

But knowledge of worldwide diversity is important for language describers also because they might otherwise unwittingly carry over the categories of their own language to the language they describe. As Mosel (2012:74) puts it: “The more researchers know about the structural diversity of the world’s languages, the better they will understand the structure of the target language, and the less they will be unconsciously influenced by the contact language or their native language.”

A closely related point is made by Rice (2006), although she uses the term “linguistic theory” for what I call comparative linguistics here. She recounts how the discussions about question-word fronting constructions in English (following Chomsky’s 1977 influential work) influenced her work on the Athapaskan language Slave (Rice 1989). Inspired by the extensive discussions surrounding the restrictions on fronting in English (e.g. the impossibility of **What_i do you believe the story that Mary bought Ø_i?*), Rice was able to investigate aspects of Slave that she would not have considered otherwise. Similarly, Himmelmann (2016:476) rightly observes that “category controversies such as the continued debate regarding the applicability and usefulness of the concept “subject” in crosslinguistic comparison have been enormously fruitful and productive...”

But again, these considerations are methodological, not theoretical in the narrow sense. There is no theoretical reason why comparison with other languages should be relevant to describing a language. Knowing about other languages is often inspiring, but it is never decisive. We can learn from the questions that our colleagues have asked about other languages, but we can never adopt their solutions wholesale, because each language has its own categories that can only be justified in language-particular terms.¹³

12. See also the collection of resources hosted at <https://www.eva.mpg.de/lingua/tools-at-lingboard/tools.php>. It seems that such questionnaires were first developed by the Leningrad Typological School (e.g. Nedjalkov & Jaxontov (eds.) 1988), and by Dahl (1985).

13. It is only if one assumes a closed set of innate categories of universal grammar that one can hypothesize that two languages have the same categories. I leave this possibility (which I discussed briefly in § 2) largely aside in this paper (I discussed it in some detail elsewhere, e.g. in Haspelmath 2019; see also Bickel 2015: § 2 on the distinction between metalanguages and explanations in terms of innate categories).

5. Scaffolds versus frameworks

As we saw in § 2, the theoretical task of p-linguistics (i.e. the study of particular languages) is to find suitable structural elements (classes, categories, rules, rule arrangements) that allow the linguist to describe the language correctly and comprehensively. One might say that the core elements of such descriptions provide a framework for the language that can be filled with more and more material, as more detail is included. For example, in a description of Russian, one might start with the discussion of the classes Noun, Adjective, and Verb, and with the various inflection classes and inflectional categories, because these occur in every sentence and are likely to cause the greatest difficulties. They could thus be said to provide a kind of structural frame for the grammar edifice. The “framework” metaphor fits nicely with the metaphors of “structures” and “building blocks” in that the entire language is analogized to a large building.

However, the word “framework” is generally used differently by linguists: It is not applied to the most important structural elements of a particular language, but it is used for “general frameworks” consisting of universal building blocks that are assumed to be part of the innate blueprint for grammars. But as I noted in § 2, the existence of such innate building blocks is quite speculative, and they cannot be used for language description in practice. In Haspelmath (2010a), I discussed the notion of a “general framework” further and argued that languages should be described in their own terms. Continuing the building metaphor, we could say that each language has its own framework. Not only the low-level building blocks and the façade differ across languages, but even the more basic structural elements may differ. For example, Krifka (1995) argues that English and Chinese differ significantly in the structures and meanings of their nominals, and one can even imagine that some languages might have a distinction between syntax and morphology, while others have very complex, “polysynthetic” words and lack syntax, and still others (“isolating languages”) have nothing but syntax.

Thus, languages differ in their smaller building blocks and their structural frames, and theoretical linguistics cannot rely on their universality (at least so far). But we saw in the preceding two sections that knowledge of the world’s languages is very useful for language describers, both for transparency of description (§ 3) and for inspiration (§ 4). Comparison is thus useful in practical terms, and to continue our building metaphor, we can liken the use of typology for description to the use of scaffolds in construction work. Without a scaffold, construction workers would not be able to construct a house, but the scaffold itself does not determine its structural frame, its building blocks, or its façade. A scaffold has immense practical utility, but once the building is finished, it is removed and is no longer needed.

Thus, we can compare questionnaires of the type discussed in Lahaussais & Vuillermet (2019) to scaffolds: A researcher uses such questionnaires to ask a wide range of questions about a language, but the questionnaires play no role in the resulting grammatical description.¹⁴ This can be turned into a slogan: Grammatical description should be scaffold-based, but framework-free.

6. Expressing similarities between structurally different systems

In the preceding two sections, we saw in what ways comparative grammar research is valuable for language description even though the description of each language can only be justified by language-internal considerations.

In this section, I will briefly recapitulate how it is possible to compare languages rigorously, i.e. not merely impressionistically or subjectively, in the absence of a closed set of innate building blocks which all languages are made up of. The widespread (and useful, see § 3) use of the same terms for different categories of different languages carries the danger of a misconception: That these terms refer to structural building blocks that are shared among the languages. As we saw, they do not – they refer to structurally distinct categories which are similar enough to be given the same names.

So how can we compare languages even if they are made up of language-particular building blocks, unique to each language? The answer is that we do not compare them in terms of their building blocks, but by means of comparative concepts (Haspelmath 2010b, 2018). For example, we can compare the words *apples*, *píngguō*, *jabloka*, and *tapuħim* in (1)–(2) because they all denote things (and are thus nouns in a comparative sense, note 10) of a specific kind, we can compare the words *three*, *sān-ge*, *tri*, and *shlosha* because they are all numerals

14. For this reason, the grammars based on the questionnaire by Comrie & Smith (1977) (e.g. Hewitt 1979 on Abkhaz, Kornfilt 1997 on Turkish) are more useful for typologists who want to find information than for readers whose main interest is the language. To understand the details of a grammar, readers first need to understand the basic language-particular “structural frame”, i.e. the key categories of its grammar around which the description must be organized. Lehmann (1989: § 3.2) saw this tension long ago when he noted that the Comrie/Smith questionnaire must be based on functional notions, but that in a comprehensive description, “a purely function-form oriented approach inevitably leads to the consequence that the different uses of a polysemous or multifunctional form are dispersed over the various chapters. For instance, in a description of English, the verb *to be* would have to be treated in ch. 1.2 (clause structure, namely copular sentences), 1.10 (possession), 1.11 (emphasis, namely clefting), 2.1.3 (verb morphology).”

of a specific cardinality,¹⁵ and so on. Comparison is also possible when the mapping is less straightforward, e.g. by means of translation questionnaires (e.g. Dahl 1985) or parallel texts (e.g. Cysouw & Wälchli 2007). In a parallel-text study based on Bible texts, the minimal comparative concept is the Bible verse. So comparative concepts may be very different from grammatical classes or categories, and for this reason, I refer to them as “concepts” (not as “comparative categories”) (Haspelmath 2018: § 2).

Morphosyntactic comparative concepts are thus defined by semantic substance, and phonological comparative concepts are defined by phonetic substance (word order may also be seen as based on pronunciation, as it is easy to observe). Quite generally, comparative concepts are substance-based, while language-particular categories are based on structural distributions.

The methodological rigour of the comparison comes about by specifying the mapping between the comparative concepts and the language-particular category in a precise way that is exactly the same for all languages. It is not admissible to use different criteria for different languages, which would lead to the kind of cherry-picking or diagnostic-fishing that Croft (2009, 2010) called “methodological opportunism”. This also means that the kind of “in-depth analyses” that are often emphasized in generative grammar and that are based on diverse language-specific considerations cannot play a role in rigorous comparison as discussed here (Haspelmath 2019). Some linguists think that large-scale comparison as in the world maps of *WALS* (Haspelmath et al. 2005; Dryer & Haspelmath 2013) is “superficial” and not very telling, but in fact it is the only way of comparing languages with their structurally unique features in a systematic and objective way. The approach is no different from other cross-cultural comparisons as practiced in anthropology and cross-cultural psychology (e.g. Ember & Ember 1998; Slingerland et al. 2020).

7. Conclusion

Grasping the relationship between language description and cross-linguistic comparison is easy and difficult at the same time.

It is easy because it has long been widely recognized that language-particular morphosyntactic categories are identified by distributional criteria, not semantic

15. In the case of *san-ge*, it is questionable whether we are dealing with a single word or two words. So strictly speaking, we should compare roots (*san* in Mandarin, *tr’-* in Russian, *shlosh-* in Hebrew). “Word” does not work as a comparative concept for rigorous comparison (Haspelmath 2011).

criteria (§ 2), and semantic and phonological categories are defined by system-internal contrasts, not by their substance. They are defined in language-particular way, and what other languages do is irrelevant. By contrast, general concepts that can be applied cross-linguistically must be of a different nature, because languages are often non-isomorphic. Languages are obviously similar phonetically and semantically, but because of the structural uniqueness of each language, the cross-linguistic similarities must be captured by concepts defined by phonetic and semantic substance, not by distributions.

But it is at the same time difficult to grasp the relationship between description and comparison, because the habit of using the same terms for different languages is so entrenched, and it is easy to confuse statements about a particular language with general statements. For example, the question “Does your language have numeral classifiers?” is necessarily general-comparative, while a statement like “Mandarin requires a Numeral Classifier after cardinal numerals like *sān*” are language-particular and thus of a different nature (requiring the distributionally defined category Numeral Classifier, not a comparative concept “numeral classifier” that is defined by semantic or phonetic substance).

The difficulty is compounded by the well-known claims of linguists working in the Chomskyan (or Bakerian) tradition that there is a smallish set of universal building blocks of the innate grammar blueprint of which all languages are made up. These claims are coherent, but so far their practical value is quite unclear.

In this paper, I have tried to show that the uncontroversially intimate relationship between language-particular categories and comparative information about the world’s languages is of a methodological kind: For transparency of description (§ 3) and for inspiration (§ 45), language describers are well-advised to acquire extensive typological knowledge. But to justify each of the categories that they set up, they must rely exclusively on language-internal criteria (§ 2). As theoretical entities, language-particular categories are independent of any comparison. When the ancient Greek and ancient Indian grammarians described their languages, they did not have any comparative knowledge, but they were still able to describe Greek and Sanskrit, respectively, in an adequate way. Comparison is very valuable for description, but not theoretically necessary, because language systems are structurally unique.

Acknowledgements

I thank a reviewer for *Asian Languages and Linguistics* for useful comments, and I also had many useful comments on Academia.edu – many thanks, in particular, to Peter Arkadiev, Willem de Reuse, Eitan Grossman, Ryan Ka Yau Lai, Alessio Muro, and Adam Tallman. The support of the European Research Council (ERC Advanced Grant 670985, Grammatical Universals) is gratefully acknowledged.

References

- Aikhenvald, Alexandra Y. (2015). *The art of grammar: A practical guide*. Oxford University Press.
- Baker, Mark C. (2001). *The atoms of language*. Basic Books.
- Berghäll, Liisa. (2015). *A grammar of Mauwake* (Studies in Diversity Linguistics). Language Science Press. <http://langsci-press.org/catalog/book/67>.
https://doi.org/10.26530/OAPEN_603339
- Bickel, Balthasar. (2015). Distributional Typology. In Heiko Narrog & Bernd Heine (eds.), *The Oxford handbook of linguistic analysis*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199677078.013.0046>
- Boas, Franz. (1896). The limitations of the comparative method of anthropology. *Science*. American Association for the Advancement of Science 4(103). 901–908.
- Boas, Franz. (1911). Introduction. In Franz Boas (ed.), *Handbook of American Indian Languages*, 1–83. Bureau of American Ethnology.
- Bochnak, M. Ryan & Matthewson, Lisa (eds.). (2015). *Methodologies in semantic fieldwork*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190212339.001.0001>
- Bond, Oliver. (2010). Language documentation and typology. *Language Documentation and Description* 7. 238–261.
- Candea, Matei. (2018). *Comparison in anthropology: The impossible method* (New Departures in Anthropology). Cambridge University Press. <https://doi.org/10.1017/9781108667609>
- Chomsky, Noam A. (1977). On wh-movement. In Akmajian, Adrian & Culicover, Peter W. & Wasow, Thomas (eds.), *Formal syntax*, 71–132. Academic Press. (Accessed March 13, 2019.)
- Comrie, Bernard & Norval Smith. (1977). Lingua descriptive studies: Questionnaire. *Lingua* 42. 1–72. [https://doi.org/10.1016/0024-3841\(77\)90063-8](https://doi.org/10.1016/0024-3841(77)90063-8)
- Comrie, Bernard. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge University Press.
- Croft, William. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press.
- Croft, William. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198299554.001.0001>
- Croft, William. (2009). Methods for finding universals in syntax. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 145–164. Springer. https://doi.org/10.1007/978-1-4020-8825-4_8
- Croft, William. (2010). Ten unwarranted assumptions in syntactic argumentation. In Kasper Boye & Elisabeth Engberg-Pedersen (eds.), *Language usage and language structure*, 313–350. De Gruyter Mouton.
- Cysouw, Michael & Bernhard Wälchli. (2007). Parallel texts: Using translational equivalents in linguistic typology. *STUF-Sprachtypologie und Universalienforschung* 60(2). 95–99.
- Dahl, Östen. (1985). *Tense and aspect systems*. Blackwell.
- Davis, Henry & Gillon, Carrie & Matthewson, Lisa. (2014). How to investigate linguistic diversity: Lessons from the Pacific Northwest. *Language* 90(4). e180–e226.
<https://doi.org/10.1353/lan.2014.0076>

- Dryer, Matthew S. & Haspelmath, Martin (eds.). (2013). *WALS Online*. Max Planck Institute for Evolutionary Anthropology. (<https://wals.info/>)
- 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. Givón*, 115–143. John Benjamins. <https://doi.org/10.1075/z.82.09dry>
- Dryer, Matthew S. (2006). Descriptive theories, explanatory theories, and basic linguistic theory. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 207–234. Mouton de Gruyter.
- Ember, Carol R. & Melvin Ember. (1998). Cross-cultural research. In H. Russell Bernard (ed.), *Handbook of methods in cultural anthropology*, 647–687. AltaMira Press.
- Epps, Patience L. & Webster, Anthony K. & Woodbury, Anthony C. (2017). A holistic humanities of speaking: Franz Boas and the continuing centrality of texts. *International Journal of American Linguistics*. The University of Chicago Press 83(1). 41–78. <https://doi.org/10.1086/689547>
- Epps, Patience. (2011). Linguistic typology and language documentation. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 634–649. Oxford University Press.
- Gil, David. (2001). Escaping Eurocentrism: Fieldwork as a process of unlearning. In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*, 102–132. Cambridge University Press. <https://doi.org/10.1017/CBO9780511810206.006>
- Gippert, Jost, Nikolaus Himmelmann & Ulrike Mosel (eds.). (2006). *Essentials of language documentation*. Mouton de Gruyter. <https://doi.org/10.1515/9783110197730>
- Goddard, Cliff & Anna Wierzbicka (eds.). (2002). *Meaning and universal grammar: Theory and empirical findings*. 2 volumes. John Benjamins. <https://doi.org/10.1075/slcs.61>
- Halle, Morris. (1959). *The sound pattern of Russian*. De Gruyter Mouton.
- Hanks, William F. & Severi, Carlo. (2014). Translating worlds: The epistemological space of translation. *HAU: Journal of Ethnographic Theory* 4(2). 1–16. <https://doi.org/10.14318/hau4.2.001>
- Haspelmath, Martin, Matthew Dryer, David Gil & Bernard Comrie (eds.). (2005). *The world atlas of language structures*. Oxford University Press.
- Haspelmath, Martin. (2007). Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11(1). 119–132. <https://doi.org/10.1515/LINGTY.2007.011>
- Haspelmath, Martin. (2010a). Framework-free grammatical theory. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 341–365. Oxford University Press.
- Haspelmath, Martin. (2010b). Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. <https://doi.org/10.1353/lan.2010.0021>
- Haspelmath, Martin. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. <https://doi.org/10.1515/flin.2011.002>
- Haspelmath, Martin. (2018). 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. De Gruyter Mouton. <https://zenodo.org/record/3519206>. <https://doi.org/10.1515/9783110607963-004>
- Haspelmath, Martin. (2019). Ergativity and depth of analysis. *Rhema* (2019)(4). 108–130. <https://doi.org/10.31862/2500-2953-2019-4-108-130>

- Haspelmath, Martin. (2021a). Towards standardization of morphosyntactic terminology for general linguistics. In Alfieri, Luca, Giorgio Arcodia & Paolo Ramat (eds.), *Linguistic categories, language description and linguistic typology*. John Benjamins.
- Haspelmath, Martin. (2021b). Word class universals and language-particular analysis. To appear.
- Hewitt, Brian George. (1979). *Abkhaz*. North Holland.
- Himmelman, Nikolaus P. (1998). Documentary and descriptive linguistics. *Linguistics* 36. 161–196. <https://doi.org/10.1515/ling.1998.36.1.161>
- Himmelman, Nikolaus P. (2016). What about typology is useful for language documentation? *Linguistic Typology* 20(3). 473–478. <https://doi.org/10.1515/lingty-2016-0020>
- Himmelman, Nikolaus. (2019). Against trivializing language description and comparison. Paper presented at the Biennial Conference of the Association for Linguistic Typology, Pavia.
- Kornfilt, Jaklin. (1997). *Turkish*. Routledge.
- Krifka, Manfred. (1995). Common nouns: A contrastive analysis of Chinese and English. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, vol. 398, 398–411. The University of Chicago Press.
- Lahaussais, Aimée & Marine Vuillermet. (2019). *Methodological tools for linguistic description and typology* (Language Documentation & Conservation Special Publication 16). University of Hawai'i Press.
- LaPolla, Randy J. & Dory Poa. (2006). On describing word order. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 269–295. Mouton de Gruyter.
- LaPolla, Randy J. (2020). Forward to the past: Modernizing linguistic typology by returning to its roots. *Asian Languages and Linguistics*. John Benjamins 1(1). 147–167. <https://doi.org/10.1075/alal.00005.lap>
- Lass, Roger. (1984). Vowel system universals and typology: Prologue to theory. *Phonology* 1. 75–111. <https://doi.org/10.1017/S0952675700000300>
- Lehmann, Christian. (1989). Language description and general comparative grammar. In Gottfried Graustein & Gerhard Leitner (eds.), *Reference grammars and modern linguistic theory* (Linguistische Arbeiten 226), 133–162. Niemeyer. http://www.christianlehmann.eu/publ/lg_descr.pdf. <https://doi.org/10.1515/9783111354590.133>
- Lehmann, Christian. (2018). Linguistic concepts and categories in language description and comparison. In Marina Chini & Pierluigi Cuzzolin (eds.), *Typology, acquisition, grammaticalization studies*, 27–50. Franco Angeli. https://www.christianlehmann.eu/publ/lehmann_ling_concepts_categories.pdf
- Lyons, John. (1968). *Introduction to theoretical linguistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139165570>
- Mithun, Marianne. (2001). Who shapes the record: The speaker and the linguist. In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*, 34–54. Cambridge University Press. <https://doi.org/10.1017/CBO9780511810206.003>
- Moravcsik, Edith A. (2003). A semantic analysis of associative plurals. *Studies in Language* 27(3). 469–503. <https://doi.org/10.1075/sl.27.3.o2mor>
- Mosel, Ulrike. (2012). Morphosyntactic analysis in the field: A guide to the guides. In Nicholas Thieberger (ed.), *The Oxford handbook of Linguistic fieldwork*, 72–89. Oxford University Press.

- Muro, Alessio. (2015). Lost in translation between typologically different grammars In Miola, Emanuele & Paolo Ramat (eds.), *Language across languages: New perspectives on translations*, 35–38. Cambridge Scholars Publishing.
- Nedjalkov, Vladimir P. & Sergei J. Jaxontov (eds.). (1988). *The typology of resultative constructions*. John Benjamins. <https://doi.org/10.1075/tsl.12>
- Plungian, Vladimir A. (2011). *Vvedenie v grammatičeskiju semantiku: Grammatičeskie značenijsa i grammatičeskie sistemy jazykov mira*. RGGU.
- Rice, Keren. (1989). *A grammar of Slave* (Mouton Grammar Library 5). Mouton de Gruyter. <https://doi.org/10.1515/9783110861822>
- Rice, Keren. (2006). Let the language tell its story? The role of linguistic theory in writing grammars. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 235–268. Mouton de Gruyter.
- Schachter, Paul. (1985). Parts-of-speech systems. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 1, 3–61. Cambridge University Press.
- Shopen, Timothy (ed.). (1985). *Language typology and syntactic description*. 3 volumes. Cambridge University Press.
- Shopen, Timothy (ed.). (2007). *Language typology and syntactic description*. 3 volumes. Cambridge University Press.
- Simpson, Adrian P. (1999). Fundamental problems in comparative phonetics and phonology: Does UPSID help to solve them? In *Proceedings of the 14th international congress of phonetic sciences*, vol. 1, 349–352.
- Slingerland, Edward, Quentin D. Atkinson, Carol R. Ember, Oliver Sheehan, Michael Muthukrishna, Joseph Bulbulia & Russell D. Gray. (2020). Coding culture: Challenges and recommendations for comparative cultural databases. *Evolutionary Human Sciences* 2. e29. <https://doi.org/10.1017/ehs.2020.30>
- Tallerman, Maggie. (1998). *Understanding syntax*. Arnold.
- Trubetzkoy, Nikolaus. (1939). *Grundzüge der Phonologie* (Travaux du Cercle Linguistique de Prague 7). Cercle Linguistique de Prague.

Address for correspondence

Martin Haspelmath
 Max Planck Institute for Evolutionary Anthropology
 Deutscher Platz 6
 D-04103 Leipzig
 Germany
haspelmath@shh.mpg.de