

## CLIPP

**Christiani Lehmanni inedita, publicanda, publicata**

### titulus

Linguistic concepts and categories  
in language description and comparison

### huius textus situs retis mundialis

[http://www.christianlehmann.eu/publ/lehmann\\_ling\\_concepts\\_categories.pdf](http://www.christianlehmann.eu/publ/lehmann_ling_concepts_categories.pdf)

### dies manuscripti postremum modificati

31.03.2017

### occasio orationis habitae

Typology, acquisition, grammaticalization studies.  
Convegno internazionale per festeggiare il trentennio di vita  
del Dottorato in Linguistica di Pavia. Pavia, 14 – 15 october,  
2016

### volumen publicationem continens

Chini, Marina & Cuzzolin, Pierluigi (eds.), *Typology, acquisition, grammaticalization studies*. Milano: Franco Angeli (Materiali Linguistici)

### annus publicationis

2017

### paginae

ignotae

# Linguistic concepts and categories in language description and comparison

Christian Lehmann  
University of Erfurt

## Abstract

The question of whether two phenomena should be subsumed under the same concept is a frequent conceptual problem in linguistics and particularly relevant in typology. Questions such as whether a certain Turkish case can be called a dative if the Latin dative is are typical of the general methodological problem of comparative concepts. Recent approaches have tended to play down the relationship of comparative to descriptive categories, as if these belonged to unrelated fields. Quite on the contrary, a fruitful relationship between comparative and descriptive linguistics presupposes that they use the same conceptual framework. This requires that the relation between a language-specific category and a corresponding comparative category be made explicit. This task has often proved elusive. Its solution presupposes the following steps:

- many linguistic concepts are prototypical rather than categorical concepts
- descriptive and comparative concepts are in a hierarchy of increasing generality and abstractness
- grammatical categories of individual languages are “hybrid” semiotic categories, while comparative categories may be purely formal, purely semantic or hybrid.

These principles are illustrated by concepts in the vicinity of numeral classification.

## 1 Introduction

Clarifying the formation of concepts in a scientific discipline is a task for epistemology and methodology, both at the general, philosophical and at the specific level which considers the situation of the discipline in question. This paper is meant as a contribution to linguistic methodology.<sup>1</sup>

### 1.1 Descriptive linguistics

In the humanities, the object area displays a certain kind of essential variation which renders it necessary to describe each variety in itself.<sup>2</sup> Since language is a problem-solving activity, and every language system is a problem-solving system, it is described so as to render manifest in which way this language solves the tasks of cognition and communication. No useful purpose is served if the description makes the language appear like another language. Quite on the contrary, the user of a description wants to know about the particular ways in which this lan-

---

<sup>1</sup> This paper has profited much from extensive discussion with Alex Burri and Edith Moravcsik. I also thank Paolo Ramat and Sonia Cristofaro for discussion at the Pavia conference 2016.

<sup>2</sup> True enough, the object area of a natural science may display variation, too. Biological species may be conceived as varieties of a genus, and supernovae and white dwarfs may be conceived as varieties of stars. There is, to my knowledge, no epistemology which would characterize kinds of variation at a general level.

guage functions. Since at least some of the linguistic elements and categories of this 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.<sup>3</sup> This is a *non sequitur*. If these are apples, they cannot be described as if they were pears; but this does not mean that it is inappropriate to subsume both under the more general concept of fruit. Likewise, the Latin personal pronoun is completely different from the Ancient Greek personal pronoun; but this has never impeded anybody, starting from the Roman grammarians, to cover the Latin variety by the same concept that the Greek grammarians had used for their variety. Nor was this scientifically inappropriate. Interlingual concepts like ‘personal pronoun’ are sufficiently abstract to comprise this kind of variation.

Quite to the contrary, the categorical imperative of language description reads:

“Describe your language in such a way that the maxim of your description could serve, at the same time, as a principle for the description of any language.”

This entails: Bring out those properties in which this language is like other languages, as well as those properties in which it differs from other languages. This requires using the same concepts as are used for the description of any other language, instead of thinking up new categories (s. §2.3 on the meaning of *category* in linguistics).

## 1.2 Comparative linguistics

Comparative linguistics compares languages just as comparative religion compares religions and comparative jurisprudence compares right systems. In the academic world, it is often considered a subdiscipline of linguistics. However, it must be clear that comparative linguistics is not a subdiscipline constituted by a certain section of the discipline’s object area, as for instance Finno-Ugric linguistics is. Nor is it constituted by a certain (kind of) theory, as for instance Montague linguistics is. Instead, it is constituted by a method. Comparison of phenomena belonging to the same level of a certain object area is a basic method in many scientific disciplines. Being a method, it has a subordinate status in the system of a science. It is not the case that comparing two elements of a certain scientific domain has a sense in itself. Instead, in the humanities, comparing two or more varieties of a given level in some object domain generally serves one of two generic purposes:

- (1) It provides insights into the range and the principles of variation obtaining in the object domain. Given an object such that essential variation in the sense indicated in §1.1 is part of its nature, then every variety represents the object, but since the object is abstract, no variety represents it fully.<sup>4</sup> Consequently, science learns about the nature of this object by comparison. In other words, comparison is a necessary method to achieve a theory of the object. In the case of linguistics, comparison of languages is the choice method for the elaboration of an empirical theory of language.
- (2) It provides insights into the functioning of one or each of the objects compared. A descriptive statement acquires its relevance on the background of the extent to which things could be otherwise. In other words, comparison helps improve the description of

---

<sup>3</sup> This is the tenor of several contributions to the discussion held on the list LingTyp from 19 to 26 January, 2016. These are currently being published in the discussion section of *Linguistic Typology* 20(2), 2016. Quotations appearing below are from the list discussion.

<sup>4</sup> The idea is rendered succinctly in the dictum (Delacroix 1924:128f) “Une langue est une variation historique sur le grand thème humain du langage.”

each of the objects compared; it enables a description in general, rather than idiosyncratic, terms.

Needless to say, in both of the respects mentioned, benefit is mutual. This is to say, the comparative method benefits from and even presupposes both a sound linguistic theory and an insightful linguistic description.<sup>5</sup>

A step in the achievement of goal #1 is the setting-up of linguistic types. Linguistic types are essentially hyperonyms of language-specific categories. For instance, both the Latin and the Ancient Greek personal pronoun instantiate the type of the personal pronoun. And since most linguistic concepts are prototypical concepts, so are types, too. Thus, the personal pronoun of one language – say English – may represent the focal instance of the prototype, while the personal pronoun of another language – say Japanese – is only marginally comprised by this concept.

## 2 Concepts

A concept is a general meaning which is used as a whole and may constitute an operand in logical (or, if taken as a mental object, in mental) operations. A concept cannot be understood in isolation, but only in its relations to other concepts in what is mostly called a conceptual network and which, depending on the uniformity and regularity of the relations involved, may take on more systematic forms, including conceptual hierarchies. These will be taken up in §2.1.

### 2.1 Relations among concepts

Empirical notions are developed and acquired on the basis of ostension. However, a concept is a scientific concept only if it is embedded in a conceptual network alias scientific theory. A scientific concept gets delimited by its position in the theory, which is the set of its relations to other concepts of the theory. Therefore, a concept is not defined in isolation. Its definition is based on its relations to neighboring concepts. Viewed in a linguistic perspective, these relations are semantic relations.<sup>6</sup> The conceptual relation which is most important in science is the relation ‘an A is a B’, called hyponymy in linguistics. Since the relation is transitive, it provides the building block for a conceptual hierarchy, viz. a taxonomy. In propositional calculus, the same relation may be expressed by the entailment: ‘necessarily, for all x, if x is A, then x is B’. This may then be extended to propositions containing relational concepts as predicates: ‘necessarily, for all x and y, if A(x, y), then B(x, y)’. The implicata of these implications are often called features of the concept A. Similarity and differences and, more generally, all kinds of relations between two concepts may then be based on their feature composition.

At the bottom of such a taxonomy, there are individuals which can bear a proper name. Thus, ‘Socrates is a man’ has a similar logical form as ‘English is a language’ and ‘the English perfect is a perfect’. We will have occasion to come back to the latter example.

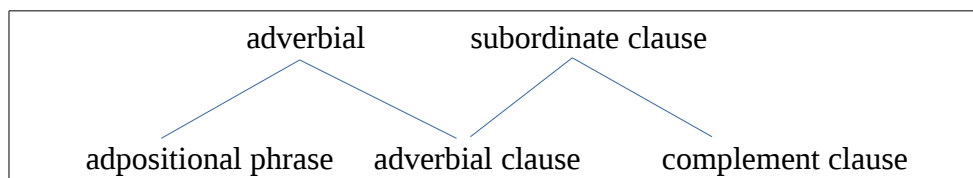
---

<sup>5</sup> Wilhelm von Humboldt (1836 [1963]:417) joins it all together in his dictum: “Um daher verschiedene Sprachen in bezug auf ihren charakteristischen Bau fruchtbar miteinander zu vergleichen, muß man der Form einer jeden derselben sorgfältig nachforschen, um sich auf diese Weise zu vergewissern, auf welche Art jede die hauptsächlichen Fragen löst, welche aller Spracherzeugung als Aufgaben vorliegen.”

<sup>6</sup> In other disciplines, such conceptual systems are sometimes called ontologies.

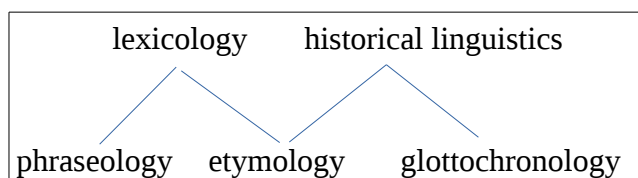
Hyponymy is a many-to-many relation.<sup>7</sup> Both an adpositional phrase and an adverbial clause are adverbials; both an adverbial clause and a complement clause are subordinate clauses. This multiple hyponymy is visualized in Diagram 1.

Diagram 1 *Multiple hyponymy with linguistic concepts*



The other semantic relation which is central to a conceptual network and possibly even more elementary than hyponymy is the part-of relation. This is, properly speaking, a relation between entities and only derivatively a relation between the concepts applying to these. Since it is transitive, too, it founds the second most important conceptual hierarchy, the meronymy.<sup>8</sup> The part-of relation is a many-to-many relation, too, as exemplified in Diagram 2.

Diagram 2 *Multiple part-of relations among linguistic concepts*



The most important meronymy in linguistics is probably the system of grammatical constructions; and it, too, is based on many-to-many relations.

Other conceptual relations which are important for linguistic methodology but which it must suffice to merely mention here include the relations 'x is a property/aspect/attribute of y', 'x manifests y' and 'x is a value of parameter y' (Lehmann 1996).

## 2.2 Kinds of concepts

Concepts may be classified by many, partly independent criteria. The most important ones for methodological purposes are treated in the following subsections.

### 2.2.1 Basic and defined concepts

An empirical theory of some object domain consists of a set of interconnected propositions. A designated subset of these are axioms; all the others are directly or indirectly based on the axioms.<sup>9</sup> The propositions make use of the concepts structuring the object domain. Again, a designated subset of these are basic concepts; all the others are defined concepts.

<sup>7</sup> This is so at least in language and linguistics. It does not apply to all taxonomies. For instance, the 'is a' relation in a botanical taxonomy is a many-to-one relation.

<sup>8</sup> Sometimes a distinction is made between a meronymy as an aspect of an ontology, and a meronymy as a lexical domain of some language.

<sup>9</sup> S. Lieb 2017, parts C and D for a formal elaboration of such a conception with respect to linguistic theories.

In linguistics, the set of **basic concepts** is heterogeneous because linguistic signs combine different epistemic domains. It may be subdivided as follows:

- (1) A certain portion of the basic concepts relate to linguistic substance. This is, on the one hand, the set of auditory impressions of speech sounds and of articulatory gestures producing these, and on the other hand, the set of entities, relations and operations which constitute the subject matter of our thinking and imagination. On the phonetic side, those are concepts like sonority, coarticulation and tone. On the semantic side, the set includes basic cognitive and communicative components and functions of grammar like control and affectedness, causation, time and place, speech act and illocutionary force. Some of these basic concepts may be defined in neighboring disciplines like phonetics and psychology.
- (2) Another subset of the basic concepts are concepts of logic, set theory and topology, which are needed in most scientific theories. These are concepts like class and element, concept, category, relation and operation, taxonomic and meronomic relations, implication, similarity and contiguity and many more. Some of these may be defined in the disciplines mentioned.
- (3) The last subset of basic concepts are semiotic concepts proper. These are concepts like expression, signification, reference, predication. They may or may not be defined in semiotics.

Concepts of the three kinds mentioned are not defined in a linguistic theory and are therefore basic from its point of view. Naturally, any scientific theory will be concerned to minimize the set of basic concepts. Moreover, for some of these, it will rely on neighboring theories for their definition, so if the theories are systematically related, these concepts may be considered defined, too. The remaining subset of basic concepts is taken as self-evident. Examples are the concepts of space, time and cause. These may be considered **atomic** or *a priori*.<sup>10</sup>

It is often not sufficiently realized that, with the exception of atomic concepts, all scientific concepts are defined concepts; in other words, they require a definition. If this is not apparent in some cases, it is only because scientists have failed to define them, to the detriment of their theories. This includes, among many others, all concepts of grammatical categories. “Pre-established categories don’t exist”, as Haspelmath 2007 insists.

It is possible that there are, in the object domain of linguistics, “natural kinds” for which the appropriate concepts are self-evident and impose themselves on the theory. As long as this is not scientifically validated, linguistic concepts are constructs. As such, they are man-made and consequently more or less appropriate. Haspelmath 2007, §2.3 mocks at “controversial category assignments” like “Is English *that* in relative clauses a pronoun or a complementizer?” because they appear to presuppose “pre-established categories”; and once these do not exist, such questions are “pointless”. However, on the one hand, the same kind of controversy arises over just any kind of non-atomic concept used in a science, including, of course, the “comparative concepts” endorsed by him. On the other, the fact that interlingual categories are not pre-established does not entail that they do not exist. Their concepts have exactly the same methodological and theoretical status as any kind of defined concept. The touchstone of the appropriateness of a scientific concept is whether its definition finds its place in a theory and is operationalizable. Some particularists seem to endorse *ad hoc* definitions.<sup>11</sup>

<sup>10</sup> Some of these may be defined in theories of mathematics or physics. However, these theories are not systematically related to linguistic theories.

<sup>11</sup> “comparative concepts are linguist-specific (in the sense that every linguist is free to define her or his own concepts),” (Haspelmath 2010[C]:674). There seems to be misunderstanding here of what

## 2.2.2 Categorical vs. prototypical concepts

Some linguistic concepts are prototypical concepts. This applies to concepts whether they are used in the description of a language or in typological generalizations.<sup>12</sup> The focal instance of a prototypical linguistic concept is the one which corresponds exactly and directly to the function of the linguistic unit in question, which would be able to fulfill functions of neighboring categories only by some transcategorization operation. In this sense, for instance, an adjective which directly functions as a modifier, but is unable to serve as a referential expression or as a predicate without some transcategorization operation, is a prototypical adjective. Given this, the Latin adjective is not a prototypical adjective, since it also serves as a referring expression without any adaptation. Neither is the German adjective a focal instance of the concept, since it serves as an adverb without any adaptation.<sup>13</sup> The English adjective comes close to the prototype, since it requires formatives – *one* and *-ly*, resp. – for the transcategorization into a noun and an adverb.<sup>14</sup>

## 2.3 Linguistic categories

It might be worth noting at the outset that a linguistic category is not a category in the Aristotelian sense. The latter is a category of what can at all be predicated on some subject in a proposition. Linguistic categories are something much more specific. Moreover, their concepts are not “pre-established” (Haspelmath 2010[C]), but defined concepts (and differ in this from Platonic, Aristotelian and Kantian categories). These definitions are the product of the analysis of the linguist (Ramat & Jezek 2009:393).

Given the heterogeneity of the object domain of linguistics, there are different kinds of linguistic categories. Only semantic categories like ‘human’, ‘instrument’ are concepts. Grammatical categories are not concepts, but classes of linguistic signs. Since a linguistic sign combines meaning with form, it is not pure meaning and therefore not a concept. Likewise, a class of signs is the common denominator of its members. As such, it combines semantic features with structural features just like a single sign does, only at a more abstract level; and therefore it cannot be a concept. This does not, of course, prevent us from speaking of the concept of a grammatical category; an entity is distinct from the concept applying to it (Frege 1892).

The concepts of grammatical categories form a conceptual network and, in particular, taxonomies and meronomies.

Diagram 3 Section of a grammatical category taxonomy

dual (number) is a number is a morphological category is a grammatical category
---

freedom of science means.

<sup>12</sup> Haspelmath (2010[i]:697) may be right in postulating that “in the end, the big picture has to be dissolved into fine-grained discrete comparative concepts.” This, however, is not the theoretical and methodological situation of linguistics in the current centuries.

<sup>13</sup> It is true that the adjective, but not the adverb, inflects. However, if this is relevant, it would rather provide an argument for basing the adjective on the adverb.

<sup>14</sup> Bhat (1994) arrives at the same conclusion, though on partly different grounds.

At the bottom of such a taxonomy, there are the grammatical categories of an individual language. Thus, the example in Diagram 3 might be extended to the left by entities like the Ancient Greek dual number. It has been repeatedly claimed that these are individuals and might therefore bear proper names (Haspelmath 2010[C], §6). This, however, is less than clear since grammatical categories of an individual language are in a taxonomy of their own, too. Thus, the Ancient Greek dual is an Ancient Greek number, and so forth.

Thus, the relation between a language-specific category like the Ancient Greek dual and an interlingual category like the dual number is one of hyponymy.<sup>15</sup> As usual, the hyperonym is more general than the hyponym. And as explained in §2.1, this relation is perfectly compatible with the possibility that a certain category of another language be a dual, too,<sup>16</sup> and with the possibility that the Ancient Greek dual be an instance of yet another interlingual category.<sup>17</sup>

## 2.4 Interlingual grammatical categories

In general comparative linguistics, we are currently exposed to a new wave of particularism.<sup>18</sup> Apparently the history of our discipline is doomed to follow the motion of a pendulum: after North American structuralism ("languages could differ from each other without limit and in

<sup>15</sup> Likewise Moravcsik 2016, §2C: "comparative concepts are taxonomically superordinate to descriptive categories."

<sup>16</sup> If the Ancient Greek dual and the Lithuanian dual are both said to be duals, then the relationship of these three concepts to each other must be made explicit. Lieb 2017, §5.5 provides a formal solution to this problem.

<sup>17</sup> Haspelmath (2010, §9) finds that "comparative concepts" and "linguistic categories" are of different kinds and therefore not in a relation of hyponymy. However, his only objection against hyponymy is the fact that a category of some language always has properties not comprised by the comparative concept. This, however, is generally the case in hyponymy, thus no reason to deny the possibility of hyponymy between comparative and descriptive concepts. Nor is it true that the language-specific category is necessarily "used in ways that are quite incompatible with the definition of the comparative concept." (p. 680) This would happen only if the comparative concept was defined in an imprudent way. Lastly, since hyponymy is a many-to-many relationship, he (l.c.) thinks that "With such rampant many-to-many relationships, a taxonomic conceptualization, while logically possible, only obscures matters." This, again, is not so. The language-particular categories used in the description of a single language are in exactly the same kind of relation. In the description of any single language, examples like Diagram 1 in §2.1 abound. This is quite generally the form of a taxonomy of linguistic concepts, be they descriptive or comparative or whatever. Making such multiple hyponymy explicit even contributes to the clarification of a concept. Cf. also Edith Moravcsik's contribution to the *lingtyp* discussion (21.01.16) "Categorization, by its very concept, does not require that the two things that are lumped together share all of their properties. We use categories so that once one property is identified for something, another one is predictable and thus its occurrence is in a sense explained. Thus a mutual or unidirectional implicational relation between two properties is sufficient to justify a category and it does not matter if in many other ways, token of the proposed category are different."

<sup>18</sup> It has become fashionable to invoke the position formulated in the oft-quoted verdict by Sapir (1921:125): "no logical scheme of the parts of speech – their number, nature, and necessary confines – is of the slightest interest to the linguist. Each language has its own scheme. Everything depends on the formal demarcations which it recognizes." Invoking this passage for a particularist position is, however, a misunderstanding. What Sapir is objecting to is a logical scheme. Rightly so. However, admitting that each language has its own scheme does not exclude the possibility that a set of such schemes come under a common denominator, just as the fact that every social community has its own marriage customs does not exclude the possibility of typologizing marriage customs.



unpredictable ways" [Martin Joos 1957]), we have had Generative Grammar ("Grammatica una et eadem est secundum substantiam in omnibus linguis, licet accidentaliter varietur" [Roger Bacon 1244]); and apparently it is now time to swing back to Joos ("Language describers have to create language-particular structural categories for their language, rather than being able to "take them off the shelf"." [Haspelmath 2007, §3]). It seems to be time to halt the pendulum in its middle position: A theory of language necessarily comprises the variation intrinsic and essential to its object, viz. languages; and the description of one of these is the more scientific the more it is based on such a theory. A description of some language which only uses its own concepts would be a contribution to no legitimate scientific activity. There are, admittedly, some published grammatical descriptions which follow the particularist maxim to a great extent. But fortunately they are few, since they are unusable.

A rigid and exclusive distributional approach to the grammar of a language has already been led *ad absurdum*: Setting it up exclusively within the confines of the language to be described results in grammatical categories which are not even language-specific, but "construction-specific" (Croft 2001:105 et pass.). There is, in fact (as Croft also shows), within the particularist approach no known method which would lead from construction-specific categories to interlingual grammatical categories.<sup>19</sup> With this, we are back to the darkest shadows of what one might have hoped was only Chomsky's caricature of "taxonomic structuralism".<sup>20</sup> Particularism is theoretically sterile. If descriptive work is to provide fruitful insights, it needs a theoretical basis.

A methodology for linguistics which denies interlingual status to grammatical categories discredits or renders impossible entire branches of research which have proved insightful in the past. Among these is general comparative linguistics when it strives to find out cross-linguistic regularities in language systems which may be formulated in implicational generalizations<sup>21</sup> of the kind 'if a language possesses category A, then it possesses category B'. Such a statement obviously presupposes the interlingual status of categories A and B.<sup>22</sup> There are also respected and fruitful endeavors in general linguistics of providing a common basis for descriptive grammars.<sup>23</sup> They even explicitly ask describers to state whether the language being described possesses or lacks a certain grammatical category. Far from being "pointless", the answer to such questions is valuable both to the user of the grammar and to the typologist, provided, of course, that the category in question has been defined appropriately at the interlingual level. Finally, it is unclear how a historical grammar of a language could conceive of such facts as that, during the history of the English language, the adjective

---

<sup>19</sup> The linguist chooses the criteria for lumping construction-specific categories into more general ones, and this choice is "opportunistic" (Croft 2001:30, 41). This may be a cynical observation statement, but can hardly be meant as a proposal for a serious scientific method. Moravcsik (2016, §1) gives a positive turn to this formulation.

<sup>20</sup> It should be noted in fairness that Croft does not stop there. He then postulates "comparative concepts" on a purely functional or a hybrid basis. However, like Haspelmath, he has no way to reconcile the "construction-specific categories" and "comparative concepts".

<sup>21</sup> or even in simpler statements of the kind 'there are no articles in Russian' (Lieb 2017, §3.5)

<sup>22</sup> Haspelmath (2007, §2.1) tries to avoid this problem of his approach by postulating a purely semantic status for the categories used in such typological statements and ascribing this methodological move to typologists who have successfully proposed them. However, there is no way of defining grammatical categories or functions on a purely semantic basis; they are necessarily "hybrid".

<sup>23</sup> like the series *Lingua Descriptive Studies* and its successors, based on Comrie & Smith 1976

lost its declension, given that from the particularist standpoint, different stages of English share nothing that might legitimately be called ‘adjective’.<sup>24</sup>

On the other hand, it should surprise nobody that concepts change as science progresses. If a new linguistic phenomenon is encountered in some language, then it is a natural and necessary question to ask whether it instantiates a concept already existent in linguistics. If the answer is negative, there are essentially two kinds of solutions:

a) One recognizes that an available concept has been defined too narrowly. One then drops a condition from its definition so that the new phenomenon fits under the concept without damage to its core or to any neighboring concept. A case in point is the concept of relative clause. Its traditional definition was based on the relative pronoun present in the subordinate clause. This proved to be inapplicable to isofunctional variants of the construction which only feature a subordinator or even nothing in its place, like the English restrictive relative clause. Dropping this condition was the right step because interlingual grammatical constructions are not defined by some item which they contain. The ensuing definition regarded the relative clause as an attributive clause such that the entity designated by its head played some role in the situation designated by the subordinate clause. This definition was then found not to apply to two kinds of constructions which otherwise fulfill the same function: On the one hand, there are head-internal relative clauses (some of them even using the relative pronoun which was the crucial mark when the concept was first formed), which do not come under any operationalizable definition of attribution. And on the other hand, most recently certain subordinate clauses have been identified which form a concept of an entity conceived as playing a certain role in the situation designated by them, but which do not necessarily form an attribute to anything. While this kind of construction has long been regarded as the substantivization of a relative clause, the newly discovered phenomena offer no basis to diagnose a secondary substantivization of what was basically an adjectival; on the contrary, the subordinate clause is basically a substantival which secondarily may be combined, as an attribute, with a head. This now leads to a definition of the relative clause as a specific kind of nominal clause – where ‘nominal’ includes ‘substantival’ and ‘adjectival’ – which is oriented towards one of its actant or circumstant positions, be this occupied or empty. Observe that by the condition of orientation, the relative clause thus conceived is yet sufficiently distinct from any other kind of subordinate clause that the theory provides. This could be, thus, the core of a prototypical concept which allows for much variation and yet is sufficiently precise to exclude phenomena which have never been regarded as relative clauses.

b) Another solution is to create a new concept. A case in point is the ergative construction. Initially (Schuchardt 1906), it was assumed to be some kind of passive construction. Subsequently, however, two properties of ergative constructions were found which exclude such a subsumption: First, in some languages, the ergative contrasts with a passive construction. Second, the ergative construction is typically the basic (active) transitive construction in a language, while a passive construction is not. Therefore one proceeded to refine the definition of the passive so that the ergative was no longer subsumable under it but became a concept of its own, with a term of its own.

Sometimes a phenomenon instantiates a new category which is in a straightforward taxonomic relation to known categories. A language may be found to make a distinction inside what used to be known as a unified category. For instance, the alienability distinction was discovered in non-SAE languages. It required little conceptual adjustment to subsume the two new categories ‘alienable possessive construction’ and ‘inalienable possessive construction’

<sup>24</sup> In the *lingtyp* discussion, Moravcsik (21 and 22/01/2016) makes a similar point with respect to the French and Italian adjective.

under the known concept ‘possessive construction’. Contrariwise, a language may be found to use the same construction for two functions that languages analyzed hitherto provide different constructions for. The general noun-modifying clause construction (Matsumoto et al. 2017:7f) is essentially a construction consisting of a nominal head modified by a clause which, however, may or may not be oriented in the sense indicated above. This is, thus, a hyperonym of the concept ‘relative clause construction’. The question of whether a language that does not formally distinguish between a relative clause and other kinds of adnominal clauses should be said to have a relative clause is like the question of whether English should be said to have an inalienable possessive construction. If, in this kind of question, a construction dedicated to a specific function is meant, the answer is ‘no’. This is the sense of the question that is usually interesting in typology. The opposite sense, viz. a construction that may cover a certain function or sense, is commonly of less use in comparative linguistics, since normally every language can somehow express a given meaning.

With such novel empirical findings, the following question arises: “[W]hen confronted with unfamiliar or previously undescribed linguistic phenomena, how do we know when to establish a new category to account for it, and when to redefine an existing one?” (Nordlinger & Sadler 2008:329).<sup>25</sup> The decision becomes rational if categories are prototypical concepts: If the core of a prototypical concept applies to the new phenomenon, then subsume it, and otherwise not.

Thus, to summarize: There is no essential difference between “comparative concepts” and “descriptive linguistic categories” (Haspelmath 2010[C]:663, 674). On the contrary, for linguistics to be able to function, the same concepts must be used in description and in comparison.<sup>26</sup> This only requires that linguistic concepts have their place in a conceptual network which, in turn, is formulated in terms of a theory of language. The concept of an interlingual category is then just more abstract than the concept of a language-specific category instantiating it.<sup>27</sup> Moreover, the interlingual concept may be a prototypical concept. Then its core features have to be identified and to be distinguished from secondary features, whose possession distinguishes the focal instance from marginal cases.

## 2.5 Definition of lexical and grammatical categories

Since the concepts of lexical and grammatical categories are hybrid in the sense of combining semantic and structural features,<sup>28</sup> the definition of an interlingual category Cg contains conditions of the following kinds:

- (1) The cognitive-communicative (or semantic) function of members of Cg:
  - a) their primary (propositional or grammatical) function

<sup>25</sup> Alex Burri (p.c.) suggests that, given the underdetermination of a scientific theory by empirical data as proposed in Quine 1951, the Nordlinger-Sadler question has no answer. Quine argues (§VI) that, since a scientific statement is interconnected with the rest of a theory, it cannot be tested in isolation, so if it is falsified by data, it is not clear which of the statements of the theory is wrong and in which sense the theory must be amended. However, the present problem concerns a set of concepts, not a set of statements. The subsumption of a phenomenon under a concept is not a falsifiable hypothesis.

<sup>26</sup> The same point is made, on an empirical basis, in Van der Auwera & Sahoo 2015. Lieb (2017, §8.1) converts it into a “basic assumption” of an informal theory of grammars, spelled out in §9.

<sup>27</sup> *A horror abstracti* is one of the features of postmodern particularism. It sometimes seems to be forgotten that science involves abstraction.

<sup>28</sup> “Therefore the categorial definition of a lexeme has to take account of both functional (semantic) and formal (morphosyntactic) criteria.” (Ramat 1999, §4)

- b) the kind of concept that they designate.
- (2) The degree of grammaticality (in the sense of grammaticalization) of Cg:
- a) with the lowest degree of grammaticality, Cg is a lexical category, i.e. either a primary or a secondary word class
  - b) with advanced degrees of grammaticality, Cg is a category of grammatical formatives or even a morphological category.

There is a certain extent of interdependence between the two kinds of conditions:

- For the primary word classes, the cognitive-communicative function is a communicative (propositional) function, viz. reference and predication. With this function, the theory associates a class of designata which the signs having this function typically represent.
- For the secondary word classes and classes of grammatical formatives, their concept is a hybrid composed of some propositional function (like modification), a certain category of concepts designated (like time) and some structural function, i.e. a function in a certain construction (Lehmann 2013).

These principles will be illustrated with the noun:

**Definition:** A noun is a word which, as a member of its class, has the primary function of heading a referential expression.

Theorem 1: The criterial context in which a noun has its primary function is in a sentence of the structure: ‘This X (is) P’, where X is the functional position for candidates and P is some predication.

Theorem 2: A sign whose primary function is to head a referential expression – and which occupies the position of X – typically designates a kind of individual physical object. Prototypical examples include ‘girl’, ‘house’, ‘moon’.

Needless to say, in an empirical science like linguistics, concepts of grammatical categories are never arrived at by deduction only. There is always feedback from empirical research, which renders revision of the concepts and their definition necessary. Some relevant cases are discussed in §2.4.

## 2.6 Operationalizing the definitions

Every scientific theory must be paired with a methodology. One of the main tasks of the latter is to operationalize the concepts of the theory.<sup>29</sup> Given a theoretical concept and a certain phenomenon, it must be possible to ascertain whether or not the latter falls under the former. The operationalization of a concept derives from it a set of criteria and a procedure to apply these to a phenomenon so that the question becomes decidable.

The set of definitions of linguistic categories serves the purpose of enabling a descriptive linguist to assign a category to every sign of the language he is describing, and enabling the comparative linguist to identify phenomena which are comparable among languages. The definition of a linguistic category Cg must enable a linguist describing language L to do two things:

- (1) Determine whether L possesses Cg.
- (2) If L possesses Cg, determine the class of signs constituting Cg.

This requires the operationalization of the definitions. This takes the following general form:

- (1) The following is presupposed:

---

<sup>29</sup> Moravcsik (2016, §2C) means the same when saying that concepts must be applicable.

- a) The definition is hybrid as explained in §2.5.
- b) Two theorems are either part of the definition or deducible in the theory to which it belongs:
  - one which determines a certain functional position P in a sentence for members of C<sub>g</sub>
  - another theorem which identifies a class of designata D<sub>g</sub> which are typically used in P.

Given this, operationalization of the concept of C<sub>g</sub> proceeds as follows:

- (2) Form sentences comprising P and insert signs designating instances of D<sub>g</sub> in it.
- (3) Assume that such signs as fit in the frame belong to C<sub>g</sub>.
- (4) The context of P constituted by such sentences is the constitutive context, on which a prototypical definition of C<sub>g</sub> in L is based.
- (5) Ascertain (by methods of distributional analysis) the distribution D<sub>b</sub> of the words used in P.
- (6) All words of L sharing D<sub>b</sub> belong to C<sub>g</sub>.
- (7) After setting up similar categories of L in the same way, in order to attain the goal of assigning a category to every sign of L, turn the definitions of the kind of # (6) into definitions of prototypical concepts as follows:
  - a) Expand D<sub>b</sub> into D<sub>b</sub>' by selecting in turn each one of the contexts constituting D<sub>b</sub> except the constitutive context and either omitting it from the definition or replacing it by another context, including contexts belonging to other distribution classes.
  - b) Check whether D<sub>b</sub>' comprises signs hitherto unclassified, and if so, subsume them under C<sub>g</sub>.
- (8) Now if two categories turn out to have the same distribution, then the language does not distinguish these categories. Merge them into one.

### 3 Specimen: enumeratives

#### 3.1 Definitions

The following subsections apply the conception of interlingual grammatical categories outlined above to the category of enumeratives. This is an *étude* which pursues two goals:

- a) It demonstrates that an interlingual grammatical category can be defined, and the definitions can be operationalized, in such a way as to tell whether a language possesses or lacks it and, in the former case, determine the members of the class in question.
- b) It examines the following implicational universal: If a language has numeral classifiers, then it has mensuratives.

The hypothesis of implication #b arises on the basis of the following observations: First, mensuratives appear to be widespread and quite common even in languages which lack numeral classifiers. Second, in languages which do have numeral classifiers, these generally share their syntactic construction with mensuratives and are often mixed up with them in descriptions. It appears that a language acquires the numeral classifier construction on the model of the measure construction.

The *étude* requires definitions of the following grammatical categories:

- cardinal numeral
- noun
- individual noun
- mass noun

- enumerative
- measure word
- numeral classifier
- mensurative.

The noun was already defined in §2.5. The remaining definitions will be provided and operationalized in the following subsections.

### 3.1.1 Numeral

**Definition:** A cardinal numeral (in the following, ‘numeral’ for short) is a word<sup>30</sup> of a class each of whose members designates one of a continuous series of integer numbers starting with 1. The prototypical numeral designates the number 3.

**Operationalization:**

- (1) Find the word designating the number 3. If there is none, the language has no numerals. Otherwise:
- (2) Determine its distribution Db.
- (3) Complete<sup>31</sup> the class with other meaningful units sharing Db.

### 3.1.2 Individual and mass nouns

The concepts of individual and mass noun are, of course, based on the concept of noun. They may then be defined as follows:

**Definitions:**

- An individual is an object whose shape is part of its essence<sup>32</sup> and not all of whose parts are alike.
- A mass is an object whose shape is not a part of its essence and all of whose parts are alike.
- An individual noun (also called count noun) is a noun designating an individual.
- A mass noun is a noun designating a mass.

The following theorems should be deducible (s. §2.5 for typical individual objects):

- It is typically the same words which found the class of nouns which also found the class of individual nouns. I.o.w., all prototypical nouns are individual nouns.
- It is typically words designating objects like ‘water’, ‘sand’, ‘salt’ which are mass nouns.

Now the procedure described for the operationalization of such concepts is applied to the object language. Depending on the language, it may turn out that individual and mass nouns differ in their distribution, and then they are distinct word classes of the language, or they have the same distribution, and then the language – although of course having words to designate individuals and words to designate masses – does not grammatically distinguish between the two kinds.

<sup>30</sup> This feature of the definition excludes the grammatical category of number.

<sup>31</sup> Not only numerals, but also other classes of a language system may be non-finite sets. Completion of such a set is a methodological operation of descriptive linguistics whose goal is to determine whether it is an open or a closed set.

<sup>32</sup> This condition is operationalized as follows: Select a candidate object, determine the word designating it, change the form of the object and check if the word designating the changed object is the same.

### 3.1.3 Measure word

**Definition:** A measure word is a word designating a unit of measure and forming a binary construction with a numeral or an enumerative phrase (defined in §3.1.4).

**Operationalization:**

- (1) Find words designating units of measuring length ('foot, span, cubit, fathom ...'), volume ('drop, pinch, handful, bucket[ful], load ...') etc.
- (2) Retain those that fulfill the definitory condition and determine their distribution Db.
- (3) Complete the class with other words sharing Db.

### 3.1.4 Enumerative

**Definition:** An enumerative is an element of a class of words or morphemes which form a binary construction with a numeral, called enumerative phrase (EP). An EP can be used as a referential expression. The EP, in turn, forms a binary construction with a noun designating what is being counted or measured. An enumerative is obligatory in the latter construction, with the exception of the operationalization of the concept of the mensurative, condition (3)c.

If a class of elements fulfills the conditions for the enumerative, but also appears on other constituents of the head noun in a nominal group, it is not a class of enumeratives (but some kind of nominal classifier or nominal class).

Preparation of **operationalization** for application in language L:

- (1) Identify the categories of individual nouns and mass nouns (§3.1.2). If L does not distinguish them formally, it may have a class of enumeratives, but no distinction between numeral classifiers and mensuratives.
- (2) Identify the category of numerals. If L lacks it, it lacks enumeratives, too.
- (3) Combine a numeral X with a noun Y as the core of a referential expression XY, as in 'she gave me XY'.
  - If the construction XY is grammatical as a referential expression for all nouns for which the combination makes sense, L lacks enumeratives.
  - If the construction requires an additional element to combine with X, this is a candidate for an enumerative ( $E_C$ ).
- (4) Omit Y from the construction [X  $E_C$  Y]. If the rest including the  $E_C$  is not a grammatical construction usable as a referential expression, the  $E_C$  is not an enumerative.
- (5) Combine an adjective as an attribute with a noun. If the  $E_C$  appears on the adjective, it is not an enumerative (and instead a member of some kind of nominal class).

The rest of the operationalization identifies either of the subclasses of numeral classifier and mensurative.

### 3.1.5 Numeral classifier

**Definition:** A numeral classifier (also called sortal classifier) is a kind of enumerative. The EP which it forms with a numeral is a classifier phrase (CP). The CP, in turn, forms a binary construction with an individual noun designating what is being counted. The class of numeral classifiers comprises at least two members the choice among which is conditioned by the counted noun.

**Operationalization** for application in language L:

- (1) Execute the preparation for enumeratives described in §3.1.4.

- (2) Form constructions  $C_s$  comprising a combination of a numeral  $X$  with an individual noun  $Y$  which mean ‘ $X$  exemplars of  $Y$ ’.
- (3) Substitute  $X$  with a member of its distribution class while holding  $Y$  constant. If there is a unit in  $C_s$  beside  $Y$  which remains constant, it is a candidate for a numeral classifier ( $NC_C$ ).
- (4) Substitute  $Y$  with a member of its distribution class while holding  $X$  constant. If the  $NC_C$  changes if acceptability is required, the  $NC_C$  is a numeral classifier, and its combination with  $X$  is a CP.

### 3.1.6 Mensurative

**Definition:** A mensurative (also called mensural classifier) is a measure word which is a kind of enumerative. The EP which it forms with a numeral is a measure phrase (MP). The MP, in turn, forms a binary construction with a mass noun designating what is being measured.

#### Operationalization:

- (1) Execute the preparation for enumeratives described in §3.1.4.
- (2) Determine the class of measure words as described in §3.1.3.
- (3) Form constructions  $C_s$  consisting of a combination  $MP_C$  of a numeral  $X$  with a measure word  $M_C$ , combined in turn with a mass noun  $Y$ , which mean ‘ $X$  units  $M_C$  of  $Y$ ’.  $M_C$  in  $C_s$  is a mensurative iff
  - a)  $MP_C$  is minimal (thus, no numeral classifier [for  $M_C$ ] intervenes)
  - b)  $C_s$  is distinct from other nominal constructions (in particular, possessive attribution)
  - c) if  $M_C$  is grammatically omissible, the meaning turns from a measure construction to a count construction (the resulting meaning is a sortal plural).
- (4) If the language has mensuratives, then any (measure) word which forms a MP with a numeral (whether or not the MP combines with a noun  $Y$ ) is a mensurative, too.

## 3.2 Enumeratives in a language sample

### 3.2.1 Enumeratives in Yucatec Maya

Individual nouns of Yucatec Maya (Mayan, Mexico) include *lool* ‘flower’ and *xch’úppal* ‘girl’. Mass nouns include *ha* ‘water’ and *bu’l* ‘beans’. Numerals include *óox-* ‘three’ and *kan-* ‘four’. The direct combination of a numeral with a noun (like *óox-lool*, *kan-bu’l*) is ungrammatical.

E1 shows a candidate for an enumerative. E2 and E3 are constructions meaning ‘ $X$  exemplars of  $Y$ ’. Between #a and #b, the numeral is replaced, but an  $NC_C$  remains beside the counted noun. Between E2 and E3, the counted noun is replaced, and the  $NC_C$  changes accordingly. Consequently, *p’éel* and *túul* are numeral classifiers.

- |    |    |   |
|----|----|---|
| E1 | a. | <i>óox-p’éel</i><br>‘three (inanimate objects)’ |
| E2 | a. | <i>óox-p’éel lool</i><br>‘three flowers’        |
|    | b. | <i>kan-p’éel lool</i><br>‘four flowers’         |
| E3 | a. | <i>óox-túul xch’úppal</i><br>‘three girls’      |



- b. kan-túul xch'úppal  
'four girls'

The language has measure words like *cháach* 'handful' and *luuch* 'cup(ful)', which are, thus mensurative candidates ( $M_C$ ). The constructions in E4 and E5 mean 'X units  $M_C$  of Y'. The  $M_C$  combines directly with the numeral, and the MP combines directly with the mass noun. Other complex nominals, in particular possessed nominals, have a completely different structure. Consequently, the  $M_C$ s are mensuratives. Finally, there are enumeratives like the one in E6 which combine with no noun. These are mensuratives, too.

- E4 a. óox-cháach bu'l  
'three handful of beans'  
b. kan-cháach bu'l  
'four handful of beans'
- E5 a. óox-luuch ha'  
'three cups of water'  
b. kan-luuch ha'  
'four cups of water'
- E6 óox-téen  
'three times'

Result: Yucatec Maya has enumeratives. It has both numeral classifiers like *p'éel* and *túul* and mensuratives like *cháach* 'handful' and *luuch* 'cup'. Their construction is generally the same. They differ in that numeral classifiers combine with individual nouns while mensuratives combine with mass nouns.

### 3.2.2 Enumeratives in Cabecar

Individual nouns of Cabecar (Chibchan, Costa Rica) include *kõchi* 'pig' and *óshkoro* 'chicken'. Mass nouns include *kichõ* 'papaya' and *kuã* 'corn'. Numerals include *mãñá-* 'three' and *tkí-* 'four'. The direct combination of a numeral with a noun (like *mãñá- kõchi*, *mãñá- kuã*) is ungrammatical.

E7 shows a candidate for an enumerative. E8 and E9 are constructions meaning 'X exemplars of Y'. Between #a and #b, the numeral is replaced, but the candidate for enumerative – now the  $NC_C$  – remains beside the counted noun. Between E8 and E9, the counted noun is replaced, and the  $NC_C$  changes accordingly. Consequently, *tãwã* and *tkã* are numeral classifiers.

- E7 *mãñá-tãwã*  
'three (elongated objects)'
- E8 a. *kõchi mãñá-tãwã*  
'three pigs'  
b. *kõchi tkí-tãwã*  
'four pigs'
- E9 a. *óshkoro mãñá-tkã*  
'three chickens'  
b. *óshkoro tkí-tkã*  
'four chickens'

The language has measure words like *táklä* ‘piece’ and *däli* ‘load’, which are, thus  $M_C$ s. The constructions in E10 and E11 mean ‘X units  $M_C$  of Y’. The  $M_C$ , however, does not directly combine with the numeral. Instead, the latter combines with the numeral classifiers already seen. The classifier phrase then forms a loose nominal group with the mass noun, as shown in E10f.

- E10      *māñá*-tkä      *táklä*   *kichó*   *kákum-á*   *i*   *ia*  
 [three-CL.PLAN piece]   *papaya*   *give-PFV*   3   *DAT*  
 ‘gave him three pieces of papaya’
- E11      *däli*   *māñá*-tkä      *kuá*   *wa*      *pshí-lě*  
 [load three-CL.PLAN]   *corn*   *INSTR*   *fill-S.ANT*  
 ‘three loads of corn (lit.: three loads filled with corn)’

Cabecar, consequently, has enumeratives only in the form of numeral classifiers. Measure words are ordinary individual nouns.

### 3.2.3 Enumeratives in German

Individual nouns of German include *Blume* ‘flower’ and *Mädchen* ‘girl’. Mass nouns include *Salz* ‘salt’ and *Wasser* ‘water’. Numerals include *drei* ‘three’ and *vier* ‘four’.

E12 shows a candidate for an enumerative ( $E_C$ ). E13f (same meanings as in E2f) are constructions meaning ‘X exemplars of Y’. The direct combination of a numeral with a noun is grammatical. There is, in fact, nothing like the  $E_C$  that could be inserted as a coconstituent of the numeral in this construction. Consequently, German lacks numeral classifiers.<sup>33</sup>

- E12      *drei Pfund*  
 ‘three pounds’
- E13    a. *drei Blumen*  
       b. *vier Blumen*
- E14    a. *drei Mädchen*  
       b. *vier Mädchen*

The language has measure words like *Pfund* ‘pound’ and *Liter* ‘liter’, which are, thus mensurative candidates ( $M_C$ ). The constructions in E15f mean ‘X units  $M_C$  of Y’. The  $M_C$  combines directly with the numeral into an MP, and the MP combines directly with the mass noun. Other complex nominals, in particular possessed nominals, have a completely different structure. Consequently, the  $M_C$ s are mensuratives. Finally, there are enumeratives like *Mal* ‘time’ in E17 which combine with no noun. These are mensuratives, too.

- E15    a. *drei Pfund Salz*  
       ‘three pounds of salt’  
       b. *vier Pfund Salz*  
       ‘four pounds of salt’
- E16    a. *drei Liter Wasser*  
       ‘three liters of water’  
       b. *vier Liter Wasser*  
       ‘four liters of water’

<sup>33</sup> One may consider defining the concept of ‘numeral classifier’ as a prototypical concept in such a way as to allow for German *Stück* and *Mann* as peripheral instances of this concept; s. Lehmann 2000.

E17        dreimal  
              ‘three times’

### 3.2.4 Enumeratives in Latin

Individual nouns of Latin include *flos* ‘flower’ and *puella* ‘girl’. Mass nouns include *sal* ‘salt’ and *aqua* ‘water’. Numerals include *tres* ‘three’ and *quattuor* ‘four’.

E18 shows an  $E_C$ . E19f (same meanings as in E2f) are constructions meaning ‘X exemplars of Y’. The direct combination of a numeral with a noun is grammatical. There is, in fact, nothing like the  $E_C$  that could be inserted as a coconstituent of the numeral in this construction. Consequently, Latin lacks numeral classifiers.

E18        tres librae  
              ‘three pounds’

E19    a. tres flores  
          b. quattuor flores

E20    a. tres puellae  
          b. quattuor puellae

The language has measure words like *libra* ‘pound’ and *sextarius* ‘pint’, which are, thus mensurative candidates ( $M_C$ ). The constructions in E21f (same meanings as E15f) mean ‘X units  $M_C$  of Y’. The  $M_C$  combines directly with the numeral, and the MP combines directly with the mass noun. These constructions, however, have the very same structure as other complex nominals, in particular possessed nominals like E23. Consequently, the  $M_C$ s are no mensuratives. The result is that Latin has no enumeratives.

E21    a. tres librae salis  
          b. quattuor librae salis

E22    a. tres sextarii aquae  
          b. quattuor sextarii aquae

E23        tres filii consulis  
              ‘three sons of the consul’s’

The gender appearing on the numeral in E21a and E22a is not an enumerative because it also appears on adjective attributes.

### 3.3 Result

The foregoing étude has shown two things:

- a) Grammatical categories can be defined at the interlingual level, and the concepts can be operationalized in such a way that they can safely be applied both in the description of a language and in the comparison of languages.
- b) The putative implicational generalization according to which possession of numeral classifiers presupposes possession of mensuratives is falsified by Cabecar.

A final methodological remark on the above definitions of several interlingual categories is necessary. Of course, it is possible that someone identifies a phenomenon in some language which should be considered a numeral classifier but fails to be covered by the definition or its operationalization. Or again, a phenomenon may come under numeral classifier by the defini-

tion offered, but should be excluded. In such a case, we would say that the definition is deficient. The natural reaction would be to amend the definition. It would not be a necessary conclusion, but would instead mean throwing out the baby with the bath water, to conclude that there are no interlingual grammatical concepts.

## 4 Conclusion

The general purpose of this contribution is to show that the description of the grammar of a language and the comparison of grammars of languages are based on one and the same linguistic theory and, consequently, employ the same conceptual apparatus. In particular, the concepts of grammatical categories are in a taxonomy such that a grammatical category of an individual language is an instance of an interlingual category, and consequently its concept is a hyponym of the interlingual concept. Since even the concepts of language-specific grammatical categories are mostly prototypical concepts, interlingual concepts are prototypical *a fortiori*. A category of a particular language may be a focal instance of an interlingual category or instantiate it only in a marginal way. A responsible application of the concept of a grammatical category to phenomena of a particular language presupposes its operationalization. This, in turn, presupposes a much more elaborate methodology than the discipline of linguistics can presently boast.

## References

- Bhat, D.N.S. 1994, *The adjectival category. Criteria for differentiation and identification*. Amsterdam & Philadelphia: J. Benjamins.
- Comrie, Bernard & Smith, Norval 1977, "Lingua Descriptive Studies: Questionnaire." *Lingua* 42:1-72.
- Croft, William 2001, *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Delacroix, Henri 1924, *Le langage et la pensée*. Paris: F. Alcan.
- Frege, Gottlob 1892, "Über Begriff und Gegenstand." *Vierteljahresschrift für wissenschaftliche Philosophie* 16:192-205.
- 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 crosslinguistic studies" *Language* 86:663-687.
- Haspelmath, Martin 2010, "The interplay between comparative concepts and descriptive categories (reply to Newmeyer)". *Language* 86:696-699.
- Ježek, Elisabetta & Ramat, Paolo 2009, "On parts-of-speech transcategorization." *Folia Linguistica* 43:391-416.
- Lehmann, Christian 1996, "Linguistische Terminologie als relationales Netz". Knobloch, Clemens & Schaefer, Burkhard (eds.), *Nomination - fachsprachlich und gemeinsprachlich*. Opladen: Westdeutscher Verlag; 215-267.
- Lehmann, Christian 2000, "On the German numeral classifier system." Schaner-Wolles, Chris & Renison, John & Neubarth, Friedrich (eds.), *Naturally! Linguistic studies in honour of Wolfgang Ulrich Dressler, presented on the occasion of his 60th birthday*. Torino: Rosenberg & Sellier; 249-253.
- Lehmann, Christian 2013, "The nature of parts of speech". *Sprachtypologie und Universalienforschung* 66/2:141-177.
- Lehmann, Christian in press., "Grammaticalization and automation". Contribution to the 23. LIPP-Symposium "Grammatikalisierung in interdisziplinärer Perspektive", Internationale Konferenz der Graduiertenschule Sprache & Literatur München, 6.-8. Juli 2016.

- Lieb, Hans-Heinrich 2017, "Describing linguistic objects in a realist way". Behme, Christina & Neef, Martin (eds.), *Essays on linguistic realism*. Amsterdam & Philadelphia: J. Benjamins; 000-000.
- Matsumoto, Yoshiko & Comrie, Bernard & Sells, Peter 2017, "Noun-modifying clause constructions in languages of Eurasia. Rethinking theoretical and geographical boundaries." Matsumoto, Yoshiko & Comrie, Bernard & Sells, Peter (eds.), *Noun-modifying clause constructions in languages of Eurasia. Rethinking theoretical and geographical boundaries*. Amsterdam & Philadelphia: J. Benjamins (Typological Studies in Language, 116); 3-21.
- Moravcsik, Edith A. 2016, "On linguistic categories". *Linguistic Typology* 20:417-426.
- Nordlinger, Rachel & Sadler, Louisa 2008, "When is a temporal marker not a tense? Reply to Tonhauser 2007". *Language* 84:325–31.
- Quine, Willard van Orman 1951, "Two dogmas of empiricism." *Philosophical Review* 60:20-43.
- Ramat, Paolo 1999, "Linguistic categories and linguists' categorizations." *Linguistics* 37: 157-180.
- Schuchardt, Hugo 1906, "Über den aktivischen und passivischen Charakter des Transitivs". *Indogermanische Forschungen* 18:528-531.
- van der Auwera, Johan & Sahoo, Kalyanamalini 2015, "On comparative concepts and descriptive categories, such as they are". *Acta Linguistica Hafniensia* 47/2:136-173.