

刘丹青 主编

吴建明 夏俐萍 副主编

# 语言类型学集刊



第二辑

BULLETIN OF LINGUISTIC TYPOLOGY

上海教育出版社



刘丹青 主编

吴建明 夏俐萍 副主编

# 语言类型学集刊

第二辑

BULLETIN OF LINGUISTIC TYPOLOGY

上海教育出版社



# 目 录

|   |                   |
|---|-------------------|
| The Debate on Comparative Categories (Concepts) in Typology |                   |
| 语言类型学界有关比较范畴(概念)的争论   |                   |
| ..... Randy J. LaPolla(罗仁地)( 1 )                            |                   |
| 寄生范畴:源于语法库藏限制条件的语义范畴 .....                                  | 刘丹青( 20 )         |
| 语言类型学视野下的汉语起点介词语义地图研究                                       |                   |
| ——以汉语方言的分析为中心 .....   | 朴正九 姜柄圭 柳秀京( 44 ) |
| 语言类型学视角下汉语方言“V快”的历时与共时考察 .....                              | 崔山佳( 88 )         |
| 北京普通话一级元音的大样本统计分析 .....                                     | 胡会娟(109)          |
| 苍南吴语位移事件的句法语义类型 .....                                       | 姜淑珍(128)          |
| 词形构造、语素库藏与语义关联:汉语方言疑问代词编码方式的                                |                   |
| 类型学研究 .....   | 盛益民(148)          |
| 从方位成分的使用看汉日英“存在”范畴原型表达的语义类型                                 |                   |
| .....   | 吴芝蒸(167)          |
| 汉语方言名词性领属结构的语义地图研究 .....                                    | 叶婧婷(177)          |
| 汉日语名词谓语句对句末名词语法化的影响 .....                                   | 黄小丽(190)          |
| 跨方言比较视角下汉语量词显赫功能及动因 .....                                   | 高亚楠(207)          |
| 从壮语看汉语史“着”表祈使语气的来源 .....                                    | 覃凤余 王全华(220)      |
| 领有动词与完成体表达的类型学研究 .....                                      | 吴春相 杜 丹(235)      |
| 名词等级对普什图语分裂施格的影响 .....                                      | 缪 敏(250)          |
| 母语类型的多维干扰   |                   |
| ——维吾尔语重音模式匹配汉语声调的机制 .....                                   | 杜兆金 陈保亚(262)      |

2 语言类型学集刊(第二辑)

|  |              |
|--|--------------|
| 汉语方言入声音节的类型学观察 .....                   | 李 兵 常 敏(292) |
| 试论湖南方言的三类体标记 .....                     | 王振宇(325)     |
| 疑问代词感叹功能的语言类型差异 .....                  | 杨 娜(350)     |
| 惯常义演变为经历义的多样性<br>——以英语、马来语和古汉语为例 ..... | 陈前瑞 杨育欣(360) |
| 语言类型学与认知语言学视角下苗汉空间词汇对比研究 .....         | 李梦娟(375)     |
| 乳山话正反问句“是不/是没 VP”及其类型归属 .....          | 张甜甜(386)     |
| 后记 .....                               | (393)        |

# The Debate on Comparative Categories(Concepts) in Typology

## 语言类型学界有关比较范畴(概念)的争论

Randy J. LaPolla (罗仁地)

### Some definitions:

|                                     |   |
|-------------------------------------|---|
| ALT                                 | Association for Linguistic Typology(语言类型学协会)  |
| LingTyp                             | Listserv list for the ALT(语言类型学协会电邮讨论网)   |
| Descriptive categories              | categories used for describing the phenomena of individual languages(描写个别语言的时候用的形态句法范畴) |
| Comparative categories/<br>concepts | categories/concepts used for comparing languages<br>(比较不同语言的时候用的语义或者语义/形态句法范畴/<br>概念)   |

### Background:

For quite a while there has been debate in the field of typology about the nature of the categories or concepts used in typological comparison and the difference between these categories or concepts and the categories used in describing individual languages.

In early 2016 there was a debate on LingTyp on this topic, and this debate was seen to be very significant, and so the Editor of *Linguistic Typology* produced a special issue of the journal *Linguistic Typology* (20.2) on this topic with position papers written by the members who had participated in the online debate. I would like to present the debate here, to clarify the is-

sues, and also point to a particular insight that came out of that debate and a following one early this year, and discuss its implications.

The full 2016 discussion is available as a doc file in the LingTyp archives at <http://listserv.linguistlist.org/pipermail/lingtyp/2016-January/thread.html> under the heading “Structural congruence as a dimension of language complexity/simplicity”.

### **The issues:**

The discussion of comparative categories began with criticisms from Jan Rijkhoff and myself about methodology in word order studies, particularly the confusion of functional/semantic and formal categories, and not addressing them both (distinguishing them) in doing comparative work (18 Jan 2016).

- Matthew Dryer (19 Jan) replied that he spent years studying the interaction of semantic and formal categories, and “[t]he result is that while the syntactic realization sometimes makes a small difference, it is overall irrelevant; by and large, generalizations over semantic categories apply the same, regardless of the syntactic realization.”
- Dan Everett (19 Jan) then said, “One of the biggest problems in this regard that I have noticed is in grammars of individual languages. Fieldworkers sometimes confuse semantic and formal categories in the grammars, classifying as a syntactic structure a semantic category.”
- I then replied (19 Jan): “Dan’s point is very important. For example, most people describing languages do not know how to distinguish agents, topics, and syntactic pivots (‘subject’), and just call anything that occurs initially as ‘subject’. Sometimes even when the linguist is clear on the difference, they still use the word ‘subject’. E.g. Y.R.Chao, in his grammar of spoken Chinese, clearly stated there is nothing like what is referred to as ‘subject’ in English, as all clauses are simply topic-comment, but he still used the term ‘subject’ for what he said was purely a topic. This has confused generations of linguists, and they call Chinese SVO, which not only implies that Chinese has such categories, but also that these categories either determine or are determined by word order. See the fol-

lowing paper arguing against the use of such shortcuts, and arguing for more careful determination of the factors determining word order in a language:

LaPolla, Randy J. & Dory Poa. 2006. On describing word order. *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix Ameka, Alan Dench, & Nicholas Evans, 269—295. Berlin: Mouton de Gruyter.

[www.ntu.edu.sg/home/randylapolla/Papers/LaPolla\\_and\\_Poa\\_2006\\_On\\_Describing\\_Word\\_Order.pdf](http://www.ntu.edu.sg/home/randylapolla/Papers/LaPolla_and_Poa_2006_On_Describing_Word_Order.pdf)

- Matthew then replied, “Randy says that calling Chinese SVO implies that Chinese has such categories. I am surprised that he would say that. I would have thought it was obvious that **classifying languages typologically does not entail that the terms employed in the typological classification correspond to categories in the language.**<sup>①</sup> Nor does it mean that these categories determine or are determined by word order. I have certainly made that clear in my work that **classifying a language as SVO makes no claim about the categories in the language, nor that these categories determine word order even if the language has such categories.**”
- I then replied (19 Jan): “Hi Matthew, Are you kidding me? Do you really think that when people read a characterisation of word order in a language which says the language is Subject-Verb-Object that readers are not going to assume that the language not only has those categories, but that they must be significant in the determination of the word order (or vice versa) for such a characterisation to be used? Why else would anyone use such a characterisation? And if it doesn’t imply the existence of these syntactic categories or their relation to word order, why use such a characterisation?... You are saying the categorical labels you use in your typological classifications have no relation to the actual typological facts of the language. So, for example, we apply the category label ‘subject’ to a language, even though we know it doesn’t have anything like a subject. Do you really want to say that? My own view has always been to stick with the actual facts of

---

① I have added bold to some sections to highlight particularly important statements.

the languages. I don't know what use a typological classification that was not based on the facts of the languages would have."

- Matthew then replied (19 Jan-partial): "As I argued in Dryer (1989), languages in which word order does not code grammatical relations and in which the word order is not based on grammatical relations but in which VO word order is more common tend to have word order properties associated with VO word order, like prepositions, while analogous languages in which OV word order is more common tend to have word order properties associated with OV word order, like postpositions. **What this means is that the GRAMMARS of what I classify as VO languages have nothing in common. It is only the languages that have something in common at the level of usage.**"
- I then replied (19 Jan-partial): "In terms of the correlations you talk about among languages that manifest what is (from my view problematically) subsumed under the VO or SVO rubric, **my view is that we should look for the reasons why, in terms of information structure, structural pivots, historical development, or whatever, the languages manifest the particular patterns they do.** Simply lumping them together under a single rubric does nothing but categorise them, and doesn't explain anything."
- Matthew then further explained (20 Jan-partial, two different posts): "... since the number of possible types is presumably infinite, typology necessarily lumps together languages which are some level are clearly different. Even if I lump together types of languages which differ in some crucial way in their grammatical system, I feel obligated to justify grouping them together. That is the basis for my grouping together languages in which there is a grammatical rule that defines the language as what I would call SVO and languages where there are no syntactic rules governing clause word order, but where the pragmatic factors governing word order and the frequency in discourse in which those different pragmatic factors arise result in more frequent SVO order at the level of frequency in usage."  
 "The class of languages I treat as SVO is defined roughly as those languages where the statistically dominant order in usage is AVP. **There is nothing that the grammars of this set of languages share; these languages resemble**

**each other only at the level of usage, not at the level of grammar.”**

- Bill Croft then came back to the original issue (21 Jan-partial):

“Dan commented that the use of terms in grammatical description is not always clear, in particular that grammatical and functional categories are confused. This is a major and serious problem. Related to it is that (hybrid) comparative concepts are confused with language-specific concepts, or are poorly defined—this is the sense in which someone uses the term ‘adjective’ or ‘ergative’, or more typically argues that a language does not have ‘adjectives’ or is not ‘ergative’, in some crosslinguistic sense. As David Beck notes, this is usually a pointless or meaningless exercise ... People very commonly interpret typologists’ talk of comparative concepts as language-specific grammatical categories or word classes...”

**Matthew takes the view that ‘grammar’ and ‘usage’ are completely separate ... and appears to say that ‘grammar’ forms a ‘system’ (in the structuralist sense?) that has nothing to do with crosslinguistic categories that typologists use to formulate language universals about things like word order or the past perfective.** I disagree with those assumptions about what I would call a grammar, that is, what a speaker knows about her language (hence the scare quotes in the preceding paragraph). The crosslinguistic diversity of grammatical behavior that led typologists to conclude that grammatical categories are language-specific is mirrored by the diversity of grammatical behavior of words and phrases across constructions that leads to the conclusion that categories are not just language specific but construction-specific (Croft 2001 etc.). **The same patterns that in one language are (or seem to be) categorical distinctions are found in other languages as token frequency differences; ... These observations break down the barrier between ‘grammar’ and ‘usage’, and also mean that language-specific categories are not fundamental building blocks of grammatical description but one side of a rich and complex relationship between constructions and the units that fill the roles in the constructions.”**

- Peter Arkadiev then replied (21 Jan): “... in my view (shared by many, I dare say), if a language has OV as only the most frequent pattern in

usage, with some alternative orders as subsidiary patterns, **this is certainly a fact about that language’s grammar, and this fact has to be reflected not only in the grammatical description of this language, but in the typological classification thereof, as well.** At least under certain granularity of typology, as allowed by our comparative concepts. Otherwise we might miss something very important.”

- Martin Haspelmath then echoed Matthew Dryer’s position in saying (23 Jan), “if grammars are analogous to genomes, the totality of our linguistic behaviour (within a speech community) is analogous to phenotypes. I would say that typologists generalize over these phenotypes—compare Matthew’s point that usage frequencies also play an important role for word order typology.”

#### **Summary of debate:**

It became clear that we all agree that there are no cross-linguistic formal categories (see e.g., Dryer 1997, 2016; Croft 2001; Haspelmath 2007, 2010b, 2017), and so language description should be inductive and based on the facts of the language, and that language-specific descriptive categories used in describing a language are unique to that language, and that comparative categories are abstractions away from the actual details of the language-specific categories. We differ though in how far we feel the abstractions can go from the actual facts of the language. We also differ in terms of what we consider to be “the facts of the language”, as Dryer and Haspelmath argue that there is a difference between the grammar of a language (the conventions the speaker follows) and the actual usage manifested in the texts of the language, essentially the Structuralist distinction between *langue*/competence vs. *parole*/performance, and argue that typology is based on *parole*/performance, and not *langue*/competence/grammar. Jan Rijkhoff, Peter Arkadiev, Bill Croft, and I independently argued that there is only what is manifested in the texts and so usage is the grammar of the language and so the facts of the language, and so there is no justification for ignoring the reasons for the patterns being the way they are.

**My reaction:**

In my paper for the *Linguistic Typology* special issue on this question, I argued that in doing both description and comparison we should work inductively, staying true to the facts of the languages as manifested in natural data, and not resort to extreme abstractions that lead to classifying languages or constructions in a way that obscures the diversity of the structures of the languages:

“Within the Structuralist paradigm and the Boasian tradition in language description, language particular categories are established on the basis of an inductive analysis of the morphosyntactic distribution of forms in texts in a particular language. The functions ascribed to the forms are inferences based on their distribution. The labels used for the categories that emerge from the analysis may be those used in the description of other languages, as long as there is enough of a ‘family resemblance’ between the two categories. That is not the same as equating the two, or imposing a category that has no justification in the language (“apriorism”), or saying that there are crosslinguistic categories, but is simply saying that the category which needs to be described to make sense of the use of the form in that language is in some way similar to a category posited for another language and so can take the same label, like a greyhound and a beagle are both given the label ‘dog’ even though there are many differences between them.” (LaPolla 2016b: 365)<sup>①</sup>

“Comparative categories are dependent on descriptive categories, as they are idealizations or prototypes formed on the basis of the family resemblances found in the descriptions. The International Phonetic Alphabet is an example of this, creating idealizations or prototypes of sounds found to be used for distinguishing meaning in language use, or, in the case of the Cardinal Vowels, sounds that can be defined in a relatively objective

---

<sup>①</sup> See also Dryer 2016: 307—308 on the use of labels for categories in language description that are the same as those used for other languages “based on striking similarity to categories in other languages for which these labels have been used”.

way so that they can be used as reference points for language-particular descriptions. As with the IPA, the grammatical comparative categories often are given the same labels as used in the descriptions, as Latin or English (or Chinese in the context of China) categories are taken as key reference points.” (LaPolla 2016b: 367)

Matthew Dryer and Martin Haspelmath argue, though, that typological comparison of languages does not depend on the individual descriptions of the languages, and is a separate enterprise altogether. Typologists are free to create categories and apply them at will to languages, using the same labels as used in many descriptions, even if the language does not manifest the particular feature identified in the label, and the reader is supposed to know that the comparative categories do not mean the same thing as the descriptive categories.

Aside from the problem of the competence vs. performance view that Dryer and Haspelmath work with, there is also a logical problem in assuming that even though the languages don't manifest the features identified in the label, they can still be given that label. In my paper I used an analogy that was discussed in the online debate, “wings” as a comparative category. From at least some perspectives saying bats and birds both have wings is not problematic, because there is enough of a family resemblance between the structure and function of the wings of birds and the membranes of bats to justify including them in the same category, but what Dryer is talking about (e.g., classifying Chinese as SVO even though the word order is controlled by information structure and not controlled by factors such as Subject and Object or A and P—which is how Dryer defines “Subject” and “Object” in Dryer (2013)—because in texts counts A's occur more frequently in initial position) is equivalent to saying that not only can birds and bats be classified as winged entities, but rockets can also be classified as such, because they have flight properties associated with winged entities. That is, ignoring the fact that the A's occur more often in initial position in Chinese because they are more often topical (and not because they are S or A) and classifying the language as SVO is the same as saying that since a bird flies because of its wings, and

a rocket flies, we can classify the rocket as a winged entity. I don't find this type of classification helpful. As argued in LaPolla & Poa 2006, it is better to look at the factors that actually determine the word order patterns in the languages, and then we can compare languages in terms of those factors.

Haspelmath's comparative concepts (e.g. 2010, 2016, 2017) are not different from the concepts we have been using all along in typology and description, except that he argues that there does not need to be a connection between what I think are the facts of the language and the way it is categorized typologically, agreeing with Dryer. This is something I have been arguing against for many years (see, e.g., LaPolla 2002, presented at a conference in 1994) because I think it is confusing to non-specialists, and because it does not produce useful results, as it makes language patterns that aren't similar look similar. For example, in word order typology, Dryer claims that the grammatical nature of a particular element makes no difference to how it patterns grammatically, but that is not true of the languages that I am familiar with. For example, it is because clausal noun modifiers in Chinese are nominal that they appear in the same construction as nominal noun modifiers and also pattern the way they do, being able to appear either before or after the noun head depending on pragmatic factors, and having no syntactic constraints on the type of semantic relation possible between the head and the modifying clause (see LaPolla 2017). In Rawang (Sino-Tibetan; Myanmar) there are both nominalized and non-nominalized clausal noun modifiers, and they pattern differently precisely because of their nominal vs. non-nominal status (see LaPolla 2008).

Although I believe the Structuralist paradigm ("basic linguistic theory") is problematic from a scientific point of view (see LaPolla 2016a), I recognize that it is still the dominant framework for language documentation and description, as well as typology, but the method must be used in a strictly inductive way, that is, no categories should be imposed on the language, and the labels used for the categories should follow the principles I discussed in Section 1 of LaPolla 2016b (partially quoted above), with each label/category explicitly defined.

Large scale language comparison, combing hundreds of grammars (of varying quality) and extracting forms that one thinks might fit one's comparative categories (regardless of what the author of the grammar might have said) can be quite problematic, and the limitations of the linguist's knowledge become all too obvious. It is much better to concentrate on languages one has a good knowledge of and contribute to typology by expanding our understanding of what is found in the languages and how we might understand it, including its historical origins. The denial of the relevance of history is another problem with the Structuralist approach: the question that started the debate was about "structural congruence as a dimension of language complexity", and the idea was to look for mystical harmony principles for the facts of word order, when in fact simple reference to the history of the language often explains the facts, e.g., if you have a possessive construction where reference to the possessor precedes reference to the possessed, and in the case of some words that construction is reanalysed as an adposition construction, as often happens, then you end up with "structural congruence" between the adposition construction and the possessive construction. See LaPolla 2002 for discussion of this and other problems with the methods and explanations used in word order typology.

In my contribution I also presented a non-Structuralist alternative view of communication and typological description, based on a theory of communication and the related cognition I have been developing for more than 20 years (see LaPolla 2015, 2016a for the most recent presentations, and LaPolla & Poa 2002 for a presentation in Chinese). In this view, communication is not based on language, but on abductive inference of the communicator's intention in performing a particular act with communicative intent, and language is seen as arising from the desire to constrain the inferential process in particular ways. Both individual language description and typology can be done by looking at whether or not the languages of interest constrain the inference of a particular aspect of meaning (which can be seen as an area of semantic space), and if so, how. For example, if we are interested in comparing languages in terms of constraining the interpretation of the inference of

the time of an action relative to the time of speaking, we can ask if the languages do it only lexically or do they do it with conventionalized morphological constructions, and, if they have conventionalized constructions for this, we can look at to what extent they “cut up” the possibilities, e.g., past vs. non-past, or three degrees of past vs. non-past, or past vs. present vs. future; and also we can ask what specific morphosyntactic form the construction takes. This gives us the full diversity of language structures, rather than obscuring that diversity.

### **The “genifier” debate:**

The issue came up again on the LingTyp list in March 2017, as Martin Haspelmath proposed a new comparative category that lumped gender systems and classifier systems into one category, called “genifiers”. Again I and others opposed such a move, as we saw it as categorisation that obscured the facts of the languages (gender marking and classifiers are doing different things). During discussion of this offline with Martin, he sent me a paper he had just finished, a draft of his reply to the papers in the *Linguistic Typology* special issue on classification. In that paper, Martin (2017) says that what he means by the use of the expression “the facts of the language” is different from my use of the expression “the facts of the language” (as in the title of my 2016 paper: “On categorisation: Stick to the facts of the language”). He says that what I am talking about is not the facts of the language, but the grammar of the language, what regulates the forms, whereas he is talking about the actual forms as spoken without regard to why they are the way they are. This again reflects his Structuralist assumption of a distinction between langue/competence vs. parole/performance, as he mentions in a footnote (Haspelmath 2017, § 10, p.19):

“... the universal in (10a) entails a statement such as (10b).

- (10) a. In almost all languages, the subject normally precedes the object when both are nominals. (Greenberg 1963, Universal 1)  
 b. In Mandarin Chinese, the subject normally precedes the object.

LaPolla (2016: § 2) objects to the claim that Chinese is an SVO language (which is a more specific claim than (10b), but otherwise very similar) because he has shown in earlier work that Chinese does not have any subject or object category, and he thinks that ‘labeling [Chinese as an SVO language] implies that these categories either determine word order or are determined by it’ (cf. LaPolla & Poa 2006). But again, this is not so. **(10b) is a correct factual statement about Mandarin Chinese (assuming that ‘subject’ means S/A, and ‘object’ means P), and it is not a rule of Mandarin grammar.**\* LaPolla may be right that ‘most people who see a description of Chinese as SVO will in fact assume that the label was given to the language because those categories are significant for determining word order in the language’ (2016: 370). But if they do, they have not understood the difference between describing a language and classifying a language from a comparative perspective. These two are different enterprises—not completely unrelated, because both are based on the facts of the language, but also not identical ... **Thus, what we compare across languages is not the grammars (which are incommensurable), but the languages at the level at which we encounter them, namely in the way speakers use them.**

\* Footnote: Confusingly, LaPolla (2016) uses the expression ‘the facts of the language’ in the sense in which I use ‘rules of the language’ (this strange terminology may be motivated by his rejection of ‘structuralism’ and the competence/performance distinction).”

In response to this, I replied (email 27 April), “In reference to what you say in § 10, my problem is that by saying what you say as a ‘factual statement’, you are not representing it in the most relevant way, which is to say topics occur before the verb, and it just happens that A and S are more often topical than not, and so end up in that position. The language is structured differently from IE languages, and yet typology in many ways is still forcing all languages into an IE structure.”

He replied (27 April-partial): “Exactly. It’s a correct factual statement, but

not a good way to formulate the rules of the language. That's why typology cannot be based on the rules of languages, but must be based on the facts of languages."

I then said (27 April): "When you say 'we need typology for an understanding of the facts of languages in functional terms' I am in full agreement, but we seem to differ somewhat on what counts as the facts of languages. For me it is more interesting to look at the functional diversity of the uses of language structures rather than just the resulting structures themselves without reference to their use or motivation. That doesn't mean you can't do comparison, it is just not simple structural comparison, it is comparing the languages in terms of the cognitive and cultural categories manifested in the individual languages."

His reply (28 April): "I guess that I'm a bit more pessimistic than many other functionalists with respect to the possibility of finding functional explanations in individual languages. Of course, one can often say interesting things, but it seems to me that these are always hypotheses that need to be tested by large-scale typology."

My reply (29 April): "**... my point was that in making the statement in terms of A/S or 'subject' it makes it seem as if that is the causal factor, so I would prefer statements that really do include the clausal factor, e.g. information structure in the case of Chinese. If a language has grammaticalised A/S pivots in a majority of constructions in the language, then one might be able to make a statement that the word order reflects this A/S pivot, leading to the word order pattern we find most often, but if there is no evidence of that, then other factors should be identified as resulting in the patterns found.**"

So here I realised the major difference between our positions was in whether we had an interest in understanding causation as a goal of typology (or any comparative linguistics) and also whether causation was an operative factor in our comparisons. I argue causation should be a factor in our comparisons, and also that understanding causation is a goal of comparative linguistics, while Haspelmath thinks otherwise:

His reply (29 April): "I never use the term 'causal factor' at the level of