

CLIPP

Christiani Lehmanni inedita, publicanda, publicata

titulus	
	On the function of numeral classifiers
huius textus situs retis mundialis	
	http://www.christianlehmann/publ/lehmann_num_class.pdf
dies manuscripti postremum modificati	
	23.09.2012
ocasio orationis habitae	
	–
volumen publicationem continens	
	Florivic, Franck (ed.), <i>Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels</i> . Lyon: École Normale Supérieure
annus publicationis	
	2010
paginae	
	435-445

On the function of numeral classifiers

Christian Lehmann

University of Erfurt

Abstract

There is a long-standing functional-cognitive answer to the question of why a language should have numeral classifiers: they individuate nominal concepts that otherwise could not be counted. A sober structural analysis of numeral classifiers in a couple of languages shows that such an hypothesis is not needed because they fulfil a purely structural function: in a language where numerals are affixes or otherwise less independent than nouns, the numeral classifier is an operator that converts a numeral into a (nominal) word. This hypothesis is in full consonance with established universals concerning the categorical status of numerals and, at the same time, explains the restriction of numeral classifiers to lower numerals in most languages.

1 Introduction

In what follows, a brief analysis of numeral classification in Yucatec Maya is offered, with some side glances at German and Chibchan languages and some hypotheses about implicational generalizations and grammaticalization concerning numeral classifiers and cardinal numerals. E1 and E2 show some nominals containing a numeral from Yucatec Maya (see § 3 below for an analysis).

- E1. a. hun-p'éel tunich
YM 'a/one stone'
- b. ka'-p'éel tunich
'two stones'
- E2. a. hun-túul xib
YM 'a/one man'
- b. ka'-túul xib
'two men'

The classifier is obligatory if the nominal is modified by a numeral; expressions such as *hun tunich* or *ka' xib* are ungrammatical. This picture is familiar from many other numeral classifier languages.

2 Functional hypotheses

Numeral classifiers are unknown in the languages of Europe and came to the attention of general linguists relatively late. When somebody who speaks a functioning language encounters, in another language, a grammatical category that he has never seen, it is an understandable, if somewhat chauvinistic reaction to ask: What do they need that category for? This is also a typical question of functional typology. There is nothing wrong with it in principle. However, the way towards a scientific answer is longer than has often been thought. The linguistically uninformed mind usually jumps to some functional explanation before having done any structural analysis. Without it, the functional explanation is necessarily speculative. This methodological – or rather, amethodological – attitude was characteristic of much of early linguistic typology since Wilhelm von Humboldt, as it could not fail to be before the advent of structural linguistics.

As for the function of numeral classifiers, it was early recognized that usually the noun to be counted selects the numeral classifier, that usually a set of nouns of the same semantic class selects the same classifier, and that the classifiers, in that sense, classify the nouns they combine with. Thus, the two numeral classifiers illustrated in E1f classify inanimate and animate nouns, respectively. That is why this grammatical category was called '(numeral) classifier' in the first place.

This gave rise to the question why counting an object should presuppose its classification. The answer was (e.g. in Lévy-Bruhl 1910) that the speakers of a numeral classifier language have not arrived at the pure concept of the number; they have not yet abstracted cardinality as the common property of all the manifold sets of objects which happen to be of the same number. That answer may have been informed by the knowledge of ancient languages and cultures like Sumerian which use several different numeral systems depending on the object to be counted. Anyway, it smacks considerably of European superiority.

In Greenberg 1972, it is observed that languages with numeral classification typically lack the category of nominal number. And if they have it – Yucatec Maya has it –, it usually does not appear in a noun phrase containing a classifier phrase, as E1f. This gives rise to the hypothesis that while a concrete individual noun in a language with nominal number is a count noun, this is not so in a numeral classifier language. Here there is no distinction between mass and count nouns. The noun does not designate an individual falling under the concept, but just the concept itself. The classifier would then have the function of individuation.¹

Now the theory of 'individuation by classification' doubtless has its appeal. It can be applied with little modification to the analysis of gender and noun class as they occur in Indo-European and Bantu languages. If a language has gender or noun class, then it has it in its pronouns (Greenberg 1963:96). The primary locus of these categories is in anaphora (cf. Greenberg 1972:33, Creissels 2006:106). The genders and noun classes are used to keep referents apart. And on the other hand, the semantics of the genders and noun classes itself is often vague at best. These grammatical categories do not so much classify referents – as the non-linguistic mind usually thinks when she thinks about gender – but instead nouns.² Thus,

¹ According to Greenberg 1972:26, "the classifier is an individualizer". Variants of this hypothesis are endorsed, among others, in Lyons 1977: 462, Seiler 1986:97, 104f and Lucy 1992.

² See Iturrioz Leza 1986 on the metalinguistic function of grammatical categories.

the function of numeral classification, just like other nominal classification, is individuation.

Unfortunately, this theory gets into conflict with languages such as Mandarin where the same classifiers that are obligatory with numerals are also obligatory with demonstratives. (Some other languages extend the use of classifiers to other constituents of noun phrases.) Now we have two kinds of numeral classifier languages: those like Yucatec Maya that require individuation by a classifier in counting, but not in exophora, and those like Mandarin Chinese that need to individuate the object both in counting and in exophora. In the former subclass, we have to assume that the demonstrative, too, has individuating force, while in the latter subclass, it doesn't and therefore needs to be accompanied by a classifier. Thus, the individuation analysis becomes circular.³

3 Structural analysis

At this point, a more thorough structural analysis is called for. At the latest in Greenberg 1972, it was observed that G1 holds:

G1. If a language uses numeral classifiers in counted nominals (Count.Nom) such as E1f, then it also has naked numeral classifier phrases (Num.Cl.P) such as E3f in the place of complete noun phrases.

- E3. a. *hun-p'éel*
 YM [[one]_{Num} – [CL.INAN]_{Num.Cl}]_{Num.Cl.P}
 'one (inanimate object)'
- b. *ka'-p'éel*
 'two (inanimate objects)'
- E4. a. *hun-túul*
 YM 'one (animate being)'

³ Lucy (1992: 43) writes: "Interpretatively, in Yucatec all nouns [...] are neutral with respect to logical unit or shape." As evidence for this claim, he (p. 74) adduces the following data (orthography adapted):

- a/one-CLFbanana(s) [numeral+general_classifier+noun]
- a. *hun-ts'ít ha's* 'one/a 1-dimensional banana (i.e. the fruit)'
 b. *hun-wáal ha's* 'one/a 2-dimensional banana (i.e. the leaf)'
 c. *hun-kíul ha's* 'one/a planted banana (i.e. the plant/tree)'
 d. *hun-kíuch ha's* 'one/a load banana (i.e. the bunch)'
 e. *hun-p'ít ha's* 'one bit banana (i.e. a bit of the fruit)'

Lucy's theory is probably an artifact of this particular example and a couple like them. First observe that different objects (rather than different aspects or conceptions of one object) are being designated by these examples. Second, the same could not be done with most other nouns. For instance, if *ha's* is replaced by the noun for 'person' (*máak*) or 'house' (*nah*), all of the examples become ungrammatical, and only one classifier (*túul* and *p'éel*, resp.) will be acceptable. Third, the theory does not explain why individuation is unnecessary and even impossible if the numeral is of Spanish origin; see below. Thus, this is more an observation about the polysemy of fruit nouns than a theory of the function of numeral classification.

- b. ka'-túul
'two (animate beings)'

Such phrases are typically used in anaphora. For instance, E4.b would be a good answer to the question 'how many cats do you have?' Here, too, the classifier is obligatory; the numeral (Num) cannot stand alone. That means that the construction of E1f is as in S1 (cf. Greenberg 1972:27f).

S1. Structure of the counted nominal (Yucatec)

[[[X]_{Num} – [Y]_{Num.Cl}]_{Num.Cl.P} [Z]_N]_{Count.Nom}

As usual, the implicational universal can be turned into a diachronic law (cf. Greenberg 1969):

- G2. If a language acquires numeral classification in constructions such as S1, it first acquires numeral classifier phrases and uses these in anaphora.

There are numeral classifier languages in which constructions of counted nominals are typically discontinuous, at least when their noun phrase is semantically indefinite. In Chibchan languages (s. Constenla et al. 1998:70 for Bribri, Quesada 2007:59f for Teribe and p. 78 for Cabecar), the usual construction is as in S2, illustrated by E5 from Teribe.

S2. Discontinuous counted nominal (Teribe)

[X]_{Nom} ... V ... [[Y]_{Num.Cl} [-Z]_{Num}]_{Cl.P}

- E5. dröng twlë-no-r pl-ara
TER machete buy-PRF-1.SG CL.LONG-one
'I bought a/one machete.' (Quesada 2000:49)

We may therefore extend G2 and postulate a diachronic development starting with an anaphoric construction in which a classifier phrase serves as an anaphor, passing via a discontinuous counted nominal as in S2 and ending up with a continuous counted nominal as in S1.

This dynamicization of the synchronic analysis serves to underline the observation that is basic in our context: It is not the noun (or nominal) that combines with a numeral classifier, but the numeral. This renders the assumption doubtful that the noun needs some operator to individuate it. It rather seems that it is the numeral that needs the classifier. (At least that much may be conceded to the original functional speculation.)

In Yucatec Maya, the vigesimal numeral system of the classical epoch fell out of use in colonial times. Today, only the numerals from one to four (often only three) are Mayan; further counting is done in Spanish, with E6 taking up on E2:

- E6. a. óox-túul xib
YM 'three men'
b. kan-túul xib = cuatro xib-o'b
'four men'
c. cinco xib-o'b
'five men'

where *-o'b* is the plural suffix. Observe that the classifier that cannot be missing from E6.a

cannot be used in E6.c. Spanish numerals in Yucatec Maya do not allow the numeral classifier. Thus, it is clear once more that it is not the Yucatec noun that needs individuation; it is the Yucatec numeral that needs the classifier. There is a clear sense that native numerals are special in a way that Spanish numerals are not.

The Yucatec numerals never occur without a classifier, and the classifiers never occur without a numeral, or so it seems. For a long time, Mayan linguistics was faced with a combination of what appeared to be two bound morphemes so that it was impossible to tell which was the host and which the affix. However, there is at least one construction that lays the classifier bare:

E7. Hay-túul-o'b? – Seis u túul-ul.
 YM how.many-CL.AN-PL six [POSS.3 CL.AN-REL]_{NP}
 ‘How many were they?’ – ‘There were six of them.’ [lit.: six (are) of them]

The answer in this dialogue again uses a numeral (in predicate function) that counts an anaphoric referent (in subject function). This time the numeral is Spanish, so it does not combine with a classifier. The anaphoric reference to the counted objects, however, is made explicit. It is the classifier that functions like an anaphoric dummy noun. The possessive construction does not need to be analyzed here (cf. Lehmann 2002). It suffices to see that the numeral classifier in E7 functions as its head, just as a noun would. (Still, this free use of the numeral classifier in Yucatec is restricted to counting; there is no construction like **le túul-o'* (DEF CL.AN-D3) ‘that one’ like *le xib-o'* (DEF man-D3) ‘that man’.)

At this point it becomes clear that the question ‘what does the language need numeral classifiers for?’ does not lead to an extralinguistic function. The primary function of this form class is purely structural: its members serve as props for the numerals, which are affixes. The primary function of numeral classifiers is to serve as dummy nouns that those numerals which are affixes can attach to.

Of course one may ask why this structural property of numerals should necessitate a form class of its own for the host function. If they are affixes, why can't they attach directly to the counted noun? In a dynamic perspective, this is not really an alternative. A numeral of nominal (rather than verbal) nature combines (as a constituent, often a modifier) with a nominal. However, in anaphora it is inconvenient to repeat the nominal all the while.⁴ Since in anaphora it suffices to identify a referent by its class, the noun that the numeral combines with is prone to grammaticalize. It thereby becomes a numeral classifier. In a dynamic perspective, a language with numeral classifiers is just further advanced in the grammaticalization of nouns that serve as hosts or heads to numerals.

Finally, this analysis leaves the question unanswered why the counted nominal containing a numeral classifier phrase is not in the plural, even if the language has nominal plural as a category. Whatever the answer to this question is, it has nothing to do with numeral classification. There are many languages such as Turkish, Quechua and Tamil, which lack numeral classifiers, but – optionally or obligatorily – use a non-plural form of the noun in the counted nominal. There is, in fact, no particular reason why it should be in the plural. If it is, as in many modern Indo-European languages, this is probably a symptom of the advanced degree of grammaticalization of nominal number in these languages.

⁴ See Lambrecht 1994:165, 201 for the topic hierarchy.

4 Classifiers and the status of numerals

Here a well-known property of numerals must be briefly recalled. From a structural point of view, numerals may belong to many different word classes. Even cardinal numerals, which appear to be mostly nominal in nature, are verbs in some languages like Yukaghir. To the extent that cardinal numerals are of nominal nature, there is a gamut between more adjectival and more substantival numerals as stated in the implicational generalization G3 (cf. Corbett 1978):

- G3. If a numeral of numerical value X is adjectival in nature, then all numerals lower than X are adjectival, too. Conversely, if a numeral of numerical value X is substantival in nature, then all numerals higher than X are substantival, too.

While the second part of G3 probably holds without exception, the first part holds under the initial condition that the numerals are at all nominal in nature. In Yucatec, the lower numerals are prefixes, the higher numerals – viz. those of Spanish origin – are words. It may thus be possible to extend the transition from adjective to noun stated in G3 in the following sense: There is an implicational scale of grammatical status of numerals as displayed in S3:

S3. *Grammatical status of numerals*

affix < adjective < noun

S3 is coupled with an implicational generalization:

- G4. Given a numeral with numerical value X occupying position Y on S3, then numerals for numbers lower than X will not occupy positions to the right of Y; and conversely, numerals for numbers higher than X will not occupy positions to the left of Y on S3.

In the case of Yucatec, the higher numerals are non-native. On this basis alone it is therefore impossible to tell whether their incompatibility with numeral classifiers is due to their higher cardinality or to their loan status. In other languages, higher cardinality is clearly sufficient to dispense with classifiers. For the Chibchan language Bribri, Constenla et al. (1998:65f) observe that the numeral classifier paradigm shrinks gradually in combination with numerals from three upwards. For Teribe, another Chibchan language, Quesada (2000:48) notes that the use of numeral classifiers decreases as numbers go up. These are just two more particular cases of a phenomenon that Greenberg had already ascertained for his sample.⁵ One may thus boldly propose the universal generalization of G5:

- G5. If, in a language, a numeral of numerical value X combines with numeral classifiers, then all numerals with numerical values lower than X also combine with numeral classifiers. Conversely, if a numeral of numerical value X does not require a numeral classifier, then numerals with numerical values higher than X do not require one, either.

From the combination of the two scalar generalizations G4 and G5, one may infer that it is the grammatical nature of the numerals that triggers the development of a class of numeral classifiers in a language (cf. Corbett 1978:367). Where numerals are (substantive) nouns, numeral classifiers cannot grow. Where numerals are adjectives, it depends on the precise nature of the adjective, i.e. on the extent that it resembles substantives. This manifests itself,

⁵ "It is particularly common for classifiers not to occur with higher units of the numerical system and their multiples, e.g. 10, 20, 60, 100, 300." (Greenberg 1972:6)

among other things, in the anaphoric construction that combines an adjective with an anaphoric (empty or dummy) head noun. Some languages like Latin require no special operation and no dummy representative of the anaphoric noun here; the adjective by itself suffices. One may assume it is implicitly substantivized, or one may say numerals, like adjectives, are more like substantival nouns in Latin than in certain other languages. The latter, like Chuvash, have special forms for numerals that are not in combination with a nominal. The case of German is interesting here. Anaphoric adjectives show declension, i.e. they agree with the anaphorically omitted noun, as in E8.

E8. Was für Eier willst du? – Gib mir mal kleine.

GERM ‘What eggs do you want? – Give me small ones.’

If the modifier is a numeral, then in Standard German it may function as an entire noun phrase in anaphora, as in E9, without any morphological modification.

E9. Wie viele Eier willst du? – Gib mir mal sechs.

GERM ‘How many eggs do you want? – Give me six.’

In colloquial German, numerals are a little less independent. There are two alternative constructions. The first employs one of two generic nouns – *Mann* ‘man > person’ and *Stück* ‘piece > thing’ – as a host to the numeral (more details in Lehmann 2000), as in E10.

E10. Wie viele Eier willst du? – Gib mir mal sechs Stück.

GERM ‘How many eggs do you want? – Give me six items.’

The other solution is to provide the numeral with an *-e* suffix, as in E11f (more examples in Drosdowski et al. 1984:278).

E11. Wie viele Eier willst du? – Gib mir mal sechse.

E12. Wann treffen wir uns? – a. Um acht Uhr.

GERM b. Um achte.

‘When shall we meet? – At eight o’clock.’

E12.a is standard and colloquial, E12.b is colloquial. At first sight, it might appear that the numeral is declined like the adjective in E8. That, however, is not so. In attributive construction, an adjective declines and shows agreement with its head noun (*kleine Eier* ‘small eggs’), while a numeral doesn’t (*sechs Eier*, **sechse Eier*). In independent use, adjectives and numerals decline differently, as may be seen in E13.

E13. a. alle neune ‘all nine (pins)’

GERM b. alle Kleinen ‘all small ones’

It appears that the *-e* suffix on numerals is a substantivizer. Thus, both of the colloquial German constructions avoid the self-standing adjective. The first provides something that may alternately be characterized as an almost full noun or an incipient numeral classifier, as the head of the nominal. The second substantivizes the adjective with a suffix which, in terms of categorical morphology, may be analyzed as a dummy head noun.

We may, thus, postulate the grammaticalization path sketched in S4.

S4. *Grammaticalization of numeral classifier*

generic noun > numeral classifier > numeral nominalizer

The two German strategies occupy the start and the end of S4. The typological relationship between Yucatec and German as regards numerals may be described thus: Given S4, the former possesses a grammatical form class occupying the center of that scale, while the latter has no corresponding form class and instead uses lexical and structural means positioned at the start and at the end of the scale.

There is no space here to analyze the function of numeral classifiers in languages where they also combine with other parts of speech, for instance with adjectives. One may, of course, ask: if the function of numeral classifiers is to serve as nominalizers, then why is there a whole paradigm of them; wouldn't one be sufficient? The answer is that this is, again, a question of grammaticalization. If nominalizers start out in anaphora as hyperonyms of their antecedent, then there is automatically such a paradigm at the beginning. In the course of grammaticalization, the paradigm shrinks. At the penultimate stage, it is usually a binary paradigm opposing animate and inanimate, or human and non-human, objects (as in Elf). Finally, the paradigm does end up as a sole morpheme that is obligatory with numerals and possibly other nominal constituents. At the beginning, variation in the classifier chosen may be used to convey some shade of meaning about the classified noun. At the end, the morpheme gives purely grammatical categorial information. Since grammaticalization is a goal-directed process, we may say that the final destination of a numeral classifier is the categorial operator.

5 Conclusion

Summarizing, we may conclude that at least in some languages that use numeral classifiers, these function as hosts to numerals that are not words. They are nominalizers in the broadest sense. They may not be in the narrow sense, because at least in some languages, the numeral classifier phrase has a somewhat different distribution from a substantival noun or nominal. In any case, this function is a purely structural one. And it suffices for understanding the use of numeral classifiers. No farther-reaching functional hypotheses about different concepts of number or of whatever is designated by a noun are necessary.

Needless to say, no general incompatibility between structural and functional analyses or superiority of the former over the latter is intended here.⁶ In a sense, a functional analysis is a higher-level analysis, since it presupposes a structural analysis. Linguistic phenomena are in a teleonomic hierarchy. A certain structural form, viewed in isolation, does not directly fulfill some cognitive or communicative function. Instead, it must first be understood as interdependent with its syntagmatic context and its paradigmatic alternatives. It may be totally determined by such language-internal factors. Only to the extent that it is not may and should we move higher up in the teleonomic hierarchy and seek a cognitive or communicative function for the phenomenon.

Abbreviations in interlinear glosses

1	first person	PL	plural
3	third person	POSS	possessive

⁶ Cf. Creissels 2006, ch. 0.3.3 on the spuriousness of the contrast 'functionalist vs. formalist'.

AN	animate	PRF	perfective
CL	classifier	REL	relational
INAN	inanimate	SG	singular
LONG	elongated object		

References

- Constenla Umaña, Adolfo & Elizondo Figueroa, Feliciano & Pereira Mora, Francisco 1998, *Curso básico de bribri*. San José, C.R.: Editorial Universidad de Costa Rica
- Corbett, Greville G. 1978, "Universals in the syntax of cardinal numbers." *Lingua* 46:355-368.
- Creissels, Denis 2006, *Syntaxe générale. Une introduction typologique*. 2 vols. Paris: Lavoisier.
- Drosdowski, Günther et al. 1984, *Duden. Grammatik der deutschen Gegenwartssprache*. Mannheim etc.: Dudenverlag (Der Duden in 10 Bänden, 4) (4th ed.).
- Greenberg, Joseph H. 1963, "Some universals of grammar with particular reference to the order of meaningful elements." Greenberg, Joseph H. (ed.), *Universals of language. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*. Cambridge, MA: MIT Press (2. ed. 1966) ; 58-90.
- Greenberg, Joseph H. 1969, "Some methods of dynamic comparison in linguistics." Puhvel, Jan (ed.), *Substance and structure of language. Lectures delivered before the Linguistic Institute of the Linguistic Society of America, University of California, Los Angeles, June 17 - August 12, 1966*. Berkeley & Los Angeles: University of California Press; 147-202.
- Greenberg, Joseph H. 1972, "Numeral classifiers and substantival number. Problems in the genesis of a linguistic type." *Working Papers on Language Universals* 9:1-39.
- Iturrioz Leza, José Luis 1986, "Semiotic levels. On the metalinguistic nature of grammatical meaning." *Función* 1:119-144.
- Lambrecht, Knud 1994, *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Lehmann, Christian 2000, "On the German numeral classifier system." Schaner-Wolles, Chris & Rennison, 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 2002, *Possession in Yucatec Maya*. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 10) (2nd revised ed.).
- Lévy-Bruhl, Lucien 1910, *Les fonctions mentales dans les sociétés inférieures*. Paris: F. Alcan.
- Lucy, John A. 1992. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press (Studies in the Social and Cultural Foundations of Language, 13).
- Lyons, John 1977, *Semantics*. 2 vols. Cambridge etc.: Cambridge University Press.
- Quesada, J. Diego 2000, *A grammar of Teribe*. München & Newcastle: LINCOM Europa.
- Quesada, Juan Diego 2007, *The Chibchan languages*. Cartago: Ed. Tecnológico de Costa Rica.
- Seiler, Hansjakob 1986, *Apprehension. Language, object, and order. Part III: The universal dimension of apprehension*. Tübingen: G. Narr (LUS, 1/III).