LINGUISTICS OF THE TIBETO-BURMAN AREA
VOLUME 13:1 — SPRING 1990

Outline of a Formal Syntax of Numerical Expressions, with Especial Reference to the Phenomenon of Numeral Classifiers

F. K. Lehman

Departments of Linguistics and Anthropology
The University of Illinois at Urbana-Champaign

0. I take as my starting point my earlier paper on so-called numeral classifiers (Lehman 1979), in which I was mainly concerned with the question whether or not classifiers are really, as is often claimed in the literature, a more or less language-and-culture-specific, comprehensive system of semantic classification[1] of the world and of experience, or rather something basically to do with the way the language handles the phenomenon of quantification. This was part of a more general line of work that is best understood by referring to my (1985b) paper, 'Cognition and Computation.' I shall not bother in the present paper to go into the substance of either of the previous papers. But it is useful, I think, to outline here the general conclusions of the earlier classifiers paper.

1. I argued that in reality classifiers did not point to a general scheme of classification. I argued this from two standpoints. First there was the straightforward substantive evidence (actually acknowledged quite generally, though just as generally brushed aside, e.g., by three of the most prominent proponents of the contrary conclusion, Becker (1975 — it is not altogether clear whether or not Becker, who takes great pains to show that a noun may be 'classified' in many different ways, should really be subject to this criticism, but he certainly does claim at least that the relation between a noun and whatever the classifier signifies is a membership/'kind of' relation), Denny (1979, 1985), and Placzek (1985a,b) that such purported comprehensive schemes are never anything like comprehensive and never rigid. That is, even for the fairly restricted subset of nouns referring to

[1] I am using the word 'classification' strictly in the technical sense of a taxonomic system of categories, of the kind in which claims that, for any item and for any such category, that the item 'is a kind of' whatever the category may be. In fact clan in this sense is not the general way in which items *cross-reference* with categories. There is a good discussion of this in Wierzbicka (1985): a knife, say, may be categorised as either a tool, a weapon, or an eating utensil, yet is not a *kind of* any of these in the sense that a human being, whatever other category it can be indexed with, is a *kind of* living being/animal (see now Keller and Lehman in press). In the looser sense, in which 'classification' need mean nothing more than the various categories a term may be indexed/cross-referenced with, of course the numeral classifiers indeed constitute a system (an open-ended one) of classification.

classes of concrete objects, the purported classification according to perceptually universal features of shape or, more generally, extension was always optional even when this was far and away the most conventional 'classification,' the default case. It is always possible to 'classify' the object class in enumeration by attending to some different feature, extensional, functional, mnemonic and so on. More importantly, the optional alternatives (in any case often a matter of deliberate and valued ingenuity on the part of adept native speakers and writers) for different classes of objects classified 'by default' with the same simply extensional classifier are rarely if ever coincident. For those idiolects, say, of Burmese in which railway trains (more commonly 'classified' as *asin:*/အစင်:) as well as rivers, pencils and the like are classified (I shall refrain from using inverted commas round this term hereafter, leaving it understood that I do not accept the idea that 'classification' is the real point of this phenomenon) as 'long, thin things' (*ahkyaung:*/အခြောင်:), the train, but hardly the pencil, is almost equally well classified as a vehicle/something to ride on (*asi:*/အစီး:). Furthermore, the so-called extensional features that motivate the theory of classifiers in question turn out not to be systematically extensional at all. And at the same time, the reason one class of objects can be classified by a given classifier may be very different from that motivating the use of the same classifier for another class of objects.

Consider, for example the common Siamese classifier *lêm*/เล่ม. It is the default classifier for books, maybe the default classifier for things like bullock carts, and for some speakers (but hardly all) also for candles, and in all these cases the motivation is far from extensional classification, or, indeed, anything we might call perceptual. It has to do simply with the idea that books are clamped together (between covers), carts are clamped together (one part to another), and candles have a wick clamped within the wax. The reasons that, in Burmese, things as diverse as stupas, folded manuscripts and fishing nets, as well as Buddhas (sacred personages) are alike classified as *ahsu*/အသု are not only wholly obscure but also necessarily inhomogeneous; for instance, the reason both Buddhas and stupas fall together here seems to be that stupas, according to Buddhist cosmology, are in principle relic-*bodies* (Pāli *kāya*) of the Buddha — one may speculate at will that folded manuscript books may be included because of the shaky supposition that they are thought primarily to be religious in content (the doctrine-body of the Lord Buddha?), but the fishing nets fall through the holes in the network of this account! The notion that this could be part of a scheme of classification rests upon the non-empirical claim that there must be a 'feature,' however abstract, however vacuous, that just this set of object-classes is felt by Burmans to share, even though all Burmans deny any such intuition; any appeal to a quite irretrievable etymological sense of this

classifier is, of course, totally irrelevant as well as unempirical. And what, pray, on this view, is to be made of the fact that the default classifier for academic subjects, problems, conferences, reasons, pastimes and 'similar' (in what sense?) abstract nouns are enumerated by means of the classifier *aya?*/အရပ် ('place' in the sense of an ordinal place in a list of, simply, 'categories')? I suppose it is possible to claim, again by vacuous generalisation, that what they share in common is the 'feature' of 'not falling under any other specific rubric,' and this is about as useful as saying, for the myriad concrete count nouns (in fact a non-fixed list of them) classifiable with *hku·*/ခု — the word means merely 'instance') that they share a classificatory-perceptual feature of being, alike, 'miscellaneous.'

But enough. This catalogue need not be further extended. The general point should be clear enough. The classifiers are actually some kind of device used in enumerative and (though not in Burmese — see Lehman 1979, 1985a) other noun-modifying expressions such as relative clauses, attributives and the like (on the demonstration that all these are indeed quantifiers of some kind, see my three papers already cited) to link the head noun with the quantifier — say, for the moment informally, an agreement device. The linkage is in the general case mnemonic only. Obviously there are some (not comprehensively many) default conventions for choosing the device, and it is equally unproblematic that for perceptually fairly uncomplicated non-abstract things, or rather the nouns naming them, the easiest things to seize upon for this purpose may be some spatial-extensional aspect of the thing in question. Nothing much more can be said in general.

Anyhow, following along in this general line of argument it is clear that I am claiming that noun classifiers/numeral classifiers (neither term is ideal, but both are fixed by tradition) are to be accounted for by a theory treating them as matters of **agreement**, much in the usual linguistic sense in which a subject, say, is said to agree with its verb in person and number, an adjective with its head noun in gender, number and/or case, and so on. As to such categories as person or number, it ought to be intuitively obvious that they cannot be appealed to as providing any scheme of classification for nouns in the intended sense, whilst any standard text book in linguistics (there is an especially good and simple treatment in Hock 1986) will demonstrate that gender is mainly morphosyntactic in function (agreement, again) arbitrary and not in general a notional scheme of classification. But, having regard specifically to classifiers, none of this can be taken as arguing against the idea that some 'aspect' of a noun class chosen for enumerative agreement marking may be so chosen as mnemonically appropriate, may even be defined, for quite culturally parochial reasons. For instance, consider the rather specifically Burmese-Buddhist ideas linking stupas and Buddhas,

mentioned above; it is sufficient for a classifier to 'remind' the hearer, or at least the clever speaker-writer (clever, literary, classifier choice is frequently the opposite of transparently motivated), of what it is that is being quantified. But the motivation may just as well be drawn from a pool of universals of human perception-cognition, as in the case of the extensional features used in the enumeration of nouns referring to objects with fairly simple and differentiated shapes. The 'feature' so chosen may be a part of the minimal lexical-semantic specification of the noun (and remember that there are more often than not a plurality of, say binary, features to use this way), e.g. 'something to ride on' for carts, or it may be part of one's encyclopaedic knowledge (knowledge-structure, in terms of a theory of cognition — cf. Keller & Lehman in press), e.g., 'something to ride on' in the case of horses. About these things I have nothing further to say in the present paper.

I have argued in my earlier classifiers paper (1979) that the theory I am now elaborating extends more naturally and homogeneously than its opponent (mainly semantic) theory of classifiers to the case of the so-called 'measure words.' I mean, of course, the words used along with quantifiers when the quantifier expression is headed by a non-count, e.g., a mass, noun: 'water, one cupful' and the like. I wrote there that the nouns being quantified, e.g., counted, were underlyingly compounds headed by a simulacrum of the measure word itself, a head systematically deleted/left empty in the sense of empty-category theory, under a quite general and independently motivated rule that omits heads of noun phrases in case (with certain limited exceptions[2]) they are 'self-classifiers,' viz., duplicated by the classifier. I shall not further elaborate that point in this paper, but I think it eminently worthwhile observing that there is evidence I inadvertently forgot about previously supporting this view. For in Burmese (Okell 1969:215) when one is counting 'rounded numbers' (order-or-magnitude numbers) of count nouns (where the classifier is the integral power of the base: tens, hundreds, thousands, etc.), one has always the option of using in the position of the noun a compound whose second (head) member is the very

---

[2] It seems to me now that the difference between true self-classifying nouns, that 'delete' in the presence of the corresponding classifier (standard unit-of-time and unit of spatial measure words, for instance) and those that do not must be that the former do not allow the word to be compounded with a following head noun of associated meaning, while the latter do. So, — *tala* တ္လ 'one month,' but *ein* အိမ် (house) *taein* တဲအိမ် 'one home;' cf. *eini-hsaun)* taya အိမ်(ဆောင်) တရာ (100 house-structures)/ *ein lahsaun* အိမ် တဆောင် 'one house.' *la·* in the sense of 'month' simply fails to compound this way: we get only *la·taya* လ တရာ '100 months.' Of course, in other than discourse-initial or topic-initial position, heads of enumerated NPs are more nearly free to be empty regardless of the distinction between self-classifiers and others, e.g. in answers to 'how much'/'how many' type questions.