

Toward a comprehensive theory of noun categorization, with special reference to Thai

Christopher I. Beckwith
Indiana University

1. Introduction

In order to review the current standard typological model of noun categorization (Dixon, 1982, 1986), or to develop a new theory, it is necessary to define the limits of the category. This paper focuses on class nouns (henceforth CLNs), a subcategory of compounds.

Compound nouns may be constructed in several ways, such as synonym or polar compounds, additive compounds, verb-noun compounds, and so on. CLNs are compound nouns formed from two elements in which one element---normally a noun or nominal stem---represents a class, and one element---frequently, but by no means always, a nominal---functions as a qualifier; in any case, the head of the compound must be a noun. For example, consider *way* in English *railway*, *byway*, *passageway*, *leeway*, *freeway*, *expressway*, *highway*, *throughway*, *runway*, *tramway*, *subway*, and so on, the class term (henceforth CLT) *-way* is the head of the compound and of the class WAY.¹ The sense of class is perhaps easier to understand if the English CLT *woman* in the CLNs *businesswoman*, *saleswoman*, *policewoman*, and so on, which belong to the class WOMAN, is compared to the English suffix *-ess* in *waitress*, *stewardess*, *actress*, *songstress*, and *mistress*, which belong to a FEMALE HUMAN or WOMAN class. In other words, the CLTs of English CLNs function morphologically and semantically not only as taxonomic class heads but also like English gender suffixes, with the difference that the number and kinds of classes that may occur in CLNs are practically unlimited.

It has been said of Thai that the categorization of CLTs and classifiers (henceforth CLFs) is "not entirely coherent" (DeLancey 1986:441), and that semantically the two types "overlap to a considerable degree" and "are clearly distinct only

as syntactic categories" (DeLancey 1986:442). This paper, which is based partly on theoretical results deriving from a study of noun categorization in Tibetan (Beckwith 1994), addresses the issue of these apparently fuzzy boundaries by investigating the internal structure of Thai class nouns, which appear to straddle the line between grammaticized and non-grammaticized (or lexical) noun categorization.²

2. Analysis

Most Thai CLNs have the surface morphological structure N + Attribute, which is normal for a left-headedness language. Classifier agreement is based, as in other classifier languages, on salient characteristics of the real-world referents of the nouns classified.³ The classifiers assigned will thus be the same for taxonomically subordinate-level nouns and for the basic level head of those same nouns. For example, consider the examples in (1), nouns with the CLT *ŋuu* 'snake'.

- (1)a *ŋuukhiaw* *nỳŋ* *tua*
 greensnake 1 CLF[animal, animal-shaped]
 'one greensnake'
 (*khiaw* 'green')

- (1)b *ŋuulyam* *nỳŋ* *tua*
 python 1 CLF[animal, animal-shaped]
 'one python'
 (*lyam* 'python')

- (1)c *ŋuuhàw* *nỳŋ* *tua*
 cobra 1 CLF[animal, animal-shaped]
 'one cobra'
 (*hàw* 'to bark')

The type of CLN in (1) is clearly taxonomic in the classic sense of Rosch (1977). Since *ŋuu* is the internal lexical head of such CLNs, I will refer to it as a 'taxonomic CLT', and to such nouns as 'taxonomic CLNs', following Iguchi (1994).⁴

Some CLNs in Thai, including several mentioned by DeLancey (1986:438-442), clearly differ from these taxonomic CLNs in their semantic structure. Consider the examples in (2), with the CLT *duaŋ* 'CLF for round shining things'.

- (2)a *duaŋcan* *nyŋ* *duaŋ*
 'moon' 1 CLF[round & shining]
 'one moon'
 (*can* 'moon')
- (2)b *duaŋtaa* *nyŋ* *duaŋ*
 'eye' 1 CLF[round & shining]
 'one eye'
 (*taa* 'eye')
- (2)c *duaŋfaj* *nyŋ* *duaŋ*
 'light' 1 CLF[round & shining]
 'one (round) light'
 (*faj* 'fire')

The examples in (2), and all other CLNs formed with *duaŋ*, take *duaŋ* as their CLF. Since the lexical heads of these CLNs are in each example clearly the second term, which is morphologically the attribute, *duaŋ* functions here not as a taxonomic CLT but as a classifying CLT. Since CLTs, unlike CLFs, are bound forms, I will refer to this kind of CLT as a 'gender CLT', and this kind of noun as a 'gender CLN', following Iguchi (1994).

Since the second term in such examples functions as internal lexical head of the CLN and takes the CLF agreement both internally (with its classifying CLT, *duaŋ*) and externally in full specifier phrases (with the classifier *duaŋ*), there are actually two CLTs in such nouns, the first term (*duaŋ*) being a gender CLT and the second a taxonomic CLT. Since the salient characteristics of a noun---or, in the case of a CLN, of its identifiable lexical head---determine CLF assignment, the head of the CLN in each of the examples in (2) is the second term. Although, taxonomically speaking, the second terms are

all basic level nouns and might be expected to have considerable variation in classifier assignment, that does not happen with these particular basic level nouns because they have already been classed together by their common morphological head, the gender CLT *duaŋ*, which as a CLF classifies each taxonomic CLT within each noun in (2). Thus, rather than variation there is instead full gender concord in the specifier phrase, as shown in the examples in (2).

Let us turn now to CLNs formed with the honorific noun *phrá?* 'lord, god, priest, Buddha image'. This noun itself takes the CLF *ʔoŋ* for honorific beings, but consider the CLNs in (3) and the CLFs they are normally assigned.

- | | | | |
|------|--|-----------------|-------------------------------------|
| (3)a | <i>phrá?câw</i>
god/ruler/king
'one god/ruler/king' | <i>nỳŋ</i>
1 | <i>ʔoŋ</i>
CLF[body (honorific)] |
| (3)b | <i>phrá?can</i>
moon
'one moon' | <i>nỳŋ</i>
1 | <i>duaŋ</i>
CLF[round & shining] |
| (3)c | <i>phrá?aathid</i>
sun
'one sun' | <i>nỳŋ</i>
1 | <i>duaŋ</i>
CLF[round & shining] |
| (3)d | <i>phrá?râadchawaŋ</i>
royal palace
'one royal.palace' | <i>nỳŋ</i>
1 | <i>lǎŋ</i>
CLF[building] |

Although example (3)a might lead one to suspect the honorific classifier *ʔoŋ* agrees with the honorific CLT *phrá?*, in fact both *ʔoŋ* and *phrá?* are in agreement with the second term, *câw*, which is the lexical head of the CLN, as in examples (3)b-d. It is manifest that in each example in (3) the CLF assignment is based not on the first term, *phrá?*, but on the second term. Thus, as in example (2), the morphological head is not the lexical head.

Consider now the examples in (4), CLNs that include

the CLT *máj* 'wood, stick'.

- (4)a *májdðɔg* *nỳŋ* *tôn*
 flowering plant 1 CLF[stalk]
 'one flowering plant (a plant known to bear flowers)'
- (4)b *májkhiid* *nỳŋ* *kâan*
 matchstick 1 CLF[stick-like things]
 'one match'
- (4)c *májkhiid* *nỳŋ* *klàg*
 matchstick 1 CLF[box]
 'one box of matches'
- (4)d *dðɔgmáj* *nỳŋ* *dðɔg*
 flower 1 CLF[flower]
 'one (specific) flower'
- (4)e *bajmáj* *nỳŋ* *baj*
 leaf 1 CLF[leaf]
 'one leaf'
- (4)f *tôn máj* *nỳŋ* *tôn*
 tree 1 CLF[stalk]
 'one tree'

Examples (4)a-c are regular taxonomic CLNs, where the morphological head is the lexical head. Examples (4)d-f are clearly different. From a taxonomic viewpoint, *máj*, the second term, represents a higher taxonomic level than the first term; *máj* is thus undoubtedly a taxonomic CLT, not a gender CLT. It does not classify the first term in these CLNs, it tells where the first term belongs in a taxonomic hierarchy. Although the first terms in examples (4)d-f are all CLFs, they are also nominal CLT stems used to form subordinate level taxonomic nouns. Therefore, here they function not as gender CLTs but as taxonomic CLTs---and the lexical heads---of these CLNs. Accordingly they take the classifier agreement. Since they do

also exist as separate classifiers, like *duaŋ* they assign themselves as classifiers of their CLNs.

There are, however, numerous cases where the CLN has been lexicalized at the word level and analyzing the constituents of the CLN will not work. Consider the examples in (5).

- (5)a *khryaŋdontrii* *ny'ŋ* *khryaŋ*
 musical.instrument 1 CLF[instrument]
 'one musical instrument'
 (*dontrii* 'music')
- (5)b *khryaŋbin* *ny'ŋ* *lam*
 flying.machine 1 CLF[cylindrical]
 'one airplane'
- (5)c *kluaŋmaj* *ny'ŋ* *dɔŋg*
 orchid 1 CLF[flower]
 'one orchid'

Example (5)a is a typical taxonomic CLN where the CLF appears to agree with itself as the lexical head. Example (5)b is not at all transparent. While it is morphologically identifiable as a CLN, it is not lexically analyzable with respect to its CLF assignment, which is probably due to another word for airplane, *ryabin* 'flying boat', because words with *rya* 'boat' as CLT head take *lam* as classifier; words for airplane thus take the same classifier.⁵ Unless one knows this connection, however, the reason for the assignment of the classifier *lam* to *khryaŋbin* is not clear. Even more opaque is example (5)c, which is morphologically a CLN built on the taxonomic CLT *kluaŋ* 'banana', with the qualifier *maj* 'wood'. The meaning 'orchid' for this CLN is not derivable from the constituents, which are therefore not analyzable. One simply must know *what it means (i.e., what it refers to) in order to assign a CLF*.

CLF agreement is thus not with nouns or CLTs themselves, but with selected salient characteristics of their real-world referents,⁶ whether or not the agreement is mappable onto the lexical head of a given CLN. This accounts for the

often considerable variation in CLF usage that has been much discussed in the literature (Becker 1986, Erbaugh 1986, Lakoff 1986, 1987, Lehman 1990, Tai & Wang 1991). This variation is generally cognitively motivated. In a recent study of Japanese classifier selection, for example, adult speakers tended to agree on classifier choices for the same noun when provided with specific situations in which the real word object would be perceived.⁷ In other words, CLF assignment and CLF categories are essentially independent of linguistic form.

3. Conclusion

With respect to Thai class nouns, then, it seems that the semantic categories of folk taxonomy, a non-grammaticalized system of noun categorization, do not extend into or overlap with the categories of grammaticized noun categorization systems, although the two types of categorization are tied together both by morphology and by the fact that each categorizes the other. It would seem that grammaticized systems of noun categorization are motivated by specific features of the particular nominal systems in which the noun categorization appears. Although the *number* of semantic categories within a given noun categorization system is constrained to some extent by pragmatic considerations, there is no formal restriction on the *kinds* of semantic categories that may be found in a given language as a whole. Thus there is, for example, no formal constraint on the development of animacy-type noun categorization in a language which has mostly classifier-type noun categorization. However, the different types tend to occur in different areas of the language, as in Thai, where natural gender occurs in the pronominal system rather than in the classifier system. Dixon's model (1982, 1986) proposes a one-to-one correspondence between language types and types of grammaticized noun categorization. It appears, instead, that there is a close correspondence between grammatical function (or grammatical category) and type of grammaticized noun categorization system.

Paradoxically, however, assignment of specific classifi-

ers to specific nouns is unconnected to any grammatical function. Classifiers refer directly to the real world features of the thing named by the noun, and to the situation or manner in which the thing is perceived; classifier categories, too, are based on real-world referents. It is this direct accessing of the real world, cutting through all other layers of language, which makes classifier categorization so intriguing. If ethologists and psychologists are able to determine through experiment what the categorization systems of various primate species are like, and then compare them to known human classifier systems, it is quite possible that something valuable might be learned about the primordial cognitive system of early humans.

Notes

- 1 The variety of qualifier elements can be surprising. In this by no means exhaustive list of examples built on *way*, the qualifier terms include nouns, a verb, adjectives, prepositions, and a prefix.
2. Other studies of Thai compounding have focused on syntactic models of compound formation (Warotamasikkhadit 1970, Placzek 1978). Keiko Iguchi, one of my graduate students, is writing a thesis dealing with class nouns' internal structure and semantic classes.
- 3 Typically physical characteristics or 'kinesthetic image schemas' (Johnson 1987, Lakoff 1987).
4. My usage of the term 'taxonomic' in several earlier papers is unfortunate. I hope the present paper will rectify it to some extent.
5. One Thai colleague suggested I find out which classifier Karen speakers use for the dragonfly, a common airplane-shaped animal, since Thai classifies all animals with *tua*, regardless of the shape. In my subsequent visit with Thai friends to a Karen village west of Chiang Mai, we saw some

dragonflies and asked a speaker of Sgaw Karen how they are classified. He replied with the classifier for flat things, /plə/; a speaker of Plang Karen, in which dialect it is pronounced /plo/, confirmed this. (Tones are not marked.)

6. This has often been stated (for example, Placzek 1978:8), but I owe the proper appreciation of the insight to Megumi Yui, one of my graduate students.

7. Unpublished research paper by Megumi Yui.

References

- Becker, A.L. 1986. The figure a classifier makes: describing a particular Burmese classifier. *Noun classes and categorization*, ed. by Colette Craig, 327-343. Amsterdam: John Benjamins.
- Beckwith, Christopher I. 1994. Categorization in Tibetan. Unpublished paper presented in Madison before the American Oriental Society.
- Craig, Colette (ed.) 1986. *Noun classes and categorization*. Amsterdam: John Benjamins.
- DeLancey, Scott. 1986. Toward a history of Tai classifier systems. *Noun classes and categorization*, ed. by Colette Craig, 437-452. Amsterdam: John Benjamins.
- Dixon, R.M.W. 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton.
- Dixon, R.M.W. 1986. Noun classes and noun classification in typological perspective. *Noun classes and categorization*, ed. by Colette Craig, 105-112. Amsterdam: John Benjamins..
- Erbaugh, Mary S. 1986. Taking stock: the development of Chinese noun classifiers historically and in young children. *Noun classes and categorization*, ed. by Colette Craig, 399-436. Amsterdam: John Benjamins.
- Iguchi, Keiko. 1994. Re-examination of the structure of nominal compounds: two basic types of class terms in

- class nouns. Unpublished research paper.
- Johnson, Mark. 1987. *The body in the mind: the bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Lakoff, George. 1986. Classifiers as a reflection of mind. *Noun classes and categorization*, ed. by Colette Craig, 13-5. Amsterdam: John Benjamins.
- Lakoff, George. 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.
- Lehman, F.K. 1990. Outline of a formal syntax of numerical expressions, with special reference to the phenomenn of numeral classifiers. *Linguistics of the Tibeto-Burman area* 13.1: 89-120.
- Placzek, James Anthony. 1978. Classifiers in standard Thai: a study of semantic relations between headwords and classifiers. Unpublished M.A. thesis. Vancouver: University of British Columbia.
- Rosch, Eleanor. 1977. Classification of real-world objects: origin and representation in cognition. *Thinking: readings in cognitive scienc*, ed. by P.N. Johnson-Laird and P.C. Willson, 212-222. Cambridge: Cambridge University Press.
- Tai, James and Liangqing Wang. 1991. A semantic study of the classifier tiao 條. *Journal of the Chinese language teachers' association* 25.1: 35-56.
- Warotamasikkkhadit, Udom. 1970. *Thai syntax*. The Hague: Mouton.