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. 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. 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 ɨuu 'snake'.

(1a) ɨuukhiaw ny השאלה tua
greensnake l CLF[animal, animal-shaped]
'one greensnake'
(khiaw 'green')

(1b) ɨuulyam ny שאלה tua
python l CLF[animal, animal-shaped]
'one python'
(lyam 'python')

(1c) ɨuhàw ny שאלה tua
cobra l 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 ɨuu 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).
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 \textit{duaŋ} 'CLF for round shining things'.

\begin{enumerate}
  \item \textit{duaŋcan} \textit{nyŋ} \textit{duaŋ}
  \begin{description}
    \item['moon'] \begin{description}
      \item[1] CLF[round & shining]
      \item['one moon'] (\textit{can} 'moon')
    \end{description}
  \end{description}

  \item \textit{duaŋtaa} \textit{nyŋ} \textit{duaŋ}
  \begin{description}
    \item['eye'] \begin{description}
      \item[1] CLF[round & shining]
      \item['one eye'] (\textit{taa} 'eye')
    \end{description}
  \end{description}

  \item \textit{duaŋfaj} \textit{nyŋ} \textit{duaŋ}
  \begin{description}
    \item['light'] \begin{description}
      \item[1] CLF[round & shining]
      \item['one (round) light'] (\textit{faj} 'fire')
    \end{description}
  \end{description}
\end{enumerate}

The examples in (2), and all other CLNs formed with \textit{duaŋ}, take \textit{duaŋ} as their CLF. Since the lexical heads of these CLNs are in each example clearly the second term, which is morphologically the attribute, \textit{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, \textit{duaŋ}) and externally in full specifier phrases (with the classifier \textit{duaŋ}), there are actually two CLTs in such nouns, the first term (\textit{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 $\text{duaŋ}$, 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 $\text{phráʔ}$ 'lord, god, priest, Buddha image'. This noun itself takes the CLF $\text{ŋoŋ}$ for honorific beings, but consider the CLNs in (3) and the CLFs they are normally assigned.

(3)a $\text{phráʔcāw}$

- god/ruler/king
- 'one god/ruler/king'

   $\begin{align*}
   &\text{nyṛŋ} \quad ?\text{oŋ} \\
   &1 \quad \text{CLF[body (honorific)]}
   \end{align*}$

(3)b $\text{phráʔcān}$

- moon
- 'one moon'

   $\begin{align*}
   &\text{nyṛŋ} \quad ?\text{oŋ} \\
   &1 \quad \text{CLF[round & shining]}
   \end{align*}$

(3)c $\text{phráʔaathid}$

- sun
- 'one sun'

   $\begin{align*}
   &\text{nyṛŋ} \quad ?\text{oŋ} \\
   &1 \quad \text{CLF[round & shining]}
   \end{align*}$

(3)d $\text{phráʔrāadchawar}$

- royal palace
- 'one royal palace'

   $\begin{align*}
   &\text{nyṛŋ} \quad \text{lāŋ} \\
   &1 \quad \text{CLF[building]}
   \end{align*}$

Although example (3)a might lead one to suspect the honorific classifier $\text{oŋ}$ agrees with the honorific CLT $\text{phráʔ}$, in fact both $\text{oŋ}$ and $\text{phráʔ}$ are in agreement with the second term, $\text{cāw}$, 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, $\text{phráʔ}$, 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