Autolexical solutions to the problem of
“parts of speech”
in Southeast Asian languages

Eric Schiller
University of Chicago

0. Introduction

The syntax of Southeast Asian languages can appear quite
difficult when observed from an ethnocentric perspective, and
this complexity is often compounded when one applies a theoretical
perspective which forces lexical items into fixed syntactic
categories determined by putative universal considerations. This
paper uses the notion of syntactic polysemy or syntactic flexibility
(Ratliff 1991) to discuss the nature of word classes in Khmer
and a few other Southeast Asian languages. Specifically, I will
concentrate on several words which appear in a wide variety of
syntactic contexts, not merely nouns and verbs, but also modals,
adverbs, prepositions, and classifiers.

1. Clarifying what we mean by “parts of speech”

By using the Autolexical technique of separating syntactic
considerations from semantic considerations (Sadock 1991), and
having a distinct inventory of word classes (or categories) at
each level, the often confusing problem of determining “parts-
of-speech” is made much easier. Let’s start with categories
which have traditionally been at least somewhat controversial:
“relator-nouns”, “classifiers”, and “coverbs”. These terms have
a tendency to be defined in purely language-specific terms, usually
by positional factors since morphology is not much help in
mainland Southeast Asian languages. For pedagogical purposes
it is often useful to determine lexical categories simply on the
basis of co-occurrence restrictions. As we shall see, however,
this approach leads to analytical problems in the languages which
allow widespread deletion, as is the case with most of the isolating
languages of Southeast Asia.

Consider the Khmer word presented in (1).
1) ទុំ<knoŋ> /knoŋ/  ‘inside’

Headley (1977) glosses it as a predicative with the meanings in, inside, within, during and gives the following examples (inter alia):

a. ប្រភេទ
   /knoŋ chnam nih/
   in year this
   ‘during this year’

b. ទុំស្រក
   /knoŋ srok/
   ‘in the country’

c. ទុំកនុង
   /coul knoŋ/
   ‘Go inside!’

d. ស្តាកអាចឈរាចារះពន្ធ
   /dak siawphau knoŋ preʔop/
   ‘to put books in a box’

Even these few examples serve to show the diversity of the uses of /knoŋ/. The semantic range is not very broad, with all senses having to do with a notion of being located inside of something, but the syntax is less clear. For the most part, it seems to be a preposition but in (1.c) either a nominal or adverbial analysis seems more appropriate. One might want to claim that (1.c) is parallel to the English translation, where one can claim that ‘inside’ is a preposition with a deleted object. Yet analyzing knoŋ as a preposition (or, as Jacob 1968 does, a pre-nominal particle) requires justification, because it is most commonly found
following the lexeme \textit{nau} in the following structure (1.e):

\begin{enumerate}
\item \textit{kruu nau kno\xspace n v\text{\text{"o}}\text{\text{"a}}t/}
\begin{quote}
\textit{be-in in temple}
\textit{‘in the temple’}
\end{quote}
\end{enumerate}

By examining this phrase alone we cannot discover the syntactic structure. And imagine the trouble we will get into if we start defining our syntactic structures in terms of the presence or absence of \textit{kno\xspace n}. A fuller sentence (1.f) does not clarify matters:

\begin{enumerate}
\item \textit{kruu nau kno\xspace n v\text{\text{"o}}\text{\text{"a}}t/}
\begin{quote}
\textit{teacher be-in in temple}
\textit{‘The teacher is in the temple’}
\end{quote}
\end{enumerate}

Although it appears that \textit{nau} is the main verb, and that \textit{kno\xspace n} heads a prepositional phrase, the addition of another word muddies the waters (1.g):

\begin{enumerate}
\item \textit{kruu deek nau kno\xspace n v\text{\text{"o}}\text{\text{"a}}t/}
\begin{quote}
\textit{teacher sleep be-in in temple}
\textit{‘The teacher sleeps in the temple’}
\end{quote}
\end{enumerate}

And as further support we can front the constituent headed by \textit{/nau/} (1.h):
h.  ដែលម្របេក្នុងប្រព្យុស

/ nəu khnəŋ vôat kruu deec /

in inside temple teacher sleep
‘In the temple, the teacher sleeps’

For this reason, (Jacob 1968) treated nəu-khəŋ as a compound preposition in such cases. (Note, by the way, that written Khmer does not separate words.) I present these facts as an introduction to the problem at hand, namely the identification of syntactic categories or parts of speech. I will not go into any further analysis here, but note that nəu was the subject of Schiller 1984, and that the analysis of khəŋ as a preposition is justified in Schiller 1991c.

Let us now turn to a case involving a classifier. Although classifiers are not as robust a category in Khmer as they are in, say, Thai, there are still many cases where they are obligatory. For counting ordinary people, the word used is nəʔ, which follows the numeral as shown in (2):

2)  ម៉ាី kWh /nəʔ/ ‘person’

សេរី < nəkk> /nəʔ/ classifier for common people’

a.  នំឈឺពួកក្រពេះនាង

/ mənuh pii nəʔ /

person 2 CL
‘two persons’

But the word is also used as a pronoun (usually second person?) as shown in (2.b) and is also the head noun in compounds (2.c,d):