

# Flying 'In' and 'Out' in Khmer and Thai

Chhany Sak-Humphry

Kitima Indrambarya\*

Stanley Starosta

University of Hawai'i at Manoa

## 1. INTRODUCTION

By definition, a word is a minimal free form. However, to identify forms as words in isolating languages which do not have inflectional morphology is not an easy task. The Saussurean sign regards sound and meaning as the basic components of a word. Semantic difference is thus a basic criterion for differentiating words of the same shape and pronunciation (Panupong, 1978, p. 217). In the case of homophones, forms like English *run* (n) and *run* (v), however, which are similar in both pronunciation and meaning, distribution must play a significant role in distinguishing lexical entries from each other. Thus, distribution is treated as one of the components of words in Lexicase theory (Starosta, 1988, in press). Panupong (1978, p. 221) summarizes the advantage of assuming such homophones as follows:

- a) To accept this analysis, it is necessary that we regard each word as having only one function. The merit lies in our being immediately able to decide to what class a word in each sentence belongs,...
- b) Not having to set up classes for polyfunctional words is much more economical.
- c) There would be no problem in labelling the words with more than one function.

This paper attempts to establish the word classes of the following forms carrying the meaning 'enter; in' and 'leave; out,' when they are preceded by main verbs glossed as 'to fly' in two Southeast Asian languages, namely Khmer and Thai. They are *caul* 'enter' and *ceñ* 'leave' in Khmer and *khâw* 'enter' and *ʔòk* 'leave' in Thai.<sup>1</sup> The following sentences containing the forms to be tested in the paper:

---

\* Now at Kasetsart University, Bangkok.

<sup>1</sup> Pending the proper assignment of word classes to these forms, we will gloss them in the first part of the paper as 'enter' and 'leave.'

**Khmer**

- (1) a. caap        haoe        *caul*        knong        trung  
          bird       fly        enter       inside       cage  
          'The bird flew into the cage.'
- (2) a. caap        haoe        *ceñ*        pii        trung  
          bird       fly        leave       from       cage  
          'The bird flew out of the cage.'

**Thai**

- (3) a. nók        bin        *khâw*        pay        nay        kron  
          bird       fly        enter       go        inside       cage  
          'The bird flew into the cage.'
- (4) a. nók        bin        ?ðək        càak        kron  
          bird       fly        leave       from       cage  
          'The bird flew out of the cage.'

In the lexicase grammatical framework, there are eight and only eight basic syntactic word classes: V (verb), N (noun), Adj (adjective), Det (determiner), Adv (adverb), P (preposition or postposition), Cnjc (conjunction), or Sprt (sentence particle) (Starosta, 1988, p. 51). In principle, based on their distributions and meanings, the forms *caul*, *ceñ*, *khâw* and ?ðək in the preceding sentences could be thought to be (1) prepositions, (2) adverbs, or (3) verbs in Khmer and Thai, as indicated by the following alternative glosses:<sup>2</sup>

**Khmer**

- (1) a. caap        haoe        *caul*        knong        trung  
          bird       fly        enter       inside       cage  
          'The bird flew into the inside of the cage.'  
          'The bird flew in to the inside of the cage.'  
          'The bird flew to enter the inside of the cage.'
- (2) a. caap haoe *ceñ*    pii        trung  
          bird fly    leave    from    cage  
          'The bird flew out of the cage.'  
          'The bird flew out from the cage.'  
          'The bird flew to leave from the cage.'

<sup>2</sup> Other parts of speech are ruled out for the following reasons: (1) adjectives and determiners may be dependents of nouns, not of verbs such as *bin* and *haoe* 'to fly,' (2) conjunctions would coordinate words of the same category, not verbs and nouns, and (3) sentence particles occur sentence-finally unless followed by another sentence particle.

## Thai

- (3) a.    *nók*    *bin*    *khâw*    *pay*    *nay*    *kron*  
          bird   fly       enter   go       inside   cage  
          'The bird flew to the inside of the cage.'  
          'The bird flew in to the inside of the cage.'  
          'The bird flew to enter the inside of the cage.'
- (4) a.    *nók*    *bin*    *?ðək*    *càak*    *kron*  
          bird   fly       leave   from   cage  
          'The bird flew from the cage.'  
          'The bird flew out from the cage.'  
          'The bird flew to leave the cage.'

The paper is divided into four sections. This first section presents an introduction. The second section discusses the criteria used to identify prepositions, verbs, and adverbs in this paper. The third section discusses how each test applies to the words in question. The fourth section presents the conclusion.

## 2. TESTING CRITERIA

The tests to be used for sentences in both Khmer and Thai are: (1) Stranding,<sup>3</sup> (2) Joint topicalization of both the word in question and the following noun phrase, referred to here as PP topicalization, (3) Choice of negation expressions, and (4) Root predicate with the choice of negation expressions used for verbs. These tests are built upon the tests used in Indrambarya (1995) and are extended here for testing other languages.<sup>4</sup>

### 2.1. Stranding

Prepositions and relator nouns cannot be stranded<sup>5</sup> while verbs and adverbs can (Indrambarya, 1995). From the point of view of lexicase theory, prepositions cannot be stranded because a prepositional phrase is an exocentric construction, that is, a construction whose head takes one or more structurally obligatory dependents, while a relator noun can't be stranded because there is no mechanism for passing

<sup>3</sup>This test is referred to as the "stranding test" rather than as the "topicalization of the following NP test" as in Indrambarya (1995), because words following the form in question could be a prepositional phrase as well as a noun phrase.

<sup>4</sup>Examples of uncontroversial words for verb, adverb, prepositions, and nouns in Khmer are discussed in Sak-Humphry (1996).

<sup>5</sup>A noun may be differentiated from a preposition in that only the former allows a determiner as its dependent (Savetamalya 1989, Indrambarya 1994, 1995).

indices across nouns to link the relator noun with its topicalized dependent in a higher clause.

2.2. PP Topicalization

A preposition and its following NP and a relator noun with its dependent are grammatical units and may be topicalized together. However, a verb and its dependent and an adverb with a following noun cannot be topicalized together (Indrambarya, 1995). This latter claim does not follow from any universal grammatical principles, but it is a consistent generalization that can be made by observing the syntactic behavior of clear cases of prepositions, relator nouns, verbs, and adverbs.

2.3. Choice of negation word

Indrambarya (1995) has observed that different word classes are negated by different negation words when serving as predicates More specifically, in Thai the negation word *mâychây* negates NP and PP predicates, while *mây* and *mâydây* negate only verbs and (non-predicate) adverbs. Thus, if a form in question may be negated in the position in which it occurs, it is possible to separate verbs and adverbs from prepositions and nouns. The choice of a negation word only serves as a one-way test, though. That is, if a form can occur with one of the negation words, we have a clue as to its part of speech. However, if the form cannot be negated at all, the test cannot tell us what part of speech the form is . This situation arises because (1) a telic verb does not allow an embedded verb to be negated at all (, 1993a, p. 58), while an adverb does not necessarily occur with any negation word (cf. Indrambarya, 1995).

In Khmer, verbs can be negated with the negation expressions *mîn*, *pom*, or *?at... (tee)* ‘not,’ while prepositions, verbs and nouns may be negated with the common negation expression *mîn mēen... tee*. Adverbs, on the other hand, cannot be negated at all. Negation in Khmer is summarized in Table 1.

Table 1. *Negation Pattern in Khmer*

	V	P	N
<i>mîn... tee</i> ‘not’	+	–	–
<i>pom... tee</i> ‘not’	+	–	–
<i>?at... tee</i> ‘not’	+	–	–
<i>mîn mēen... tee</i>	+	+	+
‘not true’			