A Categorization of thùuk in Thai: Lexicase Analysis

Saranya Savetamalya

Chiang Mai University

1. INTRODUCTION

This paper presents a syntactic analysis of the word thùuk in Thai using a lexicase framework, a type of lexicalist dependency grammar (Starosta & Hashimoto, 1984; Starosta, 1988). Within the lexicase analysis, grammatical relations are characterized by case relations, case forms, and macro role. Case relations include Agent, Patient, Correspondence, Locus, and Means. Case forms include Nominative and Accusative. Macro role is represented by Actor.

From this study, there are several different homophonous lexical entries for thùuk. These can be defined by both their semantics and syntactic distribution. thùuk can occur either as a verb or as a “particle” adverb which can be seen in the following discussion.

2. SYNTACTIC DISTRIBUTION OF thùuk

2.1. thùuk as a Transitive Verb

The first syntactic distribution of thùuk is that of a transitive verb. Here it obligatorily requires a following noun, bearing a Patient case relation as its dependent sister. There are two transitive forms of thùuk: thùuk₁ meaning ‘hit’, and thùuk₂ meaning ‘touch’. Examples:

1. lûukbôn thûuk₁ dèk
   ball      hit    child
   AGT       [+trns]  PAT
   ‘The ball hit the child.’

   In (1) thùuk₁ is a transitive verb. It requires an Agent, lûukbôn and a Patient, dèk. The reason that dèk is treated as a Patient instead of carrying some other case relation is because, according to the Patient centrality hypothesis (Starosta 1982, p. 10), a transitive verb requires both an Agent and a Patient (as in 1). If sentence (1) had no Agent, the verb would be intransitive, and would result in the ungrammatical sentence (2):

2. * lûukbôn thûuk₁
   ball      hit
   PAT       [-trns]
Therefore, dëk, which is required by the verb thùuk₁, is treated as a complement noun. As a noun, it has to carry a case relation. Which case is it? Consider the following sentence in which thùuk₁ allows one further argument:

3. lùukbôn thùuk₁ dëk thiį¹ hūa
   l[AGT] | l+trns  | l[PAT] | LOC
   l+actrl | l?[+AGT]l | l-actrl
   l?[+PAT]l
   ball  hit   child  at head
   'The ball hit the child on the head.'

hūa carries a Locus case relation, marked by the relator noun thiį. It is a locative complement rather than a Locus adjunct based on the assumption that a complement cannot be preposed, whereas an adjunct can. This results in the ungrammaticality of (3a), in which thiį hūa is a complement, and also the grammaticality of (3b) in which thiį roonrían is an adjunct:

3a. * thiį¹ hūa lùukbôn thùuk₁ dëk
    at head ball hit child
3b. thiį¹ roonrían dëk rian nàŋsuũu
    at school child study book
    'At school, the children studied.'

With regard to (3), it should be noted that it is possible for the compound thùuk+noun construction to have a verbal complement as its sister, as in (a):

a. nūu thùuk+yaphít taai
   mouse to poison die
   'A mouse was poisoned to death.'

The internal structure of this sentence is illustrated below:

```
     thùuk+yaphít
       | 2ndex |
      /---\   |
     nūu  l- trns  l taai
         l  l
         l Index  l +fint  l 3ndex  l
         l + N   l [ + Nom ] l - trns  l
         l + Nom l [ + PAT ] l  l- fint  l
         l  PAT l [ + actrl ] l  l [ +actrl ] l
         l  l [- fint ]  l  l [ +PAT ] l
```

The implied patient of taai 'die' is chained with the patient of the matrix clause.

¹Following Savetamalya (1989, p. 57), one of the subclasses of thiį is analyzed as a relator noun, marking location. It obligatorily requires a dependent noun as its sister.
"nīu 'mouse,' by the Actor Control Rule.

According to the Patient centrality hypothesis, the complement takes the Patient in its scope and, in Thai, typically occurs adjacent to the Patient. Therefore, dēk carries a Patient case relation, rather than any other case relation.

Other supporting evidence for treating thùuk₁ as a transitive verb appears in the following data:

4. lûukbon thùuk₁ hūa (khōŋ) dēk tēek
   ball hit head of child break
   'The ball hit a child's head and the head was broken.'

If thùuk₁ is treated as a transitive verb, the structure in (4) will look like the following:

```
| thùuk₁ |
  | 1[+trns] |
  | 1ndex |
  | 1[+Nom] |
  | lûukbon |
  | 1[+AGT] |
  | l3ndex |
  | l+Nom |
  | l-PAT |
  | l+actr |
  | l-dek |
  | l-trns |
  | l-fint |
  | l-COR |
  | hūa |
  | l5ndex |
  | l-PAT |
  | l-Nom |
  | l3[+PAT] |
  | tēek |
```

Let us focus on the interpretation of the verb tēek. tēek as an intransitive verb implies a Patient subject. The Patient is interpreted as hūa following the Actor Control rule, which simply states that the implied actor of an infinitival complement is coreferential with the Patient of the regent verb, formulated as follows (Starosta 1990):

1. Actor Control Rule:
   1[+actr] \ → [m[+actr]] \ m[+PAT] | l-fint |
   l-nindex |
   l

If hūa were marked by some other case relation, rather than Patient, the matrix clause would be intransitive. In that case, the implied Patient of the verb tēek would be mistakenly equated to a presumed Patient subject lûukbon. This interpretation would not be correct. The tree structure for this hypothetical interpretation is shown below:
From these arguments, the conclusion is that thùuk₁ is to be analyzed as a transitive verb and a dependent noun immediately following thùuk₁ is to be analyzed as a Patient.

On the basis of the different syntactic distributions of thùuk, mentioned above, there are three different homophonic entries for transitive thùuk. All of them require a Patient complement and have the same meaning as ‘hit,’ and since their syntactic distributions are different, each of them will be indexed differently. The first one, thùuk₁, does not allow another complement sister other than a Patient (as in 1). The second one, thùuk₂, allows one more complement sister bearing a Locus case relation (as in 2). The last one, thùuk₃, allows an infinitival complement as its dependent sister (as in 4).

Another homophonic entry, thùuk₄, occurs as a transitive verb whose meaning is ‘touch.’ Again thùuk₄ requires a Patient complement as in (6a). If there is no Patient, the sentence is ungrammatical, as in (6b):

6a. tōi thùuk₄ muuu chań
    Toy touch hand I
    | AGT | +trns | PAT |
    | +actr | ? [+AGT] | -actr |
    | ? [+PAT] |
    ‘Toy (unintentionally) touched my hand.’

6b. *tōi thùuk₄
    Toy touch
    | PAT | -trns |
    | +actr | ? [+PAT] |

Sentence (6b) is grammatical if the meaning of thùuk is ‘be correct,’ in which case thùuk would be classified differently (see section 2.3.)

2.1.1. Incorporated thùuk+Noun Construction.

A construction having thùuk followed by a noun is not always a transitive form. In the