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*: $thùuk_1$ meaning 'hit', and $thùuk_2$ meaning 'touch'. Examples:

1. lûukbɔn thùuk$_1$ dèk
   ball hit child
   AGT [+trns] PAT
   'The ball hit the child.'

In (1) $thùuk_1$ is a transitive verb. It requires an Agent, *lûukbɔn* 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. * lûukbɔn thùuk$_1$
   ball *hit*
   PAT *[-trns]*

Therefore, *dèk*, which is required by the verb *thùuk*$_1$, 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 *thùuk*$_1$ allows one further argument:

3. lûukbɔn	thùuk$_1$	dèk	thîi[1]	hŭa
|AGT |	|+trns |	|PAT|		LOC
|+actr|	|?[+AGT]|	|-actr|		
	|?[+PAT] |			
ball	hit	child	at	head

'The ball hit the child on the head.'

hŭa carries a Locus case relation, marked by the relator noun *thîi*. 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 *thîi hŭa* is a complement, and also the grammaticality of (3b) in which *thîi rooŋrian* is an adjunct:

3a. * thîi$_1$	hŭa	lûukbɔn	thùuk$_1$	dèk
at	head	ball	hit	child
3b. thîi$_1$	rooŋrian	dèk	rian	năŋsŭŭ
at	school	child	study	book

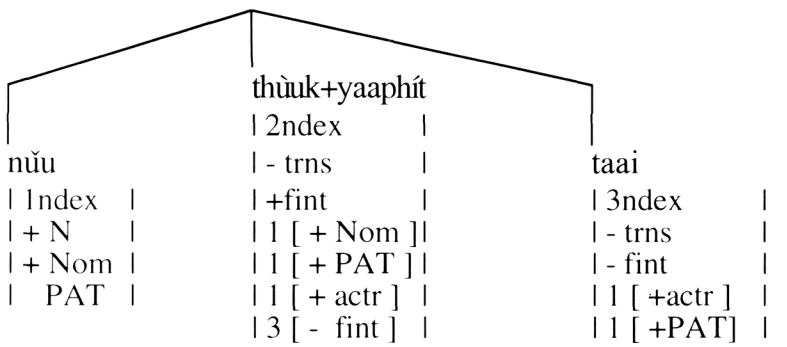
'At school, the children studied.'

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. nŭu	thùuk+yaaphít	taai
mouse	to poison	die

'A mouse was poisoned to death.'

The internal structure of this sentence is illustrated below:



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

[1]Following Savetamalya (1989, p. 57), one of the subclasses of *thîi* is analyzed as a relator noun, marking location. It obligatorily requires a dependent noun as its sister.

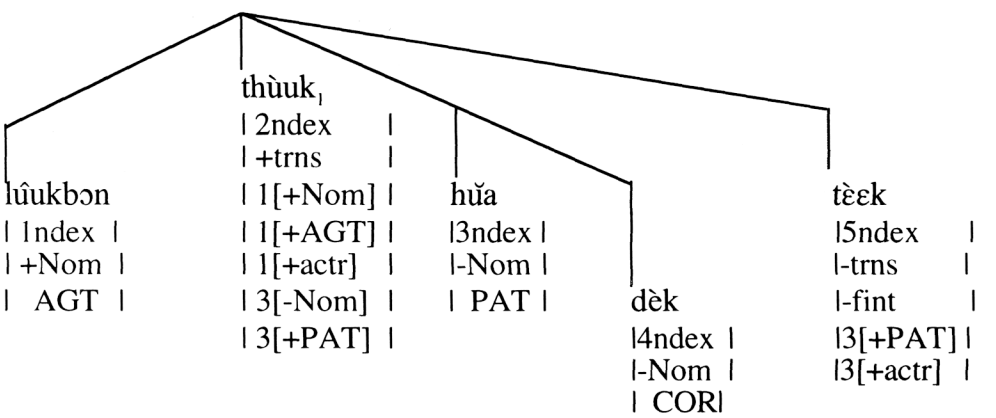
nŭu 'mouse,' by the Actor Control Rule.

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

Other supporting evidence for treating *thùuk*$_1$ as a transitive verb appears in the following data:

4. lûukbɔn thùuk$_1$ hŭa (khɔ̆ɔŋ) dèk tɛ̀ɛk
   ball hit head of child break
   'The ball hit a child's head and the head was broken.'

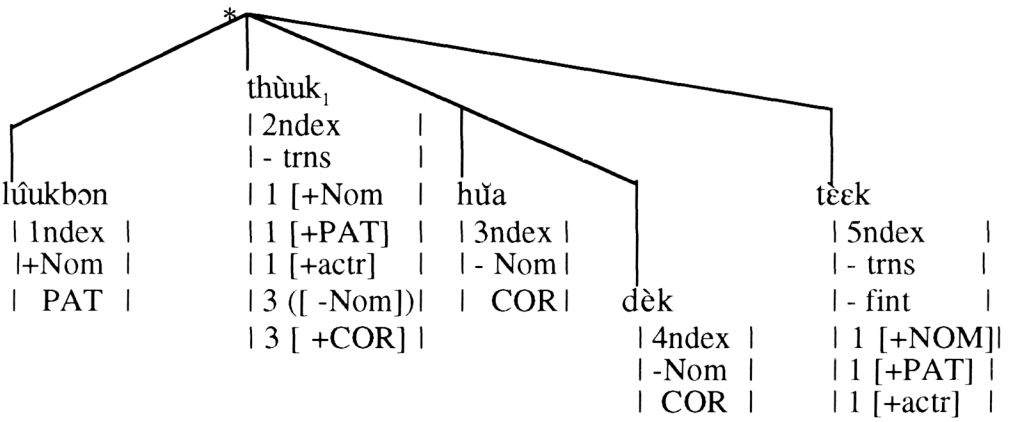
If *thùuk*$_1$ is treated as a transitive verb, the structure in (4) will look like the following:



Let us focus on the interpretation of the verb *tɛ̀ɛk*. *tɛ̀ɛk* as an intransitive verb implies a Patient subject. The Patient is interpreted as *hŭa* 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:
   |?[+actr]| --> [m[+actr]] \ | m[+PAT] |
   | -fint | | n[-fint] |
   | nndex |

If *hŭa* 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 *tɛ̀ɛk* would be mistakenly equated to a presumed Patient subject *lûukbɔn*. This interpretation would not be correct. The tree structure for this hypothetical interpretation is shown below:



From these arguments, the conclusion is that *thùuk*$_1$ is to be analyzed as a transitive verb and a dependent noun immediately following *thùuk*$_1$ is to be analyzed as a Patient.

On the basis of the different syntactic distributions of *thùuk*$_1$ mentioned above, there are three different homophonous 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, *thùuk*$_1$, does not allow another complement sister other than a Patient (as in 1). The second one, *thùuk*$_2$, allows one more complement sister bearing a Locus case relation (as in 2). The last one, *thùuk*$_3$ allows an infinitival complement as its dependent sister (as in 4).

Another homophonous entry, *thùuk*$_4$, occurs as a transitive verb whose meaning is 'touch.' Again *thùuk*$_4$ requires a Patient complement as in (6a). If there is no Patient, the sentence is ungrammatical, as in (6b):

6a. tôi thùuk$_4$ mɯɯ chăn
Toy touch hand I
| AGT | | +trns | | PAT |
| +actr | | ? [+AGT] | | -actr |
| ? [+PAT] |
'Toy (unintentionally) touched my hand.'

6b. *tôi thùuk$_4$
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