Semantic Characteristics of Thai Basic Verbs

Makoto Minegishi

mmine@aa.tufs.ac.jp

Research Institute for Languages and Cultures of Asia and Africa (ILCAA),

Tokyo University of Foreign Studies

Topics

1. Expanding the research on Thai verbs: from the examination based on ILCAA Questionnaire to that based on TNC word frequency list, counted in Sept. 2011.

2. Interim report: showing part of the results of ongoing work about semantico-syntactic features of approximately 300 Thai verbs.

Contents

1. Outline of my research on Thai verbs

2. Making a kind of Thai verb "database"

3. Previous examination based on ILCAA *Linguistic Questionnaire* (SEALS XXI at Kasetsart University)

4. Comparison of TNC top 300 verbs and ILCAAA essential 300 verbs.

Outline of our research on Thai verbs: why we are interested in Thai verbs

1. Thai: An isolating language with basic word order: SVO, NA. Verbs are two-place or one-place verbs, with exceptional three-place verbs.

2. Topic Prominence and Pro-Drop; presence or absence of nominal participant is conditioned by contexts.

3. MSEA languages tend to be described with negative words: NO inflection, NO suffix, NO case marking, NO head/dependent marking, etc.

4. "Isolated" linguists in MSEA struggle for POSITIVE features, such as topic prominence, but again, WITHOUT explicit topic marker. cf. JP *-wa* vs. *-ga*

Prominent features of Thai and other MSEA languages

1. Along with the topic prominence, Serial Verb Construction is one of the most important features of languages in Mainland SEA, thus, the most challenging, 'controversial issue' in Thai lingustics.

2. Far-Eastern point of view: comparable with Japanese 'verbal complex'; morpho-syntactic structure containing a verb followed by words, clitics, suffixes, particles, etc.

3. Q: How Thai expands predications (without overt grammatical markers)?

Ex. 'Eastern' non-hierarchical View: Serial Verbs comparable with JP accent packaging?

1. Both consist of (theoretically) unlimited number of relatively homogeneous elements (verbs or syllables)

2. Standard Japanese has a pitch accent functioning as a package for polysyllabic words: *'LHHHH....HLL'*, where *L* denotes the beginning, *H,* the position of accent nucleus that lowers the pitch of following two syllables.

3. In any level of language, to mark the beginning and the end of a sequence is important.

4. If comparable, what signals the beginning or the end of serial verbs? How about semantic features?

How to analyze verbs

1. Traditionally linguists pick up a basic verb to see its morphological features, classify it into sub-categories, and check its syntactic environments, semantic features, etc. then move on to the next one.
   Results: producing one paper for each verb.

2. In view of analyzing serial verbs, we need to have a rough sketch of many Thai verbs, as many as possible, which could be complementary to the traditional method. How about making a lexical database?

The idea of Thai verb database

- In addition to the information necessary in ordinary dictionaries, what features should be included in the Thai verb database?

- In order to apply it to the analysis of serial verbs, we need to know its semantico-syntactic features; such as 'used in

  'N1-N2-_ ' or 'N1-_N2',
  with [+Human] and/or [+Thing] subject,
  [+Voluntary], or [+Spontaneous], etc.

- Although DB is not yet sophisticated, such info. can be used for MRD in the future.

Previous work: Minegishi (SEALS XXI, 2011 at Kasetsart University)

1. Based on the trial version, examining semantico-syntactic features of Thai essential 300 verbs in ILCAA (AA-Ken)'s *Linguistic Questionnaire* (1979) (hereafter abbr. as 'AA-LQ').
2. "Essential" vocabulary of AA-LQ was selected by a group of Japanese fieldworkers who worked on Asian and African languages.

Previous results based on AA-LQ

- Thai has two major sentence types:

1. "$Noun_1$ $Noun_2$ Verb"
   [Topic-Subject-Verb (TSV)]

2. "$Noun_1$ Verb $Noun_2$"
   [Subject-Verb-Complement (SVC)]

   - Among approximately 300 verbs,
     148 verbs (49%) with Human subjects only,
     58 verbs (19%) with Thing subjects only,
     94 verbs (31%) with both Human and Thing subjects.
     96 (32%)verbs in N_1-N_2-V construction

About TNC Top 5000 frequency list

1. Top 5000 word list updated in every two or three months is available on the TNC web site.
   http://ling.arts.chula.ac.th/TNC/
   cf. Aroonmanakun, W. (2007).
2. The list provides rank, word in Thai scripts, and frequency.

Ongoing work: rebuilding the database based on TNC Top 5000 frequency list

1. We added POS info., in the very beginning V or N only; primitive criterion: what can be directly negated by *may3* 'negator' is V, otherwise N.

2. Result: too "VERBY"; more than 2000 words are initially marked 'V'. Further elaboration in subcategorizing necessary.

3. For now, focus on top 1000 word analysis.

4. Top 600 words include approximately 300 verbs, which is comparable with AA-LQ 300 verbs.

Tentative summary of 300 verbs in TNC

- Among approx. TNC 300 verbs,
  34 verbs appear in N_1-N_2-V,
  112 verbs with Human subjects only (?)
  6 verbs with Thing subjects only
  166 verbs with both Human and Thing Subjects, which include 31 verbs with a "personified" subject (?)

Tentative Comparison of 300 verbs in TNC and AA-LQ

Subject	TNC	LQ-AA
[+H] only	112	147
[+T] only	7	53
[+H] & [+T] (Personified?)	135 (31)	94

34 TNC Verbs in 'N$_1$-N$_2$-V' (cf. AA-LQ: 96)

1. 34 Verbs in 'N$_1$-N$_2$-V' Constructions
2. Most of them are stative verbs [+Spontaneous, -Kinetic, -Control], such as *khaw4 nisay5 dii* (He is good-natured.)
3. The rest includes results of changing states, or motion verbs [+/-Voluntary, +Kinetic] denoting NOT movements, but results of movements:

(a) TNC Verbs in 'N$_1$-N$_2$-V' denote 'states'

- Stative or 'Adjective' Verbs:
  hEG3 'dry', *sia5* 'rotten', *too1* 'grow', *tem1* 'full', *kEE2* 'aged', *kaw2* 'old', etc. (change) of states in some contexts.

- *dii1* 'good', *khaaw5* 'white', rOOn4 'warm' (change of states, or matter of degree in some contexts)

- 'N$_1$-N$_2$-V-N$_3$' (followed by a complement): *mWan5* 'same as', *kAAn1* 'exceed'

  khaw4   nisay5   mWan5 mEE3

  3rdPer.Pron   character same   mother

  'His character is the same as his mother's.'

(b) Verbs in 'N$_1$-N$_2$-V' denote 'states'

Results of changing states:

e.g. *khaat2* (get torn)
krapaw5 huu5   khaat2
bag   handle be-torn
'The bag has a torn handle'
(sentence), or 'a bag with a torn
handle' (compound noun)

(c) Verbs in 'N$_1$-N$_2$-V' denote 'resultative' states

Motion verbs [+/-Voluntary, +Kinetic] denoting not movements, but results of (presupposed, imaginary) movements

chuaG3-nii4 phon5lamaay4 raakhaa1 khWn3 yAq4-maak3

recently   fruits   price   rise   much

'Recently prices of fruits have risen a lot.'

(d) semantic properties of N_1 and N_2 in 'N_1-N_2-V'

N_1 may be either [+Human] or [+Thing]
     cf. organizations [+H, +T], (animals [-H, -T])

N_2 is usually a thing, or a property [+T] of N_1.

'N_1-N_2-V' describes what the topic (N_1) is like concerning its property (N_2-V).

'N_2-V' has the meaning '(minor) subject and its predicate'. (cf. double subject const. in JP)

The construction may be generalized as N_1; {N_2-V_2: (N_3-V_3: N_4-V_4: …)}.

112 TNC Verbs in 'N$_1$-V-(N$_2$)' with Human Subjects [+H] only (AA-LQ: 147)

1. Human Verbs with low transitivity

2. Kinetic action (or pose) unique to human being: *naG3* 'to sit' [+Voluntary, +Kinetic, +Control, Patient not Affected] .

3. (Physical) action with emotions, or mental activities: *hua5rOq4* 'to laugh'

4. Verbs for perception, feeling, mental activities: *hen5* 'to see', *phop4* 'to meet',

5. Verbal communication: *tOOp2* 'to answer'

6 Verbs with Thing Subjects [+T] only (AA-LQ: 53)

- 6 adjectival Verbs with [+T] subject (intransitive verbs): *Gaay3* 'easy', *khaat2* 'get torn', *kaw2* 'old', *hEG3* 'dry', *lek4* 'small', *yen1* 'cool'

- (change of) states, [+Spontaneous, -Kinetic, -Control]

- Most of them appear as N_2 (minor subject in 'N$_1$-N$_2$-V', which may be (overt Topic-less) '(T)-N$_2$-V' .

166 Verbs with either Human or Thing Subject [+H/+T]? (AA-LQ: 94)

- WORK IN PROGRESS regretted, but among them, 135 verbs are really [+H/+T], the rest 31 may be [+H] but with a "personified" [+T]subject.

- Further analysis, such as semantic subcategorization needed.

Discussions

1. Some difference so far in TNC top 300 verbs and LQ-AA essential 300 verbs.

2. Why? It may be that 300 is too small a number; suppose we analyze top 1000 and compare the proportion?

3. Even if we focus on the limited number of semantico-syntactic features for feeding DB, we need to verify the data by elicited sentence examples; it takes time and patience.

References

- Aroonmanakun, W. (2007). Creating the Thai National Corpus. *Manusaya. Special Issue*, No.13, 4-17.

- Aroonmanakun, W., Kachen Tansiri and Pairit Nittayanuparp (2009): Thai National Corpus: A Progress Report. *Proceedings of the 7th Workshop on Asian Language Resources,* ACL-IJCNLP 2009, pages 153–160, Suntec, Singapore, 6-7 August 2009.

- Institute for the Studies of Languages and Cultures of Asia and Africa (*ed*), (1979): *Linguistic Questionnaire for Asian and African Languages (Part II).*[in Japanese, Ajia-Afurika Gengo Chosahyo] Tokyo: Institute for the Studies of Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

- Minegishi, Makoto (2011): Description of Thai as an isolating language, *Social Science Information*, 50: 62-80.

- Minegishi, Makoto (2012): Voluntariness and spontaneity in Thai, *Journal of the Southeast Asian Linguistics Society (JSEALS)* Vol.4:2, pp.77-91.

Thank you for your attention!

This research is supported by JSPS KAKENHI Grant Number (B) 23300093