

Lexicological Significance of Semantic Doublets in Thai

Peansiri E. Vongvipanond

Introduction

This paper is intended to call to the attention of scholars and students of Thai the lexicology of the language.¹ This is not one of the well-trodden paths. When the concepts and theoretical framework of American structuralism were applied to the study of Thai in works such as those by Noss (1964) and Vichin (1970), not much attention was given to the study of words, because the morphological system of the language appeared then to be rather uninteresting, Thai being an isolative language. Polymorphemic words in the language are created through various compounding processes. Though it is obvious that there are many types of compounds, most works, such as those by Rachadaphan (1983) and Jurairat (1985), focus on the formal property of compounds and issues like the number of morphemes in compounds, and the grammatical category or part of speech of each of the components in a compound, as well as the compounds themselves. Semantic issues are usually not the main concern of these studies within the structuralist framework. Even when the generative-transformational theory was adopted, the situation did not improve much for the lexicological study of Thai, perhaps because the lexicon itself was not given much attention until the arrival of the word or lexical grammar theory. A small number of compounds are studied in the work of Udom (1963), and it is proposed by implication that some are ready-made units in the lexicon, while some are transformationally derived.

Many intriguing questions remain unanswered. What is the lexicological motivation or significance for each type of compound? Why are there different types of compounds? What mechanical ingenuity is devised within the lexicology of the language itself, as reflected in the existing compounds, to keep each lexicological type of compound distinct? In an attempt to seek at least a partial answer to some of these questions, a study was done based on a sample of data consisting of 800 compounds referred to

¹ Paper presented at the 18th International Conference on Sino-Tibetan Languages and Linguistics, August 27-29, 1985, Bangkok, Thailand.

in this paper as semantic doublets. The purpose of the study is to find out why speakers create this type of compound, which has rather unique characteristics. The answer cannot be definite at this stage, when not much is known about the lexicology of Thai. However, it is hoped that the paper will attract budding scholars to take a more earnest look at the lexicology of Thai.

Defining Semantic Doublets

A semantic doublet is defined here as a type of compound consisting of components that share a certain degree of semantic similarity. "Sharing a certain degree of semantic similarity" means that the components may be synchronically considered synonyms, near synonyms, or words in the same semantic field. The components in this type of compound are usually mono-lexemic; however, there are a small number of semantic doublets that contain poly-lexemic components. This latter sub-type of semantic doublet is usually a product of double or multiple semantic-doublet formation processes. Another defining characteristic of a semantic doublet is that it is usually synonymous or near-synonymous with one or both or all of its components. The synonymy here means substitutability when difference in register or speech style is not taken into consideration. With this type of internal semantic structure, one can say that semantic doublets are morphologically transparent words, according to the definition given by Ullmann (1957). That is, one can derive the meaning of each semantic doublet quite easily from the meaning of its components.

The following are examples of semantic doublets with mono-lexemic components.

a. Semantic doublets with components that are synchronic synonyms.

1. *râaŋ* + *kaay* = *râaŋkaay*
body body body
2. *kèe* + *charaa* = *kèecharaa*
aged, old aged, old aged, old

b. Semantic doublets with components that are synchronic near-synonyms.

3. *mîit* + *phráa* = *mîitphráa*
knife knife used
(generic) in farming knife (generic)
4. *tháj* + *sín* = *tháj sín*
all all through
(Adj) (Prep) (Adv)

The following are examples of semantic doublets with poly-morphemic components.

5. *rabìap* + *bèep* + *phěen* = *rabìapbèepphěen*
 regulation model plan regulation
6. *pràp* + *pruŋ* + *plian* + *pleeŋ* = *pràppruŋplianpleeŋ*
 adjust adjust change change change for the
 (seasoning) better, improve

The term semantic doublet is deliberately used in this paper to emphasize the semantic resemblance between a semantic doublet and its components, as well as the semantic resemblance between the components within a semantic doublet. The Thai word for this particular type of compound is *kham sóŋ*, which was probably used originally by Phya Anuman Rajadhon. Other terms have been used by Thai linguists. These are *khamkhûu*, or 'word pair', and *kham mii sŷykhām*, or 'word with decoration'. These terms are not adopted because they cannot reveal the significant lexical characteristics of this unique type of compound (Peansiri 1981).

Characteristics of Semantic Doublets

The study on which this paper is based began with a collection of 800 semantic doublets selected on the basis of the defining criteria outlined above. The author and two assistants, all native speakers, each went through the Royal Institute Dictionary and listed all the compounds that met the semantic doublet criteria. These three lists were compared and merged. As the team was completing the list, more semantic doublets were added, taken from each person's own competence, to clarify some of those chosen from the dictionary and to obtain the list of 800 words as planned. The analysis reveals the following formal and semantic characteristics:

Formal characteristics

The formation of a semantic doublet seems to follow the prevalent preference for the pairing rhythmic pattern; that is, the number of components in a semantic doublet is always an even number. Most doublets have two components. In cases where there are more than two, the number is usually four. Cases like this seem to be products of multiple processes, as evident in example 9 below.

7. *yák* + *yáy* = *yákyáy*
 stealthily take move stealthily take a
 a part away part away

8. *thàay* + *thee* = *thàaythee*
 transfer pour transfer a part
 a part
9. *yákyáaythàaythee* = to stealthily take away a part of something

Since most of these lexical components are monosyllabic, this pairing pattern is usually overlooked. One comes to be aware of this principle only in cases like *thanõnhõnthaaŋ* 'road' and *khamoykhacoon* 'thief, robber'. In the former, *hõnthaaŋ* itself is a semantic doublet consisting of *hõn* 'direction' and *thaaŋ* 'path', while *thanõn* is a single lexeme borrowed from Khmer and means 'road'. Since *thanõn* is bi-syllabic, one gets a double semantic doublet with four syllables. In the case of *khamooykhacoon*, it is not a double semantic doublet, for it has only two components: *khamooy* 'thief, robber' and *coon* 'thief, robber'. The added *kha* to precede *coon* is simply to achieve rhythmic balance.

A distinct difference between semantic doublets and other compounds that are not products of reduplicating processes is the fact that components of semantic compounds are members of the same grammatical categories, as can be seen in the examples given below.

Semantic doublets are not restricted to any one grammatical category. There are those that are content words and those that are function words.

NOUN:

10. *saŋ* + *panyaa* = *saŋpanyaa*
 awareness intelligence intellectual capability
11. *sáp* + *sĭn* = *sápsĭn*
 property asset asset

VERB:

12. *khàt* + *thũu* = *khàtthũu*
 polish scrub polish and scrub
13. *khúi* + *khĩa* = *khúikhĩa*
 scratch scratch scratch

ADJECTIVE:

14. *mâak* + *lăay* = *mâaklăay*
 many various various

15. *sîit* + *siaw* = *sîitsiaw*
pale sickly sickly

As for function words in the data collected, only those that are conjunctions are found.

16. *thâa* + *hàak* = *thâahàak*
if if if

17. *thǎŋ* + *mée* = *thǎŋmée*
though though though

Semantic doublets have a rather unique internal structure. The semantic relationship between the two components is neither of the additive type nor the modifying type, so one cannot very well propose that underlying each of the semantic doublets is either a conjoined construction of two members of the same category linked by an unlexicalized additive conjunction or a modifying clause, with one component being the head and the other the modifier. That is, one can not paraphrase a semantic doublet, which is a verb like 18a, with a *lé?* phrase like 18b nor with a *bèep* phrase like 18c.

18. a. *khṛṵṵp-khṛṵṵŋ* to occupy (a land); to colonize
b. *?*khṛṵṵp lé? khṛṵṵŋ* (lit.) to influence and then take over
c. *?*khṛṵṵp lé? bèep khṛṵṵŋ* (lit.) to influence in a colonizing manner

This also applies to semantic doublets that are verbs, like those in 19, and conjunctions like those in 20.

19. a. *sáp-sín* asset; property; wealth
b. *?*sáp-lé?-sín* (lit.) wealth and wealth
c. *?*sáp bèep sín* (lit.) wealth-like wealth
20. a. *thâa-hàak* if
b. *?*thâa lé? hàak* (lit.) if and if
c. *?*thâa bèep hàak* (lit.) if-like if

What is even more interesting is the fact that the degree of cohesion between the two components of most semantic doublets is rather loose. Intrusion of a lexical item, under the constraints of co-occurrence, is possible as long as it equally affects both components.

Semantic doublets that are nouns allow the intrusion of verbs of which they are the objects.

21. a. *mii riaw-reeŋmâak* to have great physical strength
 b. *mii riaw mii reeŋmâak* (same meaning as above)
22. a. *kin yùuk-yaa pay mâak* having taken a lot of medicine
 b. *kin yùuk kin yaa pay mâak* (same meaning as above)

A small number of nominal semantic doublets do not all allow such an intrusion.

23. a. *khâam phuu-khăw* to climb over the mountains
 b. *?*khâam phuu khâam khau*
24. a. *duulee ʔaakhaan-sathăanthîi* to maintain grounds and building
 b. *?*duulee ʔaakhaan duulee sathăanthîi*

Besides the intrusion of a verb, nominal semantic doublets also allow the intrusion of a preposition in the same manner.

25. a. *taam bâan-ruan* on or around houses
 b. *taam bâan taam ruan* (same meaning as above)
26. a. *dûay mon-khaathăa* with magic
 b. *dûay mon dûay khaathăa* (same meaning as above)

Just as in the case of verbs, some nominal semantic doublets, though very few, do not allow intrusion of a preposition.

27. a. *dûay ʔitthi-rít* with trained mental power
 b. **dûay ʔitthi dûay rít*

28. a. *kềề khâa-thâat* to people in one's service
 b. *kềề khâa kềề thâat* to one's servants and slaves

The same intrusion is allowed in verbs. This can be an intrusion of an auxiliary.

29. a. *cà? booy-bin* will fly
 b. *cà? booy cà? bin* will fly
30. a. *cà? khít-?àan* will come up with an idea
 b. *cà?khít cà? ?àan* will come up with an idea

It can be an intrusion of a negator.

31. a. *mây kìit-kan* not to stop or prohibit
 b. *mây kìit mây kan* not to stop or prohibit
32. a. *mâyđây phìt-phềềk* not differ from
 b. *mâyđây phìt mâyđây phềềk* not to differ from

It can also be an intrusion of another verb in the same serial verb phrase.

33. a. *chûay khám-cun* to help maintain and support
 b. *chûay khám chûay cun* to help maintain and support
34. a. *yàak phûut-khui* to want to chat with
 b. *yàak phûut yàak khui* to want to chat with

However, not all predicative or verbal semantic doublets allow such an intrusion, though this is only a small minority.

35. a. *cà? rîap-riar* will compose (a book)
 b. *?*cà?rîap cà?riar* (lit.) will polish and arrange
36. a. *mây nam-phaa* not to be attentive to
 b. *?*mây nam mây phaa* (lit.) will neither bring nor take

37. a. *chûay khum-khrɔɔŋ* to help protect
 b. *chûay khum chûay khrɔɔŋ* to help protect and dominate

Another interesting formal property of semantic doublets is the tendency towards euphony as a consequence of the alliteration of the initial consonant of the components. The following are examples.

38. *kâaw-kàay* to interfere
 39. *râap-râp* calm; undisturbed
 40. *sǎaʔòk-sǎacay* to be distressed
 41. *kàt-kin* to bite into
 42. *khǎa-khɛen* limb

Semantic characteristics

The components of a semantic doublet are, as mentioned earlier, either synonyms or near-synonyms. However, the semantic distance/proximity between the two components is not uniform for all semantic doublets. There are those that are composed of synonyms and the semantic doublets themselves, and the two lexemes that constitute their components are usually substitutable.

43. *kháw càʔ pǎian-plɛɛŋ bèɛp bâan sǎa mày*
 he will design house up anew
pǎian
plɛɛŋ
 change

He will alter the design of the house.

Many semantic doublets can substitute for only one of the components, either the first or the second. This is when the two components are only near-synonyms.

44. *kháw phûut cháʔ-ceen dii*
 he speak well
cháʔ
**ceen*
 clear

He speaks quite clearly.

45. *mii* *khon* *sũun-hǎay* *pay* *sìp* *khon*
 there-be people go ten CL
 ****sũun***
 hǎay
 missing

51. *mâŋ*(OB) + *kâŋ* = *mâŋkâŋ*
 be in a large number be in a large number or amount rich

52. *sûa* + *sâat* (TD) = *sûasâat*
 mat mat mat

The formation of semantic doublets is still a productive lexicological process. Many new words that have been coined to represent modern concepts in the various academic fields as well as in science and technology are semantic doublets.

53. *ʔâaŋ* + *ʔiŋ* = *ʔâaŋʔiŋ*
 refer to depend make reference to

54. *thòt* + *thɔɔy* = *thòtthɔɔy*
 move over retreat regress

55. *yûa* + *yúʔ* = *yûayúʔ*
 tempt urge provoke; instigate

56. *yûut* + *yùn* = *yûutyùn*
 stretch yield to touch flexible

57. *lûan* + *lɔɔy* = *lûanlɔɔy*
 move float slash-and-burn fashion

Earlier Hypotheses

Many Thai language scholars and linguists have proposed several hypotheses on the formation of these semantic doublets. Their ideas seem to fall into two hypotheses.

The Euphony Hypothesis. This hypothesis seems to have been based on the observation that alliteration and the pairing principle are prevalent in this type of compound to maintain the preferred rhythmic pattern. Implicationally, this hypothesis can be taken to point out the tendency of the language to reduce monosyllabicity.

The Practicality Hypothesis. This is a more prominent trend of thinking than the euphony hypothesis. However, many people adopt both hypotheses. According to this hypothesis, semantic doublets are formed for practical purposes. In the famous Somdej Phramaha Weerawong *Photcananukrom Phasa Thai-Isan* [Central Thai and Isan Thai Dictionary] (1972), it is said that a synonym or near-synonym is adjoined to a certain word to guarantee that readers can get its correct tone, since tones were not

usually marked in earlier writing and tones do differ among dialects. Another and also a more acceptable analysis of these compounds by Phya Anuman Rajadhon (1971) and Banchop Bhandhamedha (1971) led to an explanation that these words are formed for the following practical purposes:

Firstly, it is to translate loanwords. Examples are those given in 48-52 above.

Secondly, it is to disambiguate homophonous words. Examples are the following:

58. a. *khâa-fan* = to kill
(lit. = kill + cut into)
- b. *khâa-gûat* = value
(lit. = value + obsolete word)
- c. *khâa-thâat* = people in one's service
(lit. = servant + slaves)
59. a. *sũun-klaaŋ* = center
(lit. = center + center)
- b. *sũun-hăay* = lost; missing
(lit. = lose + lost; missing)

Thirdly, it is to differentiate the various meanings of polysemous words.

60. a. *khěŋ-rœŋ* = strong
(lit. = hard + strong)
- b. *khěŋ-krâw* = unbending
(lit. = hard + aggressive)
- c. *khěŋ-khăn* = enthusiastic; earnest
(lit. = hard + obsolete word)
- d. *khěŋ-khœŋ* = solid
(lit. = hard + solid)
- e. *khěŋ-khœun* = to resist
(lit. = hard + go against)

The Lexicological Hypothesis. This is a new hypothesis to be proposed in this paper. It is not meant to be a contradiction. Rather, it

is an attempt to add more insight to the study of this intriguing type of compound in Thai. The hypothesis is based on the lack of uniformity in the semantic resemblance between semantic doublets and their components. The meaning of semantic doublets can be classified into four types on the basis of this semantic resemblance.

a. The semantic doublet can have a “larger” or “more general” meaning.

- | | | | | | |
|-----|-------------------------|---|--------------------------|---|---|
| 61. | <i>râak</i>
root | + | <i>thăan</i>
base | = | <i>râakthăan</i>
foundation; base; basis |
| 62. | <i>yûŋ</i>
entangled | + | <i>yâak</i>
difficult | = | <i>yûŋyâak</i>
complex; complicated |
| 63. | <i>kîŋ</i>
branch | + | <i>kâan</i>
stem | = | <i>kîŋkâan</i>
limbs of tree |

b. The semantic doublet can have a “narrower” or “more specific” meaning. This usually occurs among those with components that are homophonous or polysemous words.

- | | | | | | |
|-----|---------------------|---|-----------------------|---|----------------------------------|
| 64. | <i>mêut</i>
dark | + | <i>mít</i>
black | = | <i>mêutmít</i>
pitch dark |
| 65. | <i>mêut</i>
dark | + | <i>mua</i>
unclear | = | <i>mêutmua</i>
shady, unclear |
| 66. | <i>mêut</i>
dark | + | <i>salŭa</i>
lazy | = | <i>mêutsalŭa</i>
hazy |

c. The semantic doublet can subsume the meaning of both of its components like non-reduplicative compounds in general. It is noticed that the so-called new semantic doublets are usually in this group.

- | | | | | | |
|-----|-------------------------|---|-------------------------|---|---------------------------------------|
| 67. | <i>kòt</i>
press | + | <i>khîi</i>
ride | = | <i>kòtkhîi</i>
oppress |
| 68. | <i>kon</i>
mechanism | + | <i>wíthii</i>
method | = | <i>konwíthii</i>
strategy |
| 69. | <i>lûm</i>
marshy | + | <i>lúk</i>
deep | = | <i>lûmlúk</i>
profound; insightful |

d. The semantic doublet has a distinct and independent meaning, signifying a different referent. The meanings of these semantic doublets are often

described as being “figurative” or “abstract” since they share very few semantic features with that of the components.

- | | | | | | |
|-----|------------------------|---|-----------------------------|---|---------------------------------|
| 70. | <i>râak</i>
root | + | <i>ŋâw</i>
bulb | = | <i>râakŋâw</i>
origin; cause |
| 71. | <i>phùuk</i>
tie | + | <i>khaat</i>
discontinue | = | <i>phùukkhaat</i>
monopolize |
| 72. | <i>bùk</i>
attack | + | <i>bòək</i>
open | = | <i>bùkbòək</i>
pioneer |
| 73. | <i>yôy</i>
digested | + | <i>yáp</i>
creased | = | <i>yôyyáp</i>
ruined |

In addition to these four instances in which a semantic doublet can differ from its component, one should also take into consideration the fact mentioned in the third “characteristics of semantic doublets” above about the cohesive strength between the components of a semantic doublet. Some semantic doublets exhibit a very strong cohesion and do not allow the intrusion of another lexeme or the switching of the two components. The only type of uniformity found among semantic doublets is the tendency to exhibit euphony through alliteration. These discrepancies can be interpreted as evidence of the fact that synchronically there is more than one kind of semantic doublet. There are those that were historically formed for practical purposes: to guarantee that listeners get the correct tone; to translate words borrowed from a foreign source or a different dialect; or to disambiguate homophonous and polysemous words. These semantic doublets of the earlier generations have survived but not without the effect of semantic change. Semantic doublets, once created, can become popular and thus can become competitors with words that are their components. Such a competition may lead to many consequences. Firstly, the semantic doublets may take on a new meaning that is related to and yet distinct from that of their own components and is very often abstract. In this situation, the cohesion of their components naturally becomes stronger and intrusion is not possible. Secondly, the semantic doublets may serve as the more elegant synonyms of the words that are their own components. In this type of semantic doublet, the cohesive strength is less than in the first instance. Thirdly, one or both of the components may lose out in this competition, and the semantic doublets of this type may not be morphologically transparent any more except to trained linguists. This chronological development, especially the development of semantic doublets of the first type described above may have also led to another lexical phenomenon: the creation of euphous compounds, in which words that share the same initial consonant or syllable and are similar to an extent in meaning are combined to form a new lexical item in the

language to represent new concepts, such as technical concepts borrowed from the West. These are semantic doublets like *thòtthōy* 'regressive (as in statistics)' or *dòotdiaw* 'to alienate somebody'. Meanwhile, there are new compounds that may be formed from the same pair of semantically similar words but refer to different concepts, and one finds a pair like *kìt-thúra?* and *thúra-kìt*. With this evidence, the formation of semantic doublets can be viewed historically as a practical process for guaranteeing accurate communication that later develops into a synchronic lexicological process for creating new words in the language.

This lexicological picture is even more interesting when one looks further at another type of compound consisting of words that are hyponyms, such as *khâawplaa* 'food' or, literally, 'rice-fish' and antonyms such as *chûadii* 'whatever the case may be' or, literally, 'bad-good'. This reveals that words that are semantically related are likely to be used in a good number of cases as raw materials for making new words. Semantic doublets do not come into existence simply for practical or euphonic purposes, but they, or at least some of them, do serve a truly lexicological function.

References

In English

Noss, R. B.

1964 *Thai Reference Grammar*. Washington, D.C.: Foreign Service Institute.

Udom Warotamasikkhadit

1963 "Thai Syntax: An Outline." Ph.D. dissertation, University of Texas.

Ullmann, S.

1957 *The Principles of Semantics*. Oxford: Basil Blackwell.

Vichin Panupong

1970 *Inter-Sentence Relations in Modern Conversational Thai*. Bangkok: The Siam Society.

In Thai

Anuman Rajadhon, Phya

1971 *Niruktisat*. Bangkok: Infantry Press.

Banchop Bhandhumedha

1971 *Laksana phasa Thai*. Bangkok: Khurusapha Press.

Jurairat Laksanasiri

1985 "*Khamson nai phasa Thai—samai Sukhothai læ Ayutthaya.*"
Mimeograph.

Kanchana Nacsakul

1973 "*Bæp sæmsoi nai phasa Thai læ phasa Khamen.*" Bangkok:
Seemuang Press.

Peansiri Vongvipanond

1981 *Phasa Thai 3: Khwammai nai phasa Thai (13).* Sukhothai Tham-
mathirat University Press.

Rachadaphan Setthawat

1983 "*Khamson nai phasa Thai.*" M.A. thesis, Chulalongkorn Univer-
sity.

Weerawong, Somdej Phramaha

1972 *Photcananukrom phasa Thai—Isan.* Bangkok: Watana Panich.