N+V Compound Nouns in Thai
Author(s): Worawut Warutamasintop

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
N+V Compound Nouns in Thai*

Worawut Warutamasintop
University of California, Los Angeles

Compound words in Thai, as noted by Haas (1966), are of various types, e.g. N+N, N+V, V+N, and V+V, etc. In this paper I will discuss whether N+V compound nouns
(a) are formed by transformational rules,
(b) are formed by lexical rules, or
(c) are lexical items.

I will argue that compounds of this type should be lexical items. In addition, I will discuss the cases in which some N+V compound nouns may have a phrasal origin (i.e. N + verbal modifier). Finally, I will show how the native speaker knows that the compound is related to its components, and how this knowledge should be captured in a linguistic theory.

1. N+V compound nouns

N+V compound nouns and phrases which look like compound nouns are exemplified below.

(1.1) a. khraiān-bin (machine+to fly) 'airplane'
  naām-tōk (water+to fall) 'waterfall'
  bay-phāt (blade+to spin) 'propeller'
  banday-lāān (stair+to move) 'escalator'

b. yaa-salōp (medicine+to faint) 'anesthetic'
  yaa-thāay (medicine+to have a bowel movement) 'laxative, purgative'

c. kāy-chon (chicken+to fight) 'fighting cock'
  plaa-kāt (fish+to bite) 'fighting fish'
  huu-rūnt (ear+to slide) 'zipper tab'

d. maāy-thāaw (wood+to lean on) 'walking stick'

e. nān-sāk-phiṁ (book+to print) 'newspaper'
  thūa-tāt (bean+to cut) 'peanut brittle'
  sāy-krāk (intestine+to fill) 'sausage'
  naalikaa-phōk (watch+to carry (in a pocket, etc.)) 'pocket watch'

f. naalikaa-plūk (watch+to wake) 'alarm clock'

g. bay-sēt (sheet+to finish) 'receipt'

(1.2) a. manaaw d>ng (lime+to pickle) 'pickled lime'
  krāthīam d>ng (garlic+to pickle) 'pickled garlic'
  khāaw phāt (rice+to fry) 'fried rice'
  khāaw tōm (rice+to boil) 'boiled, watery rice'

*This is a revised version of what appears in the UCLA Papers in Syntax, No. 6, 1974. Edited by Sandra A. Thompson.
The grouping of the above items is done for expository and reference purposes. Most of them are grouped according to the meaning or the paraphrased meaning of the phrases and compounds. For example, (1.1a) indicates the capability of the noun-component, (1.1b) has a causative meaning, etc. Some of them can be interpreted in more than one way. For example, (1.1c) can be understood as having either purpose or capability (of the noun-component) reading. The status of the items in (1.2), i.e. whether they are compounds or phrases, calls for discussion. The forms in (1.3a-b) consist of a noun and a so-called descriptive verb. (1.4a) contains forms which may be regarded by some speakers as compounds, but by others as polysyllabic words. The metaphorical meaning of these 'compounds' can, given some extra-linguistic knowledge, e.g. a particular Thai legend, or imagination, etc., indirectly be related to the componential parts of the 'compounds'. For example, naan-kwak is a plant whose curl leaves look like the shape of a lady's hand when it is waved palm downward. (1.4b) contains forms which may not be regarded by anybody as compounds, because their metaphorical meaning is too far fetched from the componential parts. I have tried to illustrated variety of N+V, but I do not claim to have shown them all.

Before moving on to discuss the mechanism which accounts for the compound nouns of the form N+V, we have to show that this kind of compound exists. To do this, let's consider the difference between a compound, e.g. banday-laan 'escalator (stair to slide or move)' and to lian 'table' 'to move', which is a non-compound structure. (1.5), (1.6), and (1.7) show that banday-laan, like
a noun, can take a demonstrative, a numeral, and a modifier, but to láan cannot.

(1.5) a. banday-láan nán stair move that 'that escalator'
b. *tô láan nán table move that *'that moving table'

(1.6) a. banday-láan 2 an/khrêañ Classifier '2 escalators'
b. *tô láan 2 tua Classifier *'2 moving table'

(1.7) a. banday-láan sîi dam color black 'black escalator'
b. *tô láan sîi dam *'black moving table'

(1.5) - (1.7) indicate that banday-láan does not function syntactically as a phrase consisting of a noun and a verb. However, they do not rule out the possibility that banday-láan is a noun phrase, i.e., a noun plus a modifier. The following will show that this is not the case. If láan can function as a modifier, it should be able to modify a noun in a liberal way, provided no selectional restrictions are violated and the contexts are appropriate. The fact that láan cannot be used to modify nouns such as 'table' and 'chair', even in the following situation, indicates that it cannot function as a modifier. Suppose there is an earth tremor and some tables and chairs are moving. Upon seeing the moving tables and chairs one cannot say,

(1.8) *duu tô/kâwîî láan tua nán sî look table/chair move Classifier that Particle *'Look at that moving table/chair.'

In addition, after the tremor, one cannot say (1.8) to mean 'Look at that moved table/chair' either. Since láan in banday-láan can function neither as a verb nor a modifier, it must be a part of a compound. This means that compound nouns of the form N+V exist.

2. Transformational rules, lexical rules, or compound lexical items

Earlier linguists, for example Warotamasikkhadit (1972) and Fasold (1969), derive the compounds in question by transformational rules. In addition, some authors of Thai textbooks, e.g. Yates and Tryon (1970), imply that these compounds are derived transformationally.

Warotamasikkhadit (1972), using a generalized transformation framework, derives a compound from two independent sentences. Without elaboration, he posits, for example, rules GT.26 and GT. 32 to derive sentences containing the compounds nám-khâan 'dew (water + stay, or get stuck on or in something)' and taw-riit 'an iron (stove+to iron)', as illustrated in (2.1) and (2.2). (Note: K = kh, and P = ph)
(2.1) GT.26 \( \{X+N^1(\text{Det}^1)Y\} \rightarrow X+N^1+V_i(\text{Det}^1)Y \)
\( N^2(\text{Det}^2)V_i \) Where \( N^1 = N^2 \)
naãm + kɔʔ + bay + yaâ
water hold leaf grass \( \rightarrow \) naãm+Kaãŋ+kɔʔ+bay+yaâ
'dew stays on leaves of grass'
naãm + Kaãŋ
water stay

(2.2) GT.32 \( \{X+N^1(\text{Det}^1)Y\} \rightarrow X+N^1+VB(\text{Det}^1)Y \)
\( N^2(\text{Det}^2)V_B\{dûay\}N^3(\text{Det}^3) \) Where \( N^1 = N^3 \), \( VB = V_i \), \( V_m \text{, } V_t + \text{Nom, } V_m + \text{Nom, } \text{pen + Nom} \)
kãw + mii + taw
he have stove \( \rightarrow \) Kãw+mii+taw+rît
raw + rît + Pã + dûay + taw
we iron cloth with stove

The shortcomings of this analysis are:
(a) Constraint: There doesn't seem to be a non-ad hoc way
to prevent his rules from deriving unacceptable strings. For
example, (2.3), which has exactly the same structure as (2.2),
will undergo rule GT.32, when it should not.

(2.3) Kãw + mii + màây
he have wood \( \rightarrow \) #Kãw+mii+maây+tii²
raw + tii + dèk + dûay + màây
we beat child with wood
\( \rightarrow \) #'he has a switch
(used for whipping')

The problem with rule GT.26 can be shown as follows: \( N^1 \) in
the first string in GT.26 must be allowed to be either a subject
(so that sentences like (2.1) can be derived) or an object (so
that a sentence like khãw hên naãm-khaañ 'he saw dew' can be
derived). A problem arises when we try to derive sentences like
(2.4).

(2.4) naãm bon bay bua nán kɔt csãk naãm-khaañ thî ruamtua kan
water on leaf lotus Dem born from dew Rel combine together
The water on that lotus leaf came from dew which combined together.

According to GT.26, the second string doesn't allow anything before
the noun and after the verb. Therefore, the underlying structure
of (2.4) will have to be (2.5) (disregard irrelevant transforma-
tions which may have applied).
(2.5) näâm bon bay búa nán k pérd cãak näâm thií ruamtua kan
\[N_x \ P \ N \ N \ \text{Dem} \ V \ P \ N_y \ \text{Rel} \ V \ \text{Adv}\]
water on leaf lotus That born from water which comb. to.
\[näâm \ khaâŋ\]
\[N_z \ V_1\]
water stay

First, the first string in (2.5) is semantically odd. Second, since both \(N_x\) and \(N_y\) are identical to \(N_z\), rule GT.26 can attach \(V_1\) to either \(N_x\) or \(N_y\). In the second case, we have (2.4). In the first, we have the semantically anomalous string (2.6):

(2.6) näâm-khaâŋ bon bay búa nán k pérd cãak näâm thií ruamtua kan
\[N_x \ V_1 \ N_y\]
dew on leaf lotus that born from water which comb. to.
The dew on that lotus leaf came from water which combined together.

(b) Criteria: There are no criteria for what can appear as an input to the rules. For example, instead of formalizing rule GT.32 to allow the second string in the rule to be raw riit Pâ đầuy taw "we iron cloth with 'stove'", we can formalize it in such a way that the second string can be raw/khâv/khon raw chây taw riit phâa/sâaphâa "we/he/human beings use 'stove' to iron cloth/clothing", or taw chây sâmrâp riit phâa/sâaphâa "a 'stove' is used for ironing cloth/clothing".

(c) Deletion: Deletion of meaningful items doesn't seem to be well motivated or to have independent justification. For example, all the following items in the second string of (2.2) -- the subject, the object, and the adverbial -- are subject to deletion. This kind of deletion is highly questionable.

(d) Meaning: It is not clear how the meaning of the compounds can be arrived at.

Like Warotamasikkhadit (1972), Fasold (1969) divides compounds of the type \(N+V\) into several groups. Unlike Warotamasikkhadit, however, he demonstrates how a compound itself is derived. Instead of deriving a compound from two independent sentences, he derives it from an NP containing a relative clause, according to the paraphrased meaning of the compound. For example, râa-bin (boat+to fly) 'seaplane, airplane' and naalîkaa-plûk (clock+to wake) 'alarm clock' are derived as in (2.7) and (2.9).

(2.7) BASE: \[<râa \ \text{boat} \ \text{boat} \ \text{fly} \ <S \ NP >\]
(For the benefit of the reader, I illustrate the intermediate derivations (as would be derived by Fasold himself) as follows:)

T. Rel Insertion
\[<râa \ <\text{thií} \ râa \ \text{bin} \ <S \ NP >]\
After that a general noun compound (GNC) rule, illustrated below, will turn the last structure above into a compound, i.e. ráa-bin.

(2.8) \[ \text{GNC} \quad \begin{array}{c} N \\ V \quad (N) \end{array} \quad \Rightarrow NP \]

(2.9) \[ \text{BASE: } \langle \text{naalikaa} \quad \text{khon} \quad \text{chay} \quad \text{samrap} \quad \text{kaan} \quad \text{kaan} \quad \text{pen} \quad \langle \text{naalikaa} \quad \text{pluk} \quad \text{khon} \quad \text{S} \quad \text{NP} \quad \text{S} \quad \text{NP} \quad \text{NP} \rangle \]

After a series of Relativization and Deletion rules, which need not be elaborated here, have applied to this string, the structure will be

\[ \langle \text{naalikaa} \quad \text{pluk} \quad \text{khon} \quad \text{S} \quad \text{NP} \quad \text{NP} \quad \text{NP} \rangle \]

(Actually the position occupied by N should be VP. It might be a misprint in his manuscript.)

Then Fasold (1969) states (p.116):

GNC applies to this structure to form naalikaaplukkhon, a non-occurring compound. In order to generate naalikaapluk, we will have to add a rule which applies to structures generated by GNC to delete unwanted members in cases like this. The rule is of this form:

\[ \text{Noun Compound Deletion 1. } \langle N \quad V \quad [ +N \quad [ +Pro ] \quad NP \rangle \quad NP \]

\[ 1 \quad 2 \quad 3 \quad \Rightarrow 1 \quad 2 \quad \emptyset \]

This optional rule deletes the third member of a noun compound of the form Noun-Verb-Noun if the third member is a Pro-noun like khon 'person'.

The problems encountered by Fasold's model include the following:

(a) Constraint: There does not seem to be a non-ad hoc way
of blocking the transformational rules from deriving unacceptable strings. For example, a rule which derives the acceptable compounds in (2.10a) will also derive the unacceptable ones in (2.10b), because both of them have exactly the same underlying structure.

\[
\begin{align*}
(2.10) & \quad (a) \{ \text{klu\'ay} \} \times \{ \text{kh\'aw} \} \\
& \quad \{ \text{m\'apr\'aw} \} \times \{ \text{lincii} \} \\
& \quad \{ \text{plaa} \} \\
& \quad \{ \text{sun, air, etc.} \} \\
& \quad \{ \text{copra} \} \\
& \quad \{ \text{dehydrated banana} \} \\
& \quad \{ \text{dehydrated rice} \} \\
& \quad \{ \text{dehydrated litchi} \} \\
& \quad \{ \text{dehydrated shrimp} \} \\
& \quad \{ \text{dehydrated fish} \} \\
& \quad \{ \text{to expose} \} \\
& \quad \{ \text{after being cooked} \} \\
& \quad \{ \text{man\'ut-kin-khon} \} \\
& \quad \{ \text{man-eating person} \} \\
& \quad \{ \text{man-eating plant} \} \\
& \quad \{ \text{ticket\(\to\)pawn} \} \\
& \quad \{ \text{sheet\(\to\)take leave} \} \\
& \quad \{ \text{field\(\to\)fly} \} \\
& \quad \{ \text{receipt listing} \} \\
& \quad \{ \text{application for a leave of absence} \} \\
& \quad \{ \text{airport} \} \\
\end{align*}
\]

Another example is that the Noun Compound Deletion Rule Fasold posited will also delete an item when it is not supposed to. According to this rule, the second noun of the structure Noun-Verb-Noun will be deleted if it is a Pro-noun like khon 'person'. Unfortunately, there are compounds of the form N-V-N in which the second noun is khon and it cannot be deleted, e.g. m\'an\'ut-kin-khon (man\(+\)to eat\(+\)person) 'man-eating person', and t\'om\'m\'a\'y-kin-khon (plant\(+\)to eat\(+\)person) 'man-eating plant'.

(b) Meaning - Deletion: A transformational approach which derives a compound from a structure which is based on the paraphrased meaning of the compound to be derived will encounter a problem with compounds like:

\[
\begin{align*}
(2.12) & \quad a. \quad \text{t\'u\'a\'-camnam} \quad (\text{ticket\(\to\)pawn}) \quad \text{\textquoteleft receipt listing pawned items\textquoteright} \\
& \quad b. \quad \text{bay-laa} \quad (\text{sheet\(\to\)take leave}) \quad \text{\textquoteleft application for a leave of absence\textquoteright} \\
& \quad c. \quad \text{\'a\'na\'am-bin} \quad (\text{field\(\to\)fly}) \quad \text{\textquoteleft airport\textquoteright} \\
\end{align*}
\]

For example, the meaning of (2.12b-c) is (2.12b\textquoteleft-c\textquoteleft):

\[
\begin{align*}
(2.12) & \quad b'. \quad \text{bay s\'an} \quad \text{khon} \quad \text{yas\'in} \quad \text{pha\'a} \quad \text{kh\'\'in} \quad \text{laa} \quad \text{pan} \\
& \quad \text{sheet Rel person submit for request take work leave} \\
& \quad c'. \quad \text{\'a\'na\'am} \quad \text{\'a\'n\'a\'am-bin} \quad \text{bin} \quad \text{kh\'an} \quad \text{(bin)} \quad \text{lon} \\
& \quad \text{field Rel airplane fly up fly down} \\
\end{align*}
\]
If we capture the meaning of a compound in the deep structure, (2.12b-c) will have the deep structure (2.12b'-c') respectively. This means that all the underlined items in (2.12b'-c') have to be deleted. This kind of deletion should be objected because several meaningful, unrecoverable items are deleted. On the other hand, if we disregard the meaning in the deep structure and derive, for example, (2.12b) from

(2.12) b*: **bay sâŋ laa**
   *sheet Rel to take leave

we have difficulty explaining how the actual meaning of the compounds is arrived at.

Yates and Tryon (1970) will not be discussed because they merely imply that the compounds are derived transformationally.

What is left to be decided is whether the compounds we are discussing are best described as single lexical items or as derived by lexical rules. The decision will be based on (i) the productivity of the rule which conjoins a noun with a verb, and (ii) the predictability of the meaning of the compounds.

(i) **Productivity**

If the rule which combines a noun with a verb is synchronically productive, the compound should be derived by a rule, in order to capture the correct generalizations and to keep the complexity in the lexicon from increasing. On the other hand, if the rule is not productive, or is what Cram (1972) calls 'semi-productive', the compounds should be listed in the lexicon, because there is no good way to allow the rule to apply to some cases and not to the others when in both cases the structural description of the rule is met. In addition, no generalization can be made.

(2.13) a. **nsalîkaa-phók** (watch+to carry) 'pocket watch'
   mît-phók (knife+to carry) 'pocket knife'
   pâân-phók (gun+to carry) 'pistol'
   b. *witthâyú-phók (radio+to carry) 'pocket radio'
   *nânsâî-phók (book+to carry) 'pocket book'
   c. witthâyú-krapâw (radio+pocket) 'pocket radio'

(2.14) a. **sât-liâŋ** (animal+to raise) 'domestic animal'
   b. *màa-liâŋ (dog+to raise) 'domestic dog'
   *kîy-liâŋ (chicken+to raise) 'domestic chicken'
   c. mîa-bân (dog+house) 'domestic dog'
   kîy-bân (chicken+house) 'domestic chicken'

(2.15) a. **mâây-kîiîp** (wood+to grip) 'chopsticks'
   b. tâkîap --- 'chopsticks'

(2.16) a. **din-duût** (earth+to pull) 'quicksand'
   b. *lèk/nâk-duût (iron/ore+to pull) 'magnet'
   c. mîk-laâk (mother+iron) 'magnet'

These examples indicate that the compound rule can be only semi-productive. In some cases the verbal component of a compound
has 'competition'. For example, phók 'to carry in a pocket' can be combined with 'watch', 'knife', and 'gun', but not with 'radio' and 'book', as illustrated in (2.13a-b). The second element in the compound 'pocket radio' is krápáw 'pocket', as shown in (2.13c). (As for 'pocket book', I would like to predict that nánsáː cháːbap krápáw (book Classifier pocket) = pocket edition book' or nánsáː lêm lék lék (book Class. small small) = small book' will be used, instead of *nánsáː-phók. ) likewise, the second element in 'domestic dog' and 'domestic chicken' is baan 'house' when it should have been lián 'to raise'. Finally, semantically, it is very appropriate to call 'chopsticks' and 'magnet' *máːy-khiːp (wood+to grip or take up (with forceps, chopsticks, or pliers)) and *lék/rˀx-duːt (iron/ore+to pull) respectively. But as seen in (2.15) and (2.16), these two compounds are unacceptable. Note that what makes it even more appropriate to call 'magnet' *lék/rˀx-duːt is the existence of the compound din-duːt 'quicksand' (shown in (2.16a)). The objection to deriving the compounds by lexical rules is, therefore, that adequate criteria for the application of the rules do not seem possible.

What seems to be in favor of a lexical rule hypothesis is that there are verbs such as dɔŋ 'to pickle', thɔːt 'to fry', and tôm 'to boil', etc., examples of which are given in (1.2), which occur with nouns liberally. Therefore, it seems that the compound rules are productive here. However, there is one question that has to be answered: Is a noun plus one of these verbs a compound or, as Marchand (1966) calls it, a syntactic group?

It is certainly very difficult to distinguish between a compound and a syntactic group. For example, European and American grammarians and linguists have long been trying to establish adequate criteria to distinguish between the two. Some of them resort to stress, others to spelling, etc. 5 If the situation is bad in English, it is worse in Thai. Stress is non-phonemic in Thai. To seek morphophonemic stress patterns for compounds is, therefore, out of the question. 6 Spelling offers no help because words are written next to one another without spacing. A syntactic group of two words and a compound consisting of two words look exactly the same. However, I find the following guidelines helpful.

(i) If the second element in the form N+V can't occur with nouns in a liberal way, e.g. if it can occur with only one or a limited number of nouns, it is likely that it is functioning as a componental part of a compound. For example, the verb liːm 'to move', which can be a componental part of a compound but not a modifier, can occur with bandáy 'stair', as in (1.5), but not with tô 'table' and kâːwí 'chair', as exemplified in (1.6).

(ii) If the form N+V is not understood as the sum of the two constituent elements, it must be a compound. For example, there is a conceptual difference between the compound (2.17a) and the non-compound (2.17b).
(2.17) a. phâa-lâŋ (cloth + be yellow)
b. phâa lâŋ (cloth + be yellow)

The former is conceptualized as something worn by a Buddhist monk, i.e. 'Buddhist robe'. The latter is understood as a piece of cloth with a certain shade of color (i.e. yellow), i.e. 'yellow cloth'. This conceptual distinction can be substantiated in many cases. For example, the predicate sîi sôm 'to be orange' constitutes a contradiction if its subject is the non-compound (2.17b), but not if its subject is the compound (2.17a), as illustrated by (2.18a and b).

(2.18) a. phâa-lâŋ phân nîi sîi sôm
   'This yellow cloth (=Buddhist robe) has an orange color.'
b. *phâa lâŋ phân nîi sîi sôm
   *'This yellow cloth has an orange color.'

Returning to the question whether a combination of a noun plus a verb, such as dɔŋ 'to pickle' and thɔt 'to fry', etc., is a compound, my contention is that it is not. The conclusion is based on the following observations:

(i) Syntactically, the second element (i.e. the verb) behaves like a modifying noun, e.g. mây 'wood' and lêk 'steel', etc., in that it can occur with any noun provided that no selectional restrictions are violated and/or the context is appropriate. Several of these verbs (the second elements) can occur in hypothetical forms, i.e. forms which are normally regarded as semantically weird, e.g. rɔŋ thaaw tôm 'boiled shoes' (in the expression 'During the war some soldiers had to eat boiled shoes'). Another example: Suppose that while a cook is frying something some toothpicks happen to fall into the pan; the cook may ask his friend,

(2.19) khun yâk kin mây cîmfan thɔt mây
   you want eat toothpick fry Question
   'Would you like to eat some fried toothpicks?'

(ii) Conceptually, a noun plus one of these verbs is understood the same way as a noun plus a modifying noun or a so-called descriptive verb, in that the meaning of the form is the sum of the two constituent elements.

(iii) As mentioned earlier the verbs in question can occur liberally with a noun. If a noun plus one of these verbs is a compound, there must be a very large number of this kind of compound. However, the Thai Dictionary (1950) has not listed a single N+V in which the V is the verb in question. This certainly casts doubt on the assumption that this kind of N+V is a compound, because normal dictionaries usually list a large number of compounds.

(ii) Predictability
   It is desirable to derive the compound by a rule if the meaning of the compound is always predictable on the basis of its components. In other words, if it is the case that upon knowing the meaning of
the noun and the verb which make up the compound, the speaker will always know the meaning of the compound, then the compound should be derived by a rule, in order to capture the fact that the speaker knows that the compound is related to its components, i.e. the noun and the verb. The following examples show that the meaning is not always predictable.

(2.20) a. rót-tha̯y (vehicle+to plow) 'tractor (used for plowing/farming)
b. rót-khən (vehicle+to push) 'pushcart'
(2.21) a. rooŋ-camnam (building+to pawn) 'pawnshop'
b. tuŋ-camnam (ticket+to pawn) 'receipt listing pawned items'
c. khəŋ-camnam (thing+to pawn) 'pawned items'

rót-tha̯y (vehicle+to plow) is, as described by the verbal component, 'a vehicle which is used for, or is capable of, plowing', but (normally) rót-khən (vehicle+to push) is not 'a vehicle which is used for, or is capable of, pushing'. How can, for example, the semantic rule be prevented from assigning a 'purpose' or 'capability' meaning to rót-khən? Likewise, the second component in the compounds in (2.21a-c) is the same, i.e. camnam, but each of these compounds has a different type of meaning.

Since the compounds may have various types of meaning, unless we can come up with some criteria as to how a semantic rule assigns the correct meaning to the compounds, a lexical rule approach is undesirable.

In short, the evidence concerning the productivity and the predictability argue against deriving the compounds in question by lexical rules.

Based on the two guidelines for distinguishing between a compound and a syntactic group, and my own intuition, the items in (1.1a-g) and (1.3a) are classified as compounds, and (1.2) and (1.3b) are non-compounds. The status of (1.4a-b) have already been discussed earlier.

It should be pointed out that saying that the rule is not productive does not mean that it cannot be used to derive new forms. Since new words are sometimes formed and added to the language, it is possible that any inactive morphological rules may become generative again, at least temporarily. In other words, new compounds of the type N+V may be created on the basis of the already existing forms.

Before discussing the treatment of these compounds in a grammar of Thai, I would like to mention that there seem to be cases of expressions which are in the process of becoming compound
lexical items. One such form is มูยำ 'roast pork', which is coming to be used to refer to a special kind of roast pork, i.e. the kind that has crisp skin and some layers of fat. In such a case, the phrase มูยำ 'roast pork' and the compound มูยำเรียก this particular kind of roast pork are not as easily distinguished as ผ้าลิน 'a cloth worn by a Buddhist monk' and ผ้าคอ 'a yellow cloth'. However, intuitively, the relationship seems to be similar in the two cases. We might predict that the compound มูยำ will take on more idiosyncratic properties which will distinguish it from the phrase มูยำ. Other phrases which are candidates for compound lexical items are ข้าวผัด 'fried rice' ((1.2c)), ข้าวต้ม 'boiled, watery rice' ((1.2d)), and กะยำ 'roast chicken' ((1.2h)).

3. Lexical relation

How does a Thai speaker know that the compound of the type \( N + V \) is related to the independent noun and verb which make up the compound?

First of all the speaker has an internalized knowledge of the lexical structure. In our case, he knows the internal structure of the compound. For example, given the word ข้าวผัด 'airplane', he knows that it may be broken down into two possible words, or in other words, that it consists of ข้าว 'machine' and ผัด 'to fly'. This fact may be represented as (3.1), which is an abbreviated form of a lexical redundancy rule.

(3.1)

\[
\begin{array}{ccc}
N_x & \rightarrow & N_y \\
\downarrow & & \downarrow \\
V & & V
\end{array}
\]

Second, he knows whether the meaning of the compound is related, either directly or indirectly (i.e. metaphorically), to the meaning of the two componential parts. If it is, he assumes that the compound is related to the independent words which form the compound. If not, he will not associate the compound with the independent words. The items in (1.4b) are examples of the latter case. Those in (1.4a) are on the borderline. That is, they may be regarded as compounds by some speakers, but as polysyllabic words by others, depending on extra-linguistic facts (e.g. being told, or from an imagination, etc.), and/or the particular linguistic knowledge of each individual.

Using Jackendoff's (1974) framework, I propose that the meaning relation between, for example, ข้าวผัด 'airplane', ข้าว 'machine, engine', and ผัด 'to fly' may be exemplified by (3.2).

(3.2)

\[
\begin{array}{c}
\text{+N} \\
\text{VEHICLE WHICH HAS} \\
W \text{ AND CAN Z} \\
(\text{or } W \text{ WHICH CAN Z})
\end{array}
\leftrightarrow
\begin{array}{c}
\text{+N} \\
W \\
\text{+V} \\
Z
\end{array}
\]
Concerning lexical representation, every compound of the type discussed will have a full lexical entry. For example, khrān-bin 'airplane' will have the lexical entry (3.3):

(3.3)  

\[
/khr\acute{a}n + bin/  
\]

\[+N\]

Syntactic information

Semantic information

Lexical redundancy rules, both morphological and semantic, will designate the information in the lexical entry which is predictable by the existence of the related lexical items as redundant. The predictable information designated by a lexical redundancy rule is not new information, but is the information the speaker already knows. In other words, not all the information in the lexical entry of, in this case, the compound is new information. Therefore (Jackendoff 1974, p.8):

In knowing two related lexical items one then knows less than when one knows two unrelated items of commensurate complexity.

It is claimed that this hypothesis of the compound of the type N+V, which lists the compound as a full entry, and which incorporates Jackendoff's framework of lexical redundancy rules, captures the speaker's competence concerning the relation between the compound and its components.

4. Conclusion

I have given examples of N+V compound nouns of various semantic types and argued that they should be listed in the lexicon, instead of being derived by transformational or lexical rules. Arguments against the transformational approach arise from (a) constraint on the transformational rules, (b) the criteria for the structural description of the transformational rules, (c) the deletion of meaningful items, and (d) how to arrive at the meaning of the compounds. A lexical rule approach is rejected because the rule which one may want to posit to derive the compounds can only be semi-productive. In addition, the meaning of the compounds is not always predictable from the components. I also pointed out that N+V are of two types, compound and non-compound. Two guidelines were provided for the distinction between a compound and a syntactic group of the type N+V. Furthermore, I hypothesized that some syntactic phrases may have been, or are being, lexicalized. Finally, I discussed how a Thai speaker knows that a N+V compound noun is related to the independent noun and verb which form the compound, and how this knowledge should be captured or represented in a linguistic theory.
NOTES

1. For work of other linguists and grammarians, see, e.g. Gedney (1947), Haas (1966), Noss (1964), and those referred to by Fasold (1969).

2. Semantically, *má-y-tii (wood+to beat) 'a switch used for whipping/beating' should have been a possible compound, because, at least previously, teachers and some parents in Thailand sometimes whipped their students/children with a switch. (The compound for 'a switch used for whipping/beating' is má-y-riaw (wood+to be tapering).)

3. The correct forms of these compounds are:

(2.11) má-práw-bôŋ (coconut+to be dry) 'copra'
      lín-chi-bôŋ (litchi+to be dry) 'dried litchi'
      kûn-bôŋ (shrimp+to be dry) 'dried shrimp'
      pse-bôŋ (fish+to be dry) 'dried fish'

4. For an interesting presentation of lexical rules, see Thompson (1973), in which lexical rules deriving resultative verb compounds in Mandarin Chinese are discussed.


6. Although stress doesn't always work to distinguish between a compound and a syntactic group in English, it sometimes serves as a clue, e.g. blackboard vs black board. In Thai, no such clue can be found in the stress.

7. Further discussion concerning the status of some of these forms is to follow.


REFERENCES


