Cognitive and structural aspects of abstract nominalization in Tai and some other isolating languages of Southeast Asia*

Lev N. MOREV
Institute of Oriental Studies, Russia

Abstract
Abstract nominalization has appeared in the isolating languages of Indochina as a regular grammatical event at first in Tai at most two centuries ago. It signified the pivotal shift in the verbal thinking of Tais that caused substantial diversification of Tai syntax. The Vietnamese and Khmer languages follow Tai suite. Chinese keeps aloof.

The object of examination in this paper is such a rather new phenomenon in the Tai languages as abstract nominalization, which is here understood as formation of abstract nouns from verbs without changing their semantics or, in other words, building noun forms of verbs. The main purpose of the study is to show causal relations between nominalization and syntactic structures in the Tai languages. This event is viewed in the light of cognition and of its impact on Tai syntax. At present, such a nominalization occurs in a number of Tai languages with old written traditions, as Thai (Siamese), Lao, Shan (Tai Long) in Myanmar, Tai Dehong and the other Tai dialects (in Southwestern part of Yunnan province, China) and Lue (in Southern part of Yunnan). At present, nominalization is well fledged and has clear-cut dimensions only in Thai and Lao. This being the case, the nominalization in these two languages will be taken as a graphic example of this phenomenon in Tai. As far as the other Tai languages are concerned, nominalization will be illustrated in passing due to the shortage of linguistic data. In conclusion, a surface observation of what is going on in this respect in the adjacent isolating languages, as Khmer, Vietnamese and Chinese, will be brought up.

Before addressing the investigation it is expedient to remind the readers of some important features of the languages in question, such as the hard, inflexible structure of Tai syllable ruling out the possibility of inflection and derivation in these languages, and negligible quantity of strictly formal exponents of grammatical categories or syntactic relations between words. In such a case, compounding inevitably becomes the main means of word-building and word order assumes responsibility for the relations between actants. But due to the linearity of speech the number of word orders is very limited, i.e. precedence, succession and two zero positions. Therefore the set of

---

*This article is a revision of the paper presented by the author at the Fourth ISPAL in Hanoi, December 2004.
prototypical syntactic structures in Tai theoretically is also very limited, i.e. SVO, OSV with postpositional attribute. The specific feature of Tai and of some other isolating languages in the Indochinese peninsula is that until recently the subject and other actants of utterance as a rule were displayed by substance or material words with denotative meaning. One utterance represented one event consisting of one agens (agent) and one predicate. Diachronically, Tai syntactic structures developed and varied chiefly through serialization of verbs, extension of attributes, introduction of passive constructions, coalescence of sentences and so on (for details see Diller 1988). Such a state of things existed up to the adoption of methods for abstract nominalization by means of certain morphemes. Now we move on to examination of the problem in question.

Abstract nominalization in Thai and Lao manifests itself through morphemes kaan and khwaam. The first one kaan is derived from Sanskrit kara ‘business’, ‘work’ and the other one khwaam is an indigenous word with a very vague meaning ‘matter’, ‘case’, ‘statement’, ‘speech’, ‘word’. The former is used to convert verbs, particularly action verbs, into abstract nouns, or into the names of actions. And the latter is applied to turn nonaction verbs, i.e. verbs denoting states, feelings, qualities and properties, into abstract nouns as well. In some cases both can be used. But then the meaning of compounds will be different, compare Thai/Lao kaan1-taai1 ‘dying’, ‘death’ as a name of process and khwaam1-taai1 ‘death’ as an abstract notion opposed to ‘life’, kaan-hen ‘seeing’ versus khwaam-hen ‘opinion’.

In practical grammars these two morphemes usually are treated as prefixes to form abstract nouns. But such a definition looks not quite correct, particularly with respect to kaan. The combination kaan+verb retains not only the meaning of the verb but its grammatical properties too. For example, it still governs the object directly as before: Thai/Lao kaan1 phææ2 am1-naat3 ‘spreading power;’ can be accompanied by modifier with auxiliary particle jaang2/doi1: kaan1- phææ2 am1-naat3 jaang2 wai1 ‘spreading power rapidly’. It can also nominalize a word-group or a sentence, e.g. Thai

1. kaan1-tok2 pen thaat3 khɔɔŋ5 khwaam1- praat2tha2-naa5
   Nom. fall be slave Possess. Nom. desire
   ‘becoming a hostage of desire’

Besides that, it can nominalize verb accompanied by modus or modal units, e.g. Thai

2. kaan1- ca2- jiam3-jian1 khɔɔŋ5 naa1-jok4
   Nom. Fut. visit Possess. premier
   ‘The forthcoming visit of the premier’

3. kaan2- thææk6- dai3 mii2 khwaam1 - sam1-khan2 too5
   Nom. measure can have Nom. important towards
   kaan2- dam2-long2 tfii2-vit3 khɔɔŋ1 khon2 raw2
   Nom. lead life Possess. man we
   ‘The measurability (of things) is of great importance in our life’
The collocation of *kaan* + verb sometimes undergoes semantic shift and assumes meaning different from the meaning of the initial verb. For instance, in Thai/Lao, collocation *kaan-bin* besides the meaning ‘flying’ acquired meaning ‘aviation’. So, in the first event morpheme *kaan* acts as a functional transpositor, and in the second event the same unit acts as a morphological prefix to form noun. In some cases difference between these two functions is hardly visible. But for our purpose these particulars are not so very important.

The *khwaam* word-groups are more close-packed and behave very much similar to the rest of the nouns. The incorporation of some modals between *khwaam* and verb, as in the next example from Thai, is an extremely rare exeption. It can be viewed as an exclusive event with inferior linguistic value.

4. khwaam\(^1\)- naa\(^3\)-ca\(^2\) sia\(^5\)-haai\(^5\) ko\(^2\) mai\(^3\) hen\(^5\) k\(\text{œ}t\)^{2} kyn\(^3\)
Nom. should lose Part. not see happen Dir.
‘However expected losses did not happen’

Thus, we can infer that morpheme *kaan* is a syntactic transpositor rather than a derivational prefix. As to *khwaam* it can be treated, with a proviso, as a morphological means used to form abstract nouns.

All above said about *kaan* and *khwaam* in Thai is true for Lao as well. The first acquaintance with the old texts in Lao in search of *kaan* and *khwaam* has not brought results. It seems that nominalization in Lao emerged quite recently. Lao has Thai to thank for nominalization.

The genesis of *kaan* and *khwaam* in Thai is thoroughly scrutinized in the works of Amara Prasithrathsint (Amara 1995, 1996). Her research indicates, at first morpheme *khwaam* was used as nominalizer in Thai in the end of the 13\(^{th}\) c., i.e. *khwaam\(^1\)-suk\(^2\)* ‘happiness’ in the inscription of Ram Khamhaeng, and morpheme *kaan* - in the latter half of the 17\(^{th}\) c., i.e. *kaan\(^1\)-rop\(^3\) syk\(^2\) kam\(^1\)-pan\(^1\)* ‘fighting in a battleship’ (Amara 1995;5). Up to the middle of the 19 c. both of them occurred extremely rarely. But after Siam (Thailand) embarked on modernization after the European style in the last quarter of the 19 c. Siamese establishment turned to adopt Western values including the English language. The Thais had to search for Thai equivalents to match English abstract nouns suffixed with -ing, -ment, -tion, -ness, etc. Morphemes *kaan* and *khwaam* has turned to be quite suitable for this purpose.

Since then frequency of *kaan* nominalization increased rapidly, i.e. from 0.08 occurrences (per one hundred words) in 1672 to 0.58 in 1992, meanwhile the frequency of *khwaam* nominalization oscilated unevenly, i.e. from 0.50 in 1872 rose to 1.94 in 1892, then sank to 0.46 in 1902 and the latest calculation of 1992 showed 0.56 (Amara 1995; 7). Amara noticed that the frequency of *kaan* and *khwaam* varied in the writings of different genres. She distinguished three styles: formal style (newspaper editorials), semi-formal style (narratives in fiction) and nonformal style (conversations in fiction). According to her subsequent findings, from 1872 till 1992 the frequency of *kaan* increased from 0.12 to 2.66 in editorials, from 0.07 to 0.17 in narrations
and from 0 to 0.12 in conversations; during the same period span of the frequency of *khwaam* changed from 0.77 to 1.28 in editorials, from 0.07 to 0.55 in narrations and from 0.24 to 0.12 in conversations (Amara 1996; 1213-1214).

Amara looked upon nominalization from the position of sociolinguistics and came to a conclusion that “nominalization is a linguistic device that differentiates styles in Thai” (ibid. 1207). Here, however, as it has been said above, I view the nominalization from the angle of cognition. For this purpose it is necessary to take a look at the nature of the Thai language. A survey of old writings in Tai (Thai and Lao) has attested that the ancestors of modern Thais and Lao intuitively divided their content words into two big groups: the words that could be used separately as self-sufficient entities and the words that were usually used as attributes or adjuncts to the first. Such division corresponds to the division of full words into substance words and feature words. The formal difference between the words of these two groups in Tai lies in that the latter can collocate directly with negation and the former, on the contrary, cannot do this. So, as one can easily guess, reference here is to nouns and verbs in a broad sense of the word.

A review of old writings in Thai has also demonstrated that the topic of an utterance or the subject of a sentence were usually denoted by substance words, or by the words that were able to constitute a separate object of cognition, whereas feature words could not fulfill theta roles and thus could not represent separate objects of verbal thinking. As a rule, they were used to perform predicative or attributive functions. The old manuscripts in Tai attest that Thai vocabulary of Middle Ages, except numerous religious terms, included a very small number of words with abstract meaning. The bulk of Thai substantive words belonged to the so-called specific or concrete nouns. Words of this kind mainly assumed theta roles, especially that of agents or agent of an utterance. Then this function became inherent prototypical for them.

Nominalization is not a matter of technique, i.e. to find an adequate means for, figuratively speaking, changing verb dress for noun dress. It is a matter of cognition. “The task is to strip properties and features from their carriers and imagine them as a substantive object of our thought” (Cherneiko 1997:38) and then “to shape them into such form that would permit them to occupy position of subject in the sentence.” (ibid: 60). The human mind is able to fill up such gap and to make ends meet, though it would take decades or even centuries for adaptation.

The emergence and formation of deverbal abstract nouns (*kaan/khwaam*-nouns) was a landmark in the history of Tai languages. It was a pivotal event for the Thai/Tai verbal thinking. It meant that henceforth both concrete and abstract notions could be the substantive objects of language consciousness, and their exponents, both concrete and abstract nouns, could assume theta roles. What is more, the adoption of the nominalization device had a crucial effect on Thai and Lao syntax. This kind of verbal noun can fulfill any syntactic role in the sentence the same as any other noun. Such nouns maintain the valency of original verbs, thus each of them is able to
increase the number of actants in a sentence equal to the number of its valencies and thus complicate both semantic and grammatical structures of utterance or sentence. It has manifested itself in many aspects.

First of all, nominalization made it possible to change syntactic relations between words and consequently change their syntactic positions without changing semantics of utterance. Compare two propositional structures in Thai:

5a. khææk² dɔɔn¹-thaang¹ maa¹ lææw⁴
guests travel come Perfect
‘The guests have come already’

5b. kaan¹- dɔɔn¹-thaang¹ maa¹ khoɔng⁵ kheek²
Nom. travel come Possess. guests
‘The coming of guests’

6a. myang¹ nii⁴ suai⁵ maak³
town this beautiful very
‘This town is very beautiful’

6b. khwaam¹- suai⁵ khoɔng⁵ myang¹ nii⁴
Nom. beautiful Possess. town this
‘The beauty of this town’

The propositional meanings of 5a and 5b are identical, but 5a is agens (guests) orientated and 5b is predicate (coming) orientated, whereas grammatically agens functions as attribute to predicate. Thus, the semantic structure of 5b looks like a mirror reflection of 5a. The difference between 5a and 5b is that 5b lacks sentential marker lææw⁴. But the directional morpheme maa¹ ‘to come’ is preserved, which confirms its semantic status within the utterance. Phrases 6a and 6b can be analized similarly.

Before the adoption of nominalization each Thai utterance usually represented one event and consisted of two parts: agens and predicate. The nominalization, as has been said above, made way for complicated semantic and syntactic structures. The following Thai sentence will illustrate it.

7. khaw⁵ son⁵-cai¹ tɔɔ² kaan¹- son⁵-tha⁴-naa¹
they be interested towards Nom. talk
khɔɔng⁵ raw¹ maak³
Possess. we very
‘They take a great interest in our talk’

This utterance includes two events: khaw⁵ son⁵-cai¹l “they take an interest” and raw¹ son⁵-tha⁴-naa¹l “we talk” and has two agensis: khaw⁵ “they” and raw¹ “we”. In order to combine two utterances it was necessary to convert utterance raw¹ son⁵-tha⁴-naa¹l “we talk” into kaan¹-son⁵-tha⁴-naa¹l ‘talk’, ‘conversation’ first and then make it a part of the whole by using right-side valency of the son⁵-cai¹l ‘take an interest’. Otherwise it would not be
possible to perform such an operation. This is an ordinary sample of the
nominalization device at work.

The nominalization also permits description of the complex situation
synthetically, not as a sequence of separate events. Below is an example of
synthesized sentence in Lao.

8. kaan²- het¹ khua¹ khaam⁶ nam³ het¹ hai³ thai² baan³
   Nom. make bridge cross river Caus. people village
   pai² maa² haa¹ suu⁵ kan² dai³ sa¹-duak⁵
go come seek for each other able easy
   ‘The building of a bridge across the river made contacts between
   villagers easy’

This sentence absorbed the content of two sentences connected
by taxis relations:

9. (khaw¹) het¹ khua¹ khaam⁶ nam³
   (they) make bridge cross river
   ‘(they) made a bridge across the river’

10. thai² baan³ (cyn¹) pai² maa² haa¹ suu⁵ kan²
    people village (then) go come seek for each other
    dai³ sa¹-duak⁵
    able easy
    ‘Villagers can visit each other easily’

Such was the interpretation of the above cited synthesized sentence by
the speakers of some Tai languages in Laos destitute of nominalization device
(personal field researches in Laos, 1980’s).

On the other hand, the nominalization made it possible to expand
some parts of the sentence, i.e. to substitute a verb phrase for a bare verb, as in
Thai:

11. hai³ kaan¹- tōn³-rap⁴ instead of toon³-rap⁴
    ‘to give welcome’ instead of ‘to welcome’

12. tham¹ kwaam¹- ruú⁴-cak² kan¹ instead of ruú⁴-cak² kan¹
    ‘to make the acquaintance’ instead of ‘to get acquainted’

The verbs hai³ ‘to give’, tham¹ ‘to make’ in such constructions
account for predicativity as such, while the content of the predicate manifests
itself in a verbal noun. The substitution of a binominal construction for a bare
verb does not change the semantics of the utterance, but makes it more
flexible, ready for the further modifications.

The employment of an nominalized verb instead of a bare verb also
permits reversal of the course of thinking from one object to another. Compare
two Thai sentences:
13. fuung⁵-chon¹ top² myy¹ hai³ khaw⁵
   people clap hands Dat. he
   ‘The people applauded him’

14. khaw⁵ dai³-rap⁴ kaan¹-top² myy¹ caak² fuung⁵-chon¹
   he receive Nom. clap hands from people
   ‘He deserved applause from the audience’

These two utterances share the same semantics, but differ in the orientation of process. Example 13 is aimed at “audience”, whereas example 14 is aimed at “he”. The same effect can be attained by “passivization”, but the point is that the utilization of the passive construction in Thai is restricted due to semantic reasons, i.e. it mostly bears adverse meaning. Besides this, such compounding is also used to slow down the density of information when the speaker finds it necessary.

Apart from Thai and Lao, nominalization has also been recorded in some other languages of the Tai family, as Shan (Tai Long), Tai Nya and some other Tai dialects spoken in the areas along both sides of the Chinese-Myanmar border. For this purpose all of them use indigenous words, first of all morpheme taang, whose original meaning is ‘way’, ‘route’, ‘track’, e.g. Shan taang⁴-laat⁴ ‘speech’, ‘way of speaking’ < laat³ ‘to speak’; Dehong taang⁴-phit⁶ ‘mistake’ < phit⁶ ‘to be mistaken’, Tai Nya taang⁴-han¹ ‘look’, ‘glance’ < han¹ ‘to see’, Lue taang⁴-caang⁵ ‘skill’, ‘ability’ < caang⁴ ‘know how’, ‘to be able’. It looks as if this way of nominalization in these languages is not a recent innovation, at least, it is a few centuries old. It has been attested by J.N. Cushing in his ‘Grammar of the Shan language’ (Cushing 1871). Besides this shared morpheme Lue also uses the morpheme khwaam in the same way as in Thai and Lao, e.g. khwaam²-hu⁶ ‘knowledge’ < hu ‘to know’. It seems that the usage of khwaam is a recent borrowing from Lao or Thai. At last, nowadays in Tai Dehong, Tai Nya and in the other local dialects the morpheme l̄ng² is used on a par with taang², e.g. Dehong long²-li¹ ‘goodness’ < li¹ ‘good’; Tai Nya l̄ng²-yaaap⁴ ‘hardships’ < yaap⁴ ‘difficult’, ‘hard’. The original meaning of l̄ng² is ‘matter’, ‘work’, ‘business’. It looks as if l̄ng²-words are a calque from Chinese compounds with morpheme shi, which has the same meaning as Tai morpheme l̄ng². Compare Dehong l̄ng²-li¹ ‘goodness’ with Chinese hao shi ‘good deed’.

Close examination of nominalization in these languages reveals that the words most apt to undergo nominalization are nonaction, stative verbs rather than action, dynamic verbs. The nominalized verbs for the considerable part denote physical or spiritual entities and are not sheer names of actions or states. For instance, Shan taang⁴-nung³ < nung³ ‘to wear (clothes)’ has the meaning ‘clothing’, ‘clothes’, but not ‘wearing (clothes)’; taang⁴-kin¹ < kin¹ ‘to eat’ in Lue has the meaning ‘foods’, but not ‘eating’. So, I suppose that for these languages the definition ‘substantivation’ would be more adequate than ‘nominalization’ in the sense specified above.

The above discussion was aimed at demonstrating the mechanism of abstract nominalization in the Tai languages and the effects of it.
The data presented by Amara Prasitrathisint in respect to the use of abstract nominalization in Thai confirms that its frequency is still rather low. “Even though, Amara notes, kaan and khwaam nominalizations have been used for centuries, frequent use of them is often criticized by Thai language authorities. They are regarded as unnecessarily elaborate and causing the Thai language to degenerate” (Amara 1996:1217). If that is the case, we can infer that for the Thai mind accidental concepts are still hardly perceived as substantive objects.

The phenomenon of abstract nominalization is akin not only to the Tai languages. Nowadays it is a shared feature of some other languages of East and Southeast Asia, e.g. Vietnamese, Khmer and Chinese. For this reason it would be very much to the point to make a short survey of this problem in the above mentioned languages.

At present scholarly circles are of the opinion that the Vietnamese language belongs to the Austroasiatic family which was well known for its fully fledged morphology. It means that once Vietnamese had a set of derivational affixes too. However it has not preserved them and now it resorts to compounding as a main means of word-building, including formation of abstract verbal nouns. For this purpose a number of Vietnamese or borrowed morphemes with broad or neutral meaning are used, i.e. indigenous morphemes cãi ‘thing’, ‘object’, cuọc ‘event’, ‘matter’, việc ‘work’, nỗi ‘condition’, niểm ‘feeling’, Chinese morpheme sụ́ ‘work’, ‘business’. For instance: cãi chết ‘death’ < chết ‘to die’, cuộc dũ tranh ‘struggling’ < dũ tranh ‘to struggle’, việc học ‘learning’ < học ‘to learn’, nỗi lo ‘anxiety’ < lo ‘anxious’, niểm vui ‘happiness’ < vui ‘happy’. Besides these there are some more words that are used for the purpose of nominalization occasionally. A rather long list of such morphemes and various interpretations of them (Mhitarian 1967, Bystrov 1975, Panfilov 1993, Hoang Trong Phien et al. 1970) testifies that up to now there is no common, strict linguistic rule of nominalization in Vietnamese. It is considered among Vietnamese speaking people that nominalization is characteristic for official and socio-political writings but not for fiction or everyday talk. Anyhow, nominalization is a real fact of Vietnamese, accounting for the diversification of syntactic structures in Vietnamese the same way as in Tai.

Unlike Vietnamese, the Khmer language still retains certain evidences of once well developed derivation. Modern Khmer inherited a number of former derivational affixes including infix -aʊm- and its allomorphs which had been used as means of transposition of verbs into abstract nouns, e.g. c-aʊm rout ‘reaping’ < crout ‘to reap’, c-aʊm-ih ‘riding’ < cih ‘to ride’, c-aʊm-aot ‘stupidity’ < ksaoy ‘stupid’, k-aʊm-saoi ‘weakness’ < ksoi ‘weak’. But in modern Khmer this derivational mechanism is no longer used practically. It has become a relic of the past. At present in order to convert verbs into abstract nouns the Khmers use some Sanskrit morphemes whose lexical meaning is not so transparent for the Khmer-speaking people, among them kao < kara ‘work’, ‘business’, kec < kesa ‘work’, ‘business’, ceckdey (could not find out Sanskrit counterpart).
For instance: ka-baŋviə ‘rotation’ < baŋviə ‘to rotate’, kec-
saŋkrual ‘helping’ < saŋkrual ‘to help’, seckdey-kənsak ‘cowardice’ < kənsak ‘coward’.

On the whole, nominalization in Vietnamese and Khmer develops in the same way as in Tai, i.e. by means of adaptation of indigenous or foreign morphemes to build abstract verbal nouns. This phenomenon in these two languages is rather new and unsettled, following from the multiformity and diversity of means for nominalization and from the shortage of strict rules of their usage.

The most equivocal situation in this respect is in the Chinese language. It is mostly due to the fact that the problem of parts of speech in this language is still pending. At present there are quite opposite views on this problem: some researchers distinguish the same parts of speech as in European languages (Lu Shusiang, Dragunov, Solntsev etc.), some scholars deny their existence in Chinese at all (Gao Mingkai, F. Maspero etc.). Skeptical attitudes to the parts of speech in Chinese comes from the fact that the majority of content words in Chinese do not bear any marks of a certain part of speech and are able to fulfill different syntactic functions, e.g. agens, predicate, object, attribute, while in many languages these functions are assigned to certain word classes. This phenomenon usually is defined as grammatical homonymia or polyfunctionalism. For example, lexeme gaibian can function as a verb ‘to change’, ‘to transform’ and as a noun ‘changing’, ‘transformation’. But generally speaking, Chinese tends to reduce the number of homonymous or polyfunctional words by natural distribution of syntactic functions among lexical units.

Half a century ago Chinese linguists compiled a list of the most frequently used words numbering three thousand entries. There were only about thirty words which could equally function both as nouns and as verbs. The rest of the would-be polyfunctional words were apt to occur as words of a certain lexical class (cited after Korotkov 1968;72). The attribution of noun properties to lexical unit can be also called “nominalization”. But in Chinese it manifests implicitly without displaying on the surface level. Such a state of things is quite natural for Chinese. Polyfunctionalism or grammatical homonymia, as this phenomenon has been defined by linguists, existed in Chinese long since. Korotkov supposed that such a conversion in Chinese was due to the fact that in Chinese a good number of content words “had several counterparts with similar or identical meaning that permitted the language to differentiate them not only semantically but functionally too” (ibid). It seems that in the verbal thinking or language consciousness of the Chinese there is a certain device which automatically selects one of homonymous words. But how this device works it is still a mystery of cognition.

Thus, the phenomenon of abstract nominalization or functional transposition has become a fait accompli in the most important languages of Indochina, e.g. Thai, Vietnamese, Khmer, Lao (Burmese was not considered) and now is gaining momentum rather rapidly under the influence of English globalization. Though native speakers of these languages still take nominalization as somewhat alien, nonetheless we witness a certain shift in the
verbal thinking of the speakers of the above mentioned languages. Thanks to abstract nominalization, the syntax of these languages has undergone serious modifications and is gaining some new features. Meanwhile, Chinese keeps aloof.

REFERENCES


Hoang Trong Hien. 1970. Slovoobrazovatelnye Affiksy I Poluaffiksy V Sovremennom Vietnamskom Yazyke (Derivational Affixes and Half Affixes in Modern Vietnamese). Yazyki Yugovostochnoi asi (The Languages of Southeast Asia), Moscow, Nauka Pblishing House.


Received: 12 January 2005

Institute of Oriental Studies
12 Rozhdestvenka Street
103753 Moscow
RUSSIA