Grammaticalization and the areal factor —

the perspective of East and mainland Southeast Asian languages

Walter Bisang

University of Mainz

Walter Bisang
Department of English and Linguistics
University of Mainz
Jakob-Welder-Weg 18
D-55099 Mainz
Germany

Phone:  ++49 6131 392-2778
Fax:      ++49 6131 392-3836
E-mail: wbisang@mail.uni-mainz.de

I use:
Bold for emphasis
Small capitals for marking Focus
1. Introduction - properties of grammaticalization in East and mainland Southeast Asia

In a large number of languages in East and mainland Southeast Asia, grammaticalization is characterized by the following three characteristics:

- Lack of obligatory categories and predominance of pragmatic inference even in the case of highly abstract grammatical concepts such as tense or definiteness
- Existence of rigid syntactic patterns (word-order patterns)
- No or limited coevolution of form and meaning

The definition of obligatoriness adopted in this paper is that of Lehmann (1995). A category is obligatory if the speaker is forced to specify that category by selecting a marker that belongs to it:

By this [i.e. transparadigmatic variability; W.B.] we mean the freedom of the language user with regard to the paradigm as a whole. The paradigm represents a certain category, and its members, the subcategories (or values) of that category. There may then be a certain freedom in either specifying the category by using one of its subcategories, or leaving the whole category unspecified. To the extent that the latter option becomes constrained and finally impossible, the category becomes obligatory. We shall therefore use the term 'obligatoriness' . . . (Lehmann 1995: 139)
The lack of obligatoriness is particularly remarkable in cases where the concept inferred is an abstract grammatical concept that is expressed by obligatory categories in Indo-European languages. While these functions are conventionalized in Indo-European, they are the product of pragmatic inference in many markers of East and mainland Southeast Asian languages. This is corroborated by the fact that in a number of cases one and the same marker may express different grammatical concepts in different situations or in different constructions (cf. the functions of Khmer បាល ‘come to have’ in subsection 2.2 or the functions of classifiers in Thai in subsection 2.2). As a consequence, the synchronic representation of the relation between the different functions of a grammatical marker is not that of a cline or path of grammaticalization with one function being cognitively related to its succeeding function (e.g. “object > space > time > quality” in Heine, Claudi & Hünnemeyer 1991: 65; also cf. (14)) but rather that of an initial source concept that simultaneously radiates into different directions (cf. subsection 2.1; other authors arguing against the existence of a cline in grammaticalization in East and mainland Southeast Asian languages are Ansaldo 1999 on ‘surpass’ comparatives and Enfield 2003 on ‘come to have verbs’). Even if one looks at the diachronic development of these markers it does not seem to follow a neat cline or pathway.¹ Part of this diachronic picture may not only be due to pragmatics but also to the highly complex contact situation in East and mainland Southeast Asia (cf. Enfield 2003 on ‘come to have’, also cf. subsection 4.1).

In spite of their lack of obligatoriness, grammatical markers follow very rigid word-order patterns in East and mainland Southeast Asian languages and they belong to relatively closed classes that must take certain positions within constructions. Rigid word order as
well as closed-class membership clearly indicate a high degree of grammaticalization (cf. Lehmann 1995 on paradigmaticity and syntactic variability; also cf. subsection 3.1). Thus, there is a class of markers for which there can be no doubt about its high degree of grammaticalization. These markers clearly express grammatical functions but they are not obligatory and the grammatical concepts they express are the result of inference (cf. the examples in subsections 2.1 and 2.2).

As was pointed out by Ansaldo & Lim (2004) for Sinitic languages, grammaticalization is not primarily expressed at the level of morphological reduction but rather by phonetic erosion in terms of duration and vowel quality. The two authors argue very plausibly that this is due to the discreteness of syllable boundaries and to phonotactic restraints. In my view, these phonological properties together with the broad functional spectrum of one and the same marker and the lack of obligatoriness prevent the development of morphologically integrated paradigms as we know them from Indo-European languages (cf. subsection 4.2). The consequence of both properties is that the degree of erosion attested in East and mainland Southeast Asia is not strong enough in most cases for a parallel development of meaning and form. This can be seen synchronically by looking at the data and—as a consequence—must also be true diachronically for the development of these markers.

As I would like to show in this paper, the above three properties of grammaticalization cannot be fully accounted for in any of the standard approaches to grammaticalization described in section 3. If this is true and if these properties are areal (cf. Bisang 1996), i.e., if they are widely attested in several language families spoken in East and mainland Southeast Asia, one may ask the more general question of whether the existing approaches
to grammaticalization are somehow areally biased. Could it be that certain properties have never been analysed or did never find their way into typologically-oriented questionnaires because they do not exist in familiar Indo-European languages? From such a perspective, one might assume that there exist different types of grammaticalization in which different strategies are of different importance. As I would like to show (cf. subsection 4.2), the languages to be discussed in this paper are characterized by the predominance of inference.

This paper will be structured as follows. Section 2 will present three case studies of grammaticalization that show the high importance of inference in East and mainland Southeast Asian languages. Section 3 will discuss to what extent Lehmann (1995), Heine et al. (1991), Bybee et al. (1994) and Hopper & Traugott (1993) can account for the processes of grammaticalization attested in these languages. As I shall try to show within the limits of space I have, none of these approaches fully works. Section 4 will briefly address the question of areality and the question of why the emergence of morphological paradigms is unlikely in East and mainland Southeast Asia. Finally, I will show in a brief outlook (section 5) how grammaticalization remains interesting after Newmeyer’s (1998) deconstruction if one integrates the areal perspective.

2. Examples

This section is a sketch of two phenomena that illustrate the properties of grammaticalization in East and mainland Southeast Asia. The wide range of inferences that can be triggered by one and the same marker will be discussed in subsection 2.1 on the verb
‘come to have’. The next subsection presents Thai classifiers, which represent another case of multiple inference. In this case, however, the different functions are associated with different constructions. Thus, subsection 2.2 highlights the relevance of constructions for grammaticalization. The last subsection (2.3) very briefly illustrates the rigid word-order patterns that determine the preverbal and postverbal positions of three different types of grammatical markers.

2.1. Inference of several different grammatical functions—

the case of Khmer **ban** ‘come to have’

The lack of obligatory categories and the wide range of inferences associated with highly grammaticalized items will be illustrated in this section by the verb **ban** ‘come to have’ in Khmer. As is shown by Enfield (2003, also cf. Bisang 1992, 1996), verbs with that meaning are extremely wide-spread across the languages of East and mainland Southeast Asia. Since the data on ‘come to have’-verbs are extensively discussed by Enfield (2003) and since I have discussed the data on Khmer somewhat more extensively elsewhere (Bisang 2004: 118 - 121), I will only give a summary to provide the reader with the information most relevant for this paper.

Like most ‘come to have’-verbs in mainland Southeast Asia, Khmer **ban** occurs preverbally as well as postverbally/clause-finally. Since a look at the preverbal position will be sufficient to illustrate the multifunctionality of **ban**, I will not discuss the postverbal/clause-final position.
The basic meaning of ‘come to have’-verbs is described as follows by Enfield (2003):

The predication made by the expression ‘come to have’ involves two important components, namely an event of something coming into one’s sphere of possession, and a subsequent and consequent state of possessing that thing. ...

‘Come to have’ treats the recipient (x) as the ‘primary figure’ in the event (i.e. it is encoded as grammatical subject), and there is no necessary reference to the source of transfer, nor is it specified whether the thing moves to the recipient or the recipient moves to the thing. (Enfield 2002: 38 – 39)

In the following example, Khmer ba:n ‘come to have’ is used as a full verb:

(1) Chnam nǐh yēŋ ba:n sroːŋ craːn nas.

year this we come.to.have rice a.lot very

‘This year, we came to have a lot of rice.’

If a ‘come to have’-verb occurs as a grammaticalized item in the preverbal position, its interpretation depends on a number of presuppositions. If the event denoted by the main verb is supposed to be desired the ‘come to have’-verb expresses ability or permission (2a). With other presuppositions, ‘come to have’ triggers the inference of obligation (2b), past (2c) or truth/factuality (2d):
(2) Possible inferences of ‘come to have’ in East and mainland Southeast Asian languages (revised version of Bisang 2004: 119):

a. The event E is [+desired]:
   
   $\rightarrow$ modal interpretation: ‘can’ (potential meaning: abilitative or permissive)

b. The event E is [-desired]
   
   $\rightarrow$ modal interpretation: ‘must, to have to’ (obligation)

c. In order for X to come to have E, E must have taken place:
   
   $\rightarrow$ Past (E) (particularly if E is negated)

d. In order for X to come to have E, E must be true:
   
   $\rightarrow$ truth, factuality

In Khmer, the verb $\text{ban}$ ‘come to have’ can trigger the interpretations of ability and permission (2a), of past tense (2c) and of truth or factuality (2d). Inferences in terms of obligation (2b) are marginal in Khmer. Thus, a constructed example like (3) can be interpreted as follows:

(3) $\text{ban}$ in preverbal position: ability/permission, past, truth:

$\text{khpom } \text{ban } t\text{ûu } \text{phsa}(r).$

I come.to.have go market

According to (2a): ‘I was able to go to the market.’ / ‘I was allowed to go to the market.’
According to (2c): ‘I went to the market.’

According to (2d): ‘I do go/I really go to the market.’

The different interpretations given in (2) can even be combined to a certain extent. Thus, (3) can also be translated as ‘He did go to the market’ (interpretations (2c) and (2d)).

The next example is from a novel. It consists of two instances of \( \text{ba:}n \). In its first occurrence (\( \text{ba:}n_1 \)), past tense interpretation is excluded because \( \text{ba:}n_1 \) is immediately preceded by the future marker \( \text{nùŋ} \). Since it is known from the context that the event marked by \( \text{ba:}n_1 \) is desired, the abilitative interpretation is the most likely interpretation. This also applies to the second occurrence of the ‘come to have’-verb (\( \text{ba:}n_2 \)), although a past interpretation or a factual interpretation cannot be excluded (cf. Bisang 2004: 120):

(4) Khmer \( \text{ba:}n \) in preverbal position (Bisang 1992: 414):;

\[
\begin{align*}
cû:ən-ka:l & \quad nî:ɔy \quad kɔ: \quad nûk-søŋkhu¨m \quad tha: \quad \text{ba:}n_1 \quad cû:ɔp \\
\text{sometimes} & \quad \text{she} & \quad \text{then} & \quad \text{think-hope} & \quad \text{QUOTFUT} & \quad \text{come.to.have.meet} \\
\text{borɔs} & \quad nûh & \quad tî:ɛt, & \quad \text{tae}-kɔ: & \quad pûm & \quad \text{ba:}n_2 \quad cû:ɔp & \quad dö:c & \quad \text{bɔmːnɔŋ}. & \quad \text{man} \quad \text{that} \quad \text{again} & \quad \text{but} & \quad \text{NEG} & \quad \text{come.to.have.meet} & \quad \text{as} \quad \text{wish} \\
\text{‘Sometimes she hoped to be able to meet that man again, but she wasn’t able to/} & \quad \text{DIDN’T meet [him] as she wished.’}
\end{align*}
\]

To conclude this subsection, I would like to point out some basic properties of grammaticalization that can be illustrated by Khmer \( \text{ba:}n \). There is no obligatory tense-
aspect-modality marking (TAM) in Khmer (and in other languages of East and mainland Southeast Asia). Thus, a clause such as (4) in Khmer is also grammatical (but pragmatically less adequate) if the categories marked by *baːn* are omitted. The interpretation of *baːn* depends on context. In principle, *baːn* can express all the functions listed under (2), with function (2b) being rare. Given their inferential character, the different interpretations of *baːn* should not be treated in terms of polysemy or semantic change. In addition, the various functions of *baːn* are not necessarily related to each other by a cline or a pathway, they rather seem to be inferred from the same source concept of ‘come to have’ in different contexts with different presuppositions. The diachronic development of *baːn* does not seem to follow a strict cline either—but this needs more research. Finally, *baːn* basically keeps its phonetic substance even though it expresses grammatical functions and occurs in a position for markers expressing grammatical categories.

2.2. Inference of different grammatical functions depending on the constructional context—the case of classifiers in Thai

The languages of East and mainland Southeast Asia are characterized by their transnumerality, i.e., their lack of obligatory number marking. Thus, a word like Thai *còtmaːaj* only refers to the concept of »letter« as such without any reference to number. As was pointed out by Greenberg (1974), there is a typological correlation between transnumerality and the existence of classifiers:
Numeral classifier languages generally do not have compulsory expression of nominal plurality, but at most facultative expression. (Greenberg 1974: 25)

This correlation applies to Chinese, Japanese, Thai, Hmong and a large number of other languages of East and mainland Southeast Asia. Thus, Thai còtmāaj ‘letter’ needs to cooccur with a classifier in the context of counting:

(5) a. còtmāaj sāam chabāb
    letter three CL
    ‘three letters’

b. *còtmāaj sāam
    letter three

Cross-linguistically, classifiers minimally occur with numerals. In a number of East and mainland Southeast Asian languages, they have additional functions which are all based on the function of classification, i.e., a cognitive activity whereby items of the world and their lexical representations are assigned to certain classes on the basis of certain properties or criteria. The most relevant criteria of classification are material (human, animate, inanimate, abstract), social status, physical properties (shape, consistency, size), functionality (man made) and location (buildings, places, etc.) (also cf. Denny 1976, Allan 1977, Aikhenvald 2000, Grinevald 2000).
The process of classification can be used to profile conceptual boundaries of concepts. Due to this function, classification is the basis of the two main functions involved with numeral classifiers, i.e., identification and individuation. On the one hand, classification helps identifying a certain sensory perception by using its conceptual boundaries to highlight that perception against other sensory perceptions. On the other hand, it can establish a sensory perception as an individual item by actualising its salient inherent properties which constitute it as a conceptual unit.

Apart from their function of individuation in the context of counting, Thai classifiers can assume a number of other functions related to the identificational function of classifiers. In these functions, the classifier is always optional, while it is compulsory in combination with numerals. As opposed to the various functions of Khmer *baːm* ‘come to have’, the different functions of Thai classifiers depend on the context in which they are used. The following table indicates the basic cognitive functions of the classifier, its concrete functions and the constructions with which the concrete functions are associated:

**INSERT TABLE 1 ABOUT HERE**

The individuating function of the classifier is illustrated by example (5a). The singulative interpretation of the classifier is limited to the demonstrative construction. If there is no classifier as in the case of [N DEM] the noun can be singular or plural, i.e., it remains transnumeral (6a). If the classifier is present, the construction has the structure of [N CL DEM] and a singular interpretation can be inferred (6b):
(6) Demonstrative construction [N (CL) DEM]

a. ṛót  nū
   car  this
   ‘this car’ / ‘these cars’

b. ṛót  khan  nū
   car  CL  this
   ‘this car’

In combination with stative verbs, which will be called adjectives here, the presence of the classifier triggers referential interpretation in terms of definiteness or specificity. In the adjective construction with no classifier [N ADJ], the noun is referentially neutral, i.e., it can be definite, specific, indefinite, etc. depending on the context (7a). If the classifier is present, the construction has the structure of [N CL ADJ] and the noun is interpreted as definite or specific (7b). In an example like (8) with the superlative, the classifier is obligatory because this construction presupposes a specific noun in the subject position.

(7) a. ṛót  sūi  ᶢeṇ
   car  red
   ‘a red car/the red car; the red cars / red cars’

b. ṛót  khan  sūi  ᶢeṇ
   car  CL  red
   ‘the red car/the red cars’
Finally, the classifier can be used to express contrastive focus in the demonstrative construction and in the adjective construction. This is the case if the items to be discussed are presupposed, i.e., if they have already been introduced and if the classifier has already been used. The general rule in Thai (and a number of other mainland Southeast Asian languages, cf. Bisang 1993 on Hmong) is that once a noun has been marked as definite or specific, it will remain unmarked as long as it is used in the same discourse paragraph. If this rule is broken, i.e., if a noun that has already been introduced as definite or specific is marked by the classifier again, the hearer can infer that the adjective (9) or the demonstrative (10) must be understood contrastively (for a consistent account of the use of classifiers in on-going text, cf. Becker 2005). Thus, the following example can be used in a context in which a small car and a big car are presupposed. Speaker A says that B likes the big car. B can now correct this wrong presupposition by uttering (9a) or (9b):

(9) a. Mâj châj, chɔɔp rót khan lëk mâak kwàa.

   NEG true like car CL small more

   ‘This is not true, I prefer the SMALL car.’

(8) rót khan sīi deŋ pheŋ thîi sùt.

car CL red expensive most

‘The red car is the most expensive.’
b. ช้อป รถ ฮาน ล็ก, maj ช้อป รถ ฮาน เจา.
   like car CL small NEG like car CL big

   ‘I like the SMALL car, I don’t like the BIG car.’

The same can be shown for demonstratives. If there are two cars of which one is closer to the speech-act participants than the other, a speaker can react as follows to the wrong presupposition concerning the deictic position of the car s/he prefers:

(10) a. แมจ ช่าจ, ช้อป รถ ฮาน นี่ มาก กว่า.
   NEG true like car CL this more

   ‘This is not true, I prefer THIS car [not THAT one].’

b. ช้อป รถ ฮาน นี้, แมจ ช้อป รถ ฮาน นัน.
   like car CL this NEG like car CL that

   ‘I like THIS car, I don’t like THAT car.’

To conclude this subsection, I would like to point out two fundamental properties of grammaticalization that Thai classifiers share with the TAM-marker derived from ‘come to have’-verbs (cf. subsection 2.1). There is again no gradual development from one function to the other. Depending on the syntactic context, classifiers can express a number of different functions which are all related to classification and its two main functions. In spite of their grammatical function and their occurrence in a particular syntactic position, classifiers do maximally show a small loss of phonetic substance related to tonality.
Finally, there is also a difference between ‘come to have’-verbs and Thai classifiers. While the different interpretations of ‘come to have’-verbs are not associated with different constructions, the different interpretations of Thai classifiers clearly are associated with different constructions.

2.3. The existence of rigid syntactic patterns

The existence of rigid word-order patterns will be illustrated by a brief look at the positions of TAM-markers, coverbs (COV) and directional verbs (Vd) relative to the main verb in the clause of East and mainland Southeast Asian languages. Each of these markers are products of grammaticalization. TAM-markers are verbs like Khmer *ba:m* ‘come to have’ (cf. subsection 2.1), i.e., verbs that can be used for expressing functions of tense, aspect and/or modality. Coverbs are verbs that can take the function of adpositions. Directional verbs indicate the direction of the event denoted by the main verb. Each of these markers only occurs in specific positions. In the case of Thai and Khmer, the positions for these markers are distributed as follows (for a much more detailed account on the distributional properties of these markers in Chinese, Hmong, Vietnamese, Thai and Khmer, cf. Bisang 1992, 1996):

(11) Syntactic patterns in Khmer/Thai:

```
TAM  V  Vd COV  TAM
```

The word-order pattern of (11) is illustrated by the following two examples from Khmer and Thai. The main verb is underlined, the grammaticalized items are printed in bold. There
is a preverbal TAM-marker in Khmer (12) and a postverbal TAM-marker in Thai (13). The position for directional verbs can maximally take two verbs in Thai and three verbs in Khmer. The position for coverbs always follows the position for directional verbs.

(12) Khmer:

<table>
<thead>
<tr>
<th>he</th>
<th>TAM</th>
<th>take</th>
<th>luggage</th>
<th>Vd:move.down</th>
<th>Vd:move.out</th>
<th>Vd:come</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaht</td>
<td>ba:n</td>
<td>yɔ:k</td>
<td>?ryvan</td>
<td>coh</td>
<td>cɛɲ</td>
<td>mɔ:k</td>
</tr>
</tbody>
</table>

‘He took the luggage down and out for me.’

(13) Thai:

<table>
<thead>
<tr>
<th>he</th>
<th>take</th>
<th>luggage</th>
<th>Vd:move.down</th>
<th>Vd:come</th>
<th>COV:give</th>
<th>I</th>
<th>TAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khaw</td>
<td>?aw</td>
<td>krapɔw</td>
<td>loŋ</td>
<td>maa</td>
<td>hɔy</td>
<td>phɔm</td>
<td>lɛw</td>
</tr>
</tbody>
</table>

‘He took the luggage down for me.’

The order of the grammaticalized items relative to each other in (12) and (13) is absolutely rigid. Even the order among the individual directional verbs cannot be changed. The only exception is Khmer ba:n, which can also occur in the clause final TAM-position. But this is in accordance with the word-order pattern given in (11).
3. Approaches to grammaticalization in the light of data from East and mainland SE Asia

3.1. Lehmann (1995 [1982])

Lehmann (1995) defines grammaticalization in terms of the autonomy of the linguistic sign, whereby the reduction of autonomy raises the degree of grammaticalization. Autonomy is determined by the three parameters of weight, cohesion and variability. Each of these parameters must be analysed from the paradigmatic and the syntagmatic perspective. This yields the six well-known criteria for measuring grammaticalization as illustrated in table 1:

![Insert Table 2 Here]

A look at the languages of East and mainland Southeast Asia from the perspective of these criteria leads to the paradoxical situation that their grammatical markers can express the semantics of categories which are situated at the end of grammatical clines (e.g. tense or reference, cf. subsections § 2.1 and 2.2, respectively) although their autonomy is still high in terms of the majority of the criteria discussed by Lehmann (1995). In the case of integrity (= paradigmatic weight), the grammaticalization of a linguistic sign does not run fully parallel to the reduction of phonetic integrity, i.e., there is no necessary coevolution of form and meaning. Similarly, the other two paradigmatic criteria do not produce significant effects. Paradigmatic cohesion (= paradigmaticity) only applies to the emergence of closed-class categories but it does not lead to the emergence of paradigms. Paradigmatic variability is not developed very far either, since none of the categories expressed by grammaticalized
items is obligatory (with the exception of a few well-defined instances like the obligatory use of the classifier with numerals and with the superlative in Thai). On the syntagmatic side, structural scope as suggested by Lehmann (1995) is a problematic criterion because even a highly grammaticalized item can have wide scope if this is compatible with its semantics. Bondedness (= syntagmatic cohesion) does occur with some markers (e.g. the well-known verbal markers \(-le\) and \(-zhe\) in Modern Standard Chinese) but it is certainly not a general property of grammaticalized items in the languages of East and mainland Southeast Asia. The only parameter that fully applies is syntagmatic variability.

Grammaticalized items have to follow strict word-order rules. Thus, ‘come to have’-verbs (cf. subsection 2.1) must occur in front of the main verb (or postverbal/sentence-final position). As illustrated with Thai (cf. subsection 2.2), classifiers must take specific positions in the numeral construction, the demonstrative construction and the adjective construction. Finally, TAM-markers, coverbs and directional verbs are assigned to specific positions relative to each other within the serial unit (cf. subsection 2.3). The following table summarizes to what extent the criteria for measuring the autonomy of a linguistic sign and thus its degree of grammaticalization applies to East and mainland Southeast Asian languages.

INSERT TABLE 3 ABOUT HERE
Grammaticalization in terms of Heine et al. (1991) is a creative act of problem solving by which old resources are used for the expression of new functions. Its properties are motivated by cognition as a language-external factor. The development of new functions for existing resources takes place on two levels called macrostructure and microstructure. Macrostructure is characterized by metaphoric processes, i.e., by changes of meaning from one semantic domain to another through perceptible functional or relational similarities (cf. analogy). The mapping of linguistic items from one domain of conceptualization onto another is the central strategy of grammaticalization. It takes place along universal scales or chains like the following:

(14) PERSON > OBJECT > ACTIVITY > SPACE > TIME > QUALITY

Microstructure is characterized by metonymic processes, i.e., by pragmatic inferences from context. Metonymic processes trigger context-induced reinterpretations, “whereby conversational implicatures are conventionalized to new focal senses” (Heine et al. 1991: 103). Since metaphor determines the changes from one domain to the next and since metonymy rather operates more locally as a bridge between individual domains, the strategy of metaphor is more basic than metonymy:

It would seem that metaphor and metonymy form different components of one and the same process leading from concrete to more abstract grammatical
concepts. On the one hand, this process is made up of a scale of contiguous entities that stand in a metonymic relation to one another. On the other hand, it contains a smaller number of salient and discontinuous categories such as SPACE, TIME or QUALITY. The relation between these categories ... is metaphorical but can also be described as being the result of a number of metonymic extensions. Conceivably, metonymy is the more basic component of this process in that metaphor is grounded in metonymy ... (Heine et al. 1991: 73 - 74)

If this approach is applied to the languages of East and mainland Southeast Asia, it yields at least the following two problems. (i) Metaphor is not the main strategy of grammaticalization in these languages. Pragmatic inference is more important and it does not necessarily lead to a scale as the one in (14). Even if individual lexical items may express a number of cognitive domains from this scale, they are in no hierarchical order, i.e., all of them can be inferred equally, depending on context and on the constructions they belong to. (ii) Heine et al.’s (1991) model is primarily concerned with cognition or semantics. It does not integrate the correlation between grammatical functions and their morphosyntactic expression in constructions. In order to be able to explain how certain lexical elements become productive markers of a grammatical function we need a structural framework that allows the necessary implicatures which lead to the reanalysis of a given utterance. The framework which makes this type of implicatures possible is the construction.

The criticisms concerning the approach of Heine et al. (1991) do not apply to Bybee et al. (1994). Bybee et al. attach more importance to inference than to metaphor and they integrate constructions into their model. As Bybee et al. point out, metaphor only comes in the beginning of grammaticalization:

While it is true that when one compares source concepts to related grammatical concepts one can construct a metaphorical relation between the two in many cases, our evidence suggests that the actual formation of metaphors is not the major mechanism for semantic change in grammaticization. Rather we see metaphor operating only on the more lexical end of grammaticization paths rather than propelling grams into the more and more abstract domains of grammatical meaning. (Bybee et al. 1994: 25)

Inference pervasively operates through all the stages of grammaticalization from its very beginning to its very end. It is based on Grice’s (1975) maxim of quantity, which states that the speaker does not say more than s/he must and that the hearer infers as much as s/he can.

One of the results of this delicate balance is that the hearer is obliged to extract all the meaning possible from the message, which includes all the implications that are not controversial. A semantic change can take place when a certain implication commonly arises with a certain linguistic form. That implication can
be taken as part of the inherent meaning of the form, and can even go so far as to replace the original meaning of the form. (Bybee et al. 1994: 286)

Constructions are taken to be relevant for understanding processes of grammaticalization in the model of Bybee et al. (1994). It is often the meaning of a construction or elements within it that determines the result of grammaticalization. What Heine et al. (1991) explain in terms of metaphor often turns out to be the result of inferences based on the meaning of a construction. The following quotation shows how this works for *be going to* in English:

> While it is certainly the case ... that the same schema structure for spatial ‘be going to’ is preserved in temporal ‘be going to’, it does not follow that metaphorical extension is the operative mechanism of change. Once again, the temporal meaning was present in the construction from the beginning. *We’re going to Windsor to meet the King* emphasizes the spatial but certainly makes a temporal statement as well. Again, the construction can spread gradually from cases where the spatial is important to cases where both temporal and spatial are important and finally to cases where only the temporal is relevant. (Bybee et al. 1994: 292)

The fact that the model of Bybee et al. (1994) takes inference to be more important than metaphor and the fact that it integrates constructions does not mean that this model is without problems from the perspective of East and mainland Southeast Asian languages. Two problems will be briefly mentioned at the end of this subsection.
Bybee (1985) as well as Bybee et al. (1994) take the coevolution of form and meaning for granted. As a consequence, their model cannot deal with linguistic signs that can trigger highly grammatical concepts without an equal amount of phonetic reduction. The languages of East and mainland Southeast Asia show that the claim of a “causal link between semantic and phonetic reduction” in the quotation below is not universal.

It therefore seems natural to look for a direct, and even causal, link between semantic and phonetic reduction in the evolution of grammatical material, beginning with the earliest stages of development from lexical sources and continuing throughout the subsequent developments grams undergo. Our hypothesis is that the development of grammatical material is characterized by the dynamic coevolution of meaning and form. (Bybee et al. 1994: 20)

Bybee et al. (1994) understand grammaticalization as a gradual process in terms of grammaticalization clines in which a marker develops through different functional stages (e.g. iterative > continuative > progressive > imperfective > intransitive; Bybee et al. 1994).

What makes inference interesting as a mechanism of change is the fact that inference allows the incorporation of new meaning into a gram. ... Yet the orderliness of semantic change in grammaticization and the universality of paths of change demonstrate that the process of infusion of new meaning into a gram is quite constrained. The constrained set of inferential changes that can be discovered in grammaticization are interesting in their own right, as they will
reveal us the nature of the commonly made inferences that guide speakers and hearers in conversation. (Bybee et al. 1994: 289)

As can be seen from the case of ‘come to have’-verbs (cf. subsection 2.1), a grammatical cline is not a necessary prerequisite to account for the grammatical functions they can express.

3.4. Hopper & Traugott (1993)

Hopper & Traugott’s (1993) model is based on both metaphor and metonymy. Both processes are seen as pragmatic in nature. Metonymic processes are based on conversational implicatures and are more important and prior to metaphoric processes, which are based on conventional implicatures through analogy between semantic domains. Metaphor is concerned with analogy on the paradigmatic level and operates through conceptual domains, while metonymy belongs to reanalysis or abduction at the syntagmatic level and operates through interdependent syntactic constituents. The following table provides a summary of how metonymy and metaphor are defined by Hopper & Traugott (1993):

INSERT TABLE 4 ABOUT HERE
The reason why other approaches to grammaticalization such as Heine et al. (1991) overemphasize the importance of metaphor is due to a tendency to think in terms of “lexical item > grammatical item” instead of “use of lexical item in discourse > grammatical item” (Hopper & Traugott 1993: 81). With its clear discourse basis, the model of Hopper & Traugott (1993) has a number of advantages for the understanding of grammaticalization as it is found in East and mainland Southeast Asian languages. It gives priority to metonymy and thus to inference at least at the beginning of grammaticalization processes when lexical items are enriched through conversational implicatures which later become part of the conventionalized meaning of these items. It does not depend on the coevolution of form and meaning. It does not assume that grammaticalization is necessarily gradual in terms of a continuous semantic bleaching, at least at the beginning of grammaticalization:

There is no doubt that over time, meanings tend to become weakened during the process of grammaticalization. Nevertheless, all the evidence for early stages is that initially there is a redistribution or shift, not a loss, of meaning. (Hopper & Traugott 1993: 88)

The problem with the model of Hopper & Traugott (1993) is that the above positive aspects of the model are supposed to operate only at the beginning of grammaticalization. Thus pragmatics in the sense of conversational implicatures will be replaced later by metaphoric processes and by processes of conventionalization which lead to more gradual processes and seem to fall more or less in line with assumptions concerning the coevolution of form
and meaning. Since pragmatic inferences also operate at the level of highly grammatical concepts and since we have seen that different inferences with the same lexical item do not follow a hierarchy and are not necessarily subject to the coevolution of form and meaning (cf. the case of ‘come to have’-verbs in subsection 2.1), the model of Hopper & Traugott (1993) is not fully adequate for East and mainland Southeast Asian languages.

3.5. Conclusion

As I have tried to show in my brief summary of the most prominent models of grammaticalization, none of them can fully account for grammaticalization in East and mainland Southeast Asian languages. Lehmann’s (1995) concept of autonomy only seems to work in the case of syntagmatic variability (rigid word-order patterns). Heine et al. (1991) overemphasize the importance of metaphor and neglect the relevance of constructions. Bybee et al. (1994) take the coevolution of form and meaning and the graduality of grammaticalization processes for granted. Finally, Hopper & Traugott (1993) provide all the ingredients that are needed for accounting for grammaticalization in East and mainland Southeast Asian languages but they don’t allow them to operate at later stages of grammaticalization.
4. Areality and why grammaticalization is special in East and mainland SE Asia

4.1. An areal type of grammaticalization

Grammaticalization is usually defined in two stages (i) from a lexical item to a grammaticalized item and (ii) from a grammaticalized item to a more grammaticalized item. This definition goes back to Kuryłowicz (1964) and was recently taken up again by Detges & Waltereit (2002):

Grammaticalization is a process whereby a lexical item assumes a grammatical function or whereby an already grammatical item takes on a more grammatical function. (Detges & Waltereit 2002: 188)

Many instances of grammaticalization in East and mainland SE Asian languages seem to follow a one-stage model. One can see the step from lexical item to grammaticalized item but it is often hard to clearly distinguish between more and less grammaticalized items. This situation is the result of the pervasiveness of pragmatics in the grammaticalization processes of these languages. The possibility of several pragmatic inferences with one and the same linguistic item is fairly common and there is no need for these inferences to follow a certain hierarchy (even though I admit that this is not excluded). As a consequence of this, there are instances like ‘come to have’-verbs, which can be interpreted equally in terms of ability, permission, obligation, past and emphasis of truth/factuality (cf. (2)).
In my view, the strong discourse-basis of processes of grammaticalization represents an areal type of grammaticalization which encompasses the following language families situated in East and mainland Southeast Asia:

- Mon-Khmer (branch of Austroasiatic)
- Tai (branch of Tai-Kadai)
- Sinitic (branch of Sino-Tibetan)
- Hmong-Mien

If there is a considerable number of languages in which grammaticalization also operates at levels usually associated with higher degrees of grammaticalization and if these languages form a geographical cluster in East and mainland Southeast Asia, one has to ask to what extent current approaches to grammaticalization depend on the languages and language groups they are based on. Since the languages discussed in this paper only play a marginal role in the approaches discussed in section 3, none of their proponents came across pragmatics-based processes of grammaticalization as we find them in East and mainland Southeast Asia. Even typological questionnaires seem to be biased by this lack of input. Why else would it be possible that verbs with the meaning of ‘come to have’ are hardly discussed as source concepts in generalizations concerning the development of TAM systems? From a more general perspective, one may thus wonder how many other source concepts and how many other types of grammaticalization there are to be discovered if other areas are analysed in more detail. To conclude this subsection, I would simply like to
point out that the lack of form-meaning coevolution is also attested in Slavonic languages, another area/language family that is understudied as far as grammaticalization is concerned (cf. Bisang, Himmelmann & Wiemer 2004).

4.2. How to account for the situation in East and mainland SE Asia—
on the lack of paradigms

This subsection briefly outlines why grammatical markers of East and mainland Southeast Asian languages are not organised in morphological paradigms as we known them from languages like Latin (cf. amo, amas, amat, etc. ‘I love’, ‘you love’, ‘she/he/it loves’). As I argued in the introduction, this is due to two complementary types of factors, phonological properties (Ansaldo & Lim 2004) and the broad functional spectrum of grammatical markers combined with their lack of obligatoriness. Both of these factors will be briefly explained in this subsection.

Ansaldo & Lim (2004) point out that the relative phonetic stability of grammaticalized items in East and mainland Southeast Asian languages is due to the discreteness of syllable boundaries and phonotactic restraints. They show for Cantonese and Hokkien, that grammaticalization manifests itself in the form of reduced syllable duration and vowel quality. Both of these effects have to do with the fact that grammaticalized items occur adjacent to metrically stressed syllables and thus tend to lose their own stress in such an environment. Interestingly enough, there is no significant reduction in pitch with grammaticalized items. This is explained by Ansaldo & Lim (2004) by the fact that tonal
contrast must be maintained for keeping up contrast, particularly in Sinitic languages like Cantonese and Hokkien, which distinguish three tonal registers. In Sinitic languages with only one tonal register like Modern Standard Chinese (Putonghua), reduction of pitch would not produce the same strong effect of blurring distinctive features of the language. This seems to be the reason why some morphemes in Modern Standard Chinese such as verb-final -le (derived from liǎo ‘finish’) or verb-final -zhe (derived from zhāo ‘touch, contact’) did not only lose vowel quality but reduced their diphthongs and triphthongs to the reduced tonless vowel [ə]. The correlation between the existence of a neutral tone as we find it in -le and -zhe and the presence of a single register is also pointed out by Ansaldo & Lim (2004). Facts like these clearly show that there are phonological properties that prevent the full coevolution of form and meaning that can lead to the emergence of morphological paradigms— processes that develop morphology are simply not prominent enough in most East and mainland Southeast Asian languages.

The second factor that obstructs the emergence of morphological paradigms is based on the two factors of frequency and the existence of markers belonging to clearly determined semantic domains. The central role of frequency for the development of paradigms has to be seen in the light of generality as defined by Bybee (1985) and obligatoriness (cf. quotation from Lehmann 1995 in section 1). Morphological paradigms develop from categories that are frequently used. Frequency is enhanced by semantic generality, which grants its compatibility with a large number of lexical items. If a marker is semantically general enough to be coextensive with a basic grammatical entity like noun or verb its
occurrence may become obligatory with that entity. As a consequence, it becomes even more frequent.

As we can see from Lehmann (1995: 139; cf. quotation in section 1), a paradigm consists of a certain category with its values or subcategories. Thus, paradigms refer to certain clearly defined grammatical categories such as tense (consisting of values like present, past, future) or aspect (consisting of subcategories like perfective, imperfective). The precondition for the development of categories that can be integrated into paradigms is a certain degree of homogeneity on the level of the category as a whole and semantically clear-cut definitions of its subcategories.

In East and mainland Southeast Asian languages, the broad functional spectrum of one and the same marker and the high degree of indeterminateness (lack of obligatory categories) systematically undermine the emergence of a situation in which grammatical markers are frequent and homogeneous enough to become part of a coherent paradigm. Given the optionality of grammatical markers, they are not as frequent as, e.g., tense markers in English or German even if they may express highly generalized meanings. Since the meaning of grammatical markers depends on pragmatics on all levels of grammaticalization, their functional range is not limited to a single clearly determined semantic domain. Thus, the emergence of a paradigm is rather unlikely from both perspectives, that of frequency and that of semantic homogeneity.
5. Outlook — Why is research on grammaticalization interesting?

Newmeyer (1998) deconstructs grammaticalization by showing that none of its three components of (i) downgrading analysis, (ii) semantic change and (iii) phonetic reduction is exclusively related to processes of grammaticalization. Grammaticalization is thus nothing but an epiphenomenon of these independent processes of diachronic change:

We have examined the associated set of diachronic changes that fall under the rubric of ‘grammaticalization’ and have found that no new theoretical mechanisms, nor mechanisms unique to grammaticalization itself are needed to explain them. Far from calling for a ‘new theoretical paradigm’, grammaticalization appears to be no more than a cover term for a conjunction of familiar developments from different spheres of language, none of which require or entail any of the others. (Newmeyer 1998: 295)

In East and mainland SE Asian languages, reanalysis (the occurrence in a particular position within a syntactic pattern) and semantic change (defined in terms of pragmatic inference) always cooccur, while phonetic reduction only operates to a certain degree and does not necessarily correlate with change of meaning. This pattern of interaction between the three processes of diachronic change seems to be a specific property of East and mainland Southeast Asia. In other languages and maybe in other areas there may be other patterns. If it turns out that there are such patterns of grammaticalization and that the same languages also deviate in interesting ways from standard theoretical assumptions (cf. the
discussion of subject/object asymmetry in East and mainland Southeast Asia) questions concerning the basis of grammatical properties are far from being trivial.

Endnotes

1 Even if the clines discussed by authors like Bybee, Perkins & Pagliuca (1994) are claimed to reflect diachronic stages of development, one should keep in mind that they are primarily derived from typological data based on synchronic material.

2 One of the anonymous referees suggested to treat the functions of *ban* in terms of polysemy. My point is that we are dealing with pragmatic inference here, not with semantics.
References


<table>
<thead>
<tr>
<th>Basic Function</th>
<th>Grammatical Function</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuation</td>
<td>Countability</td>
<td>Numerative Construction: [N NUM CL]</td>
</tr>
<tr>
<td>Identification</td>
<td>Singulative</td>
<td>Demonstrative Construction: [N DEM] vs. [N CL DEM]</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
<td>Adjective Construction: [N ADJ] vs. [N CL ADJ]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(definiteness/specificity: the object is accessible to the hearer)</td>
</tr>
<tr>
<td>Contrastive Focus</td>
<td></td>
<td>Adjective Construction, Demonstrative Construction (presupposed items are in contrast)</td>
</tr>
</tbody>
</table>

Table 1: Functions of the Thai classifier in different constructions
<table>
<thead>
<tr>
<th></th>
<th>paradigmatic</th>
<th>syntagmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>Integrity</td>
<td>Structural scope</td>
</tr>
<tr>
<td>cohesion</td>
<td>Paradigmaticity</td>
<td>Bondedness</td>
</tr>
<tr>
<td>variability</td>
<td>Paradigmatic variability</td>
<td>Syntagmatic variability</td>
</tr>
</tbody>
</table>

Table 2: Lehmann’s (1995: 123) criteria for measuring the autonomy of a linguistic sign
<table>
<thead>
<tr>
<th></th>
<th>paradigmatic</th>
<th>syntagmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>Integrity (-):</td>
<td>Structural scope (?):</td>
</tr>
<tr>
<td></td>
<td>No necessary coevolution</td>
<td>Grammaticalized items rarely</td>
</tr>
<tr>
<td></td>
<td>of form and meaning</td>
<td>have scope over one single word</td>
</tr>
<tr>
<td>cohesion</td>
<td>Paradigmaticity (-/+):</td>
<td>Bondedness (-):</td>
</tr>
<tr>
<td></td>
<td>No emergence of paradigms</td>
<td>Coalescence is not prominent</td>
</tr>
<tr>
<td></td>
<td>but at least closed-class categories</td>
<td></td>
</tr>
<tr>
<td>variability</td>
<td>Paradigmatic variability (-):</td>
<td>Syntagmatic variability (+):</td>
</tr>
<tr>
<td></td>
<td>No obligatory categories</td>
<td>Rigid word-order patterns</td>
</tr>
</tbody>
</table>

Table 3: Lehmann’s (1995) criteria for measuring autonomy and their relevance for East and mainland SE Asian languages
Table 4: Mechanisms of language change: metonymy and metaphor in terms of Hopper & Traugott (1993)

<table>
<thead>
<tr>
<th>Metonymy</th>
<th>Metaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Syntagmatic level</td>
<td>- Paradigmatic level</td>
</tr>
<tr>
<td>- Reanalysis (abduction)</td>
<td>- Analogy</td>
</tr>
<tr>
<td>- Conversational implicature</td>
<td>- Conventional implicature</td>
</tr>
<tr>
<td>- Operates through interdependent syntactic constituents</td>
<td>- Operates through conceptual domains</td>
</tr>
</tbody>
</table>