

- Yap, Foong Ha, Stephen Matthews & Kaoru Horie. 2004. From pronominalizer to pragmatic marker: Implications for unidirectionality from a crosslinguistic perspective. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: The nature of grammaticalization* [Typological Studies in Language 59], 137–168. Amsterdam: John Benjamins.
- Yoon, Kyung-Mi. 1995. *Kes-uy Hyengthayloncek Thukseng Yenkwu* (A study on the morphological characteristics of *-kes*). *Dong-A University Kwukekwukmwunhak Nonmwuncip* 14: 219–234.

## The grammaticalization of clausal nominalizers in Burmese

Andrew Simpson

University of Southern California

This paper is concerned with the grammaticalization of clausal nominalizers in two different but closely-related forms of Burmese, Colloquial Burmese and Literary Burmese. A contrastive overview of the morphosyntactic properties of the nominalizers *thii* and *mii* of Literary Burmese and their Colloquial Burmese counterparts *te* and *me*, together with the application of a number of tests for the identification of nominalized constructions, reveal that grammaticalization is more advanced in the colloquial language than in the literary variety: *te* and *me* have lost their nominal specifications and been reanalysed as grammatical elements of a different categorial type, instantiating verb-related mood and realis–irrealis distinctions. The comparison of the system of nominalization in the two complementary varieties of Burmese allows for insights into the evolution, spread and reinterpretation of nominalization structures within a language.

### 1. Introduction

The crosslinguistic phenomenon of nominalization is an area of morphosyntax which has attracted growing interest in both formal and functional linguistics in recent years, as data from an increasing number of languages become available and add important insights into the grammaticalization paths and syntactic structure underlying the emergence of nominalizing elements. The present paper sets out to probe this area of morphosyntactic development further with a study focussed on Burmese, a language which is particularly rich in nominalization structures and where a highly informative picture of the results of the grammaticalization of clausal and sentential nominalizers can be found through a comparison of two different though closely-related forms of the language: Colloquial Burmese and Literary Burmese. A careful examination of synchronic patterns in Colloquial and Literary Burmese provides evidence of the source and complex structure of clausal nominalizers in the language. It also indicates that, in certain cases, earlier nominalizers have undergone regrammaticalization as functional morphemes of a different formal type, no longer instantiating nominal categories, but being reanalysed as modal and complementizer-type elements. Quite generally, the patterning found in Burmese nominalization raises questions about the assumed syn-

chronic status of nominalizers in other languages, and about how the occurrence of cross-categorical reanalysis of (already) grammaticalized morphemes can effectively be detected.

The structure of the paper is as follows. Section 2 first considers how nominalizers may be identified in a language and what morphosyntactic properties and sources nominalizers seem to exhibit in general. Section 3 then turns to Burmese and attempts to establish how the current, productive clausal nominalizers in Literary Burmese have undergone grammaticalization producing complex, fused forms which occur in a range of syntactic environments. This leads on to a comparison with similar patterns in Colloquial Burmese, where the conclusion is reached that nominalization in the colloquial language has progressed a significant stage further and has resulted in the reanalysis of erstwhile nominalizers as grammatical elements of a rather different *clausal* type. Section 4 then closes the paper with a brief summary of its findings and questions which are raised by the patterning observed.

## 2. Nominalizers and their common and divergent properties: A brief overview

### 2.1 What are nominalizers and what sources do they develop from?

In the broad, developing literature concerned with nominalization, it is not uncommon to find the term *nominalizer* being made use of rather loosely to refer to elements whose status as nominalizers has not been particularly clearly established. As we begin this investigation it is therefore useful to attempt to clarify what kinds of morphemes should be genuinely labelled as nominalizers and how one might attempt to identify a particular grammatical element as a nominalizer rather than as some other kind of syntactic category.

The term *nominalizer* is a purely functional label which is appropriately used to refer to all those morphemes or words which have the specific function of creating a nominal morphosyntactic form as the result of their combination with other kinds of non-nominal input. Nominal categories, and hence the presence of functional elements which may be nominalizers, can, in turn, be identified in two basic ways: (i) through the occurrence of noun-like or nominal morphological patterns and/or (ii) via syntactic privileges otherwise commonly associated with nouns and their syntactic projections.

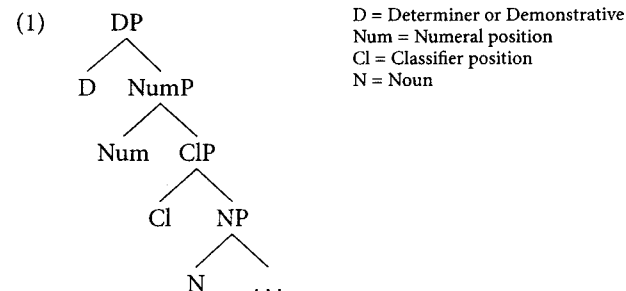
Concerning the morphological evidence for the presence of a nominal (rather than a verbal, adjectival or other) category, the potential occurrence of a range of common nominal attachments or inflections on a syntactic phrase or a combination of certain other syntactic categories with the phrase can be taken as indica-

tion that a phrase is nominal, patterning in a morphological and combinatorial way like other simple nouns and their expansions into noun-rooted phrases. The following are typical morphosyntactic indications that a syntactic constituent is nominal:

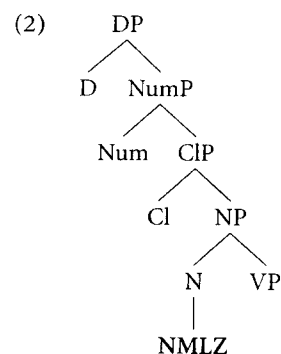
- a. Occurrence of case inflections on a constituent.
- b. Possible pluralization of the constituent.
- c. Possible enumeration of the constituent (i.e. combination of the constituent with numerals).
- d. Potential occurrence of demonstratives and adjectives with the constituent, rather than complementizers and adverbs.
- e. Use of case-marking strategies associated specifically with nouns in the marking of arguments of the head of the constituent (e.g. use of possessive or genitive case to mark the head's arguments rather than nominative or accusative case).

Syntactically, a complex constituent may be identified as a nominal phrase if it shows the distribution of other simple phrases which are clearly nominal, for example, the ability to occur in subject position or the ability to be coordinated with other clearly nominal categories. Non-nominal categories, such as verbal or adjectival phrases, are regularly excluded from such positions. However, if a verbal or adjectival phrase in combination with some additional morpheme is found to allow for occurrence in subject position or in coordination with other noun phrases, this may be taken as reasonable evidence for the nominalized status of the complex constituent and for the nominalizing function of the morpheme combined with the verb or adjective and their dependents.

Because it is the outward morphosyntactic behaviour of a constituent that identifies it as nominal, and because there are various ways in which nominal morphosyntax can be exhibited (case marking, pluralization, combination with adjectives, etc.), nominalizers may in fact originate from a range of different lexical sources. Consider a formal syntactic representation of the internal structure of a referential nominal phrase, a DP, in a head-initial language with classifiers, as given in (1).

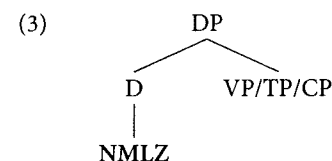


If it is assumed that the occurrence of any lower syntactic head position, such as Noun, Classifier or Numeral, will necessarily result in the projection of other, higher portions of the structure in (1), and that different parts of the structure in (1) may identify a constituent as 'nominal' in different ways (the occurrence of numerals in Num or demonstratives in D), it is possible to see that the use of any of the head constituents (D, Num, Cl or N) in (1) might in principle be used to build up a phrase with certain overt nominal properties when combined with a non-nominal complement, thus nominalizing the latter. For example, were an element of syntactic or lexical category Noun to be combined with a verbal constituent (if complex, then a verb phrase), the result would be a structure such as (2) below. In (2), a full array of nominal properties might be expected to characterize the new nominalization, given the presence of the full set of syntactic heads and projections above the Noun position. In the representation in (2), NMLZ is intended to indicate a morpheme which has the function of a nominalizer and the original syntactic category of a noun (i.e. a nominalizer derived from a noun, perhaps via bleaching of the lexical content of the noun):

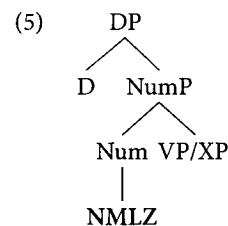
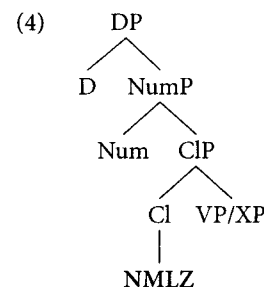


Alternatively, however, a verbal (or other non-nominal) constituent might allow for conversion into a category with certain nominal properties if combined with a nominalizer sourced from one of the higher head categories in (1), such as an element of type D (Determiner or Demonstrative). Such a constituent would be expected to allow for the syntactic distribution of other nominal expressions (e.g. unrestricted occurrence in subject position). However, it might not allow for the full array of properties commonly associated with noun phrases and their expansions into DPs due to the lack of lower portions of the nominal structure in (1), notably the Numeral, Classifier and Noun heads and their projections. Such a possibility is schematized in (3), where the nominalized constituent is indicated

as being possibly a verb phrase (VP), a tense phrase (TP) or a fully-clausal complementizer phrase (CP):



Two other intermediate possibilities are also anticipated to be available and exist, namely (i) the use of a *classifier* as a nominalizing element combining with a non-nominal complement, as in (4); and (ii) the use of a *numeral* as a nominalizer, as in (5). In both instances, the structures produced are expected to have some, but not necessarily all, typical properties of nominal projections. For example, if adjectival modification is assumed to occur via the adjunction of an adjectival phrase to a noun phrase, the absence of a noun phrase in structures (4) and (5) may be expected to correspond to a lack of adjectival modification with such nominalizations.



Interestingly, all the possible ways in which a nominalizer might theoretically be instantiated, as outlined above, do indeed seem to occur in different languages, and nominalizers grammaticalize from a variety of sources. Nominalizers ultimately derived from nouns appear to be very common in occurrence (e.g. Korean

kes, as discussed in Simpson and Wu 2001), but it is also possible to identify classifier, numeral and demonstrative or determiner sources of other nominalizers too. Burmese, for example, makes use of the numeral 'one' (*ta-*) in a range of nominalizations. Thai and Bengali, in turn, show evidence of classifier use in a nominalizing function with certain clausal constituents, whereas Chinese, Japanese and Lakota can be argued to have derived nominalizers from earlier elements of type D (cf. Simpson 2001, 2003a, 2003b; Simpson and Wu 2001 for extensive discussion).

## 2.2 Is nominalization a lexical or a syntactic process?

In addition to variation in the source category of nominalizing elements, a second important parameter of variation in the (synchronic) realization of nominalizers is whether the attachment of such elements occurs as a lexical process or is effected as part of a syntactic derivation.

Certain nominalizers seem to be very clearly phrasal attachments, combined with a constituent which is a full syntactic phrase. This is the case of the sentential and clausal nominalizers found in a significant number of languages, where a full clause is converted into a nominal argument of some other predicate by the use of an appropriate nominalizer (e.g. Japanese *no*, Korean *kes*; cf. Simpson and Wu 2001; see also the papers by Horie, Rhee, and Yap and Matthews in this volume). If sentential and other phrasal nominalizers convert a full syntactic constituent into a nominal output, such nominalizers need to be considered syntactically independent words, which combine with a complement during the syntactic derivation of a sentence. In this respect, they would be similar to determiners, complementizers and other functional elements which are understood to be discrete grammatical words rather than parts of other words (i.e. bound morphemes).

Other nominalizers, however, may appear to be affixes attached to word-level elements such as verbs, adjectives, and so on, rather than to verb phrases, adjective phrases, etc. Such nominalizations may have many more of the unpredictable properties of lexical processes (perhaps being restricted and unproductive and giving rise to allomorphic variation), and so be assumed to be purely morphological attachments, which combine with non-nominal roots or bases during the pre-syntactic creation of words.<sup>1</sup> An example of such morphological attachment would be nominalizations produced with English *-ant*, as in *servant* or *defendant*.

1. It is also possible that certain nominalizers attached to word-level elements might be syntactic attachments, combining with a word-level constituent within a syntactic structure, perhaps via cliticization or some form of (syntactic) head-movement. Various approaches to English nominalizing *-ing* assume such an analysis. Lotha (Tibeto-Burman) *a-* would be another candidate for a similar treatment (cf. Herring 1991).

## 2.3 Nominalizers and verbalizers

Given the very widespread nature of nominalizers in the world's languages, even being present in languages not considered to have morphological attachments, such as Vietnamese, it is natural to wonder whether there might be equivalents to nominalizers functioning in the verbal domain, what could be called *verbalizers*, and, if such elements exist, what they might in turn indicate about the nature of nominalizers. In other words, if nominalizers combine with non-nominal input to create words and constituents which can be utilized as noun (phrase)-like arguments, are there also functional elements which combine with non-verbal input to create constituents which can then be embedded as verbal units? And if not, why do such elements not exist?

Although there has been little discussion of such potential counterparts to nominalizers in the literature, certain reflection suggests that there are indeed verbal elements which correspond to nominalizers in their function of converting non-verbal input into a syntactically utilizable verbal form. Furthermore, it seems that such verbalizers occur both as syntactically independent words and as purely morphological attachments, in a way very similar to the occurrence of nominalizers. An example of lexical or morphological verbalization would be the application of affixes such as English *-ize* to adjectival input to form new verbs, e.g. *grammatical* — *grammaticalize*. Examples of syntactic verbalizers can be given in two quite common types. A first, crosslinguistically well attested type is the class of *light verbs*, such as (prototypically) *make* or *do*, which are used to combine with nominal input to produce verbal forms, such as Hindi *kaam karna* 'work do' = 'to work' or Japanese *benkyoo suru* 'studying do' = 'to study.' A second potential candidate for consideration as a syntactic verbalizer would be the simple occurrence of copulas, used to embed non-verbal input and create a verbal structure which allows the application of tense and aspect; e.g. *John was a great help* = *John helped a lot*.

Supposing, therefore, that there are indeed functional elements in the verbal domain equivalent to nominalizers, the study of light verbs and copulas (as well as morphological verbalizers) may be used in a comparative way to further open up and inform our understanding of the patterning and morphosyntax of nominalization. A comparison of nominalizers and verbalizers also has the potential to lead to interesting typological questions and to the issue of whether there really is the full cross-categorical equivalence which one might pre-theoretically expect in such a domain, and if not, why full equivalence does not exist.<sup>2</sup>

2. For example, while it may seem that nominalizers grammaticalize from a range of functional elements commonly associated with nouns, such as demonstratives, classifiers and numerals, it is not obvious that verbalizers are derived from higher functional elements associated with

## 2.4 Nominalization and compounding

In connection with the source of the grammaticalization of nominalizers, in certain cases (at least) it may be interesting to consider the possible connection of nominalization with compounding in a language. It is quite plausible that, in various instances, nominalizers grammaticalize from nouns which are frequently used to create nominal compounds forming a connected class of items, e.g. *-man* in *doorman*, *chairman* or *tax man*. With the occurrence of bleaching of the meaning of such nouns and a concomitant extension of the way in which they can be used to combine with other words in compounding, a simple process of nominalization may well evolve. In this regard, it is intriguing to find a correspondence between the direction of headedness in compounds in certain languages and the linear position of nominalizers relative to nominalized material. For example, languages such as English, Japanese and Burmese, which have right-headed compounds, also have nominalizers occurring to the right of the constituent or word nominalized. By contrast, languages which have left-headed compounds, such as Thai and Vietnamese, have nominalizers which occur to the left of the constituent or word nominalized. If such a patterning is non-coincidental and can be found to occur consistently in a wider sampling of languages, it may establish an interesting diachronic link resulting from processes of grammaticalization between two types of morphological operation regularly treated as being unrelated.

## 3. Nominalization in Burmese

Having considered some of the general issues involved in the study of nominalization and the grammaticalization of nominalizers in a language, we now turn to an investigation of nominalization phenomena in Burmese. The discussion will focus in particular on the sentential and clausal nominalizers present in the language, as these can be shown to reveal much about the way in which reanalysis applies to create complex new grammaticalized morphemes and words, and give rise to shifts between categorial types.<sup>3</sup> As briefly mentioned in Section 1 above,

verbs, such as tense, aspect and mood morphemes; on the contrary, the main source of verbalizers would seem to be verbs. Why this asymmetry between the source of nominalizers and that of verbalizers exists is not immediately clear.

3. In addition to the clausal and sentential nominalizers discussed in this work, Burmese also exhibits a wide range of other nominalizers, many of which combine with sub-clausal constituents. For simple reasons of space and focus, no attempt to describe these will be made here. The reader is referred to Okell (1969), Okell and Allott (2001) and Hopple (2003) for useful information and a comprehensive listing of all putative nominalizers in Burmese.

Burmese is commonly described as having two complementary forms: Colloquial Burmese and Literary Burmese. The latter is used orally in formal announcements, news broadcasting and is the most common written form of Burmese, whereas the former occurs in most spoken communication and is also sometimes found in more informal writing. The primary and most obvious difference between the two varieties of Burmese concerns the instantiation of their functional-grammatical morphemes, including the elements employed as nominalizers, and it is here that a comparison of patterns in Colloquial and Literary Burmese is often interesting.<sup>4</sup> The present examination of nominalization in Burmese will begin with patterns in Literary Burmese in Section 3.1 and then move on to Colloquial Burmese in Section 3.2.

### 3.1 Sentential nominalizers in Literary Burmese: The elements *thii* and *mii*

In Literary Burmese, the morpheme *thii* occurs in clause-final position, both in main clauses (as a sentence-final morpheme), as in example (6) below, and in clauses which are embedded as arguments of other predicates.

#### (6) Literary Burmese

U-Win-Win manee-ga yauq-laa thii.

U-Win-Win yesterday-PST arrive-come NMLZ

'U Win Win arrived yesterday.'

When *thii* is used to embed clauses as the arguments of a predicate, it is naturally accompanied by a case marker.<sup>5</sup> Examples (7) and (8) show this with the embedding of clauses as the object of a verb, while (9) and (10) illustrate the embedding of a clause in subject position. It should also be noted that the use of *thii* in examples (6) to (10) is obligatory and clauses may not occur as the arguments of verbs without this morpheme.

#### (7) Literary Burmese

canaw [U-Win-Win manee-ga yauq-laa thii]-kou caa ya thii.

I U-Win-Win yesterday-PST arrive-come NMLZ-ACC hear get NMLZ

'I heard that U Win Win arrived yesterday.'

4. A highly significant proportion of non-grammatical vocabulary is shared by Colloquial and Literary Burmese, though the latter also has available certain variant forms (frequently multisyllabic) which are often made use of in writing (cf. Okell 1969; Saw Tun 2005).

5. The case marking of nominal arguments is generally optional in Burmese, though heavily preferred in certain instances, both as parsing aids and sometimes also to signal contrastive emphasis, as in Korean (Schütze 2001). Informants indicate that the use of case particles with *thii*-embedded clauses is much preferred. For further discussion of factors triggering the use of case-marking particles in Burmese, cf. Kassevitch (2005).

- (8) Literary Burmese  
*hsain-shin-hmaa [hkalee-twee hseileiq hkou thii]-kou myin-ya thii.*  
 shopkeeper-NOM child-PL cigarette steal NMLZ-ACC see-get NMLZ  
 'The shopkeeper saw the children stealing the cigarettes.'
- (9) Literary Burmese  
*[thuu bamaa-sagaa pyaw nee thii]-hmaa theiq kaun hla thii.*  
 he Burma-language speak ASP NMLZ-NOM very good INTENS NMLZ  
 'He speaks Burmese very well.' (lit. 'His speaking Burmese is very good.')
- (10) Literary Burmese  
*[leezeiq-twin daq-poun yaiq thii]-hmaa tayaa-m-win pa.*  
 airport-within photograph take NMLZ-NOM not-be-legal POL  
 'It is illegal to take photographs inside the airport.' (lit. 'Taking photographs in the airport is illegal.')

*Thii*-suffixed clauses consequently have both the syntactic distribution (occurring in subject and object position) and marking (nominative and accusative case) of nominal phrases. The element *thii* therefore exhibits key aspects of the patterning of typical clausal nominalizers.

*Thii* is also found to occur in relative-clause structures, such as the one in (11).<sup>6</sup>

- (11) Literary Burmese  
*canaw we thii saouq*  
 I buy NMLZ book  
 'the/a book I bought'

In such an environment *thii* does not maintain its mid-level tone, but occurs instead in high creaky tone, one of the four tones which regularly occur on Burmese words. As a suprasegmental morpheme, such high creaky tone frequently signals genitive case, which is otherwise realized as a discrete morpheme *ye*, pronounced with high creaky tone.

- (12) Literary Burmese  
*canaw ye ein*  
 I GEN house  
 'my house'

6. In addition to *thii*, Literary Burmese also makes use of a morpheme *thaw* as an alternative linking element in relative clauses. This element is particularly frequent with intransitive predicates and adjectives (cf. Okell 1969 and Hopple 2003 for further discussion of *thaw*).

Sequences such as that in (12) may in fact have two pronunciations. If the possessor (here *canaw*) normally occurs with a mid-level tone, it can maintain this tone as a genitive-marked possessor. However, it is also possible for (12) to be pronounced with high creaky tone on both *ye* and the possessor *canaw*, the tone associated with genitive case spreading leftwards from *ye* to the nominal element which it marks to its left. A third pattern commonly found is for a (normally) mid-level tone possessor to occur pronounced with high creaky tone even in the full absence of genitive *ye* (*canaw*<sub>creaked</sub> *ein* 'my house'). The high creaky tone of genitive *ye* can, therefore, become disassociated from its regular host and simply occur marking an appropriate nominal element. In relative clauses such as that in (11) above, clause-final creaky tone *thii* is consequently most naturally analysed as the combination of a clausal nominalizer (mid-level tone) *thii* and genitive case, so that relative clauses in Burmese are instances of the modification of a noun by a nominalized clause, linked via genitive case.<sup>7</sup>

If we now consider what the source of the nominalizer *thii* might be, it can be observed that mid-level tone *thii* occurs elsewhere in Literary Burmese, regularly functioning as a pre-nominal *demonstrative*, as in (13).

- (13) Literary Burmese  
*thii saaouq*  
 this book  
 'this book'

Assuming this demonstrative element to be the most likely source of the nominalizer *thii* then suggests that clausal nominalizations in Literary Burmese have an internal structure such as that in (3) (parameterized in a head-final way), representing the combination of a D(eterminer)-type element with a clause to produce a DP nominal projection.<sup>8</sup>

Interestingly, mid-level *thii* also occurs in a third, rather different function in Literary Burmese, namely as a common instantiation of nominative case marking a subject. Example (14), thus, shows four different uses of *thii*: as a demonstrative, nominative case, linking a relative clause to a following nominal (with creaky tone) and in simple (non-relativized) clause-final position.<sup>9</sup>

7. Such an analysis is well-supported by the occurrence of genitive case with putative nominalizations in many instances of relative-clause marking and linking in Tibeto-Burman languages (cf., for example, Noonan 1997; Thurgood and LaPolla 2003).

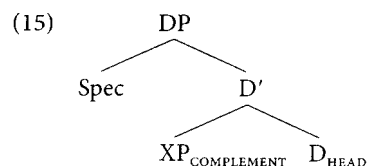
8. Cf. also Herring (1991) for evidence from other languages that clausal nominalizers may be sourced from demonstratives in Tibeto-Burman languages.

9. A teasing question is how the nominative case marker *thii* might be linked to and possibly derived from the demonstrative element *thii*. One possibility may be to consider that subject noun phrases are more frequently definite and anaphoric in reference than objects are, and so might

## (14) Literary Burmese

*thii pyiqsii-myaa-thii Daw-Hla-Mee htaa hke thii-myaa hpiq pa thii.*  
 this thing-PL-NOM Daw-Hla-Mee put ASP NMLZ-PL be POL NMLZ  
 'These are the things that Daw Hla Mee left behind.'

Assuming a demonstrative source of nominalizer *thii* to be plausible and likely, the grammaticalization of *thii* in its nominalizer function brings with it a further question relating to word order. As demonstrative *thii* occurs preceding the nominal complement with which it combines, why might it be positioned following a clausal complement when nominalizing the latter? Although no definitive answer can be provided in the absence of data on the early development of nominalizing *thii*, two speculations can be offered here. First, as Burmese is a head-final language, and *thii* as a demonstrative is likely to be positioned in a phrase-initial specifier position (SpecDP), it can be hypothesized that, when such an element grammaticalized as a nominalizer, it came to occupy the head-position of its (DP) phrase in an occurrence of the Spec-head reduction process frequently argued to characterize instances of phrasal grammaticalization (cf. Simpson and Wu 2002a and van Gelderen 2004). Given the linear organization of (elements in) specifier and phrasal head positions in a head-final, specifier-initial language such as Burmese, Spec-head reduction and grammaticalization would be expected to relocate a demonstrative from a phrase-initial specifier position to a phrase-final head-position, and so result in nominalizer *thii* coming to follow its clausal complement.



A second possibility might be to hypothesize that *thii* came to be used as a demonstrative in a *resumptive position* following a clausal subject, as occurs in certain Indic languages. This is schematized in (16).

(16) [*John likes Mary*]<sub>i</sub>, *this<sub>i</sub>* is true.

'It is true that John likes Mary.'/'That John likes Mary is true.'

Such a linear sequence might then lead to grammaticalization of the demonstrative as a nominalizer following the clause it introduces:

(17) [*John likes Mary*]-*this* is true.

attract marking with a demonstrative, encoding a familiar, [+definite] referential value.

In addition to the question of the linear order of *thii* relative to its complement clause, an interesting complication for the analysis of *thii* and of both its synchronic status and its grammaticalization comes from the fact that *thii* is actually not just a simplex nominalizer derived from a demonstrative, but also encodes realis mood, being used to combine with clauses which represent past or present actions, states and habits. Where future actions and hypothetical situations are referred to, a fully parallel set of nominalized constructions are built with the nominalizer *mii*, which differs from *thii* only in encoding irrealis mood (i.e. non-realized actions or states).<sup>10</sup> This is illustrated in (18) and (19) with clauses in subject position and nominative case, and in (20) with an accusative-marked clause in object position.

## (18) Literary Burmese

[*than shi mii*]-*hmaa-le ahman hpyiq ii.*  
 lice be NMLZ<sub>IRR</sub>-NOM-too true be NMLZ  
 'That there will be lice too is true.' (Okell and Allott 2001: 158)

## (19) Literary Burmese

[*thati pyu ya mii*]-*hmaa htaransitsataa-amyousaa-pin hpyiq*  
 attention do must NMLZ<sub>IRR</sub>-NOM transistor-type-EMPH be  
*thii.*  
 NMLZ<sub>RLS</sub>  
 'What we will have to pay attention to is the transistor type.' (Okell and Allott 2001: 158)

## (20) Literary Burmese

[*naneq ngaa naayii khan shi mii*]-*kou aloulou-hman-mi*  
 morning five o'clock about be NMLZ<sub>IRR</sub>-ACC estimate  
*thii.*  
 NMLZ<sub>RLS</sub>  
 'He estimated that it must be about five o'clock.' (Adapted from Okell and Allott 2001: 158)

Example (21) shows the occurrence of irrealis *mii* in sentence-final position, paralleling *thii* in example (6), while (22) shows its occurrence embedding a relative clause. As with *thii* when it links a relative clause to a noun, *mii* carries high creaked tone in (22) corresponding to genitive case, and is elsewhere pronounced with a mid-level tone.

10. Besides *thii* and *mii*, Literary Burmese has an element *ii* (pronounced with a high creaky tone) which occurs as a stylistic variant of *thii* and *mii* in sentence-final position or embedded quotations with no specification for (ir)realis (cf. Okell and Allott 2001: 271–272, and footnote 15 below). An example of the sentence-final use of *ii* occurs in (18).

- (21) Literary Burmese  
*maneqphan pyan laa mii.*  
 tomorrow return come NMLZ<sub>IRR</sub>  
 'He will come back tomorrow.'
- (22) Literary Burmese  
 [thuu yuu laa mii]      pyiqsii-myaa  
 he take come NMLZ<sub>IRR, GEN</sub> thing-PL  
 'the things he will bring'

A further interesting parallel in morphosyntactic patterning between *thii* and *mii* is that the latter also occurs in a pre-nominal *interrogative specifier* function, which corresponds closely to the demonstrative function of *thii*. However, whereas demonstrative *thii* is realis in picking out a referent with definite reference, interrogative *mii* is used for hypothetical reference in questioning the reference of a noun.

- (23) Literary Burmese  
*mii-thuu*  
 which-person/one  
 'who'

The clausal nominalizers *thii* and *mii* in Literary Burmese are, therefore, complex elements incorporating two parts and functions: (i) an (ir)realis component: *m-/th-* and (ii) a nominalizer component, the common or shared vowel coda: *-ii*.

Of further potential relevance here is the observation that, in addition to pre-nominal demonstrative and interrogative *thii* and *mii*, Literary Burmese also contains a simplex demonstrative element unspecified for (ir)realis distinctions whose form corresponds exactly to the nominalizing component *-ii* in *thii* and *mii*.

- (24) Literary Burmese  
*ii saaouq*  
 this book  
 'this book'

Quite possibly, then, the complex demonstrative-nominalizers *thii* and *mii* are grammaticalizations of the combination of (ir)realis together with this simplex demonstrative *ii*.<sup>11</sup> The creaky tone nominalizers occurring in relative clauses will, in turn, consist of three distinct components: (ir)realis + demonstrative-nominalizer + genitive case, and *th-/m-* + *ii* + creaky tone.

11. The simplex demonstrative *ii* is furthermore found in early Burmese inscriptions before *thii* and *mii* are attested, supporting the feeding relation hypothesized here with *thii* and *mii* being formed from *ii* together with an encoding of (ir)realis meaning. Thanks are due to a reviewer of the chapter for drawing attention to the early occurrence of *ii* in Burmese inscriptions.

### 3.2 Sentential nominalizers in Colloquial Burmese

Turning now to patterns in Colloquial Burmese, one finds that there are elements which occur in the same clause or sentence-final and relative-clause positions as the nominalizers *thii* and *mii* in Literary Burmese. Similar to Literary Burmese, the linking elements in relative-clause structures carry a creaky tone. Also parallel to the situation in Literary Burmese, there are two sets of elements which occur in these positions, one set encoding realis mood, the other encoding irrealis mood:

- a. Realis clause: *te*      realis relative clause linker: *te*<sub>+creak</sub>  
 b. Irrealis clause: *me*      irrealis relative clause linker: *me*<sub>+creak</sub>

Examples (25) to (29) below illustrate the use of these morphemes in clause-final, sentence-final and relative clause environments.

- (25) Colloquial Burmese  
*thuu maneqphan laa me htin te.*  
 he tomorrow come NMLZ think NMLZ  
 'I think he will come tomorrow.'
- (26) Colloquial Burmese  
*thuu manee-ka saaouq we te.*  
 he yesterday book buy NMLZ  
 'He bought a book yesterday.'
- (27) Colloquial Burmese  
*thuu manee-ka we te*<sub>+creak</sub> *saaouq*  
 he yesterday buy NMLZ book  
 'the book he bought yesterday'
- (28) Colloquial Burmese  
*thuu maneqphan thamain-saaouq we me.*  
 he tomorrow history-book buy NMLZ  
 'He will buy a history book tomorrow.'
- (29) Colloquial Burmese  
*thuu maneqphan we me*<sub>+creak</sub> *thamain-saaouq*  
 he tomorrow buy NMLZ history-book  
 'the history book he will buy tomorrow'

Such neat parallels between Literary and Colloquial Burmese diverge, however, in instances where clauses occur as the clearly nominalized arguments of verbs of perception and cognition in subject and object position. In such environments, where the overt occurrence of case markers confirms the nominalized status of the embedded clauses, the elements which embed clauses in Colloquial Burmese



are not, in fact, homophonous with the clause or sentence-final and relative clause morphemes *te* and *me*, but instead have the forms *taa* for realis clauses and *hmaa* for irrealis clauses. Examples (30) to (34) show the occurrence of these elements in a range of subject, object and object of postposition and (ir)realis combinations.

- (30) Colloquial Burmese  
 [thuu thoun-nayii-hmaa hote-ka htweq]-taa-kou myin-ya te.  
 he three.o'clock-at hotel-from exit-NMLZ-ACC see-get NMLZ  
 'I saw him leaving the hotel at three o'clock.'
- (31) Colloquial Burmese  
 [thin ne]-taa-ga hkeq-th-la?  
 learn ASP-NMLZ-NOM difficult-Q  
 'And how about learning it, is it difficult?'
- (32) Colloquial Burmese  
 [thuu baa-hma m pyaw]-taa-ne apyin htweq thwaa te.  
 he whatever not say-NMLZ-with outside exit go NMLZ  
 'I left as/when he did not say anything.'
- (33) Colloquial Burmese  
 [ale pyan]-hmaa-kou pyaw th-laa?  
 visit return-NMLZ-ACC say Q  
 'Did he say he was going to come for a visit?'
- (34) Colloquial Burmese  
 [thuu dii-lou louq]-hmaa-ga theiq m kaun pu, htin te.  
 he this.way do-NMLZ-NOM very not good not think NMLZ  
 'I think it is not good at all that he is going to do that.'

The most plausible analysis of the elements *taa* and *hmaa* is that they result from the collapse of *te* and *me* together with the light noun *haa*, meaning 'one, thing' (similar to English *this one, that one, the blue one*): *te* + *haa* > *taa*; *me* + *haa* > *hmaa*. Elsewhere in Burmese, the element *haa* occurs with this meaning 'one, thing' and can also optionally fuse with the demonstrative *dii* resulting in the form *daa*: *dii haa* 'this one, this thing' > *daa* 'this, that.' The clause-final elements *taa* and *hmaa* are also found to occur in headless relative clauses, as in (35) to (37).

- (35) Colloquial Burmese  
 [canaw hmaa]-taa-ga ase-in-caw pa.  
 I order-NMLZ-NOM vegetable-fried POL  
 'What I ordered was fried vegetables.' (lit. 'the one/thing I ordered')

- (36) Colloquial Burmese  
 [thuu Yangoun-hmaa we]-hmaa-ga seqbein pa.  
 he Yangoun-in buy-NMLZ-NOM bicycle POL  
 'What/the thing he is going to buy in Rangoon is a bicycle.'
- (37) Colloquial Burmese  
 [canaw yee-htaa]-taa-kou thuu theiq m caiq pu.  
 I write-ASP-NMLZ-ACC he very not like not  
 'What I wrote he really does not like.'

In such environments *taa* and *hmaa* again arguably result from the collapse of *te* and *me* and the dummy or light noun *haa*, resulting in a complex fused morpheme with the meaning 'the one, the thing that ...'

Finally, *taa* and *hmaa* are additionally found in alternation with *te* and *me* in sentence-final position, as in examples (38) and (39). In such instances, the use of *taa* and *hmaa*, rather than *te* and *me*, results in a clear difference in meaning and a cleft-like interpretation with focus on one part of the sentence as new information which is set off against a presupposed background (similar to the Chinese *shi-de* construction and Japanese *no-desu* forms; cf. Simpson and Wu 2002b; Simpson 2003a).

- (38) Colloquial Burmese  
 canaw zee-hmaa we taa.  
 I market-in buy NMLZ  
 'I bought it in the market.'/'It is in the market that I bought it.'
- (39) Colloquial Burmese  
 thuu maneqphan yauq hmaa.  
 he tomorrow arrive NMLZ  
 'He will arrive tomorrow.'/'It is tomorrow that he will arrive.'

From such a distribution and patterning, one can hypothesize a likely three-step route of grammaticalization and development of *taa* and *hmaa* from *te* and *me*: (i) from an earlier stage in which *te* and *me* occurred in all clause-final (including relative clause-final) positions, similar to the distribution of Literary Burmese *thii* and *mii*, it can be supposed that the combination of *te* and *me* with *haa* 'one' in headless relative clauses resulted in two collapsed or fused relative clause nominalizers *taa* and *hmaa*; (ii) where such *taa*- and *hmaa*-final relative clauses might have occurred as the natural object of a copula in a main clause (e.g. *This is the one/thing I bought yesterday*), significantly this would then have resulted in *taa* and *hmaa* occurring in sentence-final position, as copulas are regularly null in Colloquial Burmese; (iii) finally, from such occurrence as (relative) clausal nominalizers in sentence-final position, it can be hypothesized that *taa* and *hmaa* may

have allowed for an expansion of use as clause-final nominalizers in clauses which were not only sentence-final, but alternatively embedded as the clausal arguments of verbs of perception and cognition (and also as subjects of one-place predicates such as '(to be) easy/difficult/interesting,' etc.).

If the above can be assumed to be a reasonable path of development for *taa* and *hmaa* as general clausal nominalizers not just restricted to relative clauses, an important question now arises about the synchronic status of *te* and *me*. If *te* and *me* are the original clause-final morphemes in Colloquial Burmese (and partial inputs to the later creation of *taa* and *hmaa*) and correspond to Literary Burmese *thii* and *mii*, as seems most likely, and if the elements *thii* and *mii* in Literary Burmese can be shown to be nominalizers, a fairly natural conclusion is that *te* and *me* may have shared this function as nominalizers too, at least in their earlier stages of development. Such an assumption is supported by the observation that a genitive case creaky tone is present with *te* and *me* in their occurrence as clause-final elements in relative clauses. If (genitive) case is naturally taken to be marked only on nominal constituents, then *te*- and *me*-final relative clauses must be assumed to have been nominal(ized) clauses at least when the genitive creak was first added to *te* and *me*. Consequently, like Literary Burmese *thii* and *mii*, *te* and *me* can be concluded to have had a nominalizing function when added to clauses, in addition to expressing (ir)realis mood, at least at some point in their development. The question to be considered now is whether synchronically these elements are (still) nominalizers in the way that their frequent Literary Burmese equivalents *thii* and *mii* can be shown to be.

The answer to this question would seem to be, clearly and interestingly, 'no.' When simple tests for the nominal status of *te*- and *me*-marked clauses are employed, *te* and *me* do not in fact show signs of nominalizing the clauses with which they are combined. First of all, case markers cannot occur on clauses ending in *te* or *me*, unlike clauses ending in the clear nominalizers *taa* and *hmaa*. Compare (40) and (41) below with earlier (30), (31), (33) and (34).

- (40) Colloquial Burmese  
[U-Win-Win *manee-ga yauq te*](*\*-kou*) *pyaw te*.  
U-Win-Win yesterday arrive NMLZ-ACC say NMLZ  
'He said that U Win Win arrived yesterday.'

- (41) Colloquial Burmese  
[*canaw maneqhpan thwaa me*](*\*-kou*) *htin te*.  
I tomorrow go NMLZ-ACC think NMLZ  
'I think I will go tomorrow.'<sup>12</sup>

12. Matisoff (1972) offers an example of the attachment of accusative case *-kou* to a clause

Secondly, postpositions such as *ne* 'with' can only occur with clearly nominalized clauses ending in either *taa* or *hmaa* and not *te* and *me*:

- (42) Colloquial Burmese  
*sagaa sa pyaw laa taa/\*te ne ta-pain-neq* [...] *word begin say come NMLZ/NMLZ with at-the-same-time*  
'at the same time that he began speaking'

Thirdly, other category-sensitive patterns involving elements introducing rationale clauses similarly indicate that, while *taa* and *hmaa* do create nominal categories, *te* and *me* do not. Specifically, the Burmese words *mouq* and *caun* (meaning 'because') are elements which only allow for combination with nominal phrases and are found to naturally occur with *taa*- and *hmaa*-final (nominalized) clauses. However, they may not occur with *te*- or *me*-final clauses, indicating clearly that the latter elements (synchronically) do not nominalize the clauses with which they combine.<sup>13</sup>

- (43) Colloquial Burmese  
*ngwee lou taa/\*te caun*  
money need NMLZ/NMLZ because  
'because he needed money'
- (44) Colloquial Burmese  
*maneqhpan ethe-twee laa hmaa/\*me mouq*  
tomorrow guest-PL come NMLZ/NMLZ because  
'because there are guests coming tomorrow'

ending in *te*, namely (i) below. However, this sentence is actually not well formed and is rejected by native speakers of Burmese. First of all, the occurrence of negation should automatically displace the occurrence of *te* as in all other negative sentences. Secondly, *-kou* cannot co-occur with *te* in other, non-negative environments. Hence, even if the negation is removed, the sentence remains ungrammatical with *-kou*.

- (i) *\*[[hkinpya shi te m-shi te]-kou be-hne thi m-le?*  
you be NMLZ not-be NMLZ-ACC how know Q  
'How will I know whether you are there or not?'

13. This restriction of *mouq* and *caun* on the categorial status of their complements permitting only nouns or noun phrases would not seem to be reducible to the meaning of *mouq* and *caun* as 'because (of)'. Other lexical items in Burmese with a similar meaning, such as *lou* 'because', impose fully opposite selectional restrictions, combining only with verbal or clausal categories and not allowing noun or nominal complements. It is significant to note that *lou* does allow for combination with a clause ending in *te* or *me*, thus supporting the assumption shortly to be put forward here that such clauses are not nominalizations.

The conclusion to draw from the above is that the elements *te* and *me* in Modern Burmese do *not* create nominal categories and, therefore, synchronically are not nominalizers. However, the available evidence and the patterns surveyed from elsewhere in the language and connected with *te* and *me* clearly suggest that *te* and *me* were, at one time, part of a nominalizer paradigm similar to Literary Burmese *thii* and *mii*. Therefore, it would seem that Colloquial Burmese *te* and *me* can be seen as cases of nominalizers which have undergone a further development in their path of grammaticalization, significantly losing any nominal-related categorial specification which would support case marking and other properties of nominal(ized) constituents, and undergoing reanalysis as grammatical markers more strictly associated with verbal syntax and the representation of (ir)realis mood. If we assume that nominal syntactic patterns (such as case marking) result from the presence of a DP (cf. Section 2.1), the loss of such patterns can be interpreted as indicating that an element which previously caused a DP to be projected has undergone reanalysis over time. Where an erstwhile nominalizer thus loses its ability to project a DP, it can be suggested that it has undergone a shift in categorial identity and is no longer one of the syntactic types which normally cause a DP to be projected, that is, Determiner or Demonstrative, Numeral, Classifier or Noun (cf. Section 2.1). In such instances, it may be concluded that the nominalizer has become reanalysed as an instance of a syntactic category which is associated with verbs rather than with nouns, for example, as an occurrence of T(ense), Asp(ect), Mood, or C(omplementizer), and that this shift in identity to a non-nominal category causes the loss of nominal morphosyntax. In the case of main clause *te* and *me*, it is reasonable to assume that their categorial identity is now strictly (ir)realis Mood. Hence, their hypothesized reanalysis would be from an occurrence and function as a D-type nominalizer to a plain Mood marker combined with verbs, that is, an instance of cross-categorial regrammaticalization from one major functional domain to another (*lateral regrammaticalization*; Simpson 2003a).<sup>14</sup>

A further conclusion following from the above is that, if *te* and *me* are no longer nominalizers in Burmese, then the occurrences of *te* and *me* with a creaky tone in relative clauses can no longer be synchronically the combination of a nominalizer with genitive case.<sup>15</sup> Here again it would seem that categorial reanalysis must

14. Such a putative reanalysis might seem to represent a simplification and reduction of the featural specification of *te* and *me*: from being [+ nominal, + mood], by hypothesis, *te* and *me* become simply [+ mood].

15. Indeed, creaked *te* and *me* do not allow for separation of the genitive case from *te* and *me* via the use of independent genitive case marker (creaked) *ye*: \* [... *te/me ye*]. This suggests that creaked *te* and *me* are fully fused, new forms and are not decomposable into nominalizer and genitive case sub-parts. Also note that, if *te* and *me* do not induce nominal syntactic constituents (i.e. nominalize other phrases), then the creaky tone suprasegmental which is elsewhere an in-

have applied converting a nominalizing morpheme plus genitive case into a new, composite form which is (i) non-nominal/not a category which projects a nominal constituent, and (ii) specified for occurrence with verbs and their projections in the environment of relative clauses. If the occurrence of a creaky tone on *te* and *me* in relative clauses suggests that *te* and *me* in such clauses are not identical to *te* and *me* elsewhere in clause-final position (hypothesized to be instantiations of just Mood), then the dedicated embedding function of creaked *te* and *me* in relative clauses can be suggested to represent a reanalysis of *te* and *me* as new (relative clause) Complementizer elements. Such elements would be formally of type C(omplementizer), and, therefore, distinct from pure Mood *te* and *me*, though carrying a specification for (ir)realis mood in a way which recalls the complex occurrence of subject agreement on Complementizers in languages such as Flemish and Bavarian German (Fuss 2005).<sup>16</sup>

The syntactic patterning observed with Colloquial Burmese *te* and *me* and the way in which this is distinguished from the nominalizer forms *thii* and *mii* in Literary Burmese therefore leads to the hypothesis that a pair of erstwhile nominalizers have undergone a further, important reanalysis as elements lacking nominal specifications, and have come to instantiate new verb-related categories. Generally, then, it can be argued that the grammaticalization of elements as nominalizers does not necessarily represent a final stage of functional development and that nominalizers may, in fact, develop further into verb-related grammatical categories with formally different syntactic properties. To the extent that the distinction between clause-final *te* and *me* and relative-clause embedding creaked *te* and *me* seems to require the assumption of a categorial identity for the latter which is formally different from the former, the paradigms examined here also provide support for the existence of a special C(omplementizer) category. Such a category has often been assumed for elements such as English *that*, taken to be reanalysed as a C(omplementizer) from an earlier source as a D(emonstrative),

indication of genitive case can no longer be genitive case when present with *te* and *me* (in relative clauses), as case marking only marks nominal constituents. Finally, the conclusion that genitive case in relative clauses has been reanalysed as part of a different type of morpheme is supported by other instances in Burmese where genitive case marking seems to have clearly undergone reanalysis. In both Colloquial and Literary Burmese there are sentence-final morphemes which have developed from genitive case markers (creaked *ye* in Colloquial Burmese and creaked *ii* in Literary Burmese) and which now function as markers of (i) surprise in Colloquial Burmese and (ii) the end of the sentence in an (ir)realis neutral way in Literary Burmese. Given the complementary distribution which these elements have with other markers of the category of Mood, it is most natural to see them as reanalysed into this category and no longer performing any (genitive) case function.

16. For a view on how a single morpheme might relate to two syntactic featural specifications, such as Complementizer and Mood or Tense and Mood, cf. Simpson (2003a).

but clear evidence for such a categorial switch or reanalysis in English is not easy to identify. The Burmese patterns under study helpfully seem to suggest that the occurrence of D(emonstrative)-to-C(omplementizer) regrammaticalization and the occurrence of a formally distinct category Complementizer are indeed quite plausible.

#### 4. Summary of conclusions and consequences for further work

This study of Burmese has attempted to explore and highlight a number of properties of nominalization and its grammaticalization which can now be summarized in brief. In Section 2.1 issues concerning the sources of nominalizers were discussed and it was suggested that nominalizers may in fact arise from a variety of sources, including elements of types D, Cl, Num and N. In the course of the investigation of Literary Burmese in Section 3, evidence was presented indicating that clausal and sentential nominalizers in Burmese may be sourced from demonstratives of category D which have also grammaticalized a modal feature encoding (ir)realis distinctions. Section 2 also emphasized the need for specific morphosyntactic evidence when attempting to classify elements as nominalizers and noted that the term *nominalizer* may sometimes be used without clear confirmation of the nominalizer status of an element. In Section 3 it was then argued that, while certain morphemes in Burmese can be concluded to be nominalizers, other elements with a superficially similar distribution have in fact undergone reanalysis as elements of a different categorial type and no longer function as nominalizers in the language. Considering the syntactic realization of relative clauses in Burmese, it was shown that at least two functional morphemes are (historically) involved in the linking of a relative clause to a noun (phrase): (i) a nominalizer and (ii) genitive case. This raises an important question about the identity of linking morphemes in other similar languages (such as, for example, Chinese) where just a single element links a relative clause to a head-noun; are such elements nominalizers or occurrences of genitive case, or possibly even a grammaticalized composite of both? Finally, the examination of Burmese also showed how nominalization structures are potentially unstable and may arguably undergo reanalysis as purely verb-related forms, with nominalizers being absorbed into the functional system projected in a clause. Considerably rich in the information provided by its two systems of functional elements in the colloquial and literary forms of the language, Burmese consequently offers an interesting perspective on variation in processes of grammaticalization and on the reanalysis of grammaticalized morphemes in new directions.

#### Abbreviations

ACC	accusative	NOM	nominative
ASP	aspect	PL	plural
EMPH	emphatic	POL	politeness marker/particle
GEN	genitive	PST	past
INTENS	intensifier	Q	question marker
IRR	irrealis	RLS	realis
NMLZ	nominalizer		

#### References

- Fuss, Eric. 2005. *The rise of agreement: A formal approach to the syntax and grammaticalization of verbal agreement*. Amsterdam: John Benjamins.
- Herring, Susan C. 1991. Nominalization, relativization, and attribution in Lotha, Angami and Burmese. *Linguistics of the Tibeto-Burman Area* 14: 55–72.
- Hopple, Paulette M. 2003. The structure of nominalization in Burmese. PhD dissertation, University of Texas, Arlington.
- Kassevitch, Vadim. 2005. Syntactic and morphological markers in Burmese: Are they really optional? In Justin Watkins (ed.), *Studies in Burmese linguistics*, 87–96. Canberra: Pacific Linguistics.
- Matisoff, James A. 1972. Lahu nominalization, relativization, and genitivization. In John Kimball (ed.), *Syntax and semantics*. Vol. I. New York: Seminar Press: 237–257.
- Noonan, Michael. 1997. Versatile nominalizations. In Joan Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type. In honor of T. Givón*, 373–394. Amsterdam: John Benjamins.
- Okell, John. 1969. *A reference grammar of Colloquial Burmese*. London: OUP.
- Okell, John & Anna Allott. 2001. *Burmese/Myanmar dictionary of grammatical forms*. London: RoutledgeCurzon.
- Saw Tun. 2005. Modern Burmese writing: An examination of the status of Colloquial Burmese. In Justin Watkins (ed.), *Studies in Burmese linguistics*, 201–220. Canberra: Pacific Linguistics.
- Schütze, Carson. 2001. On Korean ‘case stacking’: The varied functions of the particles *ka* and *lul*. *The Linguistic Review* 18(3): 193–232.
- Simpson, Andrew. 2001. Definiteness agreement and the Chinese DP. *Language and Linguistics* 2(1): 125–156.
- Simpson, Andrew. 2003a. Empty determiners and nominalization in Chinese, Korean and Japanese. In Audrey Li & Andrew Simpson (eds.), *Functional structure(s), form and interpretation: Perspectives from Asian languages*, 131–161. London: RoutledgeCurzon.
- Simpson, Andrew. 2003b. On the status of ‘modifying’ DE and the structure of the Chinese DP. In Sze-Wing Tang & Luther Liu Chen-Sheng (eds.), *On the formal way to Chinese language*, 74–101. Stanford CA: CSLI.
- Simpson, Andrew & Zoe Wu. 2001. The grammaticalization of formal nouns and nominalizers in Chinese, Japanese and Korean. In Thomas E. McAuley (ed.), *Language change in East Asia*, 250–283. London: RoutledgeCurzon.

- Simpson, Andrew & Zoe Wu. 2002a. Agreement, shells and focus. *Language* 78(2): 287–313.
- Simpson, Andrew & Zoe Wu. 2002b. From D-to-T: Determiner incorporation and the creation of tense. *Journal of East Asian Linguistics* 11(2): 169–209.
- Thurgood, Graham & Randy LaPolla (eds.). 2003. *The Sino-Tibetan languages*. London: RoutledgeCurzon.
- Van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.

## The grammaticalization cline of cardinal numerals and numeral systems\*

Ferdinand von Mengden

University of Hamburg

This paper investigates the grammaticalization of cardinal numerals and numeral systems from a crosslinguistic perspective. The grammaticalization cline postulated here starts with a referential expression denoting a body part to a cardinal numeral and leads on to a functional affix which occurs in numeral systems and marks arithmetical operations in complex numeral expressions (e.g. English *-teen* in *fifteen* or *-ty* in *fifty*). The paper also shows that the numerals of the lowest cardinalities may follow a path different from that of the other simple numerals: they may evolve from pronominal elements, such as demonstratives and determiners, which often derive themselves from the cardinal numeral '1'. The fact that low numerals and pronominal elements can be both each other's sources and each other's targets seems to challenge the unidirectionality of grammaticalization.

### 1. Introduction

Cardinal numerals have received only marginal attention in the study of individual word classes, whether focussing on their morphology or on their syntax. This neglect applies particularly to studies on grammaticalization. In their *World lexicon of grammaticalization* (2002), Heine and Kuteva mention a couple of processes for which expressions for particular numerical values are claimed to be either typical sources or typical targets of the change. However, they do not include the crosslinguistically observable pathway of (cardinal) numerals as a source/target category. Two of the processes described in the *World lexicon* will be relevant for the present study: (i) the development from a word for 'hand' to the numeral '5' (Heine and Kuteva 2002: 166) and (ii) a list of nine developments in which the numeral '1' evolves into elements expressing concepts which can roughly be subsumed either as 'indefiniteness' or 'singulativity, unity' (Heine and Kuteva 2002: 219–226). In what follows, I will argue that these processes are only two aspects of a broader crosslinguistically observable developmental path in the context of

\* My thanks to Katerina Stathi, Florian Haas, Thomas Hanke and an anonymous reviewer for their critical comments on an earlier version of this paper.