

# Differential Subject Marking in Burmese

San San Hnin Tun, INaLCO Paris

Mathias Jenny, University of Zurich<sup>1</sup>

## 0. Introduction

This study investigates the uses and functions of the “subject marker” *ká* in modern Burmese (cf. Okell & Allott 2001:1f). The marker under discussion here is polysemic, also serving as postposition indicating the source of movement in space or time, labeled ‘ablative’ in the present paper. Though this function is probably related with the use as subject marker, the two functions are syntactically distinct enough to treat them as two separate morphemes. The same marker is also used to emphasize a phrase “as topic of discourse, whether subject or not” (Okell & Allott 2001:2), a function clearly related to the subject marker. The notion of ‘grammatical relations’ (GR) will be established in Burmese, based on a number of constructions. This will be done using a broader typological approach (cf. Bickel 2011). It will be seen that Burmese, though exhibiting some features of topic prominence, can be said to have a grammatical relation that can be labeled ‘subject’, which is relevant in some constructions. While in spoken Burmese subjects are regularly unmarked, some subjects in some contexts take the postpositional marker *ká*, in some cases probably obligatorily. The objective of this study is to establish the different contexts where *ká* is used to mark subjects. Based on a corpus of spoken Burmese as well as elicited data, we look not only at the discourse context, but also at the different predicate and clause types, as well as argument referents with which *ká* is likely or unlikely to occur. We claim that *ká* is not a subject marker *per se*, though it may be in the process of being formalized into one in some genres. Rather we believe that *ká* has discourse pragmatic functions. This is partly in accordance with previous studies and treatments of *ká* in the linguistic literature on Burmese, which are the topic of the second section. The two first sections, which establish the notion of GR in Burmese and look at the development and functions of (differential) subject markers in Burmese and cross-linguistically, will serve as basis for the discussion and analysis of the data presented in section 3. This discussion and analysis will be the topic of section 4. Section 5 sums up the findings and conclusions gained from the analysis of the data.

## 1. Grammatical relations in Burmese

Grammatical relations are “equivalence sets of arguments treated in the same way by some construction in a language” (Bickel 2011:402). Grammatical relations are defined by morphosyntactic properties linking an argument to its clause. This definition makes clear that

---

1 Mathias Jenny provided and elaborated on the theoretical background and the general linguistic analysis of this study, while San San Hnin Tun contributed most of the linguistic data, which is mostly taken from corpus of colloquial Burmese, consisting of 250'000 words. She also provided the statistical corpus data and gave valuable native speaker insight in the analysis of the data.

GRs are construction specific rather than language specific notions. Traditionally, grammars speak of ‘subject’ and ‘object’ as principal grammatical relations, sometimes taking them for granted as universal entities (see Bhat 1991 for a critical view against the universality of GRs). It is evident from the above statement that these notions must be assessed individually not only in each language, but also in each construction that they enter. In many cases there will be no set of arguments that behaves identically in all relevant constructions in a language. The traditional terminology is thus misleading, being applied to different subsets of arguments (cf. Bickel 2011:404). Still, for the sake of simplicity and compactness in expression, it can be useful to talk about notions such as ‘subject’ and ‘object’ if they are well defined in the constructions under consideration. The term ‘subject’ is used for the set of arguments including the single argument (S) of intransitive clauses and the (more) agent-like argument (A) of transitive clauses. Another set of arguments includes some patient-like arguments (O) of transitive clauses and all recipient/goal-like arguments (G) of ditransitive clauses. This GR can be called ‘primary object’ and is marked by the postpositional marker *ko* in Burmese (cf. Sawada 1995). The marking of O arguments in Burmese depends on their referential status, including animacy, definiteness, and topicality, a phenomenon known as ‘differential object marking’ or DOM (see Bossong 1985, 1991; Aissen 2003 among others).

In the following we will look at those constructions that select the {S, A} (or ‘subject’) relation in Burmese. In traditional European and classical Indian grammars, the notion of subject is usually related to the morphological properties of nominative case and verb agreement. On the level of syntax, subjects are important in phrase structure (word order), control and raising constructions, among others. As mentioned above, GRs, including subject, must be defined based on language internal properties. It is therefore necessary to see which (if any) constructions select the {S, A} as relevant set in Burmese.

### *Case marking*

As seen in examples (xx) and (xx) above, colloquial Burmese does make use of postpositional markers to indicate some grammatical and semantic, as well as pragmatic relations. This is no fully fledged system of case marking, though. There is nothing that could be labeled ‘nominative’ case, for example, although some markers, such as *ká* in the above examples, are compatible with the notion of subject. Objects with definite human referent are normally marked by the object marker *ko*, but this is not necessarily the case with non-human or indefinite referents. In terms of case marking, Burmese allows neutral alignment, covering S, A and O (for non-human, non-pronominal O), but not G. Case marking is thus not very useful in establishing GRs in Burmese, either, though it can give some hints in some cases, at least, even if only negative ones. The subject can be defined in terms of case marking as

*the human pronominal argument or personal name that can occur  
in preverbal position without any grammatical marker.*

It is important to include the attribute ‘human (pro)nominal or personal name’, as only

pronouns referring to humans and personal names are obligatorily marked as primary objects by postpositional *ko* if they function as O argument. Relevant examples are given in (xx) and (xx).

(xx) *θu mə-εί ne yin θύ ko pyaʔθana εα λό yá τε.*  
 3 NEG-exist stay COND 3.DEP OBJ problem seek SUB get NFUT  
 'If he wasn't there, we could cause him trouble.' (THGY\_pl)

(xx) *τσανο khələ tchi pì θwà τε.*  
 1M child carry SEQ go NFUT  
 'I carried the child as we went.' (THGY\_pl)

In example (xx), the third person pronoun *θu* is unmarked when occurring as S argument, but receives the object marker *ko* in O position in the second clause, with no overt A argument. In (xx) the first person pronoun is unmarked as A, as is the O, in spite of its human referent.

### Agreement

There is only minimal verb agreement in Burmese, namely the mostly optional plural marker *τá* in postverbal position that occurs with plural referents with either S or A roles. Plural O (or T or G) arguments never trigger this number agreement in the verb complex. This is thus a rather clear instance of a {S, A} set of arguments behaving differently from other types of arguments. Number agreement is not possible with all types of verbal predicates or collocations of arguments and predicates, though. It is at the present stage of research not entirely clear which factors lead to the acceptability or unacceptability of the plural marker on the verb, but it is likely that not only the semantics of the verbal predicate and the arguments, but also pragmatic factors are involved. In terms of verb agreement we can thus define subject as

*the NP that may trigger number agreement in the verb complex.*

Examples (xx) to (xx) below illustrate the use of the verb agreement marker *τá*.

(xx) *θu τό θwà τά τε.*  
 3 PL go PL NFUT  
 'They went.'

(xx) *θαηε.αζιν twe sa phaʔ ne τά τε.*  
 friend PL text read stay PL NFUT  
 'The friends are reading.'

(xx) *θu móun twe sà ne (\*τá) τε.*  
 3 sweets PL eat stay PL NFUT  
 'he is eating different sweets.'

### Phrase structure

In a rather consistent verb-final language with the main focus position in the immediate

preverbal position and frequent omission of known or retrievable arguments, word order is not a good indicator of GRs in Burmese. Any argument or other constituent can occur in almost any position before the verb. The position of the argument does not influence the presence or absence of case marking. Relevant examples are given in (xx) to (xx).

(xx) *phà tɛi ká phà-khəlè ko kan tɛhá lai? tɛ.*  
 frog big SBJ frog-DIM OBJ kick CAUS.fall follow NFUT  
 ‘The big frog kicked the little frog down [into the water] at once.’ (FROG02\_pl)

(xx) *phà-khəlè ko θu ká kan tɛhá lai? tɛ.*  
 frog-DIM OBJ 3 SBJ kick CAUS:fall follow NFUT  
 ‘The big from kicked the little frog down [into the water] at once.’ (FROG02\_pl)

(xx) *ʔè di hma biya tɛhau? bù wɛ tɛ.*  
 ANA this LOC beer six can buy NFUT  
 ‘There I/you/he/she/we/they bought six cans of beer.’ (THGY\_pl)

Apart from word order, case marking and agreement, GRs can also be relevant in different syntactic constructions. Bickel (2011:417ff) gives a representative list of constructions that are found to involve GRs of some type in languages across the world. In the following, we will look in more detail at those constructions that are relevant to the notion of subject in Burmese.

### 1.1 Diathesis

Two types of diathesis are found in Burmese that are relevant to the notion of GRs, namely passive (or quasi-passive), and causative constructions.

#### Passive

Passive constructions typically promote the O or T (in some cases also G) argument of a transitive clause to S of an intransitive clause, with the original A being demoted to an oblique role or dropped from the sentence. The Burmese construction that comes closest to a passive is the pattern *V khan yá* ‘be subject to V, undergo V’ (Okell & Allott 2001:29ff). The semantics of the expression *khan yá* is transparent, being made up from *khan* ‘undergo, accept’ and *yá* ‘get; can, must’. The semantic transparency of the expression results in its not being universally applicable. There is generally a connotation of negative experience attached to the *khan yá* construction, which makes it difficult to use in positive or neutral contexts. Unlike in canonical passive constructions, the original A is not necessarily demoted in Burmese and does not always receive oblique marking. The construction can be analyzed as a complex clause consisting of the matrix predicate *khan yá* and a complement clause describing the event. The complement clause exhibits reduced morphosyntactic possibilities, and the subordinate O argument is always coreferent with the matrix subject and deleted. Apparently only the matrix subject can take the marker *ká*, not the underlying A. It is also the matrix subject that triggers number agreement, never the underlying A. This can be seen as demotion in terms of case marking and agreement possibilities. Examples (xx) to (xx) illustrate the

possible passive-like constructions.

- (xx) *θu ʔame shu **khan** yá tɛ.*  
 3 mother scold accept get NFUT  
 ‘He was scolded by mother.’
- (xx) *θu **ká** khənaʔ.khənaʔ ʔame shu **khan** yá tɛ.*  
 3 SBJ often mother scold accept get NFUT  
 ‘He is often scolded by mother.’
- (xx) ***θu** **tó** ʔame shu **khan** **tsá** yá tɛ.*  
 3 PL mother scold accept PL get NFUT  
 ‘They were scolded by mother.’
- (xx) *θu **shaya** **twe** shu (**\*tsá**) **khan** yá tɛ.*  
 3 teacher PL scold PL accept get NFUT  
 ‘He was scolded by the teachers.’

Alternatively, the verb can be nominalized, with the underlying A receiving the high tone as possessive (or dependent) marker, or the underlying A can take the personal ablative marker *shi-ká*, with the predicate in its verbal form or nominalized, as seen in examples (xx) and (xx).

- (xx) *θu **ʔamé** ʔə-pyɔ̃ **khan** yá tɛ.*  
 3 mother.DEP NML-speak accept get NFUT  
 ‘He had to endure mother’s scolding.’
- (xx) *θu **shayá** **shi** **ká** (ʔə-)shu **khan** yá tɛ.*  
 3 teacher.DEP VIC ABL NML-scold accept get NFUT  
 ‘He was scolded by the teacher.’

As the underlying clause in passivization is always transitive, there is no neutralization possible here for S and A. Passive clauses are generally intransitive, resulting in a derived S, though in Burmese the quasi-passive clause can be considered transitive, resulting in a derived A from underlying O. Both S and A are therefore possible as derived arguments of the *khan yá* construction.

### Causative

The inherited morphological causative, going back to a proto-Tibeto-Burman prefix *\*s-*, which occurs as aspiration in Burmese, has lost its productivity and is fully lexicalized as a small number of verb pairs with aspirated and non-aspirated initial consonants, such as *luʔ* ‘be free, independent’ and *hluʔ* ‘set free, release’, *pyaʔ* ‘be cut off’ and *phyaʔ* ‘cut off’. A periphrastic causative, employing postverbal *se*, is in the spoken language restricted to prohibitive and desiderative contexts. The colloquially more widely used construction involves preverbal *pè* ‘give’. The causative with preverbal *pè* is considered substandard by many speakers and is not widely used in upper Myanmar, where postverbal *khàin* ‘order, command’ is more common. In both types of periphrastic causative, the causee (that is, the S or A of the non-causative

expression) receives the object marker *ko*, more regularly so in the construction with *pè* than with *se*.<sup>2</sup> The original O argument of transitive predicates remains unchanged, though there is a general avoidance of more than one element marked by *ko* in one clause. In this case it is always the causee that receives the marking. Causative expressions with preverbal *pè* are less clearly monoclausal and multiple *ko* markers are more frequent.

(xx) Periphrastic causative constructions

[CAUSER( <i>ká</i> )]	[CAUSEE <i>ko</i> ]	[O( <i>ko</i> )]	<i>pè</i> V
[CAUSER( <i>ká</i> )]	[CAUSEE( <i>ko</i> )]	[O(? <i>ko</i> )]	V <i>se</i>

Examples of the two types of causative constructions are given in (xx) to (xx).

(xx) *ʔame θəmì ko ʔəpyin hma mə=kəzà se tchin phù.*  
 mother daughter OBJ outside LOC NEG=play CAUS DES NEG  
 ‘The mother does not want her daughter to play outside.’

(xx) *shaya t̄əun.ðà twe ko ʔein pè pyan mə.*  
 teacher pupil PL OBJ house give return FUT  
 ‘The teacher will let the pupils go home.’

(xx) *ʔame t̄ənɔ́ ko (shayá ko) pè mə=pyɔ́ phù.*  
 mother 1M.DEP OBJ teacher.DEP OBJ give NEG=speak NEG  
 ‘Mother doesn’t let me tell it to the teacher.’

The inclusion of the indirect object (G = *shayá ko* ‘the teacher’) is rejected by some speakers. The preverbal negator *mə=* can occur either before the main verb or before the causative auxiliary. The word order *mə=pè pyɔ́ phù* is also common, especially in southern varieties, with no clear difference in meaning. There seems to be some regional variation in preferences in the areas where the *pè* construction is commonly used.

The role of the causee in the basic expression as S or A is irrelevant to its assignment to O in the causative expression, so the alignment is again {S, A} against O. In some cases, the causee can be marked as subject. The factors involved here are not clear at present, but it is assumed that the length of the expression, including the distance between causee and causer and word order, as well as pragmatic reasons are relevant here. The following sentence (xx) comes from the Burmese translation of the ‘The Little Prince’. Though the narrative parts are written in literary style, the dialogues are consistently written in colloquial language. The following sentence is considered correct and good language by native speakers, though they might not actually use it themselves.

(xx) *ba phyi?ló θò-khəlè ká di məʔ.lin pin t̄è twe ko*  
 what be SUB sheep-DIM SBJ this monkey.fruit tree big PL OBJ

<sup>2</sup> The causee is usually marked by *ko*. A colloquial alternative is to mark it only as dependent by the creaky tone, without the postposition *ko* (which induces the creaky tone).

*mìn sà se tchin yá ta lè.*  
 2 eat CAUS DES get NFUT.NML Q  
 ‘Why do you (have to) want the little sheep to eat the big monkey fruit trees?’  
 (Dr. Khin Lay Myint, 2004, *Mintha lei*, Yangon: Thiha Yadanar, p. 38)

In some causative constructions, the causee, although marked as primary object of the clause, retains the ability to trigger plural agreement on the verb. This can be explained as inheritance from its role as original S or A. Formally the plural marker is closer to the verb root than the causative marker, indicating that it is the plural predicate as such that is embedded in the causative construction, and the agreement refers to the embedded subject. In other cases, the agreement is not possible. The reasons for this (im)possibility of agreement is still to be determined. Relevant examples are given in (xx) and (xx).

(xx) *θe.dza pyan.lɛ sìn.zà tɛí tɛ́ se tchin pa tɛ.*  
 certain return think look PL CAUS DES POL NFUT  
 ‘[I] want [them/you pl] to think it over again to make sure.’ (SEAlang)

(xx) *tɛ̀ə̀nɔ̀ θu tó ko mə=θwà (\*tɛ́)se tchin phù.*  
 1M 3 PL OBJ NEG=GO PL CAUS DES NEG  
 ‘I don’t want them to go.’

Apart from a few unexpected examples, both (quasi-)passive and causative constructions are obviously sensitive to the {S, A} relation and help further establish the GR of subject in Burmese.

## 1.2 Conjunction reduction and non-finite constructions

Finiteness in Burmese is a feature of the clause, rather than the verb itself. An independent (finite) clause ends in one of a small set of particles indicating tense or status. The most common of these sentence final particles are *tɛ* ‘non-future; NFUT’, *mɛ* ‘future; FUT’<sup>3</sup>, *pi* ‘new situation; NSIT’<sup>4</sup> and *phù* ‘negation; NEG’. Subordinate and other non final clauses, while retaining the ability to have overt arguments and verb agreement, do not take these sentence markers and can be taken as non-finite. They end in one of a number of subordinating particles, such as *pi* ‘finish > sequential/simultaneous’, *ló* ‘that, because’<sup>5</sup>, *phó* ‘so that, in order to’, *yin* ‘if’, *yìn* ‘while’, and others. Some of these subordinators can also occur between a main verb and an auxiliary. The distinction between subordinate clauses and control constructions is therefore not straightforward and not always possible to make on purely syntactic grounds.

3 Alternative labels are realis and irrealis, respectively. A number of reasons that lie beyond the scope of the present study suggest that the analysis as tense markers is more adequate in the case of Burmese.

4 The marker for ‘new situation (NSIT)’ indicates that a situation change has occurred, usually expected, and the situation holds at the time of reference.

5 This is not to be confused with the homonymic quotative particle *ló*, which behaves syntactically differently.

Some of the non-finite subordinate constructions are sensitive to the {S, A} set of arguments, some do not show preferences of any argument over the others. No constructions are found that select exclusively S, A, or O/G, or the set {S, O/G} as relevant grammatical relation. More common is a complete neutralization without restriction, which is not unexpected in a language with frequent argument drop. Examples (xx) to (xx) illustrate conjunction reduction in biclausal expressions with different interpretations of the dropped argument, depending on the semantics of the predicate and the arguments, the position of the overt subject, the context, and, in some cases, on the nature of the linking particle. In some cases there is a strong preference for one or the other reading, which can be canceled with enough context given.

(xx) *θu khwè ko yai? pì Ø pyè θwà tɛ.*  
 3 dog OBJ beat SEQ run go NFUT  
 'He beat the dog and (?\*the dog) ran away.'

(xx) *θu khwè ko yai? ló Ø θe θwà tɛ.*  
 3 dog OBJ beat SUB die go NFUT  
 'He beat the dog and so it (less common: he) died.'

(xx) *tɛənɔ<sub>i</sub> (ká) pai?shan mə=yá yin ?ə-lou? mə=lou? phyi? phù.*  
 1<sub>M</sub> SBJ money NEG=get if NML-do NEG=do be NEG  
 'I won't work If I don't get any money.'

(xx) *tɛənɔ (ká) pai?shan mə=pè yin ?ə-lou? mə=lou? phyi? phù.*  
 1<sub>m</sub> SBJ money NEG=give if NML-do NEG=do be NEG  
 Intended: \*'I won't work if he doesn't pay me.'  
 OK: 'If I don't pay him, he won't work.'

(xx) *Ø<sub>i</sub> θwà yìn Ø<sub>i</sub> la yìn Ø<sub>i</sub> sìn.zà tɛ.*  
 go SIM come SIM think.over NFUT  
 'I think it over as I come and go.' (Okell & Allott 2001:192)

In (xx) only the interpretation with the same subject in both clauses is possible. If the subject is overtly expressed, it can occur either before the subordinate or before the matrix clause. In the position before the subordinate clause, it can be analyzed either as belonging to the matrix or the subordinate clause, and coreference of the subjects is strongly preferred.

The adverbial construction involving a (reduced) nominalized clause with the instrumental postposition *nè* expresses adverbial relations, causal or immediate-sequential. The adverbial clause always ends in the nominalized non-future marker *ta*, so that the combination *ta-né* can be taken as fixed expression with adverbial function. If there is only one overt subject occurring before the adverbial clause, the omitted subject must be coreferent with the overt one. Relevant examples are given in (xx) and (xx).

(xx) *Ø<sub>i</sub> khàlè twe θei? ne yá ta né θu tó<sub>i</sub> nau?tá tɛ.*

child PL put.to.bed stay get NFUT.NML INSTR 3 PL late NFUT  
 'They<sub>i</sub> were late because they<sub>i</sub> had to put the children to bed.'  
 (Okell & Allott 2001:112)

(xx) Ø<sub>i</sub> ʔein pyan yau? ta né Ø<sub>i</sub> thəmìn sà tɛ.  
 house return arrive NFUT.NML INSTR cooked.rice eat NFUT  
 'Right after getting back home he had a meal.' (Okell & Allott 2001:112)

The coreference is canceled if both subjects are overtly expressed, as in (xx).

(xx) yin tai?-hmú phyi? ta né hmà té lu ha  
 vehicle hit-NML be NFUT.NML INSTR wrong NFUT.DEP person TOP  
 dan.ŋwe lè ta? khan yá mɛ.  
 fine ADD attach receive get FUT  
 'If there is an accident the person who did wrong will be fined.' (SEAlang)

Typical control expressions as seen in European languages do not form a homogenous category in Burmese. This seems to be mainly due to the different stages of grammaticalization of individual auxiliaries. While some original secondary verbs have lost their independent status and have become bound operators, such as *se* 'command > CAUSATIVE', *tchin* 'want > DESIDERATIVE', others retain various degrees of independence, such as *yá* 'get > can', which is more independent syntactically than *yá* 'get > must' (s. Jenny 2009). The dependent auxiliaries appear as complex verb forms, while the latter resemble syntactically subordinate clause constructions, though the shared subject may be expressed only once in the sentence. Again, it is the {S, A} set of arguments that is relevant here, that is, the subject of the auxiliary is coreferent with the subject of the main predicate. Examples (xx) to (xx) illustrate the different types of control auxiliaries.

(xx) a. *θu θwà ló yá tɛ.* b. *θu θwà ló mə=yá phù.*  
 3 go SUB get NFUT 3 go SUB NEG=go NEG  
 'He can go.' 'He cannot go.'

(xx) a. *θu θwà yá tɛ.* b. *θu mə=θwà yá phù.*  
 3 go get NFUT 3 NEG=go get NEG  
 'He had to go.' 'He doesn't have to go.'

(xx) a. *θu θwà tchin tɛ.* b. *θu mə=θwà tchin phù.*  
 3 go DES NFUT 3 NEG=go DES NEG  
 'He wants to go.' 'He doesn't want to go.'

(xx) a. *təənɔ θwà yè tɛ.* b. *təənɔ mə=θwà yè phù/ θwà mə=yè phù.*  
 1M go dare NFUT 1M NEG=go dare NEG/ go NEG=dare NEG  
 'I dare to go.' 'I don't dare to go.'

If the auxiliary has causative (permissive or jussive) semantics, the subject of the main verb is

coreferent with the object of the auxiliary and may or may not be overtly expressed. Relevant examples are given in (xx) and (xx).

(xx) (tɕənɔ) (θú ko) θwà (phó) khàin tɕ.  
 1<sub>M</sub> 3<sub>DEP</sub> OBJ go PURP order NFUT  
 'I told him to go.'

(xx) θú ko mə=θwà se né.  
 3 OBJ NEG=go CAUS PROH  
 'Don't let him go.'

### 1.3 Other potentially relevant constructions

Several other constructions have been described in the literature as relevant to GRs, including relative expressions, which are determined by the relativization hierarchy (s. Comrie 1989:147ff; Croft 2003:147ff). Relative clauses in Burmese are not restricted to core arguments. The relativized function of the head can be anything from subject, object, recipient and possessor to almost any adjunct. Relative strategies are therefore not a useful tool to determine GRs, including S. The case is not clear at the present stage of research with quantifier floating, which is possible to some extent and seems to be more common with O as precedent than S or A. Other construction types that are sensitive to the {S, A} set of arguments are imperatives, control of reflexives, and lexical nominalization. The S or A of an imperative is usually not overtly expressed (though its expression in Burmese is more common than in European languages) and always taken to be second person, that is, the addressee. Whether overtly expressed or not, the subject of an imperative can trigger verb agreement. Number agreement on the verb seems to be more regularly expressed in imperatives than in other types of sentences. In a few rare cases, the addressee is not S or A, and a quasi-causative construction must be used in imperative and prohibitive expressions. This seems to be due to the non-agentivity of the subject of *thí* 'touch (inadvertently), come into contact with', though agentivity is not a necessary feature of subjects of imperatives. Compare the sentences in (xx) and (xx) with (xx).

(xx) a. θwà pa. b. mə=θwà né.  
 go POL NEG=go PROH  
 'Go.' 'Don't go.'

(xx) a. di sa pha? tɕá pa. b. di sa mə=pha? tɕá né.  
 this text read PL POL this text NEG=read PL PROH  
 'Read this letter.' 'Don't read this letter.'

(xx) a. <sup>?</sup>da thí pa. b. da mə=thí \*(se) né.  
 this.NML touch POL this.NML NEG=touch CAUS PROH  
 'Touch this.' 'Don't touch this.'

The antecedent of a reflexive expression in Burmese is not necessarily the subject of its immediate clause or the matrix clause, but can be a more distant NP or a covert referent.

Burmese appears to allow what has been called ‘long-distance reflexives’ (cf. Huang 1994:75ff, 2000:190ff and references therein), and the reflexive forms do not seem to be limited to subjects as antecedent, but no studies are available on this topic yet and further research is needed.

Burmese has two ways to nominalize verbal expressions. The first is clause nominalization, with all arguments retaining their marking and the whole finite clause receives the nominalizer *ha*, which merges with the finite particles *tɛ* and *mɛ* to *ta* and *hma* respectively. The nominalizer *ha* functions as an empty nominal head also in other constructions.<sup>6</sup> The clause nominalization is syntactically a relative clause with this dummy head, as in *θu pyɔ̌ tɛ ha* ‘what he says > his speaking’. While there are no morphosyntactic changes in nominalized clauses, the picture is different with nominalized predicates. Here the nominalization applies only to the verbal predicate, which loses its verbal features. Polarity, tense, status, aspect and modality are no longer available and a nominalized verb can not show number agreement. The logical S or A appears as possessor, marked by the high tone where applicable, as in *θú ʔəpyɔ̌* ‘his speaking’. The nominalized verb can itself function as subject and take the marker *ká*, as in shown (xx).

(xx) *θú ʔə-pyɔ̌ ká kàun tɛ.*  
 3.DEP NML-speak SBJ good NFUT  
 ‘His talking is good.’ (as opposed to his acting)

The assignment of possessor to the underlying arguments again chooses the {S, A} set, as opposed to the O and G argument. As sole remnant of the verbhood of the predicate, original O arguments may retain their original object marker, though usually any arguments other than subject are not expressed with nominalized predicates. In (xx) the main verb is nominalized, but the object occurs with the marker *ko*. The sentence is formally a double subject construction, so the main subject receives the marker *ká*.

(xx) *kà ko nain.ŋan.dzà.ðà twe ká pè ʔə-wɛ myà tɛ.*  
 car OBJ foreigner PL SBJ RSTR NML-buy many NFUT  
 ‘It’s foreigners that buy many cars.’ (SEAlang)

### 1.5 Summary of the notion of subject

It has been shown that the notion of subject is relevant in Burmese in a number of constructions. There appears to be no construction where other restricted neutralizations apply, such as {S, O} or others. Burmese thus shows no ergative features in any of the constructions investigated. There are a number of constructions with neutralization of argument role, such as relativization, but they include S, A and O, so they do not define argument sets. The following table summarizes the features found to be relevant for the subject in Burmese.

<sup>6</sup> The noun *ha* is translated as ‘thing’, but it’s main function is as empty nominal head, similar to English *one* in *this one, which one*, etc.

	{S, A}	O	G
1. Case:	∅ always possible sometimes <i>ká</i>	sometimes ∅ sometimes <i>ko</i>	never ∅ always <i>ko</i>
2. Agreement:	possible	never	never
3. Causatives:	becomes O	unchanged	unchanged
4. Conjunction reduction:	always possible	sometimes	never
5. Control:	always	with CAUSATIVE	never
7. Imperatives:	= addressee	≠ addressee	≠ addressee
8. Nominalization:	possessor	unchanged	n.a.

Table 1: Subject features

We now turn to the postpositional markers that have been described in the literature as marking ‘subjects’ in Burmese.

## 2. The postpositional marker *ká* and other ‘subject markers’

The postpositional markers of Burmese have received some attention in descriptive work of the language. A number of them have been described as ‘subject marker’ of some kind or another, namely *ká* and *ha* in spoken Burmese, and *θi*, *ká*, *hmá* in literary Burmese. All of these markers have other functions besides the use as ‘subject markers’. In colloquial Burmese, *ká* can be used to explicitly distinguish the subject from the object, which takes the postpositional marker *ko*. ‘Flagging’ of the GR ‘subject’ is thus one of the established functions of *ká* (Haspelmath 2005:2 and *passim*) In the following, we will look at some earlier treatments of the postpositional markers (claimed to be) involved in subject marking. The overview does not claim to be complete, but to give a representative list of earlier descriptions.

### 2.1 The Myanmar-English dictionary (Myanmar Language Commission)

*ká* (4) 1. Postpositional marker to indicate nominative case (as in *θu ká pyɔ̌* - ‘he speaks’).  
2. Postpositional marker indicating temporal case (as in *məné ká* ‘yesterday’).  
3. Postpositional marker indicating locative case (as in *màndələ̀ ká θəŋɛ.ðə̀in* ‘a friend in/from Mandalay’).

*ká* (5) [colloq] Postpositional marker indicating place (as in *màndələ̀ ká la tɛ* ‘he came from Mandalay’).

*θi* (2) ppm Word indicating nominative case (as in *maun.bá θi sa tɔ̀.zà* - ‘Maung Ba learns hard’).

*ha* (2) ppm [colloq] Same as *θi* (2) ppm (as in *θu ha lu kàun pa* ‘he is a good person’).

The Myanmar Language Commission dictionary lists the markers in question and indicates whether they belong to the colloquial or (by default) formal register, but the examples given for each marker are not conclusive as to the actual function of the marker.

## 2.2 Burmese/Myanmar dictionary of grammatical forms (Okell & Allott 2001)

Okell & Allott (2001) provide an extensive list of grammatical markers in colloquial and formal Burmese, with numerous illustrative examples and explanations, the latter rather descriptive than analytic in linguistic terms. The following list is adapted from Okell & Allott (2001), with the Burmese words transcribed according to the phonological system applied in this paper. We include in this list markers that are, according to the authors, related to topic as well as subject marking.

*ká* 1 (N~) ⇒ **(a) from, by, through N**, when N is a place or starting-point; CB, but also found in FB as an alternative to the more formal FB *N - hmá*. **(b) at N, N ago**, when N is a time, indicates point of time in the past; CB and FB. **(c)** as attribute to a noun: **N -ká N** ⇒ N from N, N at N.

*ká* 2 (N~) ⇒ **marks N as subject of sentence**; CB and FB, usually for emphasis or contrast, or to distinguish the agent from the patient. In FB some writers use *hmá* to serve the same function.

*ká* 3 ⇒ **as for Phr**, emphasises Phr as topic of discourse, whether subject or not; CB and FB. When this *ká* is used with the subject of a sentence it is difficult to distinguish it from *ká* (2) above.

*θi/ði* 3 (N~) ⇒ **marks N as subject or topic of sentence**, FB only. Cf. CB and FB *ká* which has a more contrastive effect. Regular equivalent of of Pali nominative case in Nissaya translations.

*ha* 1 (N~) ⇒ **marks the subject or topic of the sentence**, CB. Cf. CB *ká* and FB *hma*. Sometimes emphatic, sometimes heralding a generalization. [also possessive head, nominal head]

*hmá* 1 **(d) marks N as subject of sentence**, when the subject is a personal agent, or an inanimate noun standing for a personal agent. In this usage, FB *hmá* is equivalent to FB *ká*, but in CB *ká* has more contrastive force. [also restrictive marker 'just (when), only (when)'; FB also ablative marker]

*hma* 2 (N~, Phr~) **indicates topic of sentence**, usually the subject. FB, sometimes CB. [also locative marker]

*kà* 1 (Phr~) ⇒ **as for Phr, however, but**; highlighting the subject or topic of a sentence, or contrasting it to a different topic. FB, + CB *tó, ká-tó*.

It is clear from Okell & Allott's descriptions and examples, that there is some overlap and confusion among the different markers. Interestingly, two of the forms used in colloquial and formal Burmese show the same polysemy, even though the actual markers are not related, such as both *ká* in CB and FB and *hmá* in FB can be used to indicate ablative and subject/topic relations. According to Okell & Allott there is furthermore a difference between the exact functions of *ká* in CB and FB, if only in nuances. We can summarize the different markers relevant in some way to subject marking as in the following table:

Function	CB	FB
TOPIC-CONTR	<i>tʃ, ká-tʃ</i>	<i>kà</i>
TOPIC/SUBJECT	<i>ha</i>	<i>hma</i>
SUBJECT	$\emptyset$	<i>θi</i>
SUBJECT-CONTR	<i>ká</i>	<i>ká</i>
SUBJECT-CONTR		<i>hmá</i>
ABLATIVE	<i>ká</i>	<i>hmá</i>

Table 2: Topic and subject markers

According to both the Myanmar-English dictionary and Okell & Allott, *ha* belongs to the colloquial register. It is found that many speakers assign this particle to the written register and use only *ká* (or  $\emptyset$ ) in the spoken language to mark subjects and *tʃ* (or  $\emptyset$ ) for topics.

### 2.3 Sawada 1995

In a paper published in 1995, Sawada analyses the usages and functions of the particles *ko* and *ká* in colloquial Burmese (Sawada 1995). After considering locative and, by extension, temporal uses of *ká*, Sawada moves on to analyze its function as marker “attached to subjects” and the notion of subject in Burmese (Sawada 1995:175ff). Sawada is careful in making a distinction between “X can be attached to Y” and “Y is marked by X” (Sawada 1995:177ff). This distinction is important in the case of Burmese as, while the notion of subject itself is unproblematic (Sawada 1995:175f; cf. also section 1 above), it is not clear that we can take *ká* as a subject marker rather than a particle that can be attached to subjects. Clearly, in a number of given contexts the subject can be marked by *ká*, in some this is even obligatory. On the other hand, the majority of subjects in colloquial Burmese do not take any formal marking at all, while retaining other syntactic characteristics typical of subjects. In other words, the notion of subject is independent from the marker *ká*. Sawada (1995:180ff) characterizes *ká* as disambiguating device, though with low functional load, to distinguish subject and object. The latter is often marked by the marker *ko* and, where applicable, the oblique form of the nominal, namely by what Okell & Allott call “the induced creaky tone” (2001:273). Another common context where *ká* is attached to a subject is in nominal (equative) predicates (Sawada 1995:180). Colloquial Burmese does not make regular use of the copula *phyi?* in equative expressions, that is, the subject and the predicate complement are both nominal expressions. The marker *ká* is in these cases regularly (obligatorily?) attached to the subject NP. As the predicative NP is always in clause final position, the use of *ká* is not strictly necessary as disambiguating device. A third function of *ká* is to disambiguate between a matrix subject and an embedded subject (Sawada 1995:181). The common structure of embedded clauses is to occur after the matrix subject, leaving the two subjects in adjacent positions:

In cases like these it is only the matrix subject that is marked by *ká*, and it normally takes the marker in natural discourse. Again, the disambiguating functional load is not very high, as the matrix subject always occurs before the subordinate or embedded subject. Other uses of *ká* as attached to subjects are given by Sawada, partly based on descriptions by Okell (1969), including complex sentences containing many expressions (resulting in the subject being separated from its predicate by many intervening elements); where there is a possibility of ambiguity between subject and object; and when a subject in one clause is contrasted with that in another. Sawada (1995:183) concludes that “*ká* with subject has the the function of (somewhat weak) disambiguation and that of selection”.

#### 2.4 Subject markers and differential subject marking across languages

Lehmann (2002:99) mentions Burmese (together with Japanese) as a case of a language with a marked ‘nominative’ that does not go back to an earlier reconstructable ergative system. According to Lehmann, in Burmese “the nominative (both in transitive and intransitive clauses) is optionally marked by *-ka*, the ablative suffix” (2002:99). This direct development from ablative to nominative, or, in other terms, from source to subject marker, is not widely attested among the world’s languages. The expected (or normal) grammaticalization path given by Lehmann is from ablative through genitive and ergative to nominative. According to a cross-linguistic study of the origin and development of agent markers by Palancar (2002:205f), the ablative is the most common spatial category used to express the agent in passive clauses. This does not apply to the Burmese case, though, as there is nothing to suggest that the Burmese construction ever involved passivization. In a number of Tibeto-Burman (and other) languages, ablative marking is used to mark also ergative, that is the A but not the S argument (LaPolla 1995). As Burmese does not show any signs of ergativity, neither in morphology nor in syntax, this general trend in Tibeto-Burman does not explain the situation in Burmese. It is not entirely clear whether in Burmese the subject marker *ká* and the ablative marker *ká* are in fact the same etymon, or rather homophones with different origins. As ablative marker, *ká* is never directly attached to a noun denoting a human referent. In these cases the vicinitive marker *shi* must intervene, and the noun receives the high tone on the last syllable to indicate dependency. The same marker is attached to this class of nouns directly when used as indicator of a contrastive or topical subject, as in examples (xx) and (xx).

(xx) *θu ʔámé shi ká paiʔshan hnə=thaun yá tɛ.*  
 3 mother.DEP VIC ABL money two=thousand get NFUT  
 ‘He got two thousand Kyat from his mother.’

(xx) *ʔame ká shaya-má lou? tɛ.*  
 mother SBJ teacher-F do NFUT  
 ‘Mother is a teacher.’

The use of FB *hmá* in both functions seems to suggest a perceived identity at least at some point in the development of the Burmese language. Semantically, the subject can be seen as the source of the action described by the predicate (at least in the case of a dynamic predicate), which could lead to an extension of use to all kinds of subjects. Lehmann does not provide any discussion or justification of the label ‘nominative’ for the marker *ká*. While it can be attached to S and A arguments, it is not *per se* a marker of subject, as shown by Sawada (see 2.3 above). If we are to identify a ‘nominative’ case in Burmese, the most likely candidate would be FB *θi* (see above, 2.1, 2.2), which corresponds to  $\emptyset$  in CB. Overt marking is used only in certain contexts, which are to be established in the following sections. This suggests that *ká* is not a nominative marker, but is used as non-canonical subject marker (cf. Onishi 2001; de Hoop and de Swart 2008, among others). Mohanan (1995:164ff) describes subjects in Hindi which take ablative marking, but the case is obviously different from Burmese, as the choice of the subject marker is based on the verbal semantics and the construction type, namely agent in passive clauses and causee in causative constructions from transitive bases (Mohanan 1995:169). In Hindi, non-canonically marked subjects do not trigger verb agreement. The verb in these cases usually agrees with the object, or it shows no agreement with any argument in the clause.

The marker *ká* is identical in shape to the ablative marker in colloquial Burmese, as seen above. In some cases the (elaborate) ablative postposition *ká.ne* appears to have a similar discourse function as the marker *ká*, as seen in examples (xx) and (xx). The expression *pì tó* usually expresses sequentiality, but frequently has discourse pragmatic function, reinforcing the contrastive notion of the subject in this case.

(xx) *thùn.ʔeindra.bo ká.ne pì tó θətchìn θwà sho ta.*  
 PN ABL SEQ CONTR song go say NFUT.NML  
 ‘Tun Eindra Bo went to sing.’ (L206; FN.TT1)

(xx) *ʔəiʔ lè ʔəiʔ kɔ́, mí.bá twe ká.ne pì tó θəbɔ́ mə=tu phù.*  
 like ADD like TOP parents PL ABL SEQ CONTR mind NEG=same NEG

*kaun.má-lè yé mí.bá twe ká θəbɔ́ mə=tu bù.*  
 girl-DIM POSS parents PL SBJ mind NEG=same NEG  
 ‘They liked each other, but his parents did not agree. Also the girl’s parents did not agree.’ (L206; FN.TT1)

Another elaborate form of the ablative used with a subject is the expression *phəʔ ká (ne)* ‘from the side’, as in example (xx) from an overheard conversation in Yangon.

(xx) *θin.dàn khá ʔəno tó phəʔ ká.ne pè yá mə.*  
 class fee 1M PL side ABL give get FUT  
 ‘We (are the ones who) will have to pay the class fee.’

This expression seems to be the origin a construction found in some texts of modern Mon prose, which is stylistically heavily influenced by Burmese. While Mon does not normally mark GRs by prepositions, sentences like (xx) with *kəpac* ‘side’ before the subject occur in short stories written in colloquial style.

(xx) *həmùh mùʔ.paʔpèh klɛʔ mɔ̀ŋ kəwəŋ bət pəʂɔn hətaoʔ kə̀h rao,*  
 now why 2 disappear stay extent about five moon MEDL Q

***kəpac*** *wətəə hman klɛŋ pəraŋ mùə wəə kəm ha.*  
 side PN ask come news one time also Q

‘Has Wati ever asked why you have now disappeared for about five months?’

(Chan lon, chan kweh, chan awt. *The Future of Mon*, vol. 12, 2011, p. 86)

This overt subject marking in Mon seems to express contrastive or resumed subjects, but more research is needed in this respect, including the spoken varieties. The phenomenon is not widely attested in modern Mon literature and does not occur in more traditional style.

A superficially similar case is found in Japanese, where the nominative marker *ga*, used to mark “rhematic subjects” (Malchukov 2009:644; see also Ogawa 2009), goes back to an older function as genitive or attributive marker (Ogawa 2009:787; Shibatani 1990:347ff). It is still found as attributive particle (similar to *no*) in a number of fossilized expressions in modern Japanese, such as *wa ga ya* ‘my house’ and *kimi ga yo* ‘Emperor’s era’ (Shibatani 1990:347). According to Shibatani (*ibid.*), the development from attributive marker to nominative case went from subject of nominalized clauses, such as subordinate clauses, to general marker of subject, in both subordinate and independent clauses. This development can be seen as reanalysis or rebracketing of the clause structure. An illustrating example is given in (xx), from Hashimoto (1969; quoted in Shabatani 1990:351) with slightly adapted glosses.

(xx) a. *wa ga [miti]*  
 1SG GEN road  
 ‘my road’

b. *wa ga [iku miti]*  
 1SG GEN go road  
 ‘my road to go/take’

c. [*wa ga iku*] *miti*  
 1SG GEN>NOM go road  
 ‘the road that I go/take’

While the situation in Japanese appears synchronically parallel to Burmese, despite the fact that in Burmese *ká* is preferred with thematic or topical subjects and Japanese *ga* marks rhematic subjects, there is nothing in the Burmese suggesting a development from

nominalized predicates.

A connection of *ká* with the formal topic marker *kà* is also a possibility. The development from topic marker to subject marker is firmly attested cross-linguistically. Heine (2009:466) for example mentions a variety of the North-Khoisan language !Xun, where “the particle *má*, marking sentence topics, has been grammaticalized to a subject marker in specific contexts”, and the Burmese marker *ha*, which also functions as dummy nominal head and nominalizer, is described as Topic/subject marker by Okell and Allot (see above). If we take the notion of subject as grammatical(ized) conflation of topic and agent, this development path makes intuitively sense (cf. Comrie 1989:116ff; Givón 1984:139ff; 2001a:197ff). LaPolla (1995:207) states that “given the strong statistical correlation between agents and topicality, it would not be impossible for a topic marker to regrammaticalize into an agentive marker”. This is what happened in Burmese according to LaPolla, who claims the origin of the marker *ká* to lie in the topic marker *kà*.<sup>7</sup> LaPolla thus treats *ká* as an agentive marker, though it is used with all kinds of subjects. Historical considerations and comparisons are difficult in the case of Burmese, although there are numerous inscriptions from different periods going back at least to the 12th century. These older texts are all written in formal style, with the colloquial style appearing in writing only very recently. We are thus left with mere speculation when it comes to the history of grammatical markers or other aspects of the spoken language.

Apart from the uncommon grammaticalization path from ablative to nominative (source to subject) with no intermediary ergative stage as claimed by Lehmann (2002:99), the mere presence of a special marker for subjects in some contexts in Burmese requires some more investigation. While differential object marking (DOM) has received a fair amount of attention in the linguistic literature, differential subject marking (sbj) is much less common (cf. Malchukov 2006:337f; see also Bhaskararao & Subbarao 2004). There appears to be a strong connection of sbj (or rather DAM) and ergativity, though the correlation is not a clear-cut or absolute one (Malchukov 2006:340; de Hoop & de Swart 2008:3), but it is based on the “Primary Argument Immunity Principle (PAIP)”, which Malchukov (2006:340) states as “Avoid manipulating the case marking of the primary argument exclusively”. In ergative constructions, the primary argument is the {S, O} relation, in accusative constructions it is the {S, A} or subject relation. The PAIP therefore predicts that DAM in accusative constructions should only be possible if there is also DOM in the same construction.

A number of functional and semantic factors trigger sbj in different languages. One frequent phenomenon is differential agent marking (DAM), giving a special case marker to an agentive subject, whether transitive or intransitive. According to the “Relevance Principle (RelP)” (Malchukov 2006:335), features are preferentially marked on the constituent to which they are relevant. The features most relevant to the agent (or, by extension, subject) are *animacy*, *volitionality* and *kinesis* (Malchukov 2006:333). Non-canonical or special marking is thus

---

7 LaPolla does not indicate the tone of the Burmese forms and mentions only *ka*, presumably covering both *ká* and *kà*.

expected on subjects high on the animacy scale. It can also be used to indicate the volitionality of the agent, and conversely the lack thereof if the agent remains unmarked. The RelP and the PAIP can impose conflicting constraints on the case marking of arguments in some constructions. In examples (xx) and (xx), the case marking of the primary argument {S, A} is manipulated, but not the secondary argument, to indicate volitionality of the subject.

(xx) Mongsen Ao (Coupe 2007:157)

a. *a-hən a-tʃak tʃàʔ-àɿ-ùʔ.*  
 NRL-chicken NRL-paddy consume-PRES-DEC  
 'The chickens are eating paddy.'

b. *a-hən nə a-tʃak tʃàʔ-àɿ-ùʔ.*  
 NRL-chicken AGT NRL-paddy consume-PRES-DEC  
 'The chickens are eating paddy.' (implying that they are stealing it)

(xx) Mongsen Ao (Coupe 2001:160f)

a. *nì akhət.*  
 1SG cough.PST  
 'I coughed.'

b. *nì nə akhət.*  
 1SG AGT cough.PST  
 'I coughed.' (i.e. on purpose, to get your attention)

Alternatively, it is the non-volitional subject that receives the non-canonical marking, as shown in the Dhivehi example (xx) by Cain and Gair (2000), quoted in Malchukov (2006:344). Here also the verb takes the 'involitive' form, which originates in a passive form.

(xx) a. *aharen doru leppin.*  
 I door close.PAST.NON.3  
 'I closed the door.'

b. *ahannaš doru leppunu.*  
 I.DAT door close.INV.PAST  
 'I closed the door (involuntarily).'

In other constructions it is the semantics of the predicate that triggers special treatment of the subject, most commonly with experiencer and possessor subjects (see e.g. Onishi 2001 for an overview). Relevant examples are given in examples (xx) and (xx).

(xx) German

a. *Mir ist kalt.* or b. *Mich friert.*  
 1SG.DAT be.PRES.3SG cold 1SG.ACC freeze  
 'I am cold.'

(xx) Icelandic (Andrews 2001:102)

*Hana dreymdi um hafið.*  
 her.ACC dreamed about sea.DEF  
 'She dreamed about the sea.'

The syntactic behavior of non-canonical subjects differs from language to language. In the

German examples XXa. and b. above, the first person as sole argument is taken as syntactic subject, but it does not show any other of the common subject properties such as verb agreement, control and equi-deletion (\**Mich will nicht frieren* 'I don't want to be cold. '; \**Ich ging raus und war kalt* 'I went out and was cold'). In contrast, the presence of the marker *ká* in Burmese has no influence on the syntactic behavior and possibilities of the marked subject. Unlike in German and modern Indo-Aryan languages such as Hindi (s. Mohanan 1995), differential subject marking in Burmese does not result in the subject losing its subject features as described above in section 1.

In summary, one can state two common types of sbj/DAM found cross-linguistically, namely the overt marking of agentivity and volitionality, usually connected with some kind of ergative marking, on the one hand, and non-canonical marking of experiencer and possessor subjects. Which semantic, syntactic, and pragmatic parameters are relevant for triggering the marking of subjects in Burmese, as well as how the Burmese system fits in the well-known cross-linguistic patterns, are the main questions to be answered in section 4. The analysis will be based on the data given in the next section, which lists a number of contexts where subject marking is obligatory, common or (im)possible, as well as the syntactic behavior of marked as opposed to unmarked subjects. The data are taken from natural spoken language and can thus be taken as representative of colloquial Burmese as used by native speakers.

### **3. The data**

The examples of occurrences of *ká* in natural discourse in Burmese given in this section are taken mainly from a corpus of spoken Burmese based on recordings in various contexts involving various kinds of speakers of different sex, age, and social and professional backgrounds, which is supplemented with data from the SEAlang online corpus, and elicited examples by native speakers of Burmese where appropriate. The grammaticality of altered and constructed examples has been checked with different native speakers in Myanmar. The examples are arranged according to different semantic verb (or, more generally, predicate) types, referential properties of arguments, and (discourse) pragmatic contexts. As the data are taken from naturally occurring language use, it is inevitable that there is some degree of overlap between the different occurrence patterns of marked subjects. Examples are given with as much context as deemed necessary to illustrate the point under discussion. In order to determine the difference of marked and unmarked subjects, which is relevant especially in section 3.4, some of the example sentences have been manipulated and the resulting expressions checked with native speakers.

#### **3.1 Predicate classes**

We have seen above (section 2.4) that in a number of languages the verbal semantics are the main trigger for non-canonical subject marking. In this section we look at the cooccurrence possibilities of the marker *ká* with different types of predicates, both verbal and non-verbal, in

order to determine whether the presence or absence of the marker is triggered by the semantics of the predicate.

### *Transitive verbal predicates*

In transitive clauses, both A and O may be expressed or omitted, and either may receive a postpositional case marker. In (xx), only A is marked, in (xx) and (xx) both A and O receive the respective marker.

- (xx) *ʔəbá ká θəyè tɛau? pì tɔ́ θwè.pyɛ? la ta.*  
 father SBJ ghost fear SEQ CONTR shocked come NFUT.NML  
 ‘Father was afraid of ghosts and lost his mind.’ (L49; FN.behBagyi)
- (xx) *hé, boss ká tó ko θəbɔ́.tɛá ló ywè pì khɔ́ ta.*  
 INTERJ boss SBJ 1PL OBJ like SUB choose SEQ call NFUT.NML  
 ‘Hey, the boss chose me because he likes me.’ (L223; RP.KoKo)
- (xx) *ʔè=da ko θu myà twe ká bikini pyɔ́ ta pɔ́ le.*  
 ANA=this.NML OBJ 3 many PL SBJ bikini say NFUT.NML INSIST EMPH  
 ‘Now this one they call bikini, you see.’ (L29; cc.mh)

### *Intransitive verbal predicates*

Marked subjects can occur freely with intransitive predicates, both stative and dynamic, as seen in examples (xx) to (xx).

- (xx) *sìn pɔ́ hma tɛənɔ́ tó ká hmà ló mə=yá phù.*  
 stage on LOC 1M PL SBJ wrong SUB NEG=get NEG  
 ‘Once on stage, we cannot make any mistakes.’ (L32; RIT.MMA)
- (xx) *tɛámá bəwá ká ʔin.mə.tan shò tɛ.*  
 1F life SBJ extremely bad NFUT  
 ‘My life was really bad.’ (L34; INT.AT)
- (xx) *θú ʔəko ká ba ló nainɲan khyà ko thwɛ? θwà lɛ.*  
 3.DEP big.brother SBJ what SUB country other ALL exit go Q  
 ‘Why did his older brother leave for abroad?’ (L257; INT.KLM)

### *Predicate semantics*

The subject can take the marker *ká* irrespective of the verbal semantics of the predicate. As seen in examples (xx) and (xx) above, experiencer subjects can be marked as well as agents, as in (xx) above.

### *Non-verbal predicates and copulas*

The marker *ká* frequently marks the subject of non verbal predicates. In some context this use seems to be obligatory. On the other hand, *ká* is very rare with the copulas *phyi?* ‘be something’, *ɛí* ‘be somewhere’ and *hou?* ‘be so’. The latter is mainly used in questions (*hou? là*

‘is that so?’) and to negate non-verbal predicates. Often the occurrence of *ká* with a copula predicate is considered ungrammatical. There appears thus to be an asymmetry with non-verbal predicates, namely (almost) obligatory use if no copula is present, and (near) ungrammaticality if a copula is present. Relevant examples are given in (xx) to (xx).

(xx) *da ká mi khan θiʔ.ta θɔ.*  
 this.NML SBJ fire accept box key  
 ‘This is the key of the fire-proof box.’ (L101; FLMinLouq)

(xx) *di kaun ká ʔəkhàn θɔ.*  
 this body SBJ room key  
 ‘This one is the room key.’ (L101; FLMinLouq)

(xx) *ʔè di hma shainke né ho dìn pɔ accident (\*ká)*  
 ANA this LOC motorbike INSTR that PRON INSIST accident

*phyiʔ θwà ta pɔ nɔ.*  
 be go NFUT.NML INSIST ASSNT  
 ‘Now, at that place, an accident happened with a motorbike and that thing, right.’  
 (L206; FN.TT1)

(xx) *ban ká thouʔ thà ta (\*ká) tə=θàn lauʔ ɕí tɛ.*  
 bank ABL take.out deposit NFUT.NML one=million about exist NFUT  
 ‘It was about one million that he withdrew from the bank.’ (L101; FLMinLouq)

(xx) *bikini (\*ká) mə=houʔ phù.*  
 bikini NEG=be.so NEG  
 ‘It’s not a bikini.’ (L29; cc.mh)

Some examples of subjects with *ká* in clauses with copula predicates can be found, as in sentence (xx), taken from a popular text written completely in colloquial style. Notice that *ká* can occur only after the subject, not after the predicate complement.

(xx) *θú pyaʔθəna ká kó pyaʔθəna phyiʔ la tɛ.*  
 3.DEF problem SBJ body.DEF problem be come NFUT  
 ‘His problem becomes one’s own problem.’  
 (Sayadaw U Jotika, 2006, *lut-lat thaw seit*, p. 87)

It is evident from the above examples that there is no restriction of occurrence or non-occurrence of *ká* based on the semantic type of the predicate. The only emerging pattern based on predicate type is the high preference of *ká* with non-verbal predicates and high dispreference (or incompatibility) with copulas.

### 3.2 Argument types

The mapping of participant roles onto grammatical relations often interacts with the

referential properties of the participant, such as animacy and definiteness/specificity (Croft 2003:165ff). In Burmese, arguments of any referential type can be assigned the subject role, and the referential properties do not have an influence on the presence or absence of the subject marker. The subject marker *ká* can occur with inanimate subjects, as seen in (xx) and (xx). This implies that the volitionality and control over the event of the subject are not relevant to the occurrence of *ká*, a fact that is also supported by the occurrence of *ká* marked subjects with non-volitional verbs.

(xx) *θatìn twe ká thwe? la tɛ.*  
 news PL SBJ exit come NFUT  
 ‘The news came out.’ (L257; INT.KLM)

(xx) *ʔəkhú ʔətɕhein hma tɛá tɔ̀ mò ká lɛ̀ di lau? mə=hman tɔ̀ phù.*  
 now time LOC fall CONTR sky SBJ ADD this about NEG=true CONTR NEG  
 ‘Nowadays the rain is not falling that regularly anymore.’ (MKB02\_2012)

The subject marked by *ká* can be either definite or indefinite, as seen in examples (xx) to (xx) and (xx) to (xx) respectively. Definiteness is usually taken to include ‘identifiability’, ‘uniqueness’ and ‘inclusiveness’ of the referent (Lyons 1999:1ff). While in the two first examples the subjects, a personal name and a personal pronoun, are uniquely referring expressions, in the last two examples the subjects are partitive quantifiers and therefore indefinite (Lyons 1999:1000ff).

(xx) *ʔè ká máun bè hma thain.*  
 PN SBJ PN.DEP side LOC sit  
 ‘I will sit by your side.’ (L42; RP.chitGati)

(xx) *tɛəɲɔ̀ ká mə=tu ta pè pyɔ̀ lai? mɛ.*  
 1M SBJ NEG=same NFUT.NML RSTR speak follow FUT  
 ‘I will only talk about other things.’ (MKB02\_2012)

(xx) *ʔɛ=da ko θu myà twe ká bikini pyɔ̀ ta pɔ̀ le.*  
 ANA=this.NML OBJ 3 many PL SBJ bikini speak NFUT.NML INSIST EMPH  
 ‘Now this one they call bikini, you see.’ (L29; cc.mh)

(xx) *ʔè nau? tə.tɕho.tə.le ká ywa hma ʔənè sú ká lɛ.ðəmà lou? tɛ.*  
 ANA behind some SBJ village LOC few group SBJ farmer do NFUT  
 ‘Then some, a few people in the village, are farmers.’ (MKB02\_2012)

### 3.3 Context types

#### Information structure

The marker *ká* is usually described as subject and/or topic marker. Topica are variously defined in terms of ‘givenness’, ‘aboutness’, and ‘importance’ in the discourse (cf. Erteshik-Shr 2007:7ff; Lambrecht 1994:117ff). According to Givón (2001b:254), a topic has anaphoric and

cataphoric properties. Anaphorically it has referential accessibility, cataphorically it has thematic importance. One or both properties can be present in a topic, and different grammatical devices may be sensitive to one or the feature, or both. The marker *ká* can be used to mark topics, though rarely so if the topic is not at the same time subject. One of the rare occurrences in the corpus of *ká* occurring with an O argument is (xx). The additive marker *lè* here reinforces the reading of *ká* as additional topic. The collocation *ká-lè* can be seen as fixed discourse marker indicating an additional or resumptive topic, rather than a combination of independent *ká* and *lè*. Another similar common topicalizing expression which is not restricted to subjects is the combination of *ká* with the contrastive marker *tó* as *ká-tó* (*kədɔ*) to indicate contrastive topics.

- (xx) *khələ́ ká lè pyan nó tai? yá ʔòun hma le.*  
 child TOP ADD return milk CAUS.drink get more FUT.NML EMPH  
 ‘We also have to let the child drink her milk again.’ (L34; INT.AT)

While most subjects that are marked are in fact topical, the cooccurrence of the restrictive focus marker *pè* with *ká* is not ungrammatical, though examples as in (xx) seem to be rare (see also (xx) above). In (xx), the focal interrogative *bəθu* ‘who’ takes the marker *ká*.

- (xx) *louʔ-khwín po yá ta ká, ʔə-shò twe ká pè*  
 do-permission exceed get NFUT.NML SBJ NML-bad PL SBJ RSTR

*louʔ-khwín po yá tɛ.*  
 do-permission exceed get NFUT.

‘Those who get more rights to do things, it’s (only) the bad ones that get more rights to do things.’ (SEAlang)

- (xx) *təənɔ́ tó khi? mìn.ðà twe bəθu ká po tɔ́ tɛá ðə=lè.*  
 1M PL era actor PL who SBJ exceed suitable PL NFUT=Q  
 ‘Among the movie stars of our time who is better (than the others).’ (SEAlang)

Other cases of non-topical subjects marked by *ká* are first-sentence subjects in narratives. In *khanzà-thu akyaik wutthu to mya 2010*, a collection of 26 short stories originally published in different magazines, ten are written in colloquial style, both dialogues and narrative parts. Four out of the colloquial style short stories have a marked full NP subject in the first sentence. As first occurrence of a story, these subjects are brand-new and therefore not accessible, in some cases also unanchored. They therefore rank low on the topic acceptability scale (cf. Lambrecht 1994:165ff). Also cataphoric importance is not always part of their pragmatic status, as seen in example (xx). In this example the expression *ʔəlòun* ‘all’ is a universal quantifier (cf. Lyons 1999:32) and can thus be seen as a ‘permanently available topic’ (cf. Ertshik-Shir 2007:17f) which does need anaphoric accessibility. This is less true in (xx). In this case no cataphoric thematic importance is present, and the referent is brand-new, but anchored by the attributive expression ‘of the performance’.

- (xx) *ʔà.lòun ká θú ko má.khin.sì ló khɔ́ tɛá ta pa pè.*

all SBJ 3.DEP OBJ PN QUOT call PL NFUT.NML POL RSTR  
 'Everyone called her Ma Khin Si.' (Khanza-thu akyaik wutthu to mya 2010:209)

(xx) *pwè tɛ́ pəreiʔθa? ká θa θəbɔ̌.tu lɛʔ.khan mɛ sho yin*  
 performance look audience SBJ only agree accept FUT say if

*θin lɛ̀ mìn.ðà phyiʔ khwín cí pa tɛ.*  
 2 ADD actor be chance exist POL NFUT  
 'If the audience of the theater accept you, you have a chance to be an actor.'  
 (Khanza-thu akyaik wutthu to mya 2010:352)

While the marker is more common on topical subjects, topicality alone does not necessarily trigger its presence. In the following sentence, the subject is overtly marked as anaphoric. The predicate is non-verbal, which usually favors subject marking. Still the subject is not overtly marked, besides the additive-topic marker *lɛ̀*, which also marks recursive topics.

(xx) *ʔɛ̀=da lɛ̀ ʔətwe.ʔətɛoun pɔ́.*  
 ANA=this.NML ADD experience INSIST  
 'Now this is another experience.' (L32; RIT.MMA)

#### Dependent vs. matrix clauses

We have seen above (section 2.3) that, according to Sawada (1995), in complex sentences consisting of a matrix and a complement clause, it is the subject of the main clause that takes *ká*, not the subordinate subjects. There seems to be a general tendency to mark only the superordinate subject, as in example (xx). This is not an absolute rule, though, as the example (xx) shows.

(xx) [*di mɛ̀in.má (\*ká) ɲwe mɛʔ ló*] *di mɛ̀in.má ká*  
 this woman SBJ money dream SUB this woman SBJ

*θinbɔ̌.ðà tə=yau? ko lo tɛhin pa tɛ sho ló*  
 sailor one=CL OBJ want DES POL NFUT say SUB  
 'Saying that this woman wants (to marry) a sailor because she's crazy about money ...'  
 (L257; INT.KLM)

(xx) [*ho singəpu ko*] *θu ká ne.ʔaun ká thin ta ká*  
 that Singapore ALL 3 SBJ PN SBJ think NFUT.NML SBJ

[*θú θəmi ká ʔələ θwà mɛ*] *pɛ́ thin ta pɔ́.*  
 3.DEP daughter SBJ trip go FUT RSTR think NFUT.NML INSIST  
 'He, what Nay Aung thought was that his daughter was going to Singapore as a tourist.'  
 (L27; NAR.SSS2)

The structure of this sentence is not very straightforward, typical of spontaneous spoken language. It starts with the topical goal 'Singapore', then has the subject (here realized twice, first as pronoun *θu*, then the personal name) of an intended matrix clause marked with *ká*.

The matrix clause is then nominalized and marked as subject of the following clause, which contains the same nominalized predicate *thin ta* ‘thinking’. The sentence can be seen as anacolouthic construction with the second sentence (*θύ θάμι to thin ta pɔ*) consisting of the nominalized matrix (with no overt subject) and the finite subordinate clause, as juxtaposition of two finite clauses. In other cases, the subject of the subordinate clause cannot take the marker *ká*, even if there is no overt matrix subject. In (xx), *ɔame* cannot take the subject marker.

(xx) [*ɔame (\*ká)* *θe θwà toun.ká*]lè *ɛəmə Ø maun hnəmə*  
 mother SBJ die go when ADD 1F younger.brother younger.sister

*ɲà yau? ko ɛa ɛwè yá ta.*  
 five CL OBJ seek feed get NFUT.NML

‘When mother passed away I had look after the five brothers and sisters.’  
 (L34; INT.AT)

In example (xx) above (repeated here for convenience as (xxa)), the omission of *ká* makes the utterance sound unnatural in a normal discourse situation. If a matrix clause like *ɛəno mə=θí lai? phù* ‘I didn’t know that ...’ is added, *ká* in the first clause would be naturally dropped, as in (xxb).

(xx) a. *ɔábá ká θayè ɛəu? pì tɔ θwè pye? la ta.*  
 father SBJ ghost fear SEQ CONTR blood ruined come NFUT.NML  
 ‘Father was afraid of ghosts and lost his mind.’ (L49; FN.behBagyi)

(xx) a. *ɔábá (ká) θayè ɛəu? pì tɔ θwè pye? la ta*  
 father SBJ ghost fear SEQ CONTR blood ruined come NFUT.NML

*ɛəno mə=θí lai? phù.*  
 1M NEG=know follow NEG

‘I didn’t know that father was afraid of ghosts and lost his mind.’

In example (xx), only the subordinate subject receives the marker *ká*, the matrix subject is unmarked. The matrix subject is not coreferent with the subordinate subject, and both are overtly expressed.

(xx) *shaya-lè ká ɲwe hnə-ya pè tɔ*  
 teacher-DIM SBJ silver two-hundred give CONTR

*θabyɔ Ø ɲwe ko tɔ yu lai? ɛ.*  
 PN silver OBJ CONTR take follow NFUT

‘As soon as the school teacher gave him two hundred Kyat, Thabyaw immediately took the money.’ (Khanza-thu akyaik wutthu to mya 2010:37)

It is not clear at the present stage of research why in some cases the marking of both subordinate and matrix subject is accepted, while in others it is not considered grammatical.

The type of subordinate clause is likely to be relevant in this respect. The following sentences are both accepted, though in (xx) the presence of *ká* on the subordinate subject seems to be dispreferred according to native speakers consulted.

(xx) *θu ká mə=θwà yin tɛənɔ ká lè mə=θwà phù.*  
 3 SBJ NEG=go if 1M SBJ ADD NEG=go NEG  
 ‘If he isn’t going, I’m not going either.’

(xx) *ʔəphe (ká) mə=sà khin ʔəme ká lè mə=sà phù.*  
 father SBJ NEG=eat before mother SBJ ADD NEG=eat NEG  
 ‘Mother is not eating before father has eaten.’

If the subject of the matrix clause is fronted, the subordinate clause becomes center-embedded and the matrix subject preferentially seems to take the marker *ká*. Again, marking the subordinate subject (or both, or neither) is accepted by native speakers, though not with all types of subordinate clauses. Word order is obviously relevant to the presence of the subject marker, at least in some types of complex sentences.

#### Word order

In simple clauses, the order of preverbal constituents is pragmatically determined, as seen above (section 1). The subject can take the marker *ká* irrespective of its position in the clause. If a subordinate clause intervenes between the matrix subject and its predicate, *ká* becomes (near) obligatory. In the following example the subject *Khin Thanda* is separated from its predicate by an adverbial clause, and *ká* cannot be dropped.

(xx) *khin.θandá ká [θinbɔ.ðà tə=yau? phyi? hmá] θú ko lɛʔ.tha? mɛ.*  
 PN SBJ sailor one=CL be just 3.DEP OBJ marry FUT  
 ‘Khin Thanda would marry him only if he was to become a soldier.’ (L257; INT.KLM)

In the following sentence, both matrix and subordinate subjects are marked, the former both in sentence initial position and as resumptive subject after the adverbial expressions. The marker on the subordinate subject could be dropped, but not on the sentence initial matrix subject.

(xx) *θəbawa? ye ká di bɛ? hma mò ká nè tɛ ʔətwe?*  
 nature water SBJ this side LOC sky SBJ little NFUT.DEP for  
  
*θəbawa? ye ká tɔ.tɔ.tan.tan mə=lwe phù.*  
 nature water SBJ quite NEG=easy NEG

‘Natural water, because in this area it doesn’t rain much, natural water is not quite easy (to get in sufficiently).’ (MKB02\_2012)

#### Imperatives

If the subject of an imperative expression is overtly expressed, it can take the marker *ká*.

Example (xx) illustrates this use.

(xx) *maun ká kà ko cé khàn hma han.tə́.pan.tə́ né màun.*  
 PN SBJ car OBJ front room LOC confidentially INSTR drive  
 ‘Maung, in the driver seat, drive the car with style.’ (L42; RP.chitGati)

### Elliptical expressions

The marker *ká* does not occur in one-word replies, such ‘he’ as answer to the question ‘who did it’. It can therefore not be analyzed as default subject marker (or nominative case). This is in contrast to the object marker *ko*, which occurs in one-word replies, regularly on G, less commonly also on O arguments. The homonymic ablative and allative markers *ká* and *ko* show different behavior in elliptical expressions. Compare the question-answer pairs in (xx) to (xx) with (xx) to (xx).

- |      |  |  |
|------|--|--|
| (xx) | <i>di sa.ʎou? bəθú ko pè mə=lè.</i><br>this book who.DEP OBJ give FUT=Q<br>‘To who will you give this book?’ | <i>ʎəmé ko.</i><br>mother.DEP OBJ<br>‘To mother.’        |
| (xx) | <i>zè hma bəθú ko twé lè.</i><br>market LOC who.DEP OBJ meet Q<br>‘Who did you meet at the market?’          | <i>θəŋɛ.dʒìn (ko).</i><br>friend OBJ<br>‘A (my) friend.’ |
| (xx) | <i>da bəθu (ká) yè lè.</i><br>this.NML who SBJ write Q<br>‘Who wrote this?’                                  | <i>ʎəpɛ (*ká) (pɛ).</i><br>father SBJ RSTR<br>‘Father.’  |
| (xx) | <i>θu bɛ θwà lɛ.</i><br>3 which go Q<br>‘Where did he go?’   | <i>yan.goun (ko)</i><br>PN ALL<br>‘To Yangon.’           |
| (xx) | <i>θu bɛ ká la lè.</i><br>3 which ABL come Q<br>‘Where did he come from?’                                    | <i>yan.goun ká.</i><br>PN ABL<br>‘From Yangon.’          |

### Contrastive subject

The impossibility of *ká* to occur in one-word answers may have to do with its preferably marking topical, rather than focal elements, and one-word replies are always focal. This can not be the only reason, though, as focal subjects can occasionally take *ká*. If there is a contrast between two or more possible subjects, as in alternative questions, *ká* can appear on both choices or on neither, but not only one only, as in (xx). The parallelism of the expressions here is important, as the answer to this question would again not be possible with a marked subject alone (*\*ʎəme ká*).

(xx) *da bəθu yè lè, ʎəpɛ (ká) là, ʎəme (ká) là.*  
 this.NML who write Q father SBJ Q mother SBJ Q

'Who wrote this, the father or the mother?'

If contrast is a factor triggering the presence of *ká*, as seen in the parallel expressions, it is obviously not sufficient or the only factor involved. The following pair of clauses (ex. xx) contains two different subjects, both unmarked. If the subject of the first clause is marked by *ká*, the second has to take the marker as well, as in (xxb).

(xx) a. *ʔame móun.hni? ʔatɔ tɔ pè, tɛamá lai? yàun mɛ.*  
 mother batter fried fry give 1F follow sell FUT  
 'Mother, you fry the batter (for me), I go out and sell it.' (L34; INT.AT)

b. *ʔame **ká** móun.hni? ʔatɔ tɔ pè, tɛamá **ká** lai? yàun mɛ.*  
 mother SBJ batter fried fry give 1F SBJ follow sell FUT  
 'Mother fries the batter (for me), I go out and sell it.'

Another example from the same text involving two parallel constructions is given in (xx). Here the subject of the first sentence receives the marker *ká*, the subject of the second clause has the additional contrastive marker *tɔ́*, which often combines with *ká* to mark contrastive topics in general. This combination is not restricted to subjects, but can occur with any constituent of a clause. This example also shows the cooccurrence of ablative *ká* and subject *ká* in the same clause.

(xx) *θú yé ʔapɔ **ká** ha twe **ká** mwè pi pɔ́.*  
 3.DEP POSS above ABL NML PL SBJ born NSIT INSIST

*ʔau? **ká** ha twe **ká** tɔ́ mɛ=mwè θè phù pɔ́.*  
 below ABL NML PL SBJ CONTR NEG=born still NEG INSIST  
 'The older ones were born already, you see. The younger ones, on the other hand, weren't born yet.' (L34; INT.AT)

There are examples of parallel utterances with different subjects where only one takes the marker *ká*, as seen in sentence (xx). This may be due to the fact that the first clause is non-finite (dependent), though it is not subordinate to the second clause.

(xx) *ʔabá pyɔ́ θwà pi sho tɔ́*  
 father speak go finish say CONTR

*tɛanɔ **ká** mɛ=tu ta pè pyɔ́ lai? mɛ.*  
 1M SBJ NEG=same NFUT.NML RSTR speak follow FUT.  
 'If father has already talked about it, I will only talk about different things.'  
 (MKB02\_2012)

Non-contrastive subjects can be marked by *ká*, as in (xx) repeated here for convenience as (xx). The contrast expressed here is not the subject, like 'the sky, rain' as opposed to some other entity, but rather the situation 'now' as opposed to 'before'. In this case, the combination *ká-lè* can again be interpreted as additional topic marker, as seen above in ex. xx).

(xx) *ʔəkhú ʔətchein hma tɛá tɔ̀ mò ká lè di lau? mə=hman tɔ̀ phù.*  
 now time LOC fall CONTR sky SBJ ADD this about NEG=true CONTR NEG  
 ‘Nowadays the rain is not falling that regularly anymore.’ (MKB02\_2012)

The following extract shows the presence and absence of *ká* in one turn of a conversation. Not all subjects that can be understood as contrastive take the marker, and not all subjects marked by *ká* are clearly contrastive.  $\emptyset$  is added in the text where we would expect *ká*.

(xx) *ʔè ʔətchein tɛá hmá tɛəŋɔ̀ tó ká yùn.dɛ ko pyàun pì*  
 ANA time fall just 1M PL SBJ lacquerware ALL change SEQ  
  
*yùn.dɛ lou? yá pa tɛ. yùn.dɛ ká.ne hmá tɛəŋɔ̀  $\emptyset$  ʔəkhú*  
 lacquerware do get POL NFUT lacquerware from just 1M now  
  
*bədzì làin ko ʔəkhú yau? la tɛ. ʔè=dí pyin*  
 painting line ALL now arrive come NFUT ANA=this.DEP outside  
  
*mí.ðà.zú twe  $\emptyset$  lè ʔè=lo yùn.dɛ lou? tɛá tɛ.*  
 family PL ADD ANA=manner lacquerware do PL NFUT  
  
*yùn.dɛ ká tɔ̀ lu.ʔù.yè hnə=thaun cí tɛ sho yin*  
 lacquerware TOP CONTR population two=thousand exist NFUT say if  
  
*ciʔ.shɛ ká yùn.dɛ lou? tɛá tɛ.*  
 eighty SBJ lacquerware do PL NFUT.

‘At that time, when we turned to lacquerware, we had to do lacquerware. From lacquerware I now turned to painting. Besides this, the families do lacquerware like this. As for lacquerware, if we have a population of 2000, eighty do lacquerware.’ (MKB02\_2012)

While *tɛəŋɔ̀* ‘I’ on the second line is in clear contrast with *tɛəŋɔ̀ tó* ‘we’ on the first line, it is unmarked. In the following extract, the subject remains the same over two adjacent sentences. In the first sentence it is in contrast with the subject of the preceding sentence, but this does not hold in the second sentence.

(xx) *ʔè=lo ha myò twe tɛəŋɔ̀ tó ká ʔəmyɛ.tàn tchein yá tɛ.*  
 ANA=manner NML type PL 1M PL SBJ always weigh get NFUT  
  
*sin pɔ̀ hma tɛəŋɔ̀ tó ká hmà ló mə=yá phù.*  
 stage on LOC 1M PL SBJ wrong SUB NEG=get NEG.  
 ‘We always have to think carefully about this. Once on stage, we cannot make any mistakes.’ (L32; RIT.MMA)

### 3.4 Frequency and text genres

The frequency of *ká* varies among different text types. While the overall frequency in the

corpus of 251'344 words is 7'559 tokens or 3.01 %, in the category 'film script' it is only 592 tokens out of 34'079 words or 1.74%. The text genre with the highest occurrence rate of *ká* is 'narratives', which consists of 36'014 words and shows 1'406 tokens of *ká*, a relative frequency of 3.9, more than twice the frequency found in film scripts. There is a marked difference between spontaneous speech, to which narratives belong, and pre-fabricated speech, of which film scripts are part. The overall frequency in the former is 3.48%, in the latter 2.23%. The category 'narratives'

Table (xx) gives the frequency of *ká* in different text categories in the corpus. It is evident that the marker *ká* is a feature of spontaneous speech styles, but not necessarily of interactive speech, as shown by the high frequency in 'narratives', which consists of unprepared accounts of personal experiences, rather than traditional story-telling. The difference between the categories 'interview' and 'radio interview' reflects the fact that radio interviews are not totally spontaneous, as at least the questions are usually pre-fabricated, which is not necessarily the case in casual interviews. The three categories of 'pre-fabricated speech' are all in a way imitations of spontaneous speech. The difference in use of *ká* suggests a rather strong discourse pragmatic role of this marker, which is neglected in non-spontaneous genres.

TEXT CATEGORY	N OF WORDS	N OF TOKENS OF <i>KÁ</i>	FREQUENCY	RANK
<b><i>Spontaneous speech</i></b>	<b>155'614</b>	<b>5'423</b>	<b>3.48%</b>	
Narrative	36'014	1'406	3.90%	1
Interview	72'083	2'564	3.56%	2
Casual conversation	31'611	994	3.14%	3
Radio interview	15'906	459	2.89%	4
<b><i>Pre-fabricated speech</i></b>	<b>95'730</b>	<b>2'136</b>	<b>2.23%</b>	
Dialogue in fiction	20'840	539	2.59%	5
Radio play	40'811	1'005	2.46%	6
Film script	34'079	592	1.74%	7
<b>Overall</b>	<b>251'344</b>	<b>7'559</b>	<b>3.01%</b>	

Table xx: Frequency of *ká* according to text categories

#### 4. Analysis and explanation of the data

We have seen in the preceding section that the choice to mark a subject by *ká* or leave it unmarked does not depend on definable semantic or syntactic factors. There are some clear preferences, such as the marking of topical and matrix clause subjects, especially with center-embedded subordinate clauses, and the non-marking of subjects of copula predicates and focal arguments, but these are not absolute rules. Discourse pragmatics is obviously involved in triggering the presence of *ká*, especially topicality and contrast, but no single determining factor can be taken as responsible for triggering the marking. Subject marking in Burmese

cannot be accounted for in terms of bipolar distinctions and categories, but rather by taking into account a set of features and factors in, and probably beyond, the clause.

## **5. Conclusion**

## Abbreviations

ASST	assent seeking particle
DSPL	displacement in space (and time)
CONTR	contrastive
SBJ	subject marking
NML	nominal(izer)
NFUT	non-future
FUT	future
ANA	anaphoric
INSIST	insisting particle
PL	plural
ADD	additive particle
SEQ	sequential
Q	question marker
RSTR	restrictive particle
NSIT	new situation
SUB	subordinator
TOP	topic

## References

- Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory*, vol. 21/3, pp. 435-83.
- Bhaskararao, Peri & Karumuri Venkata Subbarao (eds.) 2004. *Non-nominative subjects. 2 vols.* Amsterdam/Philadelphia: John Benjamins.
- Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of linguistic typology*. Oxford: Oxford University Press, 399-444.
- Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in Neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Dieter Wanner & Douglas A. Kibbee (eds.) *New analyses in Romance linguistics*. Amsterdam/Philadelphia: John Benjamins, 143-70.
- Comrie, Bernhard. 1989. *Language universals and linguistic typology*. 2nd edition. Oxford: Blackwell.
- Coupe, A. R. 2007. *A grammar of Mongsen Ao*. Berlin/New York: Mouton de Gruyter.
- Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.
- de Hoop, Helen & Peter de Swart. 2008. Cross-linguistic variation in differential subject marking. In Helen de Hoop & Peter de Swart (eds.) *Differential subject marking*. Dordrecht: Springer, 1-16.
- Erteshik-Shir, Nomi. 2007. *Information structure. The syntax-discourse interface*. Oxford: Oxford University Press.
- Givón, Talmy. 1984. *Syntax. A functional-typological introduction. Vol. 1*. Amsterdam/Philadelphia: John Benjamins.
- Givón, Talmy. 2001a. *Syntax. Vol. I*. Amsterdam/Philadelphia: John Benjamins.
- Givón, Talmy. 2001b. *Syntax. Vol. II*. Amsterdam/Philadelphia: John Benjamins.
- Haselmath, Martin. 2005. Argument marking in ditransitive alignment types. In *Linguistic Discovery* 3.1:1-21.
- Heine, Bernd. 2009. Grammaticalization of case. In Andrej Malcuokov & Andrew Spencer (eds.). *The Oxford handbook of case*. Oxford: Oxford University Press: 459-69.
- Huang, Yan. 1994. *The syntax and pragmatics of anaphora. A study with special reference to Chinese*. Cambridge: Cambridge University Press.
- Huang, Yan. 2000. *Anaphora. A cross-linguistic study*. Oxford: Oxford University Press.
- Jenny, Mathias. 2009. Modality in Burmese: 'may' or 'must' - grammatical uses of *yá* 'get'. In

JSEALS 1, 111-126.

Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus and mental representations of discourse referents*. Cambridge: Cambridge University Press.

LaPolla, Randy. 1995. 'Ergative' marking in Tibeto-Burman. In *Senri Ethnological Studies* 41, 189-228.

Lehmann, Christian. 2002. *Thoughts on grammaticalization. Second, revised edition*. Erfurt: Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Malchukov, Andrej. 2006. Transitivity parameters and transitivity alternations. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.) *Case, valency and transitivity*. Amsterdam/Philadelphia: John Benjamins, 329-57.

Malchukov, Andrej. 2009. Rare and 'exotic' cases. In Andrej Malchukov & Andrew Spencer. *The Oxford handbook of case*. Oxford: Oxford University Press, 635-48.

Mohanan, Tara. 1995. *Argument structure in Hindi*. Cambridge: Cambridge University Press.

Myanmar Language Commission. 2006. *Myanmar-English Dictionary*. Yangon: Ministry of Education.

Ogawa, Akio. 2009. Case in a topic-prominent language. Pragmatic and syntactic functions of cases in Japanese. In Andrej Malchukov & Andrew Spencer. *The Oxford handbook of case*. Oxford: Oxford University Press, 779-788.

Okell, John. 1969. *A reference grammar of colloquial Burmese. 2 vols*. London: Oxford University Press.

Okell, John & Anna Allott. 2001. *Burmese/Myanmar dictionary of grammatical forms*. Richmond: Curzon.

Onishi, Masayuki. 2001. Introduction: Non-canonically marked subjects and objects: parameters and properties. In Alexandra Aikhenvald, R. M. W. Dixon & Masayuki Onishi (eds.) *Non-canonical marking of subjects and objects*. Amsterdam/Philadelphia: John Benjamins, 1-51.

Palancar, Enrique L 2002. *The origin of agent markers*. Berlin: Akademie Verlag.

Sawada, Hideo. 1995. The usages and functions of particles *-kou\_/-ka*. in colloquial Burmese. In *Senri Ethnological Studies* 41, 154-87.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.