

# Case frames in Sri Lanka Malay

Sebastian Nordhoff

March 24, 2011

## 1 Introduction

Sri Lanka Malay is a language with direct linking between semantic role and morphosyntactic expression. This direct link is established by means of case enclitics, and these clitics neatly map onto semantic roles. For instance, the clitic =*yang* is used to express patient, and patient can generally not be expressed by any other means. Similarly, =*nang* expresses recipient, and recipient cannot be expressed by any other means.

This is different from languages like English, where the mapping between semantic roles and morphosyntactic expression is more involved. Typically, direct objects are patients or themes in English, but this general rule does not hold in a number of cases. For instance, a patient can be found in subject position, typically reserved for agents, after passivization.

- (1) The city was destroyed by Caesar.

Or, a recipient can be found in direct object position after dative shift.

- (2) John gave me the cake.

Finally, a goal can be found in direct object position with locative alternation

- (3) John loaded the truck with hay.

In languages like English, it is therefore necessary to keep track of the transformations applied in a particular sentence in order to successfully decode the propositional content. This is not necessary in Sri Lanka Malay, where no such alternations exist. Sri Lanka Malay is a WYSIWYG language (Hengeveld, fc; Nordhoff, fc) as far as the encoding and decoding of argument structure is concerned: what you see in morphology (the enclitic) is what you get in semantics (the thematic role).

gloss verb caseframe

## 2 Basics of SLM morphosyntax

Sri Lanka Malay (SLM) is a language based on Malay vocabulary and Dravidian grammar. Its grammar differs dramatically from any other variety of Malay, so that it is considered ‘a language in its own right’ (Adelaar, 1991), rather than a dialect of Malay/Indonesian, e.g. Riau Indonesian (Gil, this volume).

Among the features which distinguish SLM from Malay/Indonesian, we find

- retroflexes, prenasalized stops, vowel length, consonant length
- relatively abundant morphology, e.g. infinitives or participles
- SOV word order
- postpositions

- preposed relative clauses
- grammaticalized indefiniteness marking
- copula constructions (Nordhoff, 2011)
- solid word class distinctions (Nordhoff, forthcominga)
- serial verbs (Nordhoff, forthcomingb)

It is thus about as different from Malay/Indonesian as English is from Marathi as far as grammar is concerned. Its lexicon, on the other hand, is solidly Malay with about 90% cognacy with other Malay varieties (Paauw, 2004).

In the context of valency, word order, head marking, and dependent marking are the most interesting domains. These shall briefly be discussed here.

## 2.1 Word order

SLM word order is generally head final, with SOV being the normal case for pragmatically neutral sentences. Note that there are no grammatical relations in SLM, so that NP NP V would actually be more appropriate than SOV. A model sentence is given in (4)

- (4) *Itthu baathu=yang incayang Seelong=dering laayeng nigiri=nang asà-baapi.*  
 DIST stone=ACC 3s.POLITE Ceylon=ABL other country=DAT CP-bring  
 ‘These stones, he brought them from Ceylon to other countries.’ (K060103nar01)

Note that the realization of four NPs is very atypical; it is more common to only have one overt NP, with the other NPs being dropped. The nature of the NPs has no bearing on the possibility to drop them; any NP can be dropped at any time. The order of the prenominal NPs is governed by pragmatic considerations of topicality. As far as morphosyntax is concerned, word order is free. In pragmatically neutral contexts, all NPs are preverbal. For focalization purposes, right extraposition of one NP is possible.

Example (4) also shows the main device SLM uses for the indication of semantic roles: postpositions. We find the accusative marker =*yang*, the dative marker =*nang* and the ablative marker =*dering*. The pronoun *incayang* ‘he’ does not take a postposition in this sentence. In the context of this paper, I will consider NPs like *incayang* to be in the nominative.

## 2.2 Head marking

While the encoding of semantic roles/valency is quite rich as far as dependents are concerned, such is not the case for verbs: there is no agreement and the existence of any diathesis marking on the verb is debatable (see below).

## 2.3 Dependent marking

The question of ‘who did what to whom’ is answered by case enclitics in Sri Lanka Malay. These morphemes are also often analyzed as postpositions, a theoretical problem common in head-final languages. Nordhoff (2009) adduces evidence for their status as clitics, rather than affixes or independent words. Table 1 gives an overview of the most common morphemes and their meanings. We will mainly be concerned with the morphemes in the first column for the remainder of this paper.

Examples (5) - (8) show prototypical instances of the use of these enclitics.

- (5) *Itthusubbath deram pada jaalang arà-kijja butthul ruuma pada=yang arà-picca-kang.*  
 therefore 3PL PL road NON.PAST-make correct house PL=ACC NON.PAST-broken-CAUS  
 ‘Therefore, they build the street, they demolish many houses.’ (K051222nar04)
- (6) *Kithang lorang=nang baaye mliiga athi-kaasi.*  
 1PL 2PL=DAT good palace IRR-give  
 ‘We will give you beautiful palaces.’ (K051213nar06)

=yang	accusative	=kapang	when
=nang	dative	=siking	because
=dering	ablative	=lanthran	because
=ka	locative	=subbath	because
		=sàsaama	comitative
		sduuduk	ablative

Table 1: Case enclitics in Sri Lanka Malay and their glosses

- (7) *Kaaya oorang=dering=jo arà-cuuri.*  
rich man=ABL=EMPH NON.PAST-steal  
‘It was from rich people that he stole.’ (K051206nar02)
- (8) *Minnyak klaapa=ka gooreng.*  
coconut.oil coconut=LOC fry  
‘Fry it in coconut oil.’ (K060103rec02)

In certain cases, the enclitics can be dropped. This is most notably found in topicalizations as in (9), where the locative marker =ka is dropped.

- (9) *Seelong=∅ samma thumpath=∅ mlaayu aada.*  
Ceylon all place Malay exist  
‘In Sri Lanka, there are Malays everywhere.’ (K051222nar04)

The accusative marker =yang is furthermore often dropped if the context of the sentence already makes it clear which referent is the agent (*baapa* ‘father’) and which referent is the patient (*pohong* ‘tree’), as reverse roles are very unlikely.

- (10) *Baapa derang=pe kubdong=ka hatthu pohong=∅ nya-poothong.*  
father 3PL=POSS garden=LOC INDEF tree PAST-cut  
‘My father cut a tree in their garden.’ (K051205nar05)

Taking a look at a number of semantic roles and a number of case enclitics, one can map the ways how they relate to each other. This is shown in Figure 2.

Above, I have argued that there is a clear and direct mapping between semantic roles and morphosyntactic expression. Taking a look at Table 2, we see that this overstates the regularities. It is common that a morpheme can be used for more than function, and, more problematically, several functions can be encoded by more than one morpheme. This is the case for the coding of patient by both =yang and =nang, as well as the coding of goal by both =nang and =ka.

Looking more closer into these cases, it appears however that they are not of the same type as passivization, dative shift or locative alternation are in English. The use of =nang for marking arguments normally seen as patients is restricted to a very small set of three verbs: *puukul* ‘hit’, *theembak* ‘shoot’, *thiikam* ‘stab’. In all of these three cases, we are dealing with transfer or impact, so that the semantic role could as well be interpreted as goal or recipient. This can be seen from the fact that the English verb *to receive* can also be used felicitously in at least two of the three contexts (*receive a shot*, *receive a blow*). The use of *receive a stab* is less idiomatic in English, but it is clear that an analysis as goal of stabbing or recipient of stabbing is entirely within the realm of possibilities, showing that what we first took as a patient is indeed not a patient in Sri Lanka Malay, but rather a recipient. This thus relieves us of the unfortunate one-to-many mapping of roles and cases. There is a lexical specification of which highly transitive verbs take a recipient (the three just discussed), and which take a patient (all the others). There is no way to alter this and to make a verb from group 1 take =yang, or to make a verb from group 2 take =nang.

The second one-to-many mapping concerns the encoding of goal. This is discussed below in Section 4.1

To sum up, even with the two minor glitches just presented, it seems fair to say that Sri Lanka Malay is very close to a direct mapping of semantic roles onto morphosyntax.

yang	nang	ka	dering
pat	loc	instr	
theme	goal	src	
	rec	path	
	ben		
	purp		
	exp		

Table 2: Mapping of case enclitics and semantic roles in Sri Lanka Malay

## 2.4 Some complications: Modals, institutions and involitive derivation

Verbs are the prototypical assigners of case in all morphosyntactic theories. In Sri Lanka Malay, there are two additional processes which lead to the expression of case on an NP, and which are unrelated to verbs. These are modals and institutional actors. A third process, involitive derivation, changes the feature [+volition] of a verb to [-volition]. As a process, a former actor will lack volition and will instead be coded as an experiencer. These three cases will now be exemplified in turn.

Sri Lanka Malay has a lexical category of modals, whose main members are *boole* ‘can’, *thàrboole* ‘cannot’, *maau* ‘want’ and *thussa* ‘want.not’. These words have in common that they assign the infinitive *mà-* to their clausal complement, and the dative *=nang* to the agent of the verb of their clausal complement. This is exemplified in (11) and (12). In (11), we see a plain clause without case marking on the NP. In (12), where a modal is added, we see that the verb is marked for infinitive with *mà-*, and the NP takes the dative marker *=nang*.

- (11) *Tony arà-nyaanyi*  
 Tony NONPAST-sing  
 ‘Tony sings.’
- (12) *Tony=nang mà-nyaanyi boole*  
 Tony=DAT INF-sing can  
 ‘Tony can sing.’

When the modal is precliticized, the infinitive marker is dropped.

- (13) *Tony=nang boole-nyaanyi*  
 Tony=DAT CAN-sing  
 ‘Tony can sing.’

In cases like (12), the coding of the actor is thus not as expected. Given that the propositional content changes as far as its epistemic value is concerned, we are, however, not dealing with a morphosyntactic alternation expressing the same semantic content, but with a difference in semantic content mirrored by a difference in morphosyntactic encoding. Still, it is important to point out that not all instances of case marking in Sri Lanka Malay are due to the 1:1-mapping argued for above.

Another case where an unrelated process obscures the 1:1-mapping are institutional actors like governments, armies or law enforcement. Institutional actors in Sri Lanka Malay take ablative marking instead of the normal  $\emptyset$  marking. This does not depend on the number of arguments the verb takes, and it does also not alter the coding of the remaining arguments. Examples (14)-(15) show the difference in encoding of non-institutional and institutional actors. In (14), we are dealing with a non-institutional actor, where we find  $\emptyset$ -marking. The institutional actor in (15) receives ablative marking. Again, it is not possible to swap these markings, as this would lead to ungrammaticality.

- (14) *Tony arà-dhaathang*  
 Tony NONPAST-COME  
 ‘Tony is coming.’
- (15) *Police=dering arà-dhaathang*  
 police=ABL NONPAST-COME  
 ‘The police are coming.’

Note that in the English translation, the difference in institutional status is also reflected, namely in the use of the plural agreement on the verb rather than the singular. Like the modals, institutional actors obscure the direct link between thematic role and morphosyntactic expression, but do not give rise to a free alternation of cases frames, which could be exploited for stylistic purposes for instance.

Finally, there is a productive derivational process in Sri Lanka Malay, the so-called ‘involitive derivation’. The prefix *kànà*, when applied to a verb, changes the feature [volition] to [–volition] and makes the new verb assign the dative to a former agent. This is shown in examples (16) - (17). In (16), the singer is volitional and is  $\emptyset$ -marked, whereas in (17), he is non-volitional, and is marked with the dative. One can argue that *kànà*-changes the semantic role a verb assigns from agent to experiencer, the latter unwillingly taking part in the action.

- (16) *Tony arà-nyaanyi*  
 Tony NONPAST-sing  
 ‘Tony is singing.’
- (17) *Tony=nang arà-kànà-nyaani*  
 Tony=DAT NONPAST-INVOL-sing  
 ‘Tony is singing involuntarily/against his will.’

To sum up this section, case marking in Sri Lanka Malay mainly depends on the lexical verb and the thematic roles it licenses. There is generally a transparent relation between the case selected and the underlying thematic role. This can be obscured by processes unrelated to argument structure, namely modals, institutional actors, and involitive derivation. We will now see how these cases are employed in predicates with between zero and four arguments.

### 3 Valency patterns

#### 3.1 One-place predicates

SLM is a role-dominated language with no grammatical relations. The semantic role of a referent is directly reflected in morphosyntax through the case marker selected (See Section 2.3). One-place-predicates can have their arguments marked with either nominative, accusative, dative, or instrumental. The discussion of case frames will therefore be restricted to these four cases here, setting aside spatial and temporal cases, among others. Cases not assigned by the verb, but by a modal or an ‘institutional actor’ are marked by an asterisk

$$(18) \left[ \left( \begin{array}{c} \text{NOM} \\ \text{DAT} \\ \text{ACC} \\ \text{DAT*} \\ \text{ABL*} \end{array} \right) \right]$$

As for one-place predicates, nominative is the most common case (19). Dative is assigned for experiencers (20), or if a modal is present in the clause (22). Accusative-marking is very restricted in one-place predicates and has only been found on one verb *thinggalam* ‘sink’ (21). The instrumental finally is not conditioned by semantic role, but by the nature of the referent: if it is an institution, such as a government or a committee, use the instrumental, otherwise, stick with the semantic role the verb assigns (23).

- (19) *Itthukapang Tony Hassan=∅ su-pii.*  
 then Tony Hassan PAST-go  
 ‘Then Tony Hassan left.’ (K060116nar09)
- (20) *Go=**dang** karang bannyak thàràsìggar.*  
 1s.familiar=dat now very sick  
 ‘I am now very sick.’ (B060115nar04)
- (21) *Titanic kappal=**yang** su-thìnggalam.*  
 Titanic ship=ACC PAST-sink  
 ‘The ship “Titanic” sank.’ (K081104eli05)
- (22) *Kiithang=**nang** [...] two o’clock=ke=sangke bole=duuduk.*  
 1PL=DAT two o’clock=SIMIL=until can-stay  
 ‘We can stay up until two o’ clock.’ (K061026rcp04)
- (23) *Police=**dering** su-dhaathang.*  
 police=abl PAST-come  
 ‘The police came.’ (K081105eli02)

Overall, nominative marking is by far the most frequent, followed by are well-developed dative for actors lacking volition (see below), mainly experiencers. The accusative is rare, and the instrumental is not conditioned by the verb.

### 3.2 Two-place predicates

With two place predicates, we have to distinguish the actor argument and the undergoer argument. The actor can be described in exactly the same way as for one-place predicates: normally, it is encoded in the nominative, but lack of volition or modals in the clause trigger dative marking. Institutional actors take the instrumental. Note that accusative marking for actors is not attested in SLM.<sup>1</sup>

The non-actor argument will normally be in the accusative, although the accusative marker =*yang* is often dropped. Recipients and beneficiaries, as well as some rare patients (of the verbs *puukul* ‘hit’, *thiikam* ‘stab’ *theembak* ‘shoot’) will also be in the dative. Askees are in the locative.

$$(24) \left[ \left\{ \begin{array}{c} \text{NOM} \\ \text{DAT*} \\ \text{ABL*} \end{array} \right\} \left\{ \begin{array}{c} \text{NOM} \\ \text{ACC} \\ \text{DAT} \\ \text{ABL} \\ \text{LOC} \end{array} \right\} \right]$$

Note that the forces operating on the actor argument (modals, institutions) are independent of the forces operating on the undergoer argument, which is only conditioned by the thematic roles the verb assigns. As a result, any of the 3×5 possible combinations in (24) is grammatical.

I will not illustrate the different possibilities for actors here, as these are identical to what has been said above for one-place arguments. As far as the undergoer argument is concerned, it can be marked as nominative (25) or accusative (26).

- (25) *Baapa<sub>A</sub> derang=pe kubbong=ka hatthu pohong<sub>P</sub>=∅ nya-poothong.*  
 father 3PL=POSS garden=LOC INDEF tree PAST-cut  
 ‘My father cut a tree in their garden.’ (K051205nar05)

<sup>1</sup>Accusative marking for actors is found in the adstrate Sinhala (Gair, 2003, 791), so that accusative marking for actors is not a completely absurd idea.

- (26) *Ithukapang lorang=pe leher<sub>P</sub>=(yang) kithang<sub>A</sub>=∅ athi-poothong.*  
 then 2PL=POSS neck=ACC 1PL IRR-cut  
 ‘Then we will cut your neck.’ (K051213nar06)

Dative marking of the undergoer argument is also found, but refers to three semantically distinct concepts: some patients are marked with the dative, as explained above. An examples for this is given in (27).

- (27) *Rose-red<sub>A</sub>=∅ buurung<sub>P</sub>=nang su-puukul.*  
 Rose-red bird=DAT PAST-hit  
 ‘Rose-red hit the bird.’ (K070000wrt04)

If a modal is present in a clause with a verb which already assigns dative to the non-actor, both actor and non-actor will be marked with the dative, giving rise to ambiguity.

- (28) *Se<sub>A</sub>=dang Farook<sub>P</sub>=nang bole=puukul.*  
 1s=DAT Farook=DAT can=hit  
 ‘I can hit Farook.’ (K081104eli05)

In those cases, the actor is normally associated with the leftmost argument, as in (28), while the undergoer is the other argument. When using pointing gestures, this can be overruled, as in (29).

- (29) *Ini kaaka<sub>A/P</sub>=nang itthu kaaka<sub>P/A</sub>=nang bole=puukul.*  
 PROX elder.brother=DAT DIST elder.brother=DAT can=hit  
 ‘This brother can hit that brother.’ ‘That brother can hit this brother.’ (K081104eli05)

A more typical use of the dative is the marking of beneficiaries, as in (30). Recipients are not found in two-place predicates, since they necessarily involve an item received, and therefore are three-place predicates.

- (30) *Derang pada=∅<sub>A</sub> arà-banthu cingala raaja=nang<sub>R</sub>.*  
 3PL PL NON.PAST-help Sinhala king=DAT  
 ‘They help the Sinhalese king.’ (K051206nar03)

Finally, goal of motion is also marked with the dative (31b).

- (31) a. *Guunung<sub>L</sub>=ka=jo kithang<sub>A</sub> arà-duuduk;*  
 mountain=LOC=EMPH 1PL NON.PAST-stay  
 ‘It is in the hills that we live,’  
 b. *guunung<sub>G</sub>=nang=jo kithang<sub>A</sub> arà-pii.*  
 mountain=DAT=EMPH 1PL NON.PAST-go  
 ‘it is to the hills that we go.’ (B060115prs01)

Example (31) also serves to illustrate the use of the locative in two-place predicates, in this case *duuduk* ‘reside’. The ablative is used with two-place predicates to encode the source of motion, as in (32).

- (32) *Spaarua<sub>A</sub> Indonesia<sub>SRC</sub>=dering dhaathang aada.*  
 some Indonesia=ABL come exist  
 ‘Some came from Indonesia.’ (K060108nar02)

Another use of the ablative, namely instrument, is rarer two-place predicates, since an instrument normally requires an additional participant to manipulate, so that we are dealing with at least three participants, which is the topic of the next section. An exception to this generalization is (33), where we find two uses of instruments in the act of playing.

- (33) *∅<sub>A</sub> Thaangang<sub>INSTR</sub>=dering buka<sub>ng</sub> kaaki<sub>INSTR</sub>=dering masà-maayeng.*  
 hand=ABL NEG.NONV leg=ABL must-play  
 ‘You must play not with the hands, but with the feet.’ (N060113nar05)

We have assumed that the actor always is in the nominative in the examples above, barring presence of modals or institutional actors. With experiencer verbs, the actor/experiencer can actually be in the dative, and the under-goer/stimulus either in the nominative or in the accusative.

- (34) *[svaara<sub>ST</sub> hatthu]=∅ derang<sub>EXP</sub>=nang su-dinngar.*  
 noise INDEF 3PL=DAT PAST-hear  
 ‘They heard a noise.’ (K070000wrt04)

- (35) *se=dang<sub>EXP</sub> ini oorang<sub>T</sub>=yang thaau*  
 1S=DAT PROX man=ACC know  
 ‘I know this man.’

To sum up, two-place predicates normally have zero-marked actors, and undergoers are either marked for accusative or dative, with locative and ablative being more marginal possibilities. In special cases, actors can be marked for dative.

### 3.3 Three-place predicates

Three-place predicates are typically predicates of transfer, i.e. giving and taking away. As such, they include an agent, a theme, and a goal (or source). The agent is typically in the nominative, although modals in the sentence can change this to dative, and institutional agents trigger instrumental marking, as always. The theme is either unmarked (=nominative) or in the accusative. Recipients are marked by the dative, and sources by the ablative. This is schematized in (36).

$$(36) \left[ \left\{ \begin{array}{c} \text{NOM} \\ \text{DAT*} \\ \text{ABL*} \end{array} \right\} \left\{ \begin{array}{c} \text{NOM} \\ \text{ACC} \end{array} \right\} \left\{ \begin{array}{c} \text{DAT} \\ \text{ABL} \end{array} \right\} \right]$$

The following examples illustrate these patterns for the verb *kaasi* ‘give’.

- (37) *Se=ppe baapa<sub>A</sub>=∅ incayang<sub>R</sub>=nang ummas<sub>T</sub>=∅ su-kaasi.*  
 1S=POSS father 3S.POLITE=DAT gold PAST-give  
 ‘My father gave him gold.’ (K070000wrt04)

- (38) *∅<sub>A</sub> incayang<sub>R</sub>=nang [appointed member=pe hathu thumpathan]<sub>T</sub>=yang anà-kaasi.*  
 3S.POLITE=DAT appointed member=POSS INDEF post=ACC government=LOC PAST-give  
 ‘(They) gave him a post as appointed member in the government.’ (N061031nar01)

### 3.4 Four-place predicates

Four-place predicates nearly exclusively involve transfer of a theme from a source to a goal. As such the typical case frame is [NOM ACC ABL DAT]. An example is (40). The accusative marker can be dropped as usual.

$$(39) \left[ \left\{ \begin{array}{c} \text{NOM} \\ \text{DAT*} \\ \text{ABL*} \end{array} \right\} \left\{ \begin{array}{c} \text{NOM} \\ \text{ACC} \end{array} \right\} \text{DAT ABL} \right]$$

- (40) *Itthu baathu<sub>T</sub>=yang incayang<sub>A</sub>=∅ Seelong<sub>SRC</sub>=dering laayeng nigiri<sub>R</sub>=nang asà-baapi.*  
 DIST stone=ACC 3S.POLITE Ceylon-ABL other country=DAT CP-bring  
 ‘These stones, he brought them from Ceylon to other countries.’ (K060103nar01)



### 3.5 Summary of valency structure

To sum up the repartition of zero, accusative, dative and instrumental, on the roles of S, A, P and R, the following can be said:

- The dative marker can be found on R and P. Additionally, it can be found on S and A if they are experiencers. Furthermore, modals can assign the dative to S or A
- The accusative marker can be found on P and in rare instances on S
- The ablative marker can be found on S and A when they are institutional. It can furthermore be found on instruments and sources, widely construed.
- Zero can be found on S, A and P. Zero is never found on R.

## 4 Alternations

### 4.1 Uncoded alternations: Case

The case to use in a given context is pretty much fixed in Sri Lanka Malay. There is normally no choice to use one marker or another. One instance where more than one case is available to encode the same content is the choice between dative and locative for goal of motion.

- (41) a. *Tony Kluum̃bu=**nang** su-pii*  
Tony Colombo=DAT PAST-go  
'Tony went to Colombo.'
- b. *Tony Kluum̃bu=**ka** su-pii*  
Tony Colombo=LOC PAST-go  
'Tony went to Colombo.'

This is a systematic alternation in Sri Lanka Malay. Goal of motion may be either expressed by =*nang*, as it were the recipient of the motion, or by =*ka*, like stative location. There seems to be free variation between the two forms. Also, there are otherwise no morphosyntactic repercussions of choosing one form over another. This contrasts with the English passive for instance, where a different encoding of the patient triggers different encoding of the agent, and a difference in voice marking as well.

A second alternation, even more restricted, is the choice between locative and ablative for the verb *mintha* 'ask'. The ablative is normally used when the item requested is tangible, whereas the locative is used when information is requested.

For all other semantic roles, the mapping is deterministic. What you see in semantics is what you get in morphosyntax, and there is no way to change this.

Depending on the analysis of the accusative marker =*yang*, one could argue that there is a alternation between  $\emptyset$ -marked nominative and =*yang*-marked accusative for certain verbs.

- (42) *Baapa derang=pe kubbong=**ka** hatthu pohong= $\emptyset$  nya-poothong.*  
father 3PL=POSS garden=LOC INDEF tree PAST-cut  
'My father cut a tree in their garden.' (K051205nar05)
- (43) *Ithukapang lorang=pe leher=(**yang**) kithang= $\emptyset$  athi-poothong.*  
then 2PL=POSS neck=ACC 1PL IRR-cut  
'Then we will cut your neck.' (K051213nar06)

An analysis involving the optionality of =*yang*, which only surfaces in order to disambiguate, seems to be the better solution, though.

## 4.2 Coded alternations: Verb

Sri Lanka Malay is not a language with strict subcategorization for the number of arguments a verb takes. It is always possible to drop arguments, and in most cases, it is possible to add arguments without further ado. There are a number of additional morphosyntactic devices which are related to the number of arguments a verb takes and which will be discussed in this section. These devices comprise the causative suffix, the prefix *kasi-* and the postverbal ‘vector verbs’ *kaasi* ‘give’ and *ambel* ‘take’.

### 4.2.1 Causative -king

The most trivial case is the causative morpheme *-king* with the allomorph *-kang*. This morpheme attaches to a verb or an adjective and introduces an additional causer, who causes the initial actor to take part in the event denoted by the predicate. Examples (44)–(46) show this for an adjective, an ‘intransitive’ verb and a ‘transitive’ verb.

- (44) *Itthuka asà-thaaro, itthu=yang arà-panas<sub>adj</sub>-king.*  
DIST=LOC CP-put DIST=ACC NON.PAST-hot-CAUS  
‘Having put (it) there, you heat it.’ (B060115rcp02)
- (45) *Inni=ka inni daalang=ka kithang aayer masà-mlidi<sub>v.intr</sub>-king.*  
PROX=LOC PROX inside=LOC 1PL water must=boil-CAUS  
‘On this, inside this, we must boil water/bring the water to a boil.’ (B060115rcp02)
- (46) *De laaye hathu nigiri=nang anà-baapi, buunung<sub>v.tr</sub>-king=nang.*  
3s.IMPOLITE other INDEF country=DAT PAST-bring kill-CAUS=DAT  
‘They brought him to another country to have him executed.’ (K051206nar02)

### 4.2.2 Prefix *kasi-*

There are two, possibly three, verbs in Sri Lanka Malay which can be argued to involve a valency-changing derivation with the element *kasi*, related to *kaasi* ‘give’ (Nordhoff, forthcomingb). These are *kasithaau* ‘inform’ from *thaau* ‘know’, *kasikaaving* ‘to give in marriage’ from *kaaving* ‘marry’, and *kasikinnal* ‘introduce’ from *kinnal* ‘know, be acquainted’. In all the three cases, a recipient is added as a new semantic role. Recipients are not among the roles the verbs without *kasi* assign. The morpheme *kasi-* is the mechanism closest to a valency-changing device Sri Lanka Malay has, but, as stated above, it is limited to a handful of verbs and probably lexicalized.

### 4.2.3 Vector verb *kaasi*

At first sight, there is another process which seems to add a beneficiary, the vector verb *kaasi* ‘give’. Note that this form is obviously etymologically related to the form just discussed, but has a long vowel and occurs after the verb it modifies, not before.

Combinations of V+*kaasi* are used to mark beneficiaries as in (47), where the fact of receiving explanations is highlighted as beneficial. Related constructions in Sinhala and Tamil have been termed ‘alterbenefactive’ (Lehmann, 1989, 227).

- (47) *Kithang=pe ini younger generation=nang=jo konnyong masà-biilang kaasi, masà-aajar.*  
1PL=POSS PROX younger generation=DAT=EMPH few must-say give must-teach  
‘It is to the younger generation that we must explain it, must teach it.’

The question which arises is whether *kaasi* changes the valency structure of the verb it modifies. Already example (47) suggests that this is unlikely. The verb *biilang* ‘say’ typically takes three arguments, a speaker, a message, and an addressee. This fact does not change through the addition of *kaasi* in (47). This is even clearer in (48). *Kaasi* could be thought to increase valency, but example (48) shows that it does in fact not have any influence on the quantity of arguments.

- (48) *Itthu muusing, [Islam igaama nya-aajar kaasi Ø] Jaapna Hindu teacher.*  
 DIST time Islam religion PAST-teach give Jaffna Hindu teacher  
 ‘At that time, those who taught Islamic religion were Hindu teachers from Jaffna.’ (K051213nar03)

The verb *aajar* ‘teach’ has already three argument before derivation, an agent, a theme, and a recipient/beneficiary. The same three roles are present after *kaasi* is added, but the beneficial component is highlighted. It is impossible to claim that *kaasi* adds an argument to *aajar*, since the number of arguments does not change. Furthermore, we also note that the quality of arguments does not change: the semantic roles are exactly identical, with the proviso that the ambiguous status of recipient/beneficiary is definitely changed to beneficiary in the *kaasi*-clause. This semantic difference, however, is not discernible in case-marking since both recipients and beneficiaries are marked by the dative.

#### 4.2.4 Vector verb *ambel*

Next to the vector verb *kaasi* ‘give’, we also find the vector verb *ambel* ‘take’, which is in a number of respects the opposite of *kaasi*. *Kaasi* highlights the alterbenefactive aspect of a certain event; *ambel* highlights the self-beneficial aspect.<sup>2</sup> An example is (49), where the normally neutral action of catching is marked as self-benefactive.

- (49) *British government Malaysia Indonesia ini nigiri pada samma anà-peegang ambel.*  
 British government Malaysia Indonesia PROX country PL all PAST-catch take.  
 ‘The British government captured Malaysia, Indonesia, all these countries.’ (K051213nar06,K081104eli06)

Compare this to (50), where the action is not beneficial, and *ambel* is absent.

- (50) *Ø Heart attack asà-peegang, baapa=le su-niñḡgal.*  
 heart attack CP-catch father=ADDIT PAST-die  
 ‘(My father) got a heart attack and died as well.’ (K051205nar05,K081104eli06)

Note that neither the number of arguments nor their morphosyntactic coding change with regard to the sentence without *ambel*, showing that self-benefactive *ambel* triggers no morphosyntactic changes.

The last process, related to self-benefactive *ambel*, is reflexive *ambel*. This is often found together with the emphatic clitic *=jo*, although this is optional.

- (51) *Incayang incayang=yang(=jo) su-buunung ambel.*  
 3S.POLITE 3S.POLITE=ACC=EMPH PAST-kill take  
 ‘He killed himself.’ (K081106eli01)

The question is again whether this changes the quantity or quality of arguments. The answer is again negative. Even when *ambel* or *=jo* are present, it is possible to express both the agent and the patient, as shown in (51). This is normally not done for obvious pragmatic reasons of economy.

We can say that the sentence in (51) has an agent NP and a patient NP just like its non-reflexive counterpart, e.g. (52).

- (52) *Kaake baapa=yang su-buunung.*  
 grandad father=DAT PAST-kill  
 ‘His grandad killed his father.’ (K081103eli04)

These NPs in (51) and (52) are marked in the same way, so that there is no reason to suspect morphosyntactic reflexes of valency structure.

The best candidates for valency-changing operations, *kasi*-V, V *kaasi*, and V *ambel* were discussed in this section. The first of these operations might be seen as valency-changing, but is restricted to three lexemes. The remaining two looked promising at first sight, but had to be discarded after closer scrutiny because they affected neither quantity nor quality of arguments.

<sup>2</sup>There are other usages of *ambel* which will not be discussed here. The reader is referred to Nordhoff (2009) and Nordhoff (forthcomingb).

### 4.3 Other alternations

Sri Lanka Malay can drop any NP if its recoverable from context. This means that it is perfectly normal for a three-place verb like *kaasi* ‘give’ to be uttered in isolation as in (53).

- (53) *su-kaasi*  
 PAST-give  
 ‘X gave Y to Z (and, you, my dear interlocutor will surely be smart enough to figure out who X, Y, and Z are)’.

It is clear that there are no slots which obligatorily have to be filled in Sri Lanka Malay. As such, the question arises how a sentence like (53) can be distinguished from a sentence like (54).

- (54) *su-uujang*  
 PAST-rain  
 ‘(It) rained.’

The overt morphosyntactic structure is exactly parallel. Where the two sentences (53) and (54) can be distinguished is by the possibility to add new referents. This is patently simple for (53), see (55), but nigh impossible for (54) barring very contrived contexts.

- (55) *Tony kaake=nang ini car=yang su-kaasi*  
 Tony grandfather=DAT PROX car=ACC past-give  
 ‘Tony gave this car to Grandfather.’

While it is very difficult to add more referents to *uujang* ‘rain’, similar problems are not found with *kaasi* ‘give’. The question arises whether the (im)possibility to add new referents might not be a better cue to case frame membership than the (im)possibility to leave out certain referents.

Taking a look at a prototypical transitive predicate like *poothong* ‘cut’, it is immediately obvious that referents bearing the dative and the ablative can be added without problems. The normal transitive sentence in (56) can be made to accommodate four NPs. The resulting sentence is (57) and is as such then indistinguishable in argument structure from (40).

- (56) *Tony ini daaging=yang arà-poothong*  
 Tony PROX meat=ACC NONPAST-cut  
 ‘Tony cuts this meat.’
- (57) *Tony piiso=dering ini daaging=yang kaake=nang arà-poothong*  
 Tony knife=ABL PROX meat=ACC grandfather=DAT NONPAST-cut  
 ‘Tony cuts this meat for his grandfather with a knife.’

NPs referring to space or time can also be added to the sentence without problems, a fact which is little surprising. All predicates can accommodate a phrase like *Kluuñbu=ka* ‘in Colombo’ or *soore=ka* ‘in the evening’. Out of the case markers surveyed in this chapter (=yang ‘acc’, =nang ‘dat’, =ka ‘loc’, =dering ‘abl’), there is only one which cannot be added to a predicate without further ado: the accusative marker =yang. Augmenting a predication with beneficiaries, instruments, locations and points in time is no problem, but adding an extra patient is extremely difficult. This clear separation, and the absence of similar patterns elsewhere, suggests that SLM verbs can be divided into two-classes: those which can take an argument with =yang (but might drop it), and those which never can take such an argument. This is the main difference made in the SLM lexicon as far as verbs go. These two classes will now be surveyed in turn.

## 5 Verb classes in Sri Lanka Malay

### 5.1 Verbs taking =yang

These are verbs which have semantic component of affectedness in them, and which thus assign a patient. Patient-taking verbs can be intransitive, like *thinggalam* ‘sink’, transitive like *poothong* ‘cut’, ditransitive like *kaasi* ‘give’ or tritransitive like *kiiring* ‘send’.

### 5.2 Verbs not taking =yang

These are verbs which either do not have any connection with affectedness at all in their meaning (*dhaathang* ‘come’), or whose affected argument is conceived of as a recipient rather than a patient in Sri Lanka.<sup>3</sup> This set comprises atransitive verbs like *uujiang* ‘rain’, intransitive verbs like *laari* ‘run’, and transitive verbs like *maaki* ‘scold’. Verbs with three or more places typically involve a patient or a theme and hence show a possibility to include =yang-phrases. Themes are normally marked by  $\emptyset$ , but by choosing a referent ranging high on the scales of animacy, definiteness, individuatedness, and topicality, =yang-marking becomes an option.

Verbs not taking =yang can be derived by the causativizer *-king* and then allow the use of =yang.

## 6 Conclusion

Grammatical information which is non-predictable and arbitrary is stored in the lexicon. The English lexicon has to store that *eat* can have one or two arguments, but that *devour* has exactly two. The intransitive option is not available for this verb. What is there to store in the SLM lexicon, and in how far does it inform our theorizing about valency? It is clear that the minimal number of argument slots is not an important category in SLM. In discourse, the number of arguments is quite low due to the frequent dropping of whole NPs. Any verb in SLM can be used with 0 realized arguments without leading to ungrammaticality. The maximum number of arguments an SLM verb can take is not important either. In English, verbs like *hesitate* do not take any non-actor argument, despite semantically plausible interpretations of a string like *I hesitate the decision* being available without causing too much of a headache. In SLM, NPs taking the dative =*nang* or the ablative =*dering* can always be added. This shows that the number of arguments any verb can take is always between 0 and at least 3.<sup>4</sup> The distinction between intransitive, transitive and ditransitive is thus not an important one in Sri Lanka Malay, and in fact, the presentation above structured into 0-place up to 4-place arguments is more an expository device than a solid fact about the language. The assignment of verbs into any of those four categories was based on their translational equivalents in English; I am not aware of any morphosyntactic test in SLM to arrive at a similar qualification.

The quantity of arguments is thus not what speakers have to store in the lexicon, since this is basically only constrained by semantic plausibility, which is found in knowledge of the world, or the ‘encyclopedia’. What must be stored in the lexicon, however, is the quality of arguments in some cases. Very often, the arguments simply match what semantics would predict (dative for recipient, ablative for instrument), but especially for the =yang-cases, there is a certain degree of arbitrariness. For instance, the intransitive verb *thinggalam* ‘sink’ takes =yang-marking, while the semantically very similar verb *jaatho* ‘fall’ does not and governs a  $\emptyset$ -marked NP, in other words a nominative. Furthermore, the transitive verb *giigith* ‘bite’ takes a =yang-marked NP, while the semantically similar *thiikam* ‘stab’ takes an NP marked for dative by =*nang*.

But does this knowledge need to be stored in the lexicon? Above, I have argued that the special dative marking for the three transitive verbs *theembak* ‘shoot’, *puukul* ‘hit’, and *thiikam* ‘stab’ is in fact not a case of lexical irregularity, but instead a case of different conceptualization of the world. As far as the worldview of the SLM speakers goes, and different from Middle European views, these actions do not take a patient, but a recipient of the damage inflicted. This is thus knowledge about the world, and not linguistic knowledge. The mapping of this world knowledge upon linguistic structures is then straightforward.

<sup>3</sup>Sinhala aligns with SLM in this respect. See Garusinghe (1962, 36,92) and (Gair, 1991, 18f). The behaviour of Tamil dialects in Sri Lanka is understudied and cannot be commented on here.

<sup>4</sup>We exclude the meteorological verb *uujiang* ‘rain’ for the moment, where one could think of a context with an instrument and a beneficiary, but will need a lot of time of convincing speakers of the acceptability of this context.

What SLM speakers need to know in order to produce the right case frames in their language is thus the knowledge of which actions require which semantic role, encyclopedic knowledge. The mapping of these semantic roles upon morphosyntax is then trivial and does not need recourse to the lexicon.

A first presentation of the content of this paper was given in Leipzig in August 2010 (Nordhoff, 2010). It was entitled ‘No valency classes in SLM’. The reason for this was that there is basically no argument/adjunct-distinction. Furthermore, everything can be done with encyclopedic knowledge in this language as far as case frames are concerned. Since then, I have realized that SLM verbs can be partitioned in meaningful classes on the basis of the cases which occur with these verbs. The repartition of verbs into these classes can lead to interesting questions, e.g. why are *banthu* ‘help’ and *puukul* ‘hit’ in the same [NOM DAT]-class? So, SLM verbs do fall into classes, but these classes are direct reflexes of semantics, and there is no need to explain a grammatical phenomenon in terms of morphosyntax or lexical subcategorization if a semantic explanation is readily available. For these reasons I conclude that the grammatical category of syntactic valency is meaningless in the context of SLM studies, but that semantic valency can fruitfully be investigated and lead to interesting results.

## References

- ADELAAR, K. A. (1991). “Some notes on the origin of Sri Lankan Malay”. In Steinhauer (1991), pp. 23–37.
- GAIR, J. W. (1991). “Subjects, Case and INFL in Sinhala”. In VERMA, M. & K. P. MOHANAN (eds.) “Experiencer Subjects in South Asian Languages”, Palo Alto: Center for the Study of Language and Information. pp. 13–41. Reprinted in Gair (1998), pp. 65–86.
- GAIR, J. W. (2003). “Sinhala”. In CARDONA, G. & D. JAIN (eds.) “The Indo-Aryan Languages”, London: Routledge. pp. 766–818.
- GARUSINGHE, D. (1962). *Sinhala, the Spoken Idiom*. München: Max Hueber.
- HENGVELD, K. (fc). “Transparency in Functional Discourse Grammar”. *Linguistics in Amsterdam*, 99(99).
- LEHMANN, T. (1989). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.
- NORDHOFF, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. Ph.D. thesis, University of Amsterdam.
- NORDHOFF, S. (2010). “No Valency Classes in Sri Lanka Malay”. Paper presented at the workshop on valency classes, Max Planck Institute for Evolutionary Anthropology, Leipzig, August 2010.
- NORDHOFF, S. (2011). “Having come to be a copula in Sri Lanka Malay – an unusual grammaticalization path”. *Folia Linguistica*, 45(1):103–126.
- NORDHOFF, S. (fc). “Transparency in Sri Lanka Malay”. *Linguistics in Amsterdam*, 99(99).
- NORDHOFF, S. (forthcominga). “Jack-of-all-trades: The Sri Lankan Malay property word”. In VAN LIER, E. & J. RIJKHOFF (eds.) “Flexible word classes: a typological study of underspecified parts-of-speech”, Oxford: OUP. Paper presented at the Workshop on flexible languages, Amsterdam.
- NORDHOFF, S. (forthcomingb). “Multi-verb constructions in Sri Lanka Malay”. *Journal of Pidgin and Creole Languages*.
- PAAUW, S. H. (2004). *A Historical Analysis of the Lexical Sources of Sri Lanka Malay*. Master’s thesis, York University.
- STEINHAUER, H. (ed.) (1991). *Papers in Austronesian linguistics, Pacific Linguistics*, vol. 1. Canberra: The Australian National University.