A grammar of
Upcountry Sri Lanka Malay

A grammar of
Upcountry Sri Lanka Malay

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 27 november 2009, te 14:00 uur

door

Sebastian Nordhoff

geboren te Erlangen, Bondsrepubliek Duitsland

Promotiecommissie

Promotor: Prof. Dr. P. C. Hengeveld

Co-promotor: Dr. U. Ansaldo

Overige Leden: Dr. K. A. Adelaar
                   Prof. Dr. J. B. den Besten
                   Prof. Dr. J. W. Gair
                   Prof. Dr. N. Himmelmann
                   Dr. N. S. H. Smith

Faculteit der Geesteswetenschappen

Contents

8   The noun phrase   415

9   The postpositional phrase   459

List of Tables

List of Figures

Abbreviations

ACMA	All Ceylon Malay Association
AGM	Annual General Meeting
ASP	assistant superintendant of police
CLI	Ceylon Light Infantry
COSLAM	Conference of Sri Lankan Malays
CRR	Ceylon Rifle Regiment
GSM	mobile phone standard
HNB	Hatton National Bank
HSBC	Hongkong and Shanghai Banking Corporation
JVP	Janatha Vimukthi Peramuna, People's Liberation Front. Part of an insurgency in the 1970s and 1980s. Now established party with seats in parliament
KMA	Kandy Malay Association
KORAMEL	Konfederasi Rakyat Melayu Langkapuri (Sri Lankan Malay Confederation)
LTTE	Liberation Tigers of Tamil Eelam, Tamil separatist movement
MP	member of parliament
PMD	Pidgin-Malay derived languages
SEA	South East Asia
SEAA	South East Asian Archipelago
SLAMAC	Sri Lankan Malay Confederation
SLM	Sri Lanka Malay
SLMC	Sri Lanka Muslim Congress
SLMT	Sri Lanka Muslim Tamil
SLP(C)	Sri Lanka Portuguese (Creole)
SSC	Secondary School Certificate
Std.	Standard
TM	Trade Malay
VOC	Vereenigde Oostindische Compagnie, Dutch East India Company
VM	Vehicular Malay
WW	world war

Table 1: Abbreviations of cultural terms

ABL	ablative		NEG	negation
ACC	accusative		NMLZR	nominalizer
ADDIT	additive		NP	noun phrase
ADHORT	adhortative		NUM	numeral
ADJ	adjective		ORD	ordinalizer
ADJCT	adjunct		PAT	patient
ADV	adverb		PL	plural
ANIM	animate		POSS	possessive, possessor
ANP	ascriptive nominal predicate		POSTMOD	postmodification
BEN	beneficiary		POSTP	postposition
CAUS	causative		PP	postpositional phrase
CLS	clause		PRED	predicate
CLT	clitic		PREMOD	premodification
COMIT	comitative		PRON	pronoun
CP	conjunctive participle		PROX	proximal
DAT	dative		PURP	purpose, purposive
DEIC	deictic		QUANT	quantifier
DEM	demonstrative		QUASIPREF	quasi-prefix
DERIV	derivational		QUOT	quotative
DIST	distal		REC	recipient
EMPH	emphatic		RELC	relative clause
EVID	evidential		RELN	relator noun
EXP	experiencer		S	singular
EXPR	expression		S	sentence
FOC	focus		SIMIL	similative
FUT	future		SOV	Subject-Object-Verb
IMP	imperative		SRC	source
INANIM	inanimate		STD	standard
INDEF	indefinite		SVC	serial verb construction
INF	infinitive		SVO	Subject-Verb-Object
INFL	inflection		TAM	tense/aspect/mood
INSTR	instrumental		UTT	utterance
INTERJ	interjection		V	verb
INTERR	interrogative		VOT	voice onset time
INVOL	involitive		VPRED	verbal predicate
IRR	irrealis		VPREF	verbal prefix
LOC	locative		VP	verb phrase
MOD	modal		WH	question word
MVC	multi-verb-construction		YN	Yes-No-question
N	noun		NONV	non-verbal
PHAT	phatic			

Table 2: Abbreviations of linguistic terms

Acknowledgments

xxiii

Part I

Background

Chapter 1

Introduction

1.1 Overview

This book is a description of Sri Lanka Malay. As its name indicates, the language is spoken in Sri Lanka. Contrary to what the name indicates, it is not a dialect of Standard Malay but a language in its own right (Adelaar 1991). It differs greatly from Standard Malay in many aspects, for example in its phonological inventory, basic constituent order, tense system, and case system, to the extent that Standard Malay and Sri Lanka Malay are mutually unintelligible.

While the lexical material is overwhelmingly of Malay origin, the grammatical structure is heavily influenced by the contact languages Sinhala (Indo-Aryan) and Tamil (Dravidian). Sri Lanka Malay can thus be said to have three 'parents' contrary to many other languages, which are clearly traceable to one parent language only.

The speakers of Sri Lanka Malay are the descendants of immigrants from what are today Indonesia and Malaysia, mainly brought by the colonial powers of the Dutch and the British between 1650 and 1850 as soldiers, exiles, convicts, and slaves. The soldiers were the most important contingent and have shaped the Sri Lankan Malay society, whose members mainly took up profession in the military, the police and related professions until very recently. During the 19th century, the Ceylon Rifle Regiment consisted exclusively of Malays, and most of the island's Malays were associated in one way or another with this regiment.

During the colonial period, the Sri Lankan Malays had close ties with the colonial powers. After the disbandment of the regiment in 1873 and the independence of Ceylon in 1948, the Malays lost the close ties with the ruling elite and saw their social position deteriorate. Nationalist policies in the aftermath of the independence led to the decline of minority languages. According to the 2001 census, there are 48.000 ethnic Malays in Sri Lanka (Smith & Paauw 2006), not all of which speak the language. Sinhala and English are replacing Sri Lanka Malay as the language of child care. The main concentrations of Malays are found in Colombo, the Upcountry towns,

and the Southern settlements of Hambantota and Kirinda. The variety described in this grammar is the Upcountry variety. There seem to be some phonological differences between this variety and the South. As for the Malay spoken in Colombo, it seems very similar to the variety described in this grammar, but I do not have enough information to establish that they can indeed be regarded as the same.[1]

Sri Lanka Malay is characterized by the left-branching word order typical of South Asia, with the arguments preceding the verb, all modifiers normally preceding the noun, and postpositions. Compared to Malay varieties elsewhere, Sri Lanka Malay has a more elaborate morphology: there are more than a dozen verbal prefixes. Closely related to this is the presence of a large number of enclitics, including more than ten case marking postpositions. Sri Lanka Malay differs from European languages in that a number of notions common in the latter are absent in Sri Lanka Malay. These absent categories include 'stress', 'subject', and 'conjunction'.

1.2   Organization of this book

This description is divided in four parts. The first part covers sociohistorical background and methodological assumptions. The second part discusses phonology. The third part takes a form-to-function approach and describes grammatical morphemes and constructions found in Sri Lanka Malay, and what functions they serve. The fourth part takes a function-to-form approach and describes how Sri Lanka Malay can communicate content, e.g. referring to time, space, entities etc., and what morphemes and constructions are used for that. The two latter parts heavily cross-reference each other.[2] The appendix contains five texts by different speakers/writers, next to some other additional material.

1.3   History of research

Little research had been undertaken on the Malay population in Sri Lanka until the publication of Bichsel-Stettler (1989) and Hussainmiya (1987, 1990). Hussainmiya (1990:7,22) informs us that Goonetileke's bibliography on Ceylon only cites nine articles whatsoever that treat Malay issues. These are mostly short (4-5 pages) and not scientific. References to Malays in general Sri Lankan history books are limited to a few sentences (Bichsel-Stettler 1989).

Bichsel-Stettler (1989) and Tapovanaye (1986, 1995) treated phonology, while diachronic issues became more important in the 1990s (Adelaar 1991, Bakker 1995, 1996, 2000a,b, 2006). Since the turn of the millennium, Sri Lanka Malay has received considerable scholarly attention by Umberto Ansaldo, Peter Bakker, BA Hussainmiya,

---

[1]Saldin (2001:34) sees all three as different dialects.

[2]A theoretical justification for this approach can be found in Nordhoff (2008).

Scott Paauw, Shihan De Silva Jayasuriya, Peter Slomanson, and Ian research has mainly focused on morphosyntax and diachrony. See the 'References' section in the appendix.

1.4 Terminology

There are a number of terminological pitfalls to deal with in the description of Sri Lanka Malay. The smallest one is whether to use Sri Lanka or Sri Lankan Malay with or without 'n'. In this grammar, I use 'Sri Lanka Malay' without 'n', but I have no strong preference. The 'Malay' part is more problematic, since the immigrants belonged to a large variety of the ethnic groups of what today is Indonesia, and only a minority were Malays. The Dutch actually referred to the immigrants as *Oosterlingen* 'Easterners' or *Javanen* 'Javanese', the latter because the port of departure was Batavia in Java. This reference to Java is still found today in the other Sri Lankan languages, which use *Jā Minissu* (Sinhala), *Cāvakar* (Tamil) or *Jāvā Manusar* (Moorish Tamil) (Hussainmiya 1990). The British however were not aware of the ethnic origin of the troops they took over from the Dutch, and named them after the language they spoke, which was based on Malay. This usage is still the norm today, and has actually made its way back into the Malay community at least in the Upcountry, where *mlaayu* 'Malay' is the normal way to designate the Sri Lanka Malays. Ansaldo (2005b) reports that in the South, Sri Lanka Malays use the term *Orang Java.* He uses *Kirinda Java* to refer to the language. The term *Java* is not found in the Upcountry, so the adoption of this practice is not possible. The academic tradition has been Sri Lanka(n) Malay (SLM), and I will stick to that tradition.

The island where the language is spoken has had several names in the course history, of which 'Ceylon' and 'Sri Lanka' are the most widely known. The Sri Lankan Malays refer to the island as *Seelon*. 'Ceylon' and 'Sri Lanka' are in complementary distribution as far as historical periods are concerned. Until 1972, the island was known to the Europeans as Ceylon, or related names in other European languages. With the creation of the republic in 1972, the term 'Sri Lanka' was coined. While 'Lanka' and related forms are local designations of the island, the 'Sri' is an addition of the late 20th century.[3] The use of 'Ceylon' for the republic is obviously anachronistic, but the use of Sri Lanka for the historical period is anachronistic as well. I will therefore use 'Ceylon' for the historical period,[4] and 'Sri Lanka' for the modern period, and for discussions which do not refer to any period in particular.

It is often necessary to refer to the places of origin of the Malay speakers. As of today, two main countries occupy the relevant area, Malaysia and Indonesia, but these

---

[3]Furthermore, the addition of 'Sri' is occasionally seen as an expression of Sinhalese nationalism. This is due to a law from 1958 requiring to include the Sinhalese character ශ්‍රී 'śrī' on number plates of cars. Tamils opposed this law and painted over the ශ්‍රී. The issue escalated and led to riots with 400 casualties and 12,000 homeless people, mostly among Tamils (Nissan & Stirrat 1990:35f).

[4]One could argue that 'Ceylon' is colonialist, but since the Sri Lankan Malays use a related term, *Seelon*, this argument does not hold in their context.

countries did not exist during the relevant period. It is often not possible or desirable to make a difference between what today is administrated by Malaysia and what is administrated by Indonesia. I will use the term "South East Asian Archipelago" (SEAA) to cover to the whole geographical region, without implying any national borders which may or may not have been relevant at a certain point of history. I occasionally also use the term "Nusantara" to refer to that region.

Chapter 2

History of the Sri Lankan Malays

Sri Lanka Malay is a contact language. The socio-cultural setting is very important for analyses of language contact (Weinreich 1953, Thomason & Kaufman 1988). This chapter shall present the different periods of contact between Malays and Lankans and analyze the socio-historical setting in each period.

The first significant number of Malays arrived in Ceylon in the 17th century. However, the contact between the islands of Ceylon and the South East Asian Archipelago (SEAA) predates this by many centuries. We can divide the history of contact between the Malays and the Lankans drawing on the following major historical events:

- The first Malays brought to Ceylon by the Dutch (~1640)

- The ousting of the Dutch by the British and the creation of a Malay regiment (~1796)

- The disbandment of said regiment (1873)

- independence of Ceylon from the United Kingdom (1948)

Five periods can thus be distinguished: a pre-Dutch period (-1640, Section 2.1 p.8), the Dutch period (1640-1796, Section 2.2 p.9), the period of the Regiment under the British (1796-1873, Section 2.3 p.15)[1], the post-regimental history and the post-colonial history (1873-today, Section 2.4 p.22).

I will briefly sketch major events during each period before discussing demographic and sociological developments that accompanied them. Domains which encompass more than one period are the cultural history (Section 2.5 p.33), the Malay settlements in Sri Lanka 2.6 (p.35), and the relation to the other Islamic group, the

---

[1]Technically speaking, the regiment was founded some years later than 1796.

'Moors' (Section 2.7 p.40), The development of the language will be discussed in Section 2.8, p.47.

2.1 Before the arrival of the Dutch

2.1.1 Political history

The first known people to settle in Ceylon were the Veddah, who live today only in remote pockets of jungle in the central highlands. It can surely be affirmed that the Veddah have not had any linguistic relation on the present day SLM because they were already pushed back into the jungles by the time the Malays entered Lanka.

The Veddah could not withstand the arrival of the populations from the Indian mainland about 500 BC (Gair & Paolillo 1997:1) that announces a second phase of settlement, resulting in the present cultures of Sinhalese and Ceylon Tamils.[2]

From the 7th century AD onwards, Arab traders settled down in India (Codrington 1926:51). The Arab traders in India became acquainted with local tongues. Some of their Tamil-speaking descendants moved to Ceylon from the seventh (Nuhman 2007:32), eighth (Hussein 2007:21) or tenth (De Silva 1981:72) century onwards. They are the ancestors of most of today's Tamil speaking Muslims in Lanka.

While we have evidence of cultural Lankan influence in the SEAA from the first millennium AD (Dupont 1959, De Casparis 1961), the earliest source of Malay culture in Sri Ceylon dates from a later period. It is found in old Sinhala literature from the periods of Polonnaruwa and the Dambadeniya (together 1098-1293) which contains words of Malay etymology. More recent literature from the 15th century states that the SEAA was renowned for spices, perfumes and other commodities (De Silva Jayasuriya 2002:11). While this testifies to some sort of contact between Ceylon and the SEAA, the impact seems to have been rather minor and restricted to some loanwords (Gunasekara 1891, Sannasgala 1976). Cultural elements that would point to a more intimate contact could be outrigger canoes (Tennent 1859).

The peaceful exchange of goods and religion found an end in 1247, when Malay forces from the peninsula made an attempt to invade Sri Lanka (De Silva 1981:67). Chandrabhānu of Tāmbralinga attacked the Polonnaruwa kingdom in Central Sri Lanka. He was defeated and withdrew to the Jaffna peninsula, where he succeeded in securing the throne for himself. Later, he tried a second invasion of Polonnaruwa, which failed and cost him his life, ending the short Malay intermezzo on the island. Today's place names *Caavakacceeri* (Javaka-settlement) and Caavakotte (Javaka Fort) witness this invasion.[3] Even if this invasion had a period of Malay rulership as a consequence, there is no link to the present day Malay communities (cf. Adelaar &

---

[2]While for the Indo-Aryans, who were to become the Sinhalese later, 500 BC seems a reasonable estimation, for the Dravidians "there is not firm evidence as to when [they] first came to the island, but come they did from very early times." (De Silva 1981:12).

[3]*Java* is taken as a generic term for the SEAA in Ceylon (Hussainmiya 1987:37).

Prentice 1996:685). An invasion that should have a much bigger impact on Lankan society was threatening from another continent: the Portuguese.

Driven by the desire for the lucrative spice trade, the Portuguese sent out expeditions to secure a base in the SEAA. In 1511 the Portuguese gained hold of Malacca on the Malaysian peninsula. By that time, Islam had made considerable inroads in SEAA and largely replaced Buddhism as the predominant religion. On their way from the Cape to the SEAA, Ceylon was a good base for the Portuguese, and in 1518 they established a fort in Colombo. Within the following decades, they gained control of the coast but could not penetrate the central highlands, where the Kandyan king would withstand European conquest until the 19th century.

A major interest of the Portuguese was to spread Christian faith. They kept other religions in very low regard, especially Islam, and brought about the expulsion of Muslims from the Kōṭṭe kingdom (De Silva 1981:102). Hussainmiya (1990:34f) therefore questions whether the Portuguese would have brought (predominantly Muslim) soldiers from the SEAA to fight for them. Bichsel-Stettler (1989:6f.) agrees that there is no historical evidence for a sizeable Malay community in Sri Lanka and doubts that the Portuguese would have tolerated the spread of Malays in Sri Lanka. However, Powell (1973:26) mentions that some *Javas* and *Kaffirs* (Africans) fought for the Sinhalese king Raju in 1587 and states that they had deserted from the Portuguese, indicating that some Malays were already found in the Portuguese forces.

2.1.2 Demographic history

The study of Malay demography before the arrival of the Dutch is hampered by the paucity of subjects to study. There was no important settlement of Malays in Ceylon before the arrival of the Dutch. Chandrabhānu's invasion did not last long enough to give birth to a Malay community in Ceylon. If any Malays were left behind after the invasion, they would have merged in their coreligionist community of Sinhalese Buddhists. The Malays brought by the Portuguese do not seem to have been very numerous, and we do not have any information whatsoever about where they had come from or what languages they might have spoken. These Malays are, however, the first to desert to the Kandyan King and may have had a certain influence on the Malay community in Kandy.

2.2 Dutch rule

The Dutch had gained considerable power and prosperity in the 17th century and projected their power to the Indian Ocean. By 1658, they had managed to oust the Portuguese from Ceylon with the help from the Kandyan King. The latter, however, had to find out that his "allies" had come to stay and that he had bartered Portuguese occupation of the coastline for Dutch occupation of the coastline (De Silva 1981:133).

2.2.1   Political history

Like the Portuguese before them, the Dutch had taken hold in the SEAA before trying to get hold of Ceylon. In their possession in the SEAA, they had met peoples with excellent soldiers, who the Dutch recruited into their service. These troops were already fighting on the Dutch side during the war with the Portuguese (Hussainmiya 1990:44). Later, they formed an important part of the Dutch military presence in Ceylon. There were garrisons in the main coastal towns of Colombo, Galle, Chilaw and Trincomalee.

The Eastern troops had several advantages over European troops: First, they were used to tropical climate and fights in the jungle (Hussainmiya 1990:46) and, second, they were not likely to side with the Sinhalese or Tamils in any future rebellion given the cultural and religious differences. Hussainmiya (1987:55) reports that the Dutch referred to these troops as *Oosterlingen* 'Easterlings' or *Javaans* 'Javanese', which is a much better name than 'Malay', which the British authorities would later choose as a generic term. Actually, the majority of the immigrants were *not* Malay, and were definitely not from the Malaysian peninsula. In the Ceylonese context, however, the word 'Malay' is reasonably unambiguous (it can only refer to the group under discussion), and it is common scientific practice to refer to this immigrant group as 'Malay' for want of a better word, which is also what I will do in the remainder of this book.

A regular supply of new recruits from Batavia was provided each year, as well as reinforcements in urgent cases (Hussainmiya 1987:48). The importance of the Malays for the Dutch army can be seen from the fact that "[a]bout two thirds of the fighting troops were Malays" in 1764 (Powell 1973:37), a year in which the absolute number of Malay soldiers also reached the record of 2500 (Hussainmiya 1990:49). This coincided with the creation of the Free Javanese Company (Hussainmiya 1990:52). From 1785 onwards, local, i.e. Ceylonese, civilian Malays were recruited into the Dutch army to replace the disbanded regiment of Free Moors (Hussainmiya 1987:54).

Beside the soldiers, who represented the numerically most important group, some other groups from the SEAA were also brought to Ceylon: exiles from rebellious provinces, convicts, and slaves.

In order to stay in control of the islands of the SEAA, the Dutch deported rebellious local regents to South Africa and Ceylon. Ceylon was preferred because it was closer to the SEAA; very unaccommodating princes would be exiled farther to Africa. The Dutch having first taken possession of the former Portuguese colonies in the Moluccas and the lesser Sunda islands, the first exiles came from there. In the late 17th century, a war of succession in Java had the Dutch deport some Javanese princes to Ceylon (Hussainmiya 1990:38).

Convicts from the SEAA were brought to Ceylon, "serving sentences according to the gravity of their crime ranging from being kept in chains to the permission to engage in handicraft" (Hussainmiya 1990:44). In 1782, the Dutch allowed minor offenders to join the ranks of the military. Other convicts were set free and were allowed to settle in Ceylon. The former convicts thus formed part of the early Malay population in

Ceylon and the 'Malay' community of Sri Lanka can be said to owe its origins in part to these Malay convict settlers.

A third group were slaves and servants brought mainly from the Moluccas and Lesser Sunda islands to serve the aristocratic exiles and the Dutch. Like the convicts, some of them joined the army (Hussainmiya 1990:48). They were not a very numerous group and did not constitute a distinguishable separate community.

Particularly after the middle of the 18th century, the proportion of free Malays grew. From 1769 onwards, the free Malays had their own category for tax collecting purposes (Bichsel-Stettler 1989:9), which indicates that their number could not be neglected. This comprised discharged soldiers and descendants of political exiles. Their occupations included gardening, rattan weaving and petty trade. Some were also employed as domestic servants. These Malays counted as foreigners in Ceylon and had to perform *Uliyam* (soccage),[4] a duty which the political exiles and soldiers were exempt from (Hussainmiya 1990:48f).

2.2.2 Demographic history

The origin of the four different groups of immigrants (Exiles, convicts, slaves, and soldiers) will be discussed in this section.

The first exiles came from the Moluccas and the Lesser Sunda Islands. Later, there were also exiled Javanese princes (Hussainmiya 1990:38f). Hussainmiya further cites Bacan, Tidore, Timor, Madura and Sumatra. The Javanese succession wars in the late 17th century brought more Javanese exiles to Ceylon. Most of the exiles had their families with them at deportation time or they joined them later. Hussainmiya (1990:39) gives 200 as a likely number for the exiles at the end of the 18th century, about 10% of the Malay population at the time. Most of the exiles lived in the 4 main coastal towns Colombo, Galle, Trincomalee and Jaffna.

While we have a certain knowledge of how many exiles came from which region, the case is more difficult for the convicts and the slaves (Hussainmiya 1987:47). The number of slaves is uncertain but small, and they came from the Moluccas and the Lesser Sunda islands (Hussainmiya 1990:48); the provenance of the convicts is unknown. In 1751 there were 131 convicts serving their sentence; the number of convicts having joined the army or regained freedom is unclear.

The most important contingent of Malays brought by the Dutch were, however, the soldiers. They were mainly recruited in Batavia (Jakarta). At that time, Batavia was a port city with many different peoples mingling. There are records of the following ethnic groups serving in the military: Ambonese, Bandanese, Balinese, Buginese, Javanese, Madurese, Malays (Bichsel-Stettler 1989, Raben 2000). In Ceylon, they were referred to as *Javaans*, but we see now that this is dramatically simplified given the ethnic diversity of the newcomers. These "Neo-Indonesians", as Hussainmiya (1987:57) calls them, formed the base for the Malay society in Ceylon. Already in

---

[4]For more on soccage, see Ludowyk (1966:31), Codrington (1926:149f), De Silva (1981:164).

Batavia, they are likely to have developed a group identity while waiting to be sent to the battlefield (Vlekke 1943:174). This identity and cohesion would have been further strengthened in the unknown foreign country.

In Batavia, there were different *kampungs* for the different ethnic groups (de Haan 1935:451ff). But this does not mean that there was no contact between the different groups. The *kampungs* were not autarkic,[5] and one had to interact with other communities and make oneself understood. To this end, inhabitants employed a language of interethnic communication, a Malay-based lingua franca in their case (de Haan (1935:470), Raben (2000:102)). All over the SEAA, the use of some lingua franca is common, which varies from region to region (Adelaar & Prentice 1996). The nature of this language will be discussed in more detail in Section 2.8.1.

Most of the nationalities present in Batavia were Muslims, but not all of them, as Hussainmiya (1987:58) states:

> In Dutch times, not all the easterners who came to Sri Lanka were the followers of Islam. It is particularly difficult to establish the religious background of the Amboinese [sic], Balinese, and even Javanese, because among the first group, there were a considerable number of Christians, while most of the Balinese belonged to Hindu Religion [. . .] Nevertheless, it is quite clear that a large majority of the early Malays were followers of Islam. (Hussainmiya 1987:58)

The number of Malays serving varied in function of the military needs of the Dutch. Under normal conditions, there were 800 men in service, divided into 10 companies. But sometimes the Dutch need for soldiers was so dramatic that the *kampungs* were virtually depopulated (Hussainmiya 1987:56). In 1764 the number of Malay soldiers rose from 800 to 2500 in anticipation of a war with the Kandyan king. In 1795/6 there were probably around 1400 Malays serving, comprising recruited civilians. For 1800, Hussainmiya (1990:49f) estimates the number of Malays in Ceylon at around 2400, possibly their highest figure (Hussainmiya 1987:54) because the Dutch were anticipating a British attack and had stocked up.

To sum up, during the Dutch period, there were two main groups. First there was the high prestige group of the exiles from the Moluccas and the Lesser Sunda Islands. This group made up about 10% of the Malay population. Second, there were the soldiers, who came from diverse ethnic backgrounds, but who most likely communicated in some kind of vehicular Malay already developed back in Batavia. These soldiers formed the most important group, and they are likely to have absorbed the smaller groups of the slaves and convicts, who could not affirm their own identity given their low prestige and often joined the ranks of the soldiers where they had to adapt to the language already in use.

---

[5]Raben (2000:97) criticizes the uncritical interpretation of Dutch rules regarding segregation by de Haan (1935) and argues that "a yawning gulf existed between law and implementation." Also see Abeyasekere (1987:31).

During Dutch rule, a shift in the composition of the Malay population in Ceylon took place. While there were many exiles from the Moluccas and Lesser Sundas at first, they were quickly outnumbered by exiles from the Javanese succession war and the Batavian soldiers. It seems that the Eastern community was absorbed by the Javanese community at the close of the 18th century. Immigrants of the late 18th century would find an established society of "Ceylon Malays" with strong roots in Batavia, into which they would quickly integrate (Hussainmiya 1987:55-57).

2.2.3 Sociological history

Hussainmiya (1987, 1990) provides some sociological information for the Malays under Dutch rule, but at times this information is contradictory. These contradictions concern the degree of isolation of the exiles, the extent of contact with the local population, and the possibility to practice their faith.

The two main groups of the exiles and the soldiers formed the bulk of the Malay population. Slaves and convicts were much less important in number and had regular connection with the soldiers, whose ranks they could join to gain their freedom (Hussainmiya 1990:48). It appears that the exiles formed an elite of spiritual leadership, while the soldiers and the associated groups contributing most of the souls.

The exiles were not allowed to communicate with each other, but in how far this was enforced remains questionable given that marriages between children of exiles were contracted and attendance at celebrations such as births was also high (Hussainmiya 1990:43f). Some exiles were given ranks in the military in order to raise the moral of the Malay troops (Hussainmiya 1990:50), linking the exiles with the soldier population, and probably also communicating with each other.

While the exiles and the soldiers would see their stay in Ceylon as temporary, the free Malays had no plans to return to the SEAA. They are likely to have invested more in building up an independent existence in Ceylon. Free Malays engaged in commercial activities such as gardening, rattan weaving and petty trade (Hussainmiya 1987:53). This brought them into contact with the other Malay groups, but also with the other ethnic groups in Ceylon.

As far as the contact between the Malays and Lankan peoples is concerned, the situation varies. As for the exiles, normally, only the head of a family was considered such; the other members of the family could move about freely. Given the low number of exiles and the interdiction to communicate with other exiles, contact with locals seems probable. On the other hand, the exiles had no plans of staying in Ceylon for a long time, which makes them less likely to intermingle with the local population (cf. Muysken 2001:160). Contact could also be inhibited by the Malay's residence inside the Dutch forts or in houses guarded by sentries (Hussainmiya 1987:43). As for the soldiers, they fought for the Dutch against the local population. This makes friendly contact unlikely. Different religious beliefs did not help the issue. Hussainmiya (cf. 1987:51) adduces that the Sinhalese even feared the Malays, and it seems reasonable to assume that the same should be true for the Hindu Tamils. The Malays seem to

have participated in the local cultural life, though; cockfighting was very popular, for instance (Pieris 1918:131).

The business within a Malay regiment could be carried out in Malay or in Dutch. Provisions and housing were also provided by the Army. Existing sources suggest that the need for interaction with the local population only arose in order to find a suitable wife.[6] Even this would not be necessary if the soldier had brought a wife from the SEAA, which was common (cf. Hussainmiya 1987:50).

Sources on religious life are contradictory. On the one hand, the Dutch government in Ceylon had banned public Islamic religious ceremonies, which was why they were often held in secrecy (Hussainmiya 1987:46). On the other hand, we have sources stating that from 1783 onwards, Malay mosques existed (Bichsel-Stettler 1989:9). Bichsel-Stettler (1989:11) affirms that the Malays were a clearly different community with distinct ethnic identity and own places of worship.

2.2.4   Early Intermarriage

It seems that already in the early ages of Malay presence in Ceylon, the Malays intermarried with the local population. Christopher Schweitzer, an employee of the VOC notes:

> The wives, who in part are Ambonese, in part Sinhalese and Malabarian may not say anything [against the stripping of their ornaments]. (Schweitzer 1931[1680]:106), [my translation][7]

and also shortly after the British conquest, the commandant of Kalutara would write:

> As far as I can learn there is little damage to be apprehended from the Malays in this district who by intermarriage with the natives have become in some degree naturalized and are indeed much quieter, and more industrious than the Sinhalese or Moors. (Hussainmiya 1990:61)

Still, it should be noted that during the Dutch period "the Eastern soldiers, when they embarked for Ceylon, brought their womenfolk along with them. It is not known in what proportion such women from the East-Indies came to the island during the Dutch period. Later the British authorities actively encouraged the foreign Malay recruits to bring along their families." (Hussainmiya 1990:47).

A critical assessment of the extent of and evidence for intermarriage over the whole period of SLM history can be found below in Section 2.7.

---

[6]However, official sources tend to exclude unofficial activities such as smuggling (Pieris 1918:68), gambling (Pieris (1918:131), Schweitzer (1931[1680]:106)), and prostitution, which might have brought the groups into contact without this being reflected in official documents.

[7][...] die Weiber, welche theils Ambonesin, Singulesin und Malabarin, dörffen nichts darzusagen.

2.3 British rule

2.3.1 Political history

The British were the third European power to rule Ceylon. They took over from the Dutch in 1796. The Malay soldiers defended fiercely and are even said to have been the only worthwhile enemy of the British (cf. Powell 1973:42,44). This earned them the respect of the British. Many of the Malay prisoners of war taken by the British returned to Ceylon shortly after that, serving in the British native army.

2.3.1.1 The Kandyan Wars

The first British governor of Ceylon, Governor North was sympathetic to the Malays and valued their military skills highly. He created a regiment of Malays, which changed names several times during its existence: Malay Corps, His Majesty's Malay Regiment, the First Regiment, and finally the Ceylon Rifle Regiment. At that time, the British regarded the Malays as an exclusive soldier class. Even civilian Malays were registered in order to enlist them later (Hussainmiya 1990:61). The Regiment became the central institution of Malay life. Younger boys could serve as half-pay boys, a name which indicated that they earned exactly half the wage of an older soldier (Hussainmiya 1990:97), and an invalids' company took care of the ones that had been less fortunate in serving His Majesty. Hussainmiya (1990:73) indicates that about 75% of the Malay population served in the Regiment during that period. From 1802 onwards, retired Malays were sent to the South and given land to colonize: first Hambantota, later Kirinda and Palatupana (Hussainmiya 1990:63).

The British kept on depending on Malays for their supply of native soldiers, because the Sinhalese and Tamils proved to be poor soldiers (Hussainmiya 1990:24). Most of the able-bodied Malays were recruited into the army, and, as a consequence, there were not many civilian Malays left (Hussainmiya 1990:73). But their number could not meet the needs of the British. Recruitment from abroad was encouraged and companies were formed in Hambantota and Tangalle (Hussainmiya 1990:64f). In 1800, recruitment from Penang in Malaysia was started. This was the first time that Malays from the Peninsula came to Ceylon. This recruitment policy gave new life to the society of Malays remaining in Ceylon (Hussainmiya 1990:73), with a total of 172 new recruits in 1801 and 1803.

The exiled princes were not recruited but given an allowance because North estimated them important for the social cohesion of the Malay population for the spiritual and cultural leadership they provided (Hussainmiya 1990:79).

When the British attacked Kandy in 1803, many Malays deserted and joined the Kandyan king fraternizing with the Malays who had already deserted during Dutch times. Codrington (1926:166) gives the number of 450 deserters out of a total 700 Malay soldiers. Following this war, the Regiment only numbered 600 men (Hussainmiya 1990:72) compared to 1400 at the end of the Dutch reign. Governor North left

the island shortly after in 1805 and was replaced by Governor Maitland. The latter was aware of the Malay desertion during the Kandyan war and blamed the British defeat on them. He wanted to get rid of the Malay soldiers as soon as possible. His plans were to abolish the whole regiment, but he failed even in dissolving the invalid Malays' company, which he said to be expensive and unserviceable (Hussainmiya 1990:67,78). Not only the soldiers were affected by his resentment; he also sent back some political exiles, thus depriving the Malays of their cultural leaders (Hussainmiya 1990:79).

The next lasting governor, Robert Brownrigg,[8] held the Malays in high esteem. As a first action, he recruited the 150 Malays that Maitland had not cared to draft. Seeing that the numbers of Malays were running low as a consequence of his predecessor's policy, he started a recruiting mission in Java and recruited 412 soldiers in 1811, accompanied by 214 women and 208 children (see Table 2.1 on page 18). On the next military campaign against Kandy, he counted on his Malay Regiment to win the war, a hope which was fulfilled (Hussainmiya 1990:68,81f).

Under Governor Brownrigg, the Regiment further extended the domains in Malay life for which it was responsible. A Royal Military school was founded where the Malays were taught Literary Malay and religion (Hussainmiya 1990:96).

In 1816, the British protectorate over Indonesia stopped, and recruiting from Java was no longer possible. The recruiting would subsequently be done on the Malaysian peninsula.

2.3.1.2 Malays in Kandy

The central kingdom of Kandy had been a refuge for Malays deserting or escaping from the colonial powers ever since the Portuguese times. Malays fighting for the Sinhalese king Rahju against the Portuguese in 1587 are mentioned by Powell (1973:26). During Dutch and British rule, more Malays had escaped the colonial forces and fled to Kandy. The first Malay reported to have done so during Dutch times was a royal exile who slipped into Kandy in 1742 (Hussainmiya 1990:69). In 1765/66 the Dutch General Baron van Eck left behind a large number of Malay soldiers when retreating to Colombo (Powell 1973:37f). Generally, the deserted soldiers and runaway slaves were held in high esteem by the Kandyan king (Hussainmiya 1990:78). The British could not prevent desertions to Kandy either (Hussainmiya 1990:67).

The Kandy Malays remained true to their military inclination in Kandy as well and served as soldiers, guards, and even as bodyguards for the king. After the British-Kandyan wars, the Kandyan king Wickremasinghe was eager to assure their loyalty. But Wickremasinghe lost his temper at the end of his reign (Hussainmiya 1990:71) and the loyalty of the 200–300 Kandy Malays was split in the 1815 war with the British (Hussainmiya 1990:82f). The Kandyans lost the war, sealing the fate of the last Sinhalese kingdom in Lanka (De Silva 1981:220ff).

---

[8]Glossing over the brief intermezzo by John Wilson 1811-1812.

2.3.1.3   The Ceylon Rifle Regiment

After the complete conquest of Ceylon by the British in 1815, the Malay Regiment enjoyed a high status. In 1827, the Ceylon Rifle Regiment (CRR) was founded. Despite the change of the name, the conditions of service remained pretty much the same for the Malay soldiers. But soon after the formation of the CRR, there was a shortage in recruits from the local Malays (Hussainmiya 1990:83f). The Malays seem to have changed their attitude towards military career. An important aspect was the reduction of pay. Other reasons were the enlistment for life, which was not popular, and the service with Sepoys and Kaffirs, two groups which the Malays held in low esteem. Furthermore, there were now many possibilities of employment in other professions due to the growing economy, such as government, police and plantations. So the military no longer had a monopoly on the Malay work force. Failing to attract local Malays, the British again tried to recruit from abroad and opened a recruiting station in Penang in 1833 (see Table 2.1 on page 18). This station proved successful in the beginning, but this would change. The recruiting station shifted to Singapore in 1840, was stopped in 1842, revived in 1845 and finally abandoned in 1848 due to its meager success.

The Malays recruited during this period are said to be of "low quality" compared to the native Malays and the recruits from Java in 1813 and 1816, and rarely could secure higher ranks (Hussainmiya 1990:95). During the time of the CRR, Ceylon saw a small but steady influx from the Malay peninsula of about 50 recruits a year. This influx is thus an order of magnitude lower than the 412 soldiers recruited in 1811 alone.

But Ceylon was pacified and there was not much to do for the CRR. To make use of the men under arms, a Malay detachment was sent to serve in Hong Kong from 1847 to 1854. Service in Hong Kong was not popular and casualties were high (Hussainmiya 1990:99). As a consequence of this, the already low popularity of the CRR with the Malays fell further: "By the 1850s it became relatively difficult to find Malays willing to serve in the CRR" (Hussainmiya 1990:88).

In 1848, the Regiment was engaged against a rebellion. This was its last military action. The regiment continued to shrink. In 1849 the Regiment was short of 483 men, in 1851, 22 companies were reduced to 18 companies, and in 1854, to 14, of which 11 were Malay. In 1865, the detachments in Hambantota, Badulla, Kurunegala and Jaffna were withdrawn (Hussainmiya 1990:101). The police department took over some of the duties from the CRR and many Malays switched over. From 1870 onwards, the Malay element was dominant in the police (Hussainmiya 1990:102f.).

With no warfare at all in 25 years, it did not seem acceptable to finance the CRR any further and the Regiment was disbanded in 1873, leaving the remaining 700 Malays to take up new occupations (Hussainmiya 1990:99). 300 of them are known to have joined the police. Few discharged soldiers went back to the SEAA. Even after the disbandment of the Regiment in 1873, most Malays stayed in Ceylon (Hussainmiya 1990:96).

year	number	commentary
1801	81	recruits with their families sent by the Raja of Cochin (Malays employed by the Dutch who had defected to the Raja)
1803	82	recruits from Penang, some accompanied by their families and
	9	boys
1811	412	soldiers from Java accompanied by
	214	women
	208	children
1813	?	recruits from Madura (no number given)
1816	228	Javanese from Semarang and Gresek accompanied by their families
1819	102	recruits from Penang
1830	50	recruits from the Straits of Malacca and Singapore
1834	100	recruited in the Straits Settlement
1835	64	recruited in the Straits Settlement
1836	50	recruited in the Straits Settlement
1837	37	recruited in the Straits Settlement
1838	29	recruited in the Straits Settlement
1839	37	recruited in the Straits Settlement
1840	32	recruited in the Straits Settlement
1841	33	recruited in the Straits Settlement
1845-46	73	recruited in the Straits of Malacca and Singapore
1856-57	7	recruited in Borneo, Pahang, Trengganu, and Kelantan

Table 2.1: Arrival of recruits during the British Period based on (Hussainmiya 1987:60ff) and Bichsel-Stettler (1989:18). Note that the total number of Malays recruited from abroad during North's government is not known.

2.3.2   Demographic history

By the end of the Dutch rule, the Malays appear "to have been on a firm footing in the island much so that the British administrators, with very few exceptions, identified and referred to them only as Malays [and not by more precise denominations SN]" (Hussainmiya 1987:55). After a period of consolidation following the British takeover, the British took the Malays in service and, after all, not much had changed for them: they continued to serve in the army. The total number of Malays under arms was 1000 in January of 1800 before the recruiting from abroad started, then 1200 in 1801 (Hussainmiya 1990:62). The new recruits were allowed to bring their wives and children (Hussainmiya (1990:66), Hussainmiya (1987:12)). The influx of recruits from abroad can be seen in Table 2.1. The new recruits were the more important as Governor Maitland had sent back a good deal of the royal exiles in 1808, thus leaving the Malays

in Ceylon without their "cream" and ending the days of aristocratic influence. The newcomers of 1811 made up about a quarter of the Malay population.[9]

The subsequent recruitments were done in the Malay peninsula and were a lot less massive. The recruits came in a much more trickling way (see Table 2.1), with peaks in 1819 and 1834 attaining 100. After the Regiment had plunged to 1000 men in 1850, the rules of the regiment were changed and henceforth allowed retirement after 15 years of service. Only a few Malays made use of this and went back to Singapore or Penang. In Colombo and Kandy the Malays were numerous enough to populate entire quarters of the town: Katukalle in Kandy and Slave Island in Colombo (Hussainmiya 1990:111f).

The disbandment of the regiment did not have a great impact on the composition of the Malay population in Sri Lanka. Only 80 Malays went back to Singapore (Hussainmiya 1990:104). But there were no more new recruits from the East bringing fresh blood and new cultural developments to Ceylon. This lacking cohesion was detrimental to the Malay society. And while the number could be maintained, the degree of internal organization and cohesion was a lot smaller than during the CRR's days. The families had lost their great point of common interest. The new jobs in the police or on the plantations involved work on a smaller scale, in which one would meet a lot less Malays on an everyday basis than before. Even if many Malays served in the police, it was not an exclusive Malay domain like the Regiment, and Sinhala or Tamil officers would have forced the use of English, Sinhalese or Tamil in the office. Furthermore, the critical mass for maintaining a language might not have been attained by the smaller police stations.

To sum up, during the time of the Regiment, the Malays saw an important portion of their fellows leave: the aristocrats sent back in 1808. Governor Maitland did not recruit more Malays, which had the number of enlisted Malays plunge to 800. Governor Brownrigg wanted a strong Malay Regiment and perpetrated a massive recruitment, first from Java, and after the British had lost Java, from Penang and Singapore. Until 1819, 1016 new recruits joined the ranks of the Sri Lanka Malays, about evenly from Java and the peninsula. It was more and more difficult to find recruits in the East. The number of soldiers decreased. After 25 years without bellicose activity, the CRR was disbanded. While this did not entail a back-migration to the SEAA, it stopped the arrival of "fresh blood" to Ceylon and deprived the Malays of an important domain for using the Malay language.

---

[9]The number of Malay soldiers was 1200 in 1801, ca. 1280 if we count the 1803 recruits. Given the near-total recruitment of the Malays through the Dutch, and the British policy to also accept civilians in the ranks of the Regiment, we can say that the total number of Malay males combining soldiers and free Malays cannot be much higher than 1600 in 1805 when Gvr. North left. (This does not include women and children.) Gvr. Maitland had this number reduce further through sending the exiles back to Java and not replacing the casualties. If we assume that 200 exiles were sent back with their families, then a reasonable guess for the male Malay population in 1811, when the first recruits from Java came, is maybe 1300. Comparing this to 412 recruits arriving in 1811 with their families, the newcomers made up about a quarter of the Ceylonese Malay population.

2.3.3   Sociological History

The Sri Lanka Malays can be said to have developed a distinct identity in the beginning of the 19th century, different from both Lankans and Malays in the SEAA (Hussainmiya 1987:11). Their small community was organized around the military service, with the exiles being cultural and spiritual leaders, who sometimes also served in the regiment. The exiles formed an important part of the Malay society, and they "lent a certain amount of cohesion, reflected glory and guidance to its members" (Hussainmiya 1990:79), but a good part of them was sent back in 1808. The military officers then assumed the role of community leaders. They were well respected in the Regiment, even among British officers. Until about 1850 the Regiment formed a central part of the Malay society with everything revolving around it. Security, solidarity and understanding among the families characterized the life in the cantonment. The houses were built close to the military administration center, as was the mosque, where worships were held in the Malay language. Children could attend the school of the Regiment and become half-pay boys later before entering the ranks of the CRR, whereas old people were taken care of by the Invalids' Company. Marriage within the Regiment was frequent and the social status of a community member could be evaluated by the number and prestige of officers attending his wedding (Hussainmiya 1987, 1990).

"The Regiment constituted the centre of Malay-life in Sri Lanka throughout the British rule during the 19th century" (Hussainmiya 1987:93). The Malays were highly regarded in the regiment, also by the British officers. A career from a half-pay boy to a captain was possible, contrary to European officers, who were taken from the military academy and never from the low ranks (Hussainmiya 1990:122). The Regiment was bound by strong kinship ties. Many Malay families depended on the Regiment as their only source of income (Hussainmiya 1990:78) and soldiers joined because of family tradition. The Regiment took care not only of the active soldiers, but also of retired soldiers, who were either given land in the south or cash. Taking land in the south does not seem to have been very popular with the Malays because it would have removed them from the community. Unless the land was close to the cantonment (for instance *Kampung Pensen* in Kandy), the Malay preferred the money. While the Regiment provided cultural and religious support for the Malays, the actual living conditions of the soldiers seem to have been pretty poor. In the South, things seem to have been worse and soldiers are said to have lived in "wretched huts" (Hussainmiya 1990:115f).

In 1812, the first military school for Malay boys was founded. Their number rose to 5 by 1860 (Hussainmiya 1990:99). The goal of the school was to provide good education to the ranks of the Malays. Officers could order their soldiers to attend school. Even if parents were not in the military, they could send their children to this regimental school, which many did. The education provided was excellent and included Malay language, Tamil and English, but no Christian religious education (Hussainmiya 1990:96). Saldin (2001:25) states that English began to be spoken in Malay homes at that point in time due to economic necessity, although it is unclear

on what sources he bases that statement. In the 19th century, the literacy rate among Malays was the highest among all ethnic group in Ceylon (Marga Institute, Sri Lanka Center for Development Studies (1988:48), cited in Bichsel-Stettler (1989:17)). There are some numbers of the students attending the regimental schools:

> In 1835 there were 149 Malay boys studying at military school belonging to the CRR [...] in 1873, when the CRR was disbanded it was said that there were 269 half pay boys serving the regiment. (Hussainmiya 1990:97)

In 1872, shortly before the Regiment was disbanded, 391 soldiers were enrolled (Bichsel-Stettler 1989:17). If we consider the falling number of soldiers in the late 19th century, it is impressive that there were still that many students in the Regimental schools. This must be due to civilian Malays sending their children to the Regimental schools as well.

The attractiveness of the Regiment began to decline in the 1820s. After 1850, with low pay and service in Hong Kong making the Regiment unpopular, the number of free Malays (*Orang Priman*) went up. Civilian Malays would engage in the following occupations: civilian government departments, private business establishments, private European plantations, civil police department, district government agent's establishment, coffee and sugar plantation overseers, conductors, security guards, tea-makers, servants, gardeners, rattan weavers, clerks, peons, and book-keepers (Hussainmiya 1990:90).

Police service proved particularly popular, maybe because of the similarity to military work. 90% of the police sergeants in Colombo were Malay in 1833. After the coffee blight in 1872, tea cultivation began in Ceylon. The Malays found work in the new estates and in the railway construction companies that connected the highland to Colombo. But contrary to the work in the Colombo police force, the Malays were more of a minority in the countryside. For instance, they were outnumbered by the Indian Coolies in the tea estates (Hussainmiya 1990:90f).

The British fostered the religion of the Malays and provided a good infrastructure. There were special Malay mosques next to the cantonments. This aided the cohesion of the community. Services in the Moor mosques were unpopular because the service there was held in Tamil and Arabic, which did not suit the Malays. This also sheds doubt on the proposed strong link with the Tamil Moor community.

> The soldiers insisted on having mosques situated in the vicinity of their living quarters. Otherwise they had to travel far out of their camps to attend collective worships in the mosques belonging to the Moors. The Malays were at a disadvantage in attending the Moor mosques because the sermons and other religious proceedings were held, not in their mother tongue, Malay, but in Tamil and Arabic. (Hussainmiya 1990:124)

The mosques did not only serve as a place for religious meetings. Public meetings could also be held in mosques, also in Malay language. Mosques were founded by

Malays in Colombo, Hambantota, Kandy, Kurunegala, Kinniya, Gantulawa, Galle, Kalpitiya, Badulla, Kirinda (Hussainmiya (1990:124), Hussein (2007:449), see Figure 2.4).

The cultural leadership of the Regiment was challenged towards the end of the CRR (1873). In an argument about the position of the Imam in the Malay mosques, the soldiers could not impose their candidate, and a certain kind of schism occurred with military Malays and civilians attending different services (Hussainmiya 1990:125).

The Malays had regular contact with the SEAA during the 19th century. While they had a distinct identity, they had not forgotten their roots, and the new recruits arriving mainly in the first half of the century assured that the ties would not weaken. This would continue until the death of the last retired soldiers of the Regiment (Hussainmiya 1987:13). The recruiting missions were often accompanied by Malay officers and "some members of such recruiting missions stayed long enough in the peninsula, at times more than two to three years at a stretch, to bring back not only new developments in Malay culture, but also Malay literate texts and manuscripts to be distributed in their community" (Hussainmiya 1987:13). Hussainmiya also affirms that the Malays enjoyed visiting Penang, Malacca and Singapore. It is not entirely clear whether he sees this as some kind of tourism, or whether he refers to the recruiting officers.

2.4 After the Regiment

2.4.1 Political history

The disbandment of the Regiment had a major impact on the Malays. Many of them now served in the police, but contrary to the Regiment, the police was organized in small stations all over the country. They lost a lot of their social cohesion as compared to the regimental days. Occupations reported by Saldin (2003:10) include: jail guards; prison overseers; firemen;[10] overseers in the inland tea plantations, which were just being created as a result of the coffee blight; and railway workers. All these jobs had in common that the tight cohesion of the Regiment was lost and social life was no longer centered around the barracks. Smith (2003) comments that this disbandment cannot be seen as catastrophic, but it appears that the loss of social cohesion could still be seen as the major change within the social history of the Malays. Hussainmiya illustrates this:

> Nevertheless, as for the Malays there was a major difference between serving in the Police Department and serving in the Regiment. As policemen, the Malays were required to serve in the most remote corners in the island, mostly in small units, and therefore their community came to

---

[10]Hussein (2007:432) states that 75% of the police force, 90% of the staff of prison services, and 100% of the Colombo Fire Brigade Quarters were Malays.

> be disbursed in far away places unlike during the regimental days, when
> they could live in comparatively large numbers in the few military can-
> tonments, which in fact was a prime factor that helped in strengthening
> the social cohesiveness of the community. (Hussainmiya 1990:194)

Hussainmiya's object of study being the Regiment, his report ends with the af-
termath of the disbandment. We have very little information about the events in the
remainder of the 19th century. On a more general scale, the turn of the century saw
the rise of the movement for independence as well as the rise of communalism of the
different ethnic groups in Sri Lanka (De Silva 1981, Hellmann-Rajanayagam 1990,
Nissan & Stirrat 1990). Sinhalese nationalism resented Christianity, Westernization
and British rule. At the same time, the (Hindu and Christian) Tamils called in their
rights, and the Moors asserted themselves as well (Nuhman 2007:12). The Malay
community, too, formed a political association in Colombo for the first time in 1922
(Saldin 2001:19), the *All Ceylon Malay Association* (ACMA), which spread to other
parts with Malay population (Hussainmiya 1987:21).

> The political-identity crisis for the Malays also brought in its wake a
> strong desire to revive the community's cultural past which had been in
> the process of slowly being discarded. [...] The Malays' desire to assert
> their special heritage has been pursued relentlessly since then by various
> means, by organizing public lectures on their history, by activating spe-
> cial committees to conduct archival research on their past to search and
> revive their original Malay Jawi, and Javanese script, to publish books in
> Malay and to conduct special classes in Malay language etc. (Hussain-
> miya 1987:16–18)

Still, more and more mosque services switched from Malay to Tamil (Bichsel-
Stettler 1989:23).

To counter being subsumed into the "Muslim" group, the Malays demanded rep-
resentation in parliament (Hussainmiya 1987:15f). They even denied having any links
with the Moors at all (Hussainmiya 1987:19). From 1924 to 1965, the Malays had a
deputy in parliament, who was not elected by universal suffrage, but rather appointed
by the government to represent the minority of the Malays, strengthening their ethnic
cohesion (Hussainmiya 1987:20). The Malays' representative was TB Jayah, whose
legacy is still valued today. TB Jayah was also part of the delegation negotiating the
terms of independence with the British, obtained in 1948. After the 1956 election,
which replaced the European-oriented UNP with the rurally based Sinhala MEP, the
Sinhala parties used their majority in parliament to vote a number of laws favouring
Sinhala. Among these was the infamous 'Sinhala only' law, making Sinhala the sole
official language of the country in 1956, which was resented by the other communi-
ties and fueled the ethnic tensions (Nissan & Stirrat 1990:35). As for the Malays, the
mandatory use of Sinhala in the education system meant that they had to devote more
time to that language than before, when they had mainly functioned in SLM at home

and English in education. This caused a sort of chain reaction, where Sinhala pushed
English out of the educational systems, and the Malays, who wanted to preserve their
superior command of English as compared to the other groups, chose to educate their
children in English, which in turn pushed SLM out of that domain (Bichsel-Stettler
1989, Saldin 2001, Ansaldo & Lim 2006, Lim & Ansaldo 2007).

While many Malays had taken up work in the tea estates after the decline of the
regiment (Saldin 2003:39), in the latter half of the 20th century, many Malays quit the
tea estates and moved to Kandy or other Upcountry towns. This might be related to the
nationalization of the tea estates in the 1970s, which led to the good jobs being given
to Sinhalese (Nissan & Stirrat 1990:37). From the 1960s onwards, the ACMA declined
in importance, possibly a signal of an identity crisis among the Malays (Hussainmiya
1987:21).

The ethnic tensions between Tamils and Sinhalese led to the outbreak of open
civil war in the 1980s, which in turn led to a rise of ethnic consciousness in the other
groups and culminated in the creation of the SLAMAC (Sri Lankan Malay Confeder-
ation) on 18 August 1985 (Hussainmiya 1987:22).[11] SLAMAC hosted the 2nd Malay
world symposium in the same year (Bichsel-Stettler 1989:40). As of today, SLAMAC is
defunct (Saldin 2001:23), and its functions have been taken over by COSLAM (Confer-
ence of Sri Lankan Malays). SLAMAC was founded four years later than the Sri Lanka
Muslim Congress (SLMC), which is mainly supported by Moors. The Moors were able
to turn the SLMC into a political party with some influence. This influence led to recent
statistics subsuming both Malays and Moors under the ethnic [sic] group "Muslims".
This makes future representation of the Malays in parliament unlikely, since their eth-
nic group of "Muslims" is already represented in parliament by the SLMC, even if
all of the members are Moors, whose interests very often diverge from the Malays'.
This arrogation of voice of all Muslims by the Moors has led to irritation within the
Malay community, who do not feel that the Moors can speak on their behalf, albeit
that they are Muslims (Nuhman 2007:24f). The up-to-now last Malay MP, elected in
1989 (MH Amith), is then indeed not a member of the SLMC, but of the United National
Party (UNP).

Bichsel-Stettler (1989:38) resumes the political dilemma of the Malays in Sri
Lanka as follows:

> The Malays are in the uncomfortable position of being a minuscule mi-
> nority within a much larger and more influential minority [the Moors,
> S.N.] in a political climate unfavorable to minority interests. (Bichsel-
> Stettler 1989:38)

2.4.2   Social history

While Hussainmiya provides ample detail about the sociohistoric conditions in which
the Malays evolved in the 19th century, information of comparable quality is not avail-

---

[11]Saldin (2001:22) has 1984.

able for the 20th century because the Ceylon Rifle Regiment, Hussainmiya's object of study, had been disbanded by then. In what follows, I will sketch the sociohistorical and demographic history of the Upcountry Malays based on extended discussions with several older members of the Malay communities from Kandy, Gampola, Nawalapitiya, and Badulla. I also draw on similar interviews conducted by Bichsel-Stettler (1989).

The end of the Regiment also meant the end of the regimental schools, and thereby the end of education in Malay.[12] The Malays switched to English or Tamil education instead (Hussainmiya (1987:88), Saldin (2003:14)).[13]

Assessments of the importance of this event for the evolution of the language vary, as the following two quotes illustrate.

> The potentially disruptive effects of the disbandment of the regiment produced its aftermath within almost one generation in the community [...] [U]nlike in the Regiment when the Malys could live in closely-knit cantonments which reinforced their social and geographical cohesiveness, other sectors which employed the Malays, particularly the police service caused dispersion of the community. Above all, in the new areas of employment, there was no need for the Malays to learn their own language. (Hussainmiya 1990:155)

> However wrenching this may have been for the lives of individual Sri Lankan Malays, [the disbandment of the CRR] can hardly be considered the type of revolutionary change required for rapid contact-induced change. (Smith 2003:19)

The disbandment was the most catastrophic event of Malay history in Ceylon. Whether this has had a decisive impact on the language or not can only be decided if we compare the shape of the language at two points in time: before (t_0) and after (t_1) the event. Unfortunately, we have no information at all about the shape of the colloquial language before 1900. This entails that a comparison between the language at the points in time t_0 and t_1 is impossible. This question simply cannot be decided based on the known material.

Still, sociolinguistic theory suggests that weakening of the social network incurred by the disbandment of the Regiment probably had a major impact on the Malay language. Milroy & Gordon (2003:124) remind us:

> Networks chiefly constituted of strong ties support minority languages resisting institutional pressures to language shift; but when these networks weaken, language shift is likely to take place. (Milroy & Gordon 2003:124)

---

[12]A private girls' school was run for some years after the end of the Regiment (Hussainmiya 1987:102).
[13]Interestingly, the Moors, who are said to be culturally close to the Malays, refused English education to avoid being proselytized (Nuhman 2007:57).

The generation of Malays born around 1880 (thus the grandparents of the oldest speakers) chiefly worked in the army, law enforcement, the railway, or as overseers on the estates, where Malays formed 10% of this group (Bichsel-Stettler 1989:26) (See Table 2.2 for an overview). In those occupations, there was little need for Malay education, and command of English and a local language like Sinhala or Tamil was more important (Hussainmiya 1990:147). The children and grandchildren went to mixed schools, where Malays, Sinhalese, Moors, Burghers and Tamils[14] were educated together. Generally, Malays either opted for Tamil-medium or, most often, for English-medium schools (Bichsel-Stettler 1989:24). Sinhala-medium was not chosen until the middle of the century (Bichsel-Stettler 1989:27f). In one school, in the first grades, classes were taught in Tamil with one subject being taught in English, while in the advanced grades, English was the medium of instruction (K051213nar02). Many schools were run by Irish nuns, where education was in English.[15] Command of other languages was only necessary for working Malays. Since women very often stayed at home and cared for the children, it is possible that they did not have extended contact to the other ethnic groups at all. Bichsel-Stettler (1989:22) even speculates that at the beginning of the 20th century, there might still have been some monolingual Malay women (also see Saldin 2001:25f).

Before independence, the Upcountry Malays had close ties with the Burghers on the estates. During this period, many Malays dropped their traditional name prefixes (Saldin 2003). Malay lifestyle resembled more the Burghers', and was less traditionalist than the lifestyle of the Moors. This is still the case today with the Malays being much more lenient with regard to the literal following of Islam teachings than the Moors, who tend to be more observant. Examples of this are the wearing of the veil, which is more common among Moor women than among Malay women, and the size of the veil, which tends to be less for Malays, if present at all. Regular attendance of religious services seems to be more important among the Moors than among the Malays. An additional factor which might influence this is that mosques operate in Tamil, which is understood by all Moors, but not by the younger generation of Malays.

Sri Lankan Malays are commonly organized in local organizations (*Sri Lanka Malay Association of Colombo, Uva Malay Association*, etc.), which cultivate various aspects of Malay culture (song, dance, sports). Songs and dances are normally adapted from either Sinhalese, Tamil or Malaysian/Indonesian records or videos. Sinhalese and Tamil are translated into Sri Lanka Malay, but Malaysian/Indonesian is not. These organizations seem to follow a British model and protocol with treasurers, patrons, minutes, and elections of boards based on proposition and seconding. Most of these associations are affiliated with the umbrella organization SLAMAC (Sri Lanka Malay Confederation)/COSLAM (Conference of Sri Lankan Malays).

After independence from the UK, language policies favouring Sinhala made more and more Malays choose Sinhala-medium schools if they could not afford expensive

---

[14]Estate Tamils (Indian Tamils) or Ceylon Tamils.

[15]A curious fact is that in one school, Islamic education was taught by Tamils from Jaffna, who were Hindus. Apparently, this did not pose a problem for either of the two communities.

Professional		**Industrial**	
government or local government	769	mason	106
		draper/cloth dealer	6
physician/surgeon or medical practitioner	5	tailor/millmen	38
		butcher/meat salesmen	21
Domestic		poultry and egg seller	1
barber	8	fishmonger	9
Commercial		paddy/grain seller	3
shopkeeper	19	baker, bread/cake seller	97
boutique keeper	62	vegetable dealer	3
general trader	83	coconut/copra seller	2
petty trader	22	coffee picker	6
boatmen	3	tobacco/cigar maker or dealer	1
Agricultural			
estate conductor	78	betel/arecanut seller	2
coffee garden conductor	8	timber feller	8
horse/cattle trader	2	salt dealer	1
fisherman	10	**Jewellery and Gem Trade**	
cultivator	40	goldsmith/silversmith	1
coffee garden cultivator	8	gem digger	5

Table 2.2: Occupations of Malays in 1881 following Marga Institute, Sri Lanka Center for Development Studies (1988:51), cited in Bichsel-Stettler (1989:26).
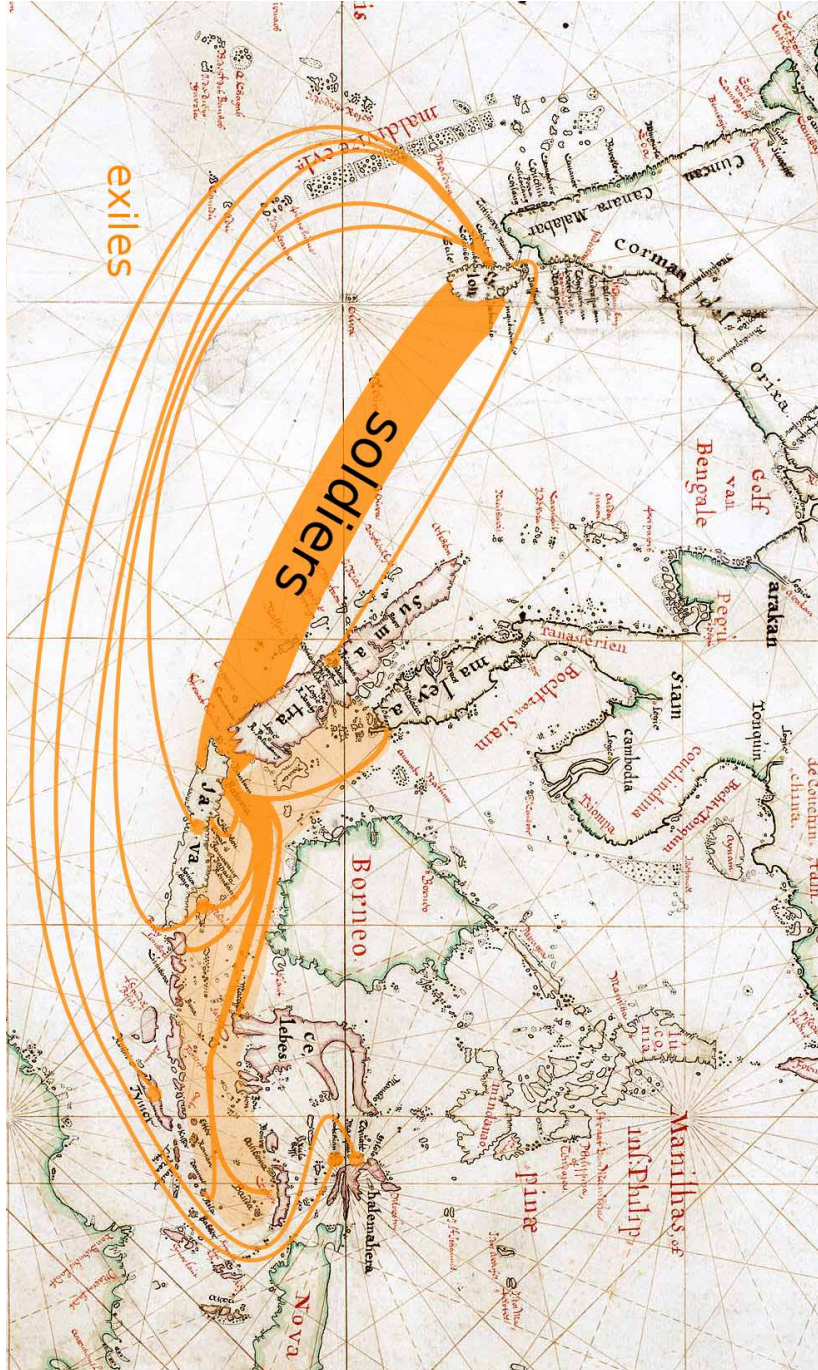
Figure 2.1: The provenance of the Malay immigrants during Dutch rule. The size of the arrows corresponds to the relative strengths of the contingents. The routes represented on the graphics are not necessarily those employed by the immigrants, with the exception of the stop-over of soldiers in Batavia/Jakarta. Small arrows represent exiles, whose precise numbers are not known.
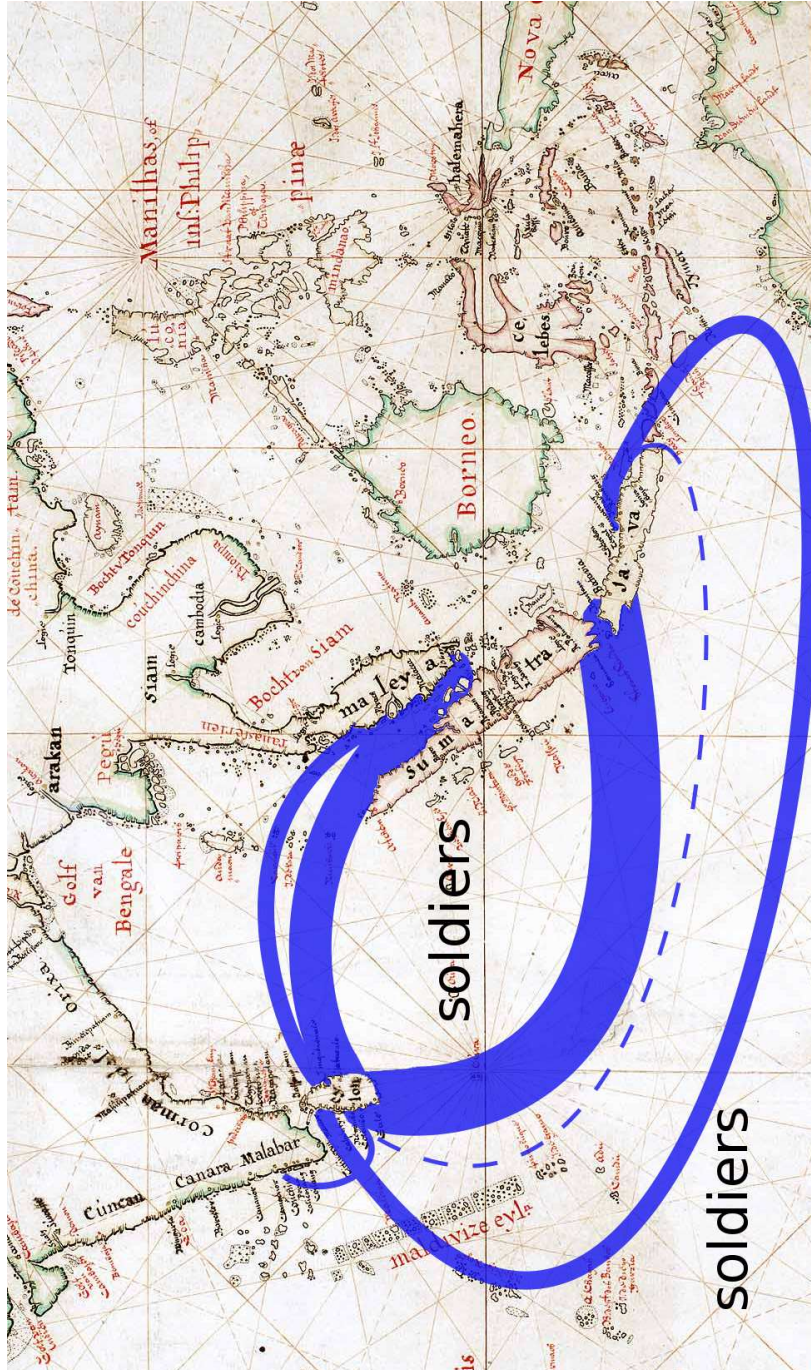
Figure 2.2: The provenance of the Malay immigrants during British rule. The first soldiers came from Jakarta; later Penang and the Straits settlement on the continent became more important. The size of the arrows corresponds to the number of immigrants.
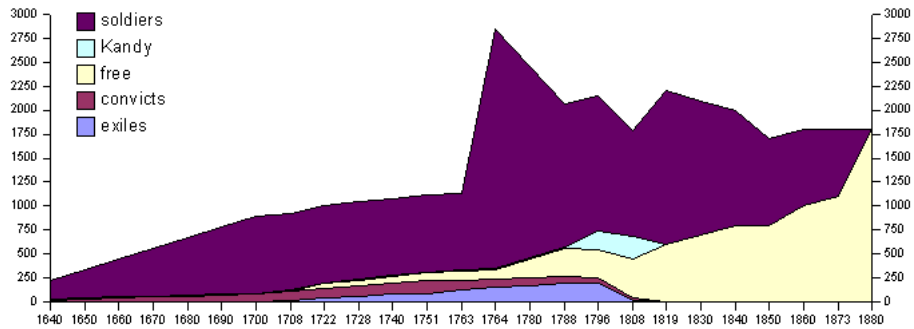
Figure 2.3: The development of the Malay population in Ceylon over time. Data are scarce and unreliable. Data for soldiers are good from 1764 onwards. The other groups are mostly based on two or three years where information is available and linear growth or decrease to/from these years. Furthermore, there is virtually no information about women, so the diagram is to be taken with a huge grain of salt. Sure points are the beginning and end of the exiles population, the short interlude of Kandy Malays and the growth of the free Malay population after 1796 to the detriment of the soldier population.
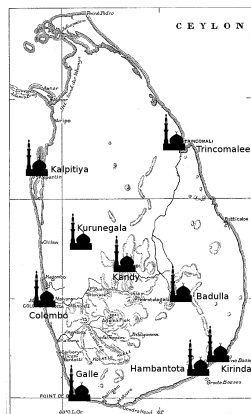
Figure 2.4: Location of Malay mosques in the 19th century.

private schools, which continued to operate in English. With the Malays wanting to keep the comparative advantage they had because of their language skills, many opted for English as the language of the home, to complement Sinhala as the language of education. This pushed Malay out of the equation (Bichsel-Stettler 1989:29f). Saldin (2003:76) comments:

> When we were children, we spoke Malay at home but, spoke English at school, as that was the medium of instruction. Our children were born to a different world where, Sinhala was the official language and the medium of instruction. They had to study in Sinhala. [...] Sheila insisted that we speak English at home for the benefit of the children [...] The Malay families that did not have an academic background in English spoke Malay at first. When the children went to school and learnt in Sinhala, they continued to use Sinhala in their homes.[16]

While in former times, women were not necessarily conversant in English, through the rise of female education in the 1930s, more and more women learned English. This language then supplanted Malay also in the household. The following quotations from Saldin's biography attest to this.

> We used to speak in Malay in our family household but, after I married, I rarely had a chance to speak in Malay. Malay was an alien tongue at Taprobane. (Saldin 2003:76)

> Malay was hardly spoken in the homes of the members of the Malay Uniques Club. (Saldin 2003:83)

> [In the Malay association] Most of the members of the older generation knew their Malay but due to various reasons were reluctant to speak it. (Saldin 2003:94)

The numerical dominance of the Moors leads to the use of Tamil in the mosques, which many Malays do not understand (Saldin 2003:86). However, the Wekande Mosque in Slave Islands delivers sermons in Malay on one Friday each month (Saldin 2003:86,113). Furthermore, there are some Sinhala services now, at least in Kandy.

Around 1957 Malay was introduced as a school subject, but not for long (Saldin 2003:118). Since 1993, it is again possible to take Malay as a subject in O/L and A/L exams (Saldin 2001:29). There is no Malay course as such, but the prospective student is required to work on his own with a manual conceived by BDK Saldin, who also grades the exam.

The current sociolinguistic situation for the Malays is such that the most important language to master is English, which is needed for economic progress. The second most important language is Sinhala, which is the official language, and the majority

---

[16]Also see Hussainmiya (1990), Ansaldo & Lim (2006), Lim & Ansaldo (2007), Ansaldo (2008).

language in the parts of the country where the Malays live. Knowledge of Tamil is less important, and many younger Malays in Kandy do not know Tamil anymore. Still, Malay children in the smaller Upcountry towns acquire Tamil as well. Malay, finally, is also not regularly transmitted to the younger generations. The older generations of Malays derive an enormous pride from their quadrilingualism (Ansaldo & Lim 2006, Lim & Ansaldo 2007, Ansaldo 2008) and it seems that it is their ability to communicate well with every ethnic group which conveys them their ethnolinguistic identity (Bichsel-Stettler 1989, Ansaldo 2009b). Frequent allusions are made to this fact: "we know all four languages", "we are the only ones to understand everybody", etc. This is understandable from the fact that the Malays are a tiny minority in Sri Lanka and can simply not afford particularism. The only way for them to get along is to adapt to as many contexts as possible. While the older generations have very good command of grammar, style and register in all four languages, this is not necessarily the case with the younger generation, who often only have full command of Sinhala and a local variety of English. Knowledge of Tamil is often absent, and full command of Sri Lanka Malay (which would permit to convey any message) is as well. Still, the Malays have the highest literacy rate in Sri Lanka (93.2%), second only to the Burghers (98.2%) (Bichsel-Stettler 1989:32).

In the course of the twentieth century, the main fields of occupation of the Malays did not change, and various jobs in the maintenance of public order like the police, jail guards, firemen or soldiers remained popular. Given that after independence the Sinhalese acceded to the ruling positions in these areas, the Malays' loyalty now resides with them, and support for the actions of the Sri Lankan government and the Sri Lankan Army (both dominated by Sinhalese) in the civil war is widespread.[17] However, jobs in the government sector were predominantly given to Sinhalese from the 1960s onwards (Bichsel-Stettler (1989:29,33f), Nissan & Stirrat (1990:37), Osman & Sourjah (2008:55)), a fact which is resented by the Malays, who had had a long tradition as public servants, and the English language skills necessary for that post before the rise of Sinhala nationalism.

In recent years, however, white-collar activities in the commercial sector seem to be increasing at the expense of the traditional occupations (Saldin 2001:48). Saldin argues that in these areas, Malays have an edge due to their relatively superior command of English. Also, temporary work in the Middle East is common (Bichsel-Stettler 1989:34), actually since the 1970s (Saldin 2001, Osman & Sourjah 2008), as is emigration, mainly to Australia and Canada, where emigrant Sri Lanka Malay associations can be found.[18] In the 1990s, there was an initiative for massive emigration to Malaysia, the *Malay Action Front*, but it failed to make a noticeable impact (Saldin 2001:45ff).

The Malays are quite well integrated into the Sri Lankan society and participate

---

[17]I do want to emphasize the complexity of the Sri Lankan society and stress that a simplistic equation government=Sinhala is not warranted. The government is not entirely constituted of Sinhalese, and its actions do not have the support of all Sinhalese.

[18]SLAMA in Australia `http://www.geocities.com/slama_web/`, SLAMAT in Toronto.

in all walks of life up to the topmost strata, including supreme court judges (Nuhman 2007:24), national sportsmen, ministers, film stars, singers, media presenters, important journalists (Osman & Sourjah 2008:53), or high representatives of Sri Lankan Airlines. This contrasts with the assessment by Hussainmiya (1990) (probably gained in the 1970s) that

> there are no noteworthy Malay businessmen and therefore no surprise if one cannot find wealthy people among them. In this respect, the Malays find themselves on the lowest rung of the economic ladder as opposed to other Muslim groups in the island, such as the Tamil-speaking Moors and the traditional business communities of 'Borahs' and 'Memons'.

While it is true that the Malays are less inclined to commerce than other parts of Lankan society, there are a number of wealthy Malays in Colombo, but their wealth normally stems from dependent work. This is in line with Bichsel-Stettler's analysis of the Malays' economical situation in the beginning of the 20th century:

> The Malays were not rich entrepreneurs, nor were they strongly represented in prestigious fields such as law or medicine; nevertheless, as a group they led a solid existence in urban centers. (Bichsel-Stettler 1989:27)

Today, it is possible to marry members of other communities, but the spouses are then required to convert to Islam if they were following another religion before (Lim & Ansaldo 2006). As a rough estimate, about 10% of marriages involve a member of the Moor community and 5% involve a member of the Sinhala Buddhist community. Marriages to either Hindus or Christian are very uncommon. Among the higher strata of society, intermarriage may be more common. Saldin (2003:83f) remarks that at one stage in the nineties the key office bearers of the Malay Club were children of non-Malays or were married to non-Malays. This is in opposition to former times, when, according to Saldin (2003:64), mixed marriages led to ostracization.

As of today, it appears that the Malay culture is clearly Lankan. Funeral and wedding customs are almost identical to the Moors (Saldin 2003:1,79), with a combination of Muslim, general South Asian and Western practices, while the customs relating to a girl's growing up have the greatest similarity with the Sinhalese (Saldin 2003:35).

2.5 Cultural history

In the course of history, Malay poets have produced a good number of literary works in Sri Lanka, mostly during the period of the CRR.[19]

---

[19]There is no indication of literary activity preceding the Regiment. If there was literary activity, it could not have been widespread and would be confined to religious texts (Hussainmiya 1990:143).

During the Regiment, artistic work as well as newspapers were produced. Artistic work comprised adaptions of traditional topoi as well as the creation of own poems. Genres found in Malay literatures are: religious kitabs; books on magic, sorcery and divination; catechism and prayer; family trees; and personal memoirs (Hussainmiya 1990:135). There is a penchant towards religious, particularly mystic literature made for specialists, emphasizing the observance of *Sharia'* and the quest for *ma'rifat* (gnosis). The genre of 'Syair' enjoyed particular popularity (Hussainmiya 1990:139f). Hussainmiya further notes that historiography of the SEAA dynasties is completely absent. This can be understood as an affirmation of an own Sri Lanka Malay identity. On the other hand, there are two works that are unique to Ceylon: 'Hikayat Raden Bagus Gusti' and 'Hikayat Indera Kuraisy', whether they are originally from Ceylon or just did not survive in the rest of the Malay speaking world (Hussainmiya 1990:135f).

Literature was written in High Malay, not in Sri Lanka Malay or other varieties of Malay. The script used was a modified form of Arabic without dots, Jawi script, also called Gundul.[20] The name *gundul* has a Javanese etymology, which might indicate that the exiles from Java had brought literature to Ceylon (Hussainmiya 1990:134ff).

Sri Lanka Malay literature was not isolated. Grammars and dictionaries were exchanged between Ceylon and the SEAA (Hussainmiya 1990:145). Despite some localisms, the Sri Lankan Malay literature did not differ much from those works (Hussainmiya 1990:149). Two newspapers were issued in Ceylon, *Alamat Langkapuri* and *Wajah Selong*. The former was popular with the established Malay population while the latter was preferred by the more recent arrivals (Hussainmiya 1990:156).

Literary interest was fostered by the British by the creation of a library-like recreation room for Malay officers (Hussainmiya 1990:146). Malay literature is not only to be read but also to be recited. The Regiment provided a frame for this. With the disbandment of the Regiment, this kind of literary (inter)action became much more difficult. The poets moved away in search of work and it was difficult to have them meet. By consequence, literary activity stopped in the beginning of the 20th century. The last officers educated in the regimental schools died out, and the younger members of the community did not know how to write Malay. In their new civil occupation, knowledge of English and Sinhala or Tamil was more important. The first literary work in Roman script was published in 1906, but it is of inferior quality (Hussainmiya 1990:149), and its appearance was ephemeral. As long as the literature persisted, the Malays were aware of their origins in the SEAA, whereas when the literary tradition died with the disbandment of the Regiment, the links with the SEAA went more and more into oblivion (Hussainmiya 1987:13f). Towards the turn of the millennium, several Sri Lankan Malays profited from the time allowed by their retirement to write books about Malay culture and language. These books are written in English. Examples include Saldin (2001, 2003) and Thaliph (2005).

Literary activity seems to have been confined to Colombo and Kandy (Hussain-

---

[20]*Gundul* means "hairless", and the dots were seen as hairs.

miya 1990:146). *Kampung Katukalle* and *Kampung Pensen* in Kandy are mentioned frequently as places of copying the manuscripts.

As of today, Sri Lanka Malay is mainly a spoken language (cf. Hussainmiya 1990:134). Occasionally, some songs or other artistic material are written down in Latin script, Sinhala script, or Tamil script. The regular radio program *Suara Melayu* (established 1957, Saldin 2001:28), for instance, contains a popular radio drama, which is enacted by normally three persons reading out the text written in Tamil script. Another example is a text on the life of the prophet Mohammed, found in Kandy, which is written in Sinhala. On the first Friday of each month, and on some Islamic holidays, the state television broadcasts a Malay program. Bichsel-Stettler (1989:47) reports that letters to the producers of the radio drama are written in Malay using Sinhala, Tamil or Latin script. Starting in the 1980s, there have been some attempts to set up language classes for the Sri Lankan Malays, which do not seem to have gained wide popularity (Saldin 2001:39ff). The Malaysian representation in Sri Lanka is very eager to promote Malaysian language and culture among the Malays abroad, while the Indonesian representation shows less enthusiasm in this regard.

There are a number of traditional songs called *panthongs* in Sri Lanka. These used to be accompanied by the violin and a kind of tambourine called *rabaana*. The last person in Kandy to accompany the *panthongs* on the violin died some years ago, while the last Kandyan to perform the chant and rabaana died in 2008. Newer Sri Lanka Malay songs are adaptions of Sinhala, Tamil, Indian or English tunes and follow general Indian/Western styles. Some of the Sri Lankan Malays in Kandy have a vast array of modern musical instruments and perform(ed) in bands for weddings and other festivities. Depending on the ethnicity of the bride and groom, they would adapt the set for the evening to Sinhala, Tamil, or English songs, with the occasional Hindi in between.

2.6 Malay settlements

Being soldiers, the Malays had always lived close to the centers of colonial power when they were not engaged in military campaigns. The native Lankan population was much more dispersed, and even today, Sri Lanka is not a very much urbanized country.[21] Moreover, the exiles were confined to the centers of colonial power because they had to be under control of the authorities. This urban base distinguishes the Malays from the Sinhalese and Tamil Hindus, and to a lesser degree from the Moors.

During Dutch times, Malay regiments were stationed in the main coastal towns of Colombo, Galle and Trincomalee (Hussainmiya 1990:50). Kandy was still independent during that time, but had a certain population of deserted Malays. After the British had vanquished the Dutch, and later the Kandyans, the main Malay settlements were Colombo and Kandy. The first non-military settlement was founded in

---

[21]Ranked #181 among 199 surveyed countries, `http://www.nationmaster.com/graph-T/peo_per_liv_in_urb_are` retrieved on Sun Jun 28 09:06:52 CEST 2009.

1802: Hambantota. Retired Malays were given land to cultivate there. This colony was followed shortly by Kirinda and Palatupana not far away.

Besides Colombo, Kandy and the South, there were further detachments in the towns already inhabited by Malays during Dutch times, and in Badulla, which had been under Kandyan rulership. Following the 1848 rebellion, special detachments were set up in Kurunegala and Matale. In 1860, there were five companies in Colombo, three in Kandy, two in Galle, two in Trincomalee, one in Badulla, and the Sepoy company in Jaffna. In addition, detachments of soldiers from Colombo or Kandy or Galle were sent to Puttalam, Kurunegala, Hambantota and Chilaw. Figure 2.5 shows the towns where the Malays were placed in 1860. In 1865, the Military Commission recommended the closing of all CRR stations for reasons of economy, except those of Colombo, Kandy, Galle, and Trincomalee (Hussainmiya 1990:101). Hussein (2007:425ff) states that

> Malay settlements in the other upcountry districts such as Badulla and Bandarawela seem to have originated with the lure of employment in the plantation estates of these areas. These appear to have been principally peopled by disbanded or retired soldiers of the Malay garrison stationed at Kandy. More recently, Malays settlements have arisen from colonization schemes sponsored by the state. The earliest such scheme was Beragama Malay Colony founded in 1930, and this was followed by Bolane Malay Colony off Ambalantota. Recent years have also seen a drift of Malays from Colombo to Wattala, particularly to the relatively new settlements such as Akbar Town in Hunupitiya named after a well known Malay Mas Thajun Akbar. (Hussein 2007:425f)

The former Malay populations of Kalpitiya and Trincomalee (and possibly Jaffna) seem to have absorbed in the surrounding Moor population (Hussein 2007:37,426f).

According to the censuses of 1981 and 2001, the Malay population concentrates in the population centers given in table 2.3.

The variety of SLM described in this grammar is spoken in the central Upcountry town of Kandy. Kandy was the last seat of the Sinhalese kings, and the last kingdom to surrender to the British. It is the cultural centre of the Sinhalese and houses the Temple of the Tooth containing the most important Buddhist relic in Sri Lanka, a tooth of the Buddha. Kandy is the second largest city of Sri Lanka, with a population of about 100,000. Historical population figures given by the Municipal Council[22] are 2500 for the early 19th century and 16,881 for 1871.

Kandy lies at an altitude of 500m, at the foot of the central tea cultivating region. The main road (and railway) accessing that region leaves from Peradeniya, at about 5 km distance from Kandy. The first Malays to arrive in Kandy were soldiers deserted from the colonial armies. This happened during Dutch times. It is not entirely clear whether the Portuguese as well had already Malays under their service

---

[22]`http://www.kandycity.org/kandycity-geogra.html`, retrieved on Wed Mar 25 11:34:49 CET 2009.
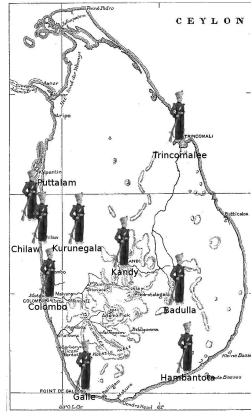
Figure 2.5: Malay companies or detachments in 1860

	Town	Population 1981	Population 2001
1	Colombo	20,041	19,428
2	Gampaha	8,077	12,333
3	Hambantota	4,380	7,385
4	Kandy	2,648	2,210
5	Badulla	1,300	1,390
6	Kurunegala	1,201	1,123
7	Nuwara Eliya	1,113	n/a

Table 2.3: Population centers with Malay population. Census data stem from Bichsel-Stettler (1989) for 1981 and from Hussein (2007:423) for 2001. Notable is the gain of population in Gampaha district, where many Malay families from Colombo moved to. The population loss in Colombo was countered by immigration from other areas. The rise of Malay population in Hambantota is surprising and cannot be explained here.
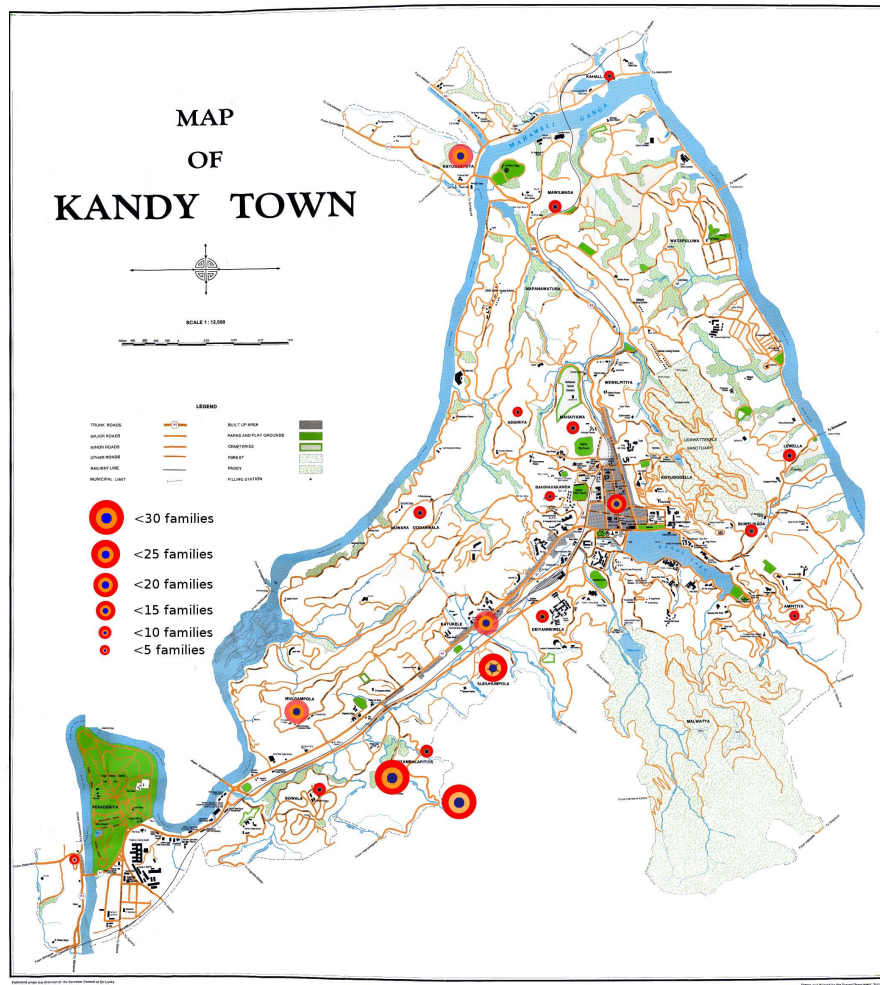
Figure 2.6: Map of Kandy town. Total around 250 families.

(see 2.3.1.2). The Malay population centered around the quarters of Bogambara and Katukalle, where there were also Malay mosques (which still bear that name today). Especially the Old Peradeniya Road leading through Katukalle and Mulgampola contained many Malay households. However, this area has developed economically and many new shops have opened there. This has led many Malays to sell their property there and to move to more rural areas. The roads leaving from Peradeniya Road southbound into the hill area through Mulgampola, Rosawatta and Heerassagala also contain many Malay settlements. But today, Malays are not confined to one area and can be found interspersed in all quarters of Kandy. The Malay population in Kandy is about 2,500, making up 2.5% of the total population. Even in quarters with relatively many Malays, the percentage would not be higher than perhaps 5%, making the Malays a not very visible minority. In fact, there are other inhabitants of Kandy who are not aware of the existence of a Malay minority in their town at all. Figure 2.6 gives an overview of the repartition of the Malay population in Kandy.

The main upcountry towns of Kandy, Nawalpitiya, Nuwara Eliya, Hatton and Badulla were by no means isolated, and there were frequent contacts among inhabitants of different towns, which can be seen by the very common marriages involving partners from different towns. Marrying into a family of another town was a common practice (and still is today). This might have helped preserve Upcountry Malay as a coherent variety with little internal diversification. As an interesting side effect, the idiolectal diversity we find in the Upcountry seems to be tied not to region, but to family lines. For instance, one speaker from Kandy uses [aəʔɖuːʔuʔ] for the present tense from *aràduuduk* 'is.sitting' even in careful speech, which is normally pronounced [arəɖuːɖuk]. This makes more sense if one knows that she was born in Badulla, but that her father was born in Hambantota in the South, where such contractions seem to be more common. The rest of her idiolect does not differ much from other idiolects from Kandy, but this one feature betrays the origin of her family in the South two generations earlier.

Contact with Colombo was also very frequent (at least after independence), and most Malays from Kandy have relatives in Colombo. Other than that, one speaker from Kandy had her mother from Negombo, to the North of Colombo, and another's mother is from Puttalam. The wide-spread kinship ties of the Malays prove that in the last three generations, the Malay network covered all of the Upcountry and the western coast with Colombo and Negombo. It is less common to have relatives in the South (Hambantota, Kirinda); this community seems to form a separate network (Ansaldo & Lim 2006, Lim & Ansaldo 2007). There are some Malays in Galle, who still speak Malay, but who seem to be quite isolated from the rest. Figure 2.7 gives an impressionistic illustration of the perceived strength of ties. Further, more sophisticated demographic research must show whether the postulated strengths of ties hold up to statistical analyses.

The Malays in Kandy have long been organized in cultural associations, the oldest one of which is the Kandy Malay association. Recent years have seen a number of schisms and splits among the Kandy Malays, leading to the foundation of the

Figure 2.7: Impressionistic strength of contacts between Kandy and other towns, and between Badulla and the South

Hill Country Malay Club and the *Hill Country Malay Association*. The Malays in Kandy take part in island-wide cultural activities, such as singing, dancing, and sport events. In distinction to certain strata of Colombo society (Ansaldo & Lim 2006, Lim & Ansaldo 2006), there seems to be little interest in teaching Malaysian Malay in Kandy, and attempts to speak Malaysian Malay are occasionally derided as *Thuan thuan language*, imitating the reduplicating pattern found in Malaysia, but not in Sri Lanka.

While in the surrounding Upcountry towns, the Malay language seems to be regularly transmitted to the younger generation, this seems to be less the case in Kandy, where English and Sinhala have made inroads into the domain of child care.

2.7   Relation to the Moors

In all academic studies about SLM, with the exception of De Silva Jayasuriya (2002) and Ansaldo's work, a strong emphasis on the contact between the Malays and the Moors[23] is found. This has been called the 'Tamil bias' (Ansaldo 2005b, 2008). The proponents of a close link between Malays and Moors argue that the two minorities shared Muslim faith and were thus natural allies in a predominantly Buddhist and Hindu country under Christian colonial reign. This is also used to explain the linguis-

---

[23]Some Moors resent the term 'Moor' for their ethnic group and prefer 'Conakar/Sōnahar'. 'Moor' goes back to the Portuguese who designated Arabs as 'Moros', while 'Conakar' goes back to Tamils who designed first Greeks and then Arabs that way (Hussein 2007:2), so neither of those is a native term (Nuhman 2007:45). In the course of history, there has been a heated debated among this community whether one should identify as 'Moor' or 'Muslim' (Nuhman 2007:42ff). I follow academic usage in employing the term 'Moor', which furthermore has the advantage of being less ambiguous than 'Muslim'.

tic shape of SLM, because Moor women married to Malay men would have acquired an imperfect Malay that they passed on to their children. This imperfect acquisition would be at the base of modern SLM. Ansaldo (2005b, 2008) criticizes that the historical evidence provided for this scenario is not up to scientific standards, and that oral history of the Sri Lankan Malays suggests that extensive intermarriage with the Moors did not happen.[24]

On analyzing the references in the literature one finds that the claim of close contact to the Moors is mainly based on Hussainmiya (1987, 1990). The theory does not lack plausibility and is supported by some historical sources. But Hussainmiya often argues on intuitive grounds and the sources for certain of his claims are not given or vague. He himself (1987:45) says that "the details of the nature of mutual contacts between these people cannot be documented." Let's take a close look at what he presents.

The argument is based on:

- the gain in Islamic knowledge

- the Moors helped a Malay prisoner

- Christopher Schweitzer's report

- the Dutch *tombos* and Sri Lankan *Kadutams*

- a cultural comparison with Malays in South Africa

Hussainmiya (1990:43) argues that Malays in Ceylon must have had contact with great Islamic scholars. Thus, a certain exile named Radin Adipati Natakusuma was banned to Ceylon in 1743. When he returned to Jogyakarta in 1758, he was made chief of the religious officials there. A similar case is that of Wirakusuma, who became the leader of a religious group on returning. Other exiles are said to have become pupils of the Islamic teachers Sayyid Musa Ngidrus and Ibrahim Asmara in Ceylon. This is seen as evidence for contact between the exiles and Islamic scholars in Ceylon. Whether these scholars were Moors is a different question. Wieringa (1997) informs us that Ibrahim Asmara was an Islamic apostle from Champa (today's Vietnam), and as such it is unlikely that he was a Sri Lankan Moor. It remains furthermore questionable whether the normal soldiers would have undertaken thorough religious studies in Ceylon.

The second piece of evidence are minutes from the Dutch Political Council Colombo, suspecting the Moors of aiding exiled Malays in communicating with each other.

The third piece of evidence is a report of a German employee of the VOC, Christopher Schweitzer, from 1680, taken as evidence for Malay-Moor intermarriage by Smith & Paauw (2006).

---

[24]Ian Smith (p.c.) remarks that Malays have a political interest in being seen as separate from the Moors (see Section 2.4.1 above), and that, therefore, they might be downplaying their Moor heritage.

> Their language is Ambonese, but the majority also speak Malay, Sinhala,
> Portuguese and Dutch. [...] The wives, who in part are Ambonese, in
> part Sinhalese and Malabarian may not say anything [against the stripping
> of their ornaments]. (Schweitzer 1931[1680]:106) [my translation][25]

It is difficult to find evidence for contact between the Malays and the Moors in
this passage. There are of course the Malabarian wives, with *Malabarian* being an
old word for "person speaking a South Indian language" (mostly Tamil) (cf. Rogers
1990:92). But nothing is said about their religion, so Hindu Tamils cannot be excluded
(cf. Hussainmiya 1987:68). There are also Sinhalese and Ambonese wives, and the
soldiers are not even mentioned to speak Tamil,[26] only Malay, Sinhala and the colonial
tongues. If one has to rely on this fragment of text, then it surely does not support a
close contact between Malay soldiers and the Moors, because this would have resulted
in knowledge of Tamil rather than Sinhala.

The fourth piece of evidence are the Dutch *tombos*, particularly the head tombo
available at the Sri Lanka National Archive 1/3758, in which a number of cases of
intermarriage are said to be reported. Unfortunately, Hussainmiya does not give the
numbers, so it is difficult to evaluate this claim. He (1987:25) also cites Lankan mar-
riage registers (*Kadutams*) that are in his possession and that also list intermarriages.
These registers could not be consulted for this book.

Ansaldo (2008) checked these claims against the photocopies of the Sri Lankan
National Archive stored in the Dutch Nationaal Archief in the Hague. He reports that
the tombos are inconclusive at best:

> 1. The records for the period up to 1796 are damaged by water, making
> significant parts of the entries illegible. The most revealing information
> for identification here are the signatures of the parties. There is how-
> ever hardly any information of ethnic group, which makes it difficult to
> identify Malay/Indonesian and Moors given that both groups share the
> practice of adopting Arabic names. In a particularly interesting section in
> the tombos dedicated to mixed marriages (cf. Hussainmiya 1987), only
> five of 238 entries clearly refer to individuals of Javanese origin: of these,
> two records refer to Javanese-Moor marriages, one to a Javanese-Javanese
> marriage, and the remaining two are unclear.
> 2. The following period until 1919, albeit under British rule and there-
> fore less interesting for our claim, shows a more structured archiving sys-
> tem where indication of race is given. Where legible, this reveals still a
> majority of Western marriages, a growing number of marriages between
> Eurasians and Burghers (locally born of Dutch/Western heritage), and
> between Burghers. There are two clear entries involving Malays, one

---

[25][...] ihre Sprach ist Ambonesisch können aber meistentheils ach Malley- Singules- Portuges- und Hol-
ländisch [...] die Weiber, welche theils Ambonesin, Singulesin und Malabarin, dörffen nichts darzusagen.

[26]Interestingly, Hussainmiya (1990:45) affirms that "soldiers also spoke Sinhala and Tamil." How he
finds this in Schweitzer's text is not clear to me.

married to a Eurasian (between 1867-1897), and one to a Burgher (1885-1897). From 1897 onwards, race is clearly specified; of 196 entries only one is Malay. (Ansaldo 2008)

Given the little intelligence which could be gathered, Ansaldo (2008) conducted some pilot interviews with Colombo Malays, who professed not being aware of intensive Moor intermarriage in their ancestry. The time-depth, and breadth of these interviews is limited of course, but they can be complemented by native amateur genealogy. Burah (2006) has compiled a book about his genealogy, tracing his origins back to the Javanese exiles. This book contains 303 instances of marriages where both partners are mentioned. One of the partners is in the lineage of the exiled princes and is therefore Malay. The group membership of the other spouse varies. While the ethnicities of the spouses are not mentioned, their names often clearly indicate their background. In Sri Lanka, names can indicate which group the bearer belongs to. European first names are indicative of Burghers or other Christians, e.g *Coralie Nadine Caspersz* (Burgher) or *Wilson Gunaratna* (Sinhalese Christian). Sinhalese last names (like *Gunaratna*) indicate that the bearer is either a Sinhalese Buddhist or a Sinhalese Christian. Tamil last names (like *Arulambalam*) can either point to Tamil Hindus or Tamil Christians. Many general purpose Muslim names exist which are borne by Moors and Malays alike (cf. Hussainmiya 1987, Ansaldo 2008), e.g. *Fatima*, but some names are borne exclusively by Malays.[27] Cases in point are the first names *Tuan* and *Gnei* or the last names *Dole* or *Sourjah*. I am not aware of any names which are exclusively borne by Moors. Finally, there are some names in the list which cannot be classified in this schema, very often because they are not complete (e.g. *Mina*). Table 2.4 breaks down the spouses of the descendants of the exiles. A list of the marriages retrieved is found in the appendix (p. 795).

	Malay	general Muslim	Sinhalese	Tamil	Burgher	Unclear
absolute	189	73	14	3	9	15
relative	62.4%	24.1%	4.6%	1.0%	3.0%	4.6%
relative without unclear cases	65.6%	25.3%	4.8%	1.0%	3.1%	/

Table 2.4: Provenience of spouses of Malays in Burah (2006). Total number of marriages: 303

The column 'general Muslim' mentioned above must be split between Malays and Moors. How this split should be done is unclear, but we can set an upper limit to

---

[27]Hussein (2007:436) contains a list of common Malay names. Also see Hussainmiya (1987:8)

intermarriage with Moors: No more than a quarter of all contracted marriages involved a Moor. This holds true under the assumption that all instances classified as 'general Muslim' actually refer to Moors. At least 2/3 of all Malay marriages did not involve a partner of another ethnic group. The upper bound for in-group marriages is 91.3%, adding up 'Malay', 'general Muslim', and 'Unclear'. Intermarriage with the Buddhist and Christian community is present but not common; both in comparable proportion of about 4%. Intermarriage with the Tamil Hindu community is quasi non-existent. Marriage with non-Muslims is found in about 9% of the cases, but the names given to children of such unions indicate that those families followed Muslim (and indeed Malay) naming practices for their offspring.

The genealogy of the exiles treated in Burah (2006) spans about 200 years. It starts with the exiles in the 18th century and goes through various positions in the Regiment, ending in civil professions in the 20th century. It appears that the offspring of the exiles touches all important domains for Malay culture, and it is unlikely that the marriage pattern followed by the individuals surveyed here is completely unrepresentative of the community as a whole. It seems fair to assume that between 10-15% of all marriages involved a Moor partner.[28] While this reflects the privileged positions of the Moors as compared to the other groups, it is not obvious that this rate of intermarriage would have had a significant effect on the evolution of the language.

Based on Burah (2006), the extent of intercommunal marriage totals about 20%. As far as the regiment Malays are concerned, the percentage might have been even lower. The British Major Skinner "attributed the financial difficulties of Malays to continuous inter-marriage in the regiment, and the resultant growth of 'swarming connection' of soldiers" (Hussainmiya 1990:119). The use of 'intermarriage' in this context is not about marrying *outside* the regiment, but rather *inside* the regiment, i.e. with other Malays.

The fifth piece of evidence is of a rather speculative nature and compares the fate of the Malays sent to South Africa with the Sri Lanka Malays. The Malays in South Africa have completely lost their original language and culture, which is attributed to the absence of a coreligionist community such as the Moors in Lanka, which would support them. From the fact that the Malay culture has not been lost in Sri Lanka, Hussainmiya (1990:53) concludes that this must be due to the Muslim community of the Moors. While such an analysis is possible, it has to be mentioned that the disappearance of the Malay element in South Africa has surely more causes than just the presence or absence of another minority, and that even the presence of a strong Moor community in Ceylon has not prevented Malay culture from being lost in the 20th and 21st centuries. Such a monocausal explanation should be handled with care.

---

[28]This percentage seems to be valid for the present day as well. Bichsel-Stettler (1989:39) reports that among the Malays she interviewed, there had been 12 cases of intermarriage with Moors, either by them (22 consultants) or by their parents (44, if none of the consultants where siblings). The ratio is thus 12/66=18.2%. However, the pattern does not seem to be uniform: Nuhman (2007:417) reports that in 1956 and 1960 more Malays married outside their community than within. Ansaldo (2008), on the other hand, reports that "of approximately 50 families interviewed in total, only two revealed genealogies including Moor-Malay intermarriage." More demographic research is thus needed.

To sum up, we have the transfer of Islamic knowledge to some exiled princes; Dutch suspicions; dubious attested intermarriages in the *tombos*; Schweitzer's report, which casts doubt on the relation between Malay soldiers and Moors; and a speculation about South African Malays, which is, well, a speculation. This is far from being solid evidence for a close Malay-Moor relationship or linguistic influence from Tamil. Future research should be aware of the shakiness of these assumptions and take a more sceptical approach, like Ansaldo (2005b).

After discussing the points that are said to be evidence for a close relation between Malays and Moors, I would like to briefly shed light on some more facts which make pervasive intermarriage unlikely: At least in the Upcountry, many Malays still have distinctively South East Asian features, and can easily be told apart from Sinhalese, Tamils, and Moors, which look more South Asian (Hussein 2007:416). If intermarriage had been that rampant, we would expect that the South East Asian features would have been submerged by now. While South East Asian physical features are easy to be found in the Upcountry, this does not mean that all Malays in the Upcountry look like Indonesians. The spectrum stretches from prototypically Indonesian to prototypically Lankan, with all combinations in between. My point is not that no intermarriage took place, only that it was less pervasive than assumed in the literature. This is reflected in conflicting testimonials from foreigners having come to Sri Lanka at different points in time. Percival (1803:115) remarks:

> Although the Malays intermarry with the Moors and other castes particularly in Ceylon and by this means acquire a much darker colour than is natural to a Malay; still their characteristic features are so striking predominant, that they cannot be mistaken.

Over a century later, a visitor from abroad noted

> Meeting Ceylon Malays there one cannot help noticing that some of them have the features of Javanese while others look like Malays and their personal names incline to both Javanese and Malay. (Said 1926 , cited in Hussainmiya 1987:34)

Again half a century later, the former Malaysian Prime Minister Tunku Abdul Rahman remarked in the Malaysian newspaper *Star* in June 1981:

> They look more like Indians [. . .] than Malays and their language is strongly influenced by the Indian dialect. [. . .] [A]fter generations of intermarriage, it is hard to pick one from the other Malays, or the Moors, except when they themselves announce their racial identity. (Hussainmiya 1987:7)

These quotes are of course impressionistic, but show that different people had different impressions about Sri Lankan Malay appearances. If anything can be drawn

from those quotes, it is that the Malays became more Indian-like over the course of time. While in the early 19th century, their "characteristic features are so striking predominant", and in the early 20th century they "have the features of Javanese", in the late 20th century "they look more like Indians". These quotes do not support the view that heavy intermarriage was going on in the early days. However, it is most likely that these three different people each met only a tiny subset of Sri Lankan Malays, and that their reports are valid for that subset only. Hence, they do not reflect a global appreciation, and cannot be used as exact sources. As stated above, even today the Sri Lankan Malay community comprises everything from very Indian to very Indonesian, and depending on the Malay one meets, his ethnic identity is easy or impossible to ascertain. It is likely that Percival met Malays of a more Indonesian type, while Tunku met Malays of a more Indian type, but this should not lead us to hasty generalizations.

At least in Kandy, the relations between the Malays and the Sinhalese are better than with the Moors. Many Malays do not speak Tamil anymore, while this is still the language preferred by most Moors.[29] The Malays are traditionally loyal to the powers that be (Osman & Sourjah (2008:52), Saldin (2001:48)), first the colonial powers, then the Sinhalese (cf. Bichsel-Stettler 1989:35). Since the Moors were not associated with these powers, they are not held in very high regard. This former association with the colonial powers has also led the Malays to adopt a more 'Western' lifestyle (cf. Saldin 2001:41). All senior Malays asked about their youth affirmed that they used to hang out with Burghers, leading a 'happy-go-lucky'-lifestyle (Hussainmiya 1990:27), which is unusual with Moors.[30] Today, most Malays have a very liberal interpretation of Islam, while the Moors tend to be more conservative. This difference in lifestyle again makes overwhelming intermarriage unlikely. Saldin (2003:35) reports that as far as education is concerned, Malays seem to be closer to Sinhala practices.

After screening the evidence, let us now turn to the plausibility of early contact with the Moors. If we imagine the situation in the early days of the Malays in Ceylon, would contact between Malays and Moors be plausible? Would exiles be interested in meeting Islamic scholars? This cannot be excluded, but it does not seem likely that theology was one of their major preoccupations. Islamic scholarship was even less of an argument for the soldiers, convicts and slaves, whose interest in advanced theological discussion can be estimated to have been rather low. The latter groups had, on the other hand, interest in finding a Muslim wife, if they had not brought their family (Hussainmiya 1990:47). However, intermarriage to non-Muslims is already

---

[29]Some Moors speak Sinhala (Nuhman (2007:62), Hussein (2007:46)).

[30]A penchant for gambling is already noted in early reports, e.g. Schweitzer (1931[1680]:106) or Pieris (1918:131). The literary activities of the Malays may add a more intellectual overtone to this lifestyle, but this would not make them resemble the Moors more. To the contrary, Moors were known as a "backward community" (Hussainmiya 1987:95), unlikely to engage in literary pastime. Furthermore, Malays had the highest literacy among all ethnic groups in Sri Lanka, while Moors had the lowest (Marga Institute, Sri Lanka Center for Development Studies 1988:48), cited in Bichsel-Stettler (1989:17). As for the present day conflation of Malays and Moors in government statistics, Bichsel-Stettler (1989:36) remarks: "[A] comparison of the statistics given for the Moors with that for the Malays in the tables in this paper shows clearly how greatly the communities differ."

present in the early days, as shown in Burah (2006).

A further problem for contact between Malays and Moors could be that the colonial powers of the Portuguese and the Dutch tried to contain Muslim influence. In 1626, the Portuguese expelled the Moors (De Silva 1972:84ff),[31] who would settle around Batticaloa, then under Kandyan rule (Codrington 1926:113). The Dutch restricted their share of coastal trade, and forced many Moors into agricultural life in the East of Ceylon (Hussainmiya 1990:42).

The Malays, on the other hand, were always close to the centers of colonial administration (Ansaldo 2009b), either because they had to be guarded, in the case of the exiles, or because they were the guards of the colonial power in the case of the soldiers.

After the middle of the 18th century, the percentage of free Malays grew (Hussainmiya 1990:49), and free Malay could easily enter in contact with the Moors. But the Malays liked to stay close to their community,[32] and the Dutch had just pushed the Moors away from the colonial cities (Batticaloa is an exception to this, and maybe Trincomalee).[33]

Starting in 1783 the Malays had their own mosques, which continued to provide services in Malay afterwards.

> As the Malay community grew in numbers, and with the patronage extended by the colonial rulers, their religious needs were fulfilled according to Malay traditions. As the Malays claimed, special mosques for their congregations were set up in the Malay-majority townships, while they were ministered by special Malay *Khatibs* or priests (as styled by the Malays). The Malays also had their own religious *Kitabs* and legal texts written in [t]he Malay language. (Hussainmiya 1987:19)

Religion was clearly a domain where the Malays affirmed their identity. Of course, the Moors were fellow Muslims, but it seems that the mosque was not so much a place of interethnic contact but a place of affirmation of group identity (Hussainmiya 1987). Influence through religion could have taken place before 1783 or after 1930, when Tamil began to replace Malay as the language of the mosque (Bichsel-Stettler 1989:23).

2.8 Language issues

After looking at the sociohistorical development of the SLM community, it is now possible to investigate some questions of a more linguistic nature. The first such question is the origin of SLM. Other questions concern the development of the language

---

[31]Also see Vlekke (1943:72f) for the motivations of the Portuguese.

[32]Ansaldo (2009b) even argues that *endo*gamy was a defining feature of the SLM community at that point in time, rather that *exo*gamy as claimed by Smith et al. (2004) and Smith & Paauw (2006).

[33]It is interesting that the British later would refrain from stationing a Malay Regiment in Batticaloa.

in Ceylon, neighbouring languages, and the relative importance of different contact languages over time.

2.8.1   Sources in the SEAA

What are the linguistic ancestors of Sri Lanka Malay? It is clear that the Sri Lankan Malays ultimately came from South East Asia, but which languages or dialects from which islands contributed to its genesis? People from many different islands arrived in Sri Lanka as soldiers after having spent a variable period in Batavia/Jakarta before sailing off to a battlefield. During the stay in and around Batavia, this soldier population began to acquire a distinct identity, distinct from the Javanese and Sundanese of Java (Vlekke 1943:174).[34] Smith (2003:14) argues that no single mother tongue could impose itself, and the lingua franca of the SEAA, a variety of Malay, was chosen for interethnic communication.

This variety has been given various names, among which Vehicular Malay (Smith 2003), Trade Malay (Ansaldo 2005b) and Bazaar Malay need to be mentioned.[35] Like most trade languages, Vehicular/Trade Malay was not monolithic and would be adapted to the communicative needs of speakers of different ethnic groups in different ports, with different influences from native languages. The sociolinguistic situation of Batavia/Jakarta is traditionally very complex. Vlekke (1943:132) states that "in seventeenth century Batavian [. . .] [i]n some respects, the confusion of languages was like that of Babylon". Adelaar & Prentice (1996) list the following relevant varieties for the Jakarta region: *Java Malay, Java Port Malay, Jakarta Malay (Betawi), Jakartan Indonesian,* and *Tangsi Malay*. It is beyond the scope of this study to investigate the sociohistorical developments of these varieties, but see Grijns (1991) or Paauw (2004, 2008) for some discussion.

This complexity suggests that there were many input languages which participated in the formation of SLM. The different homelands of the Malays arriving have contributed to a greater or lesser extent to SLM as we know it today. Bakker (1996, 2006) claims that "it is fairly well known what Malay looked like when it was imported to the island by the Dutch". This opinion is maybe a bit optimistic given that the exact nature of the language (or languages) that the new immigrants spoke is far from clear to the remaining researchers. The total absence of written records from that period does not make the task any easier. What can be done is to look at several features of present day SLM and trace them back to Malay varieties of the SEAA. This endeavour has been undertaken by Adelaar (1991). He finds that Moluccan, Bazaar Malay, Baba Malay and Jakartanese have contributed most to SLM, besides the notwithstanding influence of Tamil and Sinhala (Table 2.5). Paauw (2004) critically reassesses the data and analysis by Adelaar and comes to somewhat different results. He finds that many of the features of SLM are found in several varieties of Vehicular Malay (VM),

---

[34]Raben (2000:99) actually supports the idea that group identity in Indonesian societies of the time was generally *not* created through ethnicity, but rather through "personal attachments and charismatic power."

[35]It should be noted that Bazaar Malay can also refer to other varieties (Paauw 2003).

so that they cannot be attributed to any single variety. Furthermore, he traces more of the features to adstrate influence from Tamil, and discards most of the Moluccan influence, as well as all of Baba and Bazaar influence.

Paauw (2004) further undertook an etymological analysis of the lexicon. His findings are presented in table 2.6. In this table, origin *Sri Lanka* means all words that joined the vocabulary after the language had reached Ceylon, and includes Tamil, Sinhalese, Portuguese, Dutch and English words. *Common Malay* indicates words understood in every part of the Malay world, whereas *Indonesian* words mostly come from Betawi or Javanese and are not understood outside Indonesia. *Malaysia*, finally, indicates words from the peninsula.

Paauw (2004) concludes that "the Malay variety most likely responsible for the lexification of Sri Lanka Malay was a variety originating in Jakarta, and that Moluccan Malay and the Malay of the Malay peninsula also contributed, in a more limited manner." Two features which Adelaar and Paauw have not addressed actually corroborate this analysis: SLM has words with schwa in the final syllable (e.g. [ʈaːnəm] 'plant'), typical of Jakarta (Adelaar 1985), and SLM drops schwa in antepenultimate position ([criːʈa] 'story' compare Std. Malay *cərita*). This is also common in Jakarta (Uri Tadmor, p.c.). Phonological influence from Malaysia is noted by Paauw (2004) in the lowering of high vowels in the last syllable as compared to Indonesian, e.g. [maːsok] 'enter', compare Indonesian *masuk*, although Paauw (2008:99) states that in Eastern contact varieties of Malay this lowering can also be found.

2.8.2 Three theories of genesis

It is a fact that Sri Lanka Malay as spoken today is strikingly different from any other Malay variety. It must have undergone radical language change in the last 350 years (Ansaldo & Nordhoff 2009). There are competing theories as to how this language change came about, and by what it was triggered. The idea defended by Hussainmiya, Smith and Paauw in their publications (Smith 2003, Smith et al. 2004, Smith & Paauw 2006) is that SLM is the product of intermarriage between Malay men and Moor wives, who spoke Tamil. The Tamil speakers would then try to acquire the language of their soldier husbands, very much like the slave populations in the Caribbean tried to acquire a European language. The language learners would transfer grammatical structures of their original language into the target language. This means that a strong Tamil substrate would remain in the variety spoken by the wives, which would then be passed on to the children, and by and by generalize to the speech of the whole community. This scenario of genesis is then very similar to the genesis scenario proposed for Sri Lanka Creole Portuguese (Smith 1979), where the indigenous Tamil population of Batticaloa tried to acquire Portuguese and imposed Tamil grammatical structures on that language. This theory of genesis fits well with the Relexification Theory (Muysken 1981, Lefebvre 2001) of Creole genesis, where we find a combination of the grammar of the substrate with the lexicon of the superstrate. In Lefebvre's account of Haitian Creole, the West-African language Gbe would be the substrate and

	Standard Malay	Sri Lanka Malay	Adelaar	Paauw
1)	*h>	h*h> Ø	Mol, Baz, Baba, Jak	general VM
2)		retroflex series	(Tamil and Sinh.)	Tamil
3)		(contrastive) vowel length	(Tamil and Sinh.)	Tamil
4)		(contrastive) consonant gemination	(Tamil and Sinh.)	Tamil
5)	m/n/ŋ/	ŋ	Mol	*Tamil
6)	ə	j/u or ə varying with j/u	Mol	*Tamil
7)	retention of most of the inherited morphology	loss of most of the inherited morphology	Mol, Baz, Baba [...]	general VM
8)	locative preposition +noun phrase	noun phrase + linker + locative preposition	Baz	Tamil
9)	noun + determiner	determiner+ noun preposition	Mol, Baz, Baba, Jak	general VM
10)	possessed + possessor	possessor + linker + possessed	Mol, Baz, Baba, elsewhere	general VM
11)	noun + adjective	adjective + noun	(Tamil and Sinh.)	Tamil
12)	prepositions	postpositions	(Tamil and Sinh.)	Tamil
13)	subject-verb-object	subject-object-verb	(Tamil and Sinh.)	Tamil
14)	ada denoting existence of a noun	ada/ərə: progressive aspect of a verb	Mol, Baz, Baba	*Tamil
15)		negators: tər/tra	Mol, Baba	Mol
16)	full tense-mood-aspect adverbials	full and reduced tense-mood-aspect adverbials	Mol, Baba	general VM
17)	plural personal pronouns are independent lexemes	plural personal pronouns are historically compound forms with *oraŋ	Mol, Baz, Baba	general VM
18)		1st and 2nd personal pronouns borrowed from Hokkien Chinese	Baz, Baba, Jak	Jakarta Malay
19)		plural marker pada	Jak	Jakarta Malay

Table 2.5: Features distinguishing Sri Lanka Malay from Standard Malay, and their origins according to Adelaar (1991) and Paauw (2004). Features which Paauw argues to be misanalyzed by Adelaar are marked with an asterisk *.

origin	percentage
Sri Lanka	10,0%
Common Malay	55,7%
Indonesia	11,7%
Malaysia	0,6%
Loan Words	21,7 %

Table 2.6: Origins of the SLM lexicon according to Paauw (2004)

French the superstrate; for SLM, Tamil would be the substrate and the language of the immigrants the superstrate.

A somewhat different position is taken by Bakker (2000a,b, 2006), who argues that the variety spoken in Sri Lanka was not heavily influenced by the local languages until the end of the 19th century, citing documents written by military officers which comply with the general Malaysian/Indonesian standard. Only after the disbandment of the regiment in 1873 would there be heavy language change toward Tamil (and possibly Sinhala), which would take place within one generation.

Ansaldo (2005b, 2008, 2009b) finally proposes that the ecology of Sri Lanka Malay has more in common with settings of metatypy as described by Ross (1996, 1997, 2007). SLM would not be the result of intermarriage and relexification, as proposed by Smith et al. Rather, SLM would be the product of the immigrants' having to handle three languages at the same time: Malay, Sinhala, and Tamil. Given that the grammatical markup of Sinhala and Tamil is very similar, and to lessen the cognitive burden of handling too many different constructions, SLM speakers would gradually emulate the Lankan structures they use everyday in Sinhala and Tamil in their own language as well. SLM would be the in-group language (*esoteric* in Ross's terms), which is influenced by the language(s) of wider communication (*exoteric*), in this case Sinhala and Tamil. This is analogous to the grammatical change of Takia (*esoteric*) towards Waskia (*exoteric*) in the Madang province in Papua New Guinea, as described by Ross (1996).

All these three theories of genesis are tied in one way or another to the historical facts, but they make different assumptions about the historical sociolinguistic setting. Some of these assumptions are explicit and some are implicit, some are supported by evidence, while others are not. In the following sections, I will discuss the socio-historical setting which is assumed for each of the three theories. This will be followed by an explanation of the mechanism which would have led to the change of the language. I will then discuss the internal plausibility of the setting and mechanism assumed, before I turn to the historical evidence presented by the supporters of the particular theory. This evidence is sociological on the one hand, and linguistic on the other. In the expository parts, I will not engage in criticism of the proposed theories. In the sections after the exposition, both linguistic and sociological evidence presented will be discussed, critically evaluated, and weighed against other pieces of evidence

not necessarily discussed by the supporters. Finally, every theory will be checked for its place in current theorizing on language contact. Does it comply with generally contact linguistic theory, or does it go against some basic assumptions and principles in the field?

For expository reasons, I will start with Bakker, then turn to the Tamil substrate hypothesis as advanced by Hussainmiya, Smith and Paauw, and discuss metatypy as proposed by Ansaldo in the end.

2.8.2.1   The rapid conversion hypothesis

This argument was the first to be published (Bakker 1995). It involves two components: SLM has converged towards Tamil (and possibly Sinhala), and this change happened rapidly. At first, this hypothesis argued that change took place between 1870 and 1906, but later versions (Bakker 2006) are more lenient as to the precise start and end dates of the rapid change. In the earlier versions of this hypothesis, convergence was the mechanism at work, while in later work, metatypy is also mentioned as a possibility. In this section, I will limit myself to the discussion of convergence. Metatypy will be discussed more extensively in Section 2.8.2.3.

Argument   Bakker first mentions Sri Lanka Malay in Bakker (1995:17f), but the discussion is too brief to really distill any claims about its genesis (besides the analogy to Sri Lankan Portuguese). In Bakker (2000a,b), he develops that

> [Sri Lanka Portuguese and Sri Lanka Malay] were documented in the early part of the last [=19th SN] century and in both cases it is clear that the languages were creolized at that point in time: they lost all person inflection, almost all derivation, and all morphological irregularities. Instead, these languages appeared to have preverbal markers for tense, mood and aspect, a system which may be considered typical for creole languages (Holm 1988). Word order is rather rigid SVO and the languages are typologically close to isolating. In short, they look like a prototypical creole. This, however, is only true for older stages of these languages.

He then goes on to argue that modern SLP and modern SLM have become more complex and have departed from that creole typology. Bakker argues that "the whole process took place in one or two generations."[36] In Bakker (2000b:607), he concludes

> In the case of SLM we must [...] assume a few decades at most for the radical changes, as earlier documents show a more ordinary form of Malay.

> Thus, 'convergence intertwining' may change a language structurally and typologically in just a few decades from isolating to agglutinative.

---

[36]also in Bakker (2006:153).

Bakker (2006:152) gives some more indications as to the sources of his claims about earlier stages of the language: "early sources (from 1806 or 1820 to the 1930s) show a form of Malay with little grammatical influence from Tamil." Furthermore, he claims that Hussainmiya (1987) contains some instances of Tamilized features in early texts, which he discounts as minor.

Bakker's conclusion is somewhat puzzling:

> The two main opinions on the matter of *when* the convergence happened both have to accept that it is a very rapid process. If it happened soon after arrival, [. . .] the convergence or metatypy must have happened within one or a few generations. Alternatively, if it happened in the first half of the twentieth century [. . .], it must also have happened within one or two generations.

The third possibility, that the change happened between arrival and the second half of the 20th century, as a continuous process, is not addressed.

Sociolinguistic setting assumed   Bakker's hypothesis makes few sociolinguistic assumptions. The main and most important assumption is that Tamil was present as an important adstrate during the whole period, while Sinhala influence is also possible (Bakker 1995:46).

Mechanism   Bakker assumes convergence, or metatypy as defined in Foley (1986) and Ross (1996, 2001).

Plausibility   If the premisses are as Bakker suggests, the developmental scenario he sketches is plausible, although not everyone will agree that language change can happen that quickly. Ansaldo (2005b, 2008) and Ansaldo & Matthews (2007), for instance, argue that rapid change is not a possibility for the restructuring of the case system we find in SLM. They argue that a restructuring of a whole functional system (like case) cannot be accommodated under the rapid conversion hypothesis and must necessarily be gradual, requiring more time than allowed for by Bakker.

Historical evidence   Bakker refers to Hussainmiya's work for the historical evidence he needs, but does not indicate what parts, chapters or pages contain that evidence. It is likely that he refers to the *Syair Kisahnya Khabar Orang Wolenter Benggali* (Hussainmiya 1987:86), from 1820.[37] The Syair should show that the Malays used a creole-like language in the beginning of the 19th century. As pointed out by Smith & Paauw (2006:178), this text is in Literary Malay, which, while sharing some features

---

[37]The nature of the earlier documents cited by Bakker, I have not been able to ascertain. Bakker (p.c.) affirms that he based his work exclusively on Hussainmiya. Hussainmiya (1987:86ff) contains a list of all early texts. None of them predates 1820.

with Creoles (no person inflection, SVO word order), is generally not considered a creole. The linguistic composition of this text can be explained by the British having brought Malay teachers from the Peninsula, who taught Literary Malay to the officers, and is surely not a result of creolization in Ceylon. Smith & Paauw (2006) argue that during the period of the Regiment, diglossia existed, with Literary Malay as the high variety and Sri Lanka Malay as the low variety, mirroring the general sociolinguistic pattern found in South Asia in general, and in Sinhala and Tamil in particular.[38] This is further corroborated by the fact that newspapers edited in Ceylon had subscribers in the SEAA (Hussainmiya 1987:89), who apparently could understand the language of the publication (Saldin 2001:24).

Bakker argues that the language emerged rapidly. He gives two possibilities for the rapid emergence: either shortly after arrival, or in the 20th century. For the 'early' hypothesis, he does not present any historical evidence. He argues that if diglossia involving 'traditional' and 'Tamilized' varieties existed in the 19th century, then the Tamilized variety must have developed very quickly, within one or two generations. It is not entirely clear to me why this would be necessary. For diglossia to be present in 1850, any time between 1650 and 1840 would be possible as a 'start date' for the Tamilized variety, and more than two generations could be encompassed. As for the 'late' hypothesis, he argues that until the middle of the 19th century, the Malays' language was very similar to other Malay varieties. He bases this claim on the existence of texts from that period which do not show Lankan features. Other texts he refers to are from the beginning of the 20th century, and show Lankan features. Thus, the language change must have happened between these two points in time. Note that for the 'early' hypothesis, Bakker assumes diglossia at 1850, while for the 'late' hypothesis, he does not. Depending on the hypothesis to defend, his interpretation of the sociohistorical setting in the 1850s is different.

Bakker seems to have an 'all-or-nothing' approach to diglossia in the 19th century. One variety must be High Malay, the other one must be either SLM as we know it today, or 'Proto-SLM'. However, the possibility of gradual development of SLM cannot be ruled out. It is possible that 'Early Low SLM', spoken during Dutch times, gradually evolved into 'Modern Low SLM', spoken today, through the intermediate stage of 'Middle Low SLM', spoken during British times. This variety would then be in a diglossic situation with High Malay, used in the British Regiment schools and the literary works. In order to argue for rapid conversion, one has to show that either 'Early Low SLM'='Middle Low SLM' with rapid conversion to 'Late Low SLM'. Or one has to show that 'Middle Low SLM'='Late Low SLM', with rapid conversion taking place before, ideally within one or two generations. No proof for either of these equations has been brought forward.

---

[38]Whether the restricted use of Literary Malays by some officers can be called 'diglossia' is open to debate, but it is certain that Literary Malay was not a language of widespread use in 19th century Ceylon.

Linguistic evidence Bakker shows that SLM shares many structures with Tamil, and also mentions similarities with Sinhala in passing. These structures back up his claim of convergence. Bakker does not use linguistic evidence to prove his claim of rapid development.

Counter-evidence Since there is no real evidence for Bakker's theory, counter-evidence to refute his claims is not necessary, strictly speaking. Counter-evidence against Bakker's claims can be divided into three domains: 1) against convergence, 2) against early rapid change, and 3) against late rapid change. As for 1), the claim that SLM grammatical structure is close to Tamil structure[39] seems pretty much uncontroversial, although some non-Tamil features remain in SLM (Slomanson 2006). As for 2), since there are no documents available, it is very difficult to make any claims about the structure of SLM preceding 1800. 3), finally, is the only domain where true counter-evidence can be found. Smith & Paauw (2006) point out that the documents from the middle of the 19th century Bakker refers to are written in Literary Malay by officers who had had schooling in that variety. They argue that at that point in time, there was a diglossic situation, with SLM as the low variety and Literary Malay as the high variety. All literary production from that period is in Literary Malay and cannot be used to make any claims about the structure of the colloquial language. Another argument against the 'late rapid change' hypothesis is the absence of strong dialectal differences in SLM. If the claim is that SLM changed after the disbandment of the regiment which entailed catastrophic changes, we would expect that these changes would only affect the speakers who witnessed them. Speakers in remote places like Hambantota (settled in 1802), with little contact to the other communities, should not change their language (cf. Smith 2003). We would thus expect a linguistic difference between the 'catastrophic communities' and the others. While there is some dialectal difference between Hambantota and the other dialects, it is mostly phonological. In any case, it does not point towards the Colombo community having had a significantly different development from the Kirinda community. If we find the same linguistic structures in Colombo and Hambantota, it makes no sense to attribute the structure of the Colombo dialect to some social changes which could never have affected Hambantota.

Place in contact language theory The mechanisms of convergence and metatypy are well-established in contact linguistics, although the speed of development Bakker assumes is not a common feature of theories of language contact, and indeed metatypy is generally argued to be a gradual process (Ross 2007:127).

2.8.2.2 The Tamil substrate hypothesis

This hypothesis is implicitly present in Hussainmiya (1990) and worked out in detail by Smith et al. (2004), Paauw (2004) and Smith & Paauw (2006). It assumes that the

---

[39]And Sinhala structure for that matter.

development of SLM was analogous to the development of Sri Lanka Portuguese, and analogous to the development of Creole languages in other parts of the world. Speakers of a substrate language (Tamil in this case) try to acquire a superstrate language (Malay), but transpose the grammatical structures of their native language onto the language to be acquired.

Argument   The Tamil substrate hypothesis states that "SLM developed as a result of intermarriage between Malay men and Moor women, vehicular Malay thus served as its lexifier and Sri Lanka Muslim Tamil (SLMT) as its principal substrate language" (Smith et al. 2004). This is explicated in more detail in the following two passages.

> From their arrival, [the Malays] associated closely with the established Tamil-speaking Sri Lanka 'Moor' community, with whom they shared the Muslim religion [. . .]. As the soldiers were single males for the most part, a high degree of intermarriage with the Moor community was inevitable. Through this interaction between the two communities, a new language arose with Tamilized structure and with a Malay lexicon. (Smith & Paauw 2006:160)

> From the earliest days in Sri Lanka, 'Malays' has a close relationship with the Tamil-speaking 'Moors' [. . .]. Intermarriage was first noted by a German observer, Christopher Schweitzer [. . .], who also commented on the multilingualism of the Malay soldiers (Hussainmiya 1987:49–50), and was confirmed by Dutch *thombos* (administrative records) of the eighteenth century (Hussainmiya 1987:52). An indigenized variety of Malay must have arisen among the children of these marriages, whose fathers spoke VM but may have even been familiar with Tamil through their interactions with the Moor community, and whose mothers spoke Tamil natively and must have attempted to learn Malays without the benefit of instruction. Tamil-speaking mothers must also have tried to make Malay the language of the home; otherwise it would be difficult to imagine the language being maintained by subsequent generations. (Smith & Paauw 2006:176f)

Sociolinguistic setting assumed   For this hypothesis, three sociological assumptions are fundamental: a) there was heavy intermarriage between Malays and Moors; b) Tamil was the substrate and Malay the superstrate, i.e. Malay had more prestige than Tamil; and c) the children of bilingual parents acquired the non-native variety of the mother, rather than the native variety of the father. Furthermore, Smith and colleagues discard Sinhala as an important factor. Why this is done is not entirely clear. Presumably it is because the recombination of the two components grammar and lexicon, which is the basis of the relexification model, is difficult to implement if the number of involved languages is not two.

Mechanism   Although they do not state so explicitly, Smith and colleagues seem to adhere to a substratist school of creole studies. The idea is that grammatical structure from a substrate is carried over into the superstrate when the latter is learned. The new language is thus a recombination of substrate grammar and superstrate lexicon. Formalizations of this approach can be found for instance in Muysken (1981) or Lefebvre (2001).

Plausibility   The substratist/relexification theory is well-tested in many contact settings, and the argument as such is plausible if the assumptions are correct. However, the statement "Tamil-speaking mothers must also have tried to make Malay the language of the home; otherwise it would be difficult to imagine the language being maintained by subsequent generations" (Smith & Paauw 2006:177) is not completely plausible, as Ansaldo (2008) notes. He says that "there [would] have been [no] plausible reason for Tamils/Sinhalese to restructure their own varieties in acquiring SLM; they were, after all, speakers of larger, socially more prestigious languages in which the SLM speakers would have been quite competent." To return to the original idea by Smith & Paauw (2006): Why should Tamil mothers try to make Malay the language of the home? What incentive could they have had? They were speakers of a majority language after all. If, as Smith et al. argue, the Tamil mothers' keeping to Tamil would have posed a problem for the explanation of the maintenance of the language, there is an easier solution than assuming that they chose Malay as a home language: the amount of Tamil mothers was much more limited than what Smith et al. assume. As such, they could not exert as important an influence. We will see in the following sections that the historical evidence suggests, indeed, that the extent of intermarriage has been overstated in the literature (Also see Section 2.7).

Sociological evidence   The historical evidence for intermarriage is said to come from Schweitzer's report and from the Dutch tombos. No evidence is presented as to the minority immigrant language being the superstrate and the local majority language the substrate. No argument is made why children would acquire the mother's non-native variety.

Linguistic evidence   Smith (2003) notes that Sinhala and Tamil are very similar in their markup, and notes left-branching word order, quotatives, case markers, and similar nominal and verbal categories. Smith notes that Hussainmiya claimed Tamil influence for SLM structures which could also have been of Sinhala influence, and that Jayasuriya claimed Sinhala influence for structures which could as well have been of Tamil influence. As a consequence, he investigates areas where Sinhala and Tamil differ, and finds that SLM is closer to Tamil in the areas he investigated, namely definiteness, number, animacy, and the accusative case. These findings are also used in later publications (Smith et al. 2004, Smith & Paauw 2006). This would support a more intensive contact with Tamil speakers.

Smith et al. (2004) argue that, in the domain of phonology, SLM has a constraint which bans retroflex stops in initial position, just like Tamil, but unlike Sinhala.

Counter-evidence   The hypothesis hinges on the assumption of heavy intermarriage with the Moor community. As argued above in Section 2.7, the evidence for intermarriage has been grossly overstated. This is not to say that there was no intermarriage. I just fundamentally call into question all the evidence which has been presented for *heavy* intermarriage, which generally does not adhere to the standards of the field. For instance, in Smith & Paauw (2006:176) we find again "intermarriage was first noted by a German observer, Christoph Schwei[t]zer." It is true that Schweitzer talked about intermarriage in this passage, but this merits a full quotation to illustrate the quality of sourcing used to defend the Moor intermarriage hypothesis:

> Their language is Ambonese, but the majority also speak Malay, Sinhala, Portuguese and Dutch. [. . .] The wives, who in part are Ambonese, in part Sinhalese and Malabarian may not say anything [against the stripping of their ornaments]. (Schweitzer 1931[1680]:106), [my translation][40]

There is no explicit mention of Moors in this passage, but Malabarians (a term referring to South Indians, including Tamils) are mentioned, however without mentioning the religion, which could be Hindu or Islam (Moor). As for the languages, we find Malay, Sinhalese, Portuguese and Dutch. Tamil is not mentioned. Now, this is only one passage from a report made by a non-linguist at one point in time. We cannot exclude that there was intermarriage with Moors, and we cannot exclude that the soldiers' wives spoke Tamil. The source quality is not sufficient for that. But even less can we claim, based on this source, that there *was* intermarriage with Moors or that the Malays did speak Tamil. This is, however, what Smith et al. do.

Another problematic case are the Dutch tombos, which are said to contain evidence for heavy intermarriage. Ansaldo (2008) checked these records and found only two cases of intermarriage for the Dutch period. This is surely not enough to support claims of heavy intermarriage. For more discussion about the poor scientific practice with regard to claims about early intermarriage, refer to Section 2.7.

The second assumption is that Tamil was the substrate and Malay the superstrate. This is clearly an attempt to apply the substratist framework to the Lankan setting, and that framework requires that the grammar come from the substrate. Hence, Tamil must be the substrate. This is a linguistic argumentation. If we square this linguistic argumentation with sociology, things get problematic. The substrate is supposed to have low prestige, the superstrate high-prestige (Arends et al. 1995:99). In the colonial setting, the most likely candidates for superstrate would have been the colonial languages Dutch and English (Ansaldo 2005b, 2008).[41] Next would be languages of

---

[40][. . .] ihre Sprach ist Ambonesisch können aber meistentheils ach Malley- Singules- Portuges- und Holländisch [. . .] die Weiber, welche theils Ambonesin, Singulesin und Malabarin, dörffen nichts darzusagen.

[41]See Saldin (2001:25) for a similar account.

law, religion, commerce, and art. Law would probably be a European language. Religion would be Tamil for the Muslims. Commerce and art would be Tamil or Sinhala. The only sociolinguistically important domain from where Malay could have claimed prestige would be the military. But even there, the colonial languages were present, and used in the higher ranks, so that there is no domain in which Malay would have been the dominant language (Ansaldo 2008). It is therefore difficult to conceive how the Tamils could have interpreted the Malay language as superior, hence superstratal, and a worthwhile target language to acquire. This applies even more since the Tamil language itself enjoyed high prestige, being the language of the rulers.

> Kandy was ruled in its last years by a dynasty of Tamil-speaking kings, the Nayakkaras from Madurai in south India. [. . .] [S]ome of the Kandyan chiefs who were signatories to the Kandyan convention (drawn up when the kingdom fell to the British) signed their names in Tamil script rather than Sinhala. (Nissan & Stirrat 1990:23f)

All this shows that speakers of Tamil had little reason to see their language as devoid of prestige (i.e. substratal) as compared to Malay.[42] Ansaldo (2008) furthermore criticizes the use of the term 'superstrate' for the language spoken by the immigrants, as this was a lingua franca with very low prestige to begin with, which does not easily lend itself to superstrate function.

The third assumption is that in bilingual marriages, the children learn the imperfect variety of the husband's language as spoken by the wife, rather than the fluent language of the wife and/or the fluent language of the husband. This would presuppose that the Tamil mothers spoke 'broken' Malay to their children, rather than Tamil. The input in 'broken Malay' by Moor mothers would have had to be more important than the input in 'good Malay' by the husbands and other fluent speakers in the extended family. This is not very convincing as a scenario.

Linguistic counter-evidence   Smith (2003) identified four areas where SLM was closer to Tamil than to Sinhala: definiteness, number, animacy, and the accusative. Sinhala has obligatory marking of definiteness, while Tamil does not. Smith claims that SLM aligns with Tamil here, but Sections 6.4.1 (p. 319) and 10.4.1 (p. 472) of this grammar clearly show that indefiniteness marking is obligatory, and thus that SLM is closer to Sinhala here than to Tamil. As for number marking, this is obligatory in Sinhala, and optional on Tamil inanimates (Lehmann 1989:20). Smith claims that SLM aligns with Tamil in this domain, but it turns out that number marking is also optional with animates in SLM (Section 15.6.1.1, p. 639), so that this seems to be a retention of the historic liberal marking of number of Malay varieties rather than

---

[42]The prestige situation may have been different in the genesis of Sri Lanka Portuguese. Portuguese as the language of a colonial power is more likely to have enjoyed high prestige. The Malays however were not in the same position of power as the Portuguese, and it seems clearly inappropriate to equate the sociolinguistic position of the Portuguese with the sociolinguistic position of the Malays.

copying of the Tamil pattern. As far as the interplay of accusative and animacy is concerned, Smith claims that SLM aligns with Tamil in this domain. In this grammar, it is found that SLM takes a middle position between Sinhala and Tamil as far as accusative marking is concerned. Sinhala has heavy restrictions on accusative marking, Tamil makes more use of the accusative in clearly demarcated areas, and SLM is very permissive. SLM can use the accusative in areas where Sinhala cannot (inanimates), but accusative marking is hardly ever obligatory. This is different from what is found in Tamil, where definite objects are obligatorily marked for accusative. See Section 6.4.4.1 (p. 328) for a discussion of the accusative in SLM.

As far as phonological evidence is concerned, the Tamil constraint against retroflex onsets claimed by Smith is not found as such in SLM. There are both dental onsets ([d̪aːt̪aŋ] 'come') and retroflex onsets ([ɖaːpur] 'oven') in SLM, which would point to Sinhala influence in phonology rather than Tamil influence. Actually, voiced dental onsets are a lot less common than voiced retroflex onsets (9 against 42 in Saldin's (2007) dictionary). Furthermore, Tamil also bans voiced stops from absolute onsets, which is not what we find in SLM. The most salient feature of Tamil phonology, the complete absence of voicing distinctions, has not had any repercussions on SLM phonology, where all traditional voicing contrasts are maintained. If there was indeed heavy Tamil L2-learner influence in SLM, we would expect initial devoicing, devoicing of geminates, and voicing of singleton voiceless stops. None of this is found. If we leave the domain of constraints and turn to segmental phonology, it turns out that the phoneme inventories of SLM and Sinhala are nearly identical, while Tamil is divergent in having no voiced stops, no prenasalized stops, no schwa, but showing retroflex sonorants, which are not found in SLM or Sinhala (Table 2.7).

To sum up, in the domain of morphosyntax we have Sinhala influence in indefiniteness marking, retention of traditional optional number marking, and a muddled accusative system, which is somewhere between Sinhala and Tamil. In phonology, the phoneme inventories do not suggest Tamil influence, and the purported constraints on the occurrence of initial retroflexes do not hold either. In short, there is no clear sign of exclusive Tamil influence.

Place in language contact theory   This hypothesis is well inscribed in substratist theories of creole genesis, although the application of the terms 'substrate' and 'superstrate' to the Sri Lankan setting is problematic (Ansaldo 2008).

Furthermore, rather than the standard split grammar/lexicon, SLM seems to have undergone a more fine-grained mixture of domains (Ansaldo 2008). Parts of verbal morphology are clearly Malay (Slomanson 2006), other parts of the grammar are clearly Lankan, and phonology is a mix of Sinhala and Malay features (see relevant chapters in this book).

	labial	dental	alveolar	retroflex	palatal	velar	glottal
stops							
voiceless	p	t̪		ʈ	c	k	
voiced	b	d̪		ʈ	ɟ	g	
prenasalized	^{m}b			^{n}ɖ	^{n}ɟ	^{ŋ}g	
nasals	m		n		ɲ	ŋ	
fricatives	(f)		(s)(z) (ʃ)				(h)
approximants	ʋ				j		
liquids			r, l				

(a) SLM consonant phonemes

	labial	dental	alveolar	retroflex	palatal	velar	glottal
stops							
voiceless	p	t̪		ʈ	c	k	
voiced	b	d̪		ɖ	ɟ	g	
prenasalized	^{m}b	^{n}d̪		^{n}ɖ		^{ŋ}g	
nasals	m		n		ɲ	ŋ	
fricatives	(f)		(s)(z) (ʃ)				(h)
approximants	ʋ				j		
liquids			r, l				

(b) Sinhala consonant phonemes based on and adapted from Gair (2003), Karunatillake (2004)

	labial	dental	alveolar	retroflex	palatal	velar	glottal
stops	p	t̪	t̲	ʈ	c(ɟ)	k	
nasals	m		n	ɳ	[ɲ]		
fricatives	(f)		[s](z) (ʃ)				(h)
approximants	ʋ				j		
liquids			r, l	ɭ			

(c) Sri Lankan Tamil consonant phonemes based on and adapted from Zvelebil (1966), Suseendirarajah (1973a), Selvarajagopal (1979), Lehmann (1989), Schiffman (1999).

Table 2.7: Consonant inventories in the adstrates. Parentheses indicate phonemes only found in loanwords. Brackets indicate positional variants which are sometimes included as phonemes.

Inventories of SLM and Sinhala are quite similar, while Tamil is divergent. These charts use traditional notation; in the phonology section of this grammar (p. 92), I propose a reanalysis of the retroflexes.

2.8.2.3   The metatypy hypothesis

This hypothesis is a variation of the rapid convergence hypothesis discussed above, with the modification that no claims about the speed of development are made, and that Sinhala and Tamil are both accepted as possible sources of grammatical features. Ansaldo (2005b) contains the first formulation of that hypothesis, while the term 'metatypy' is first used in Bakker (2006). The idea is fleshed out in more detail in Ansaldo (2008, 2009b) and Ansaldo & Nordhoff (submitted). This hypothesis states that Sinhala and Tamil were adstrates, which both exerted pressure on SLM. The Malays had to be conversant in these two languages of wider communication, which led to the copying of the Lankan structures into Malay grammar.

Argument   Malay was an esoteric language for the immigrants, in the sense of Ross (1996), while Sinhala and Tamil were exoteric languages, i.e. languages of wider communication. Sinhala was used to interact with the Sinhalese population, which, at that point in time in the relevant regions would have made up about 80% of the population. Tamil was used to interact with the rest of the population, but was used more frequently than the pure numbers would suggest because of religious affinities. The grammatical structures of Sinhala and Tamil are very similar.[43] Their typological profiles are close to identical, as noted by Smith (2003). The Malay language was not a standardized or *focused* variety in the sense of LePage & Tabouret-Keller (1985). Rather, it was a *diffuse* variety, i.e. a variety where the actual realization of an utterance contained little sociological information as to membership of a linguistic group. The immigrants were all from different regions, had different mother tongues, and did not attach much importance to the trade language they used to communicate with each other.[44] As such, this trade language was very variable, and very malleable. This malleable language was then confronted with two adstrates of similar typological markup, which imposed that markup on the immigrants' language. However, while grammar was malleable, this was not the case of the lexicon, which remained by and large Malayic. It can be speculated that the grammar of that language at that point in time was diffuse, with little 'acts of identity' (LePage & Tabouret-Keller 1985) associated with it, while it was the lexicon that served as an emblem of group membership.

Sociolinguistic setting assumed   This hypothesis assumes that Malay was an esoteric language and at least one Lankan language was a relevant exoteric language (Tamil in the case of Bakker 1995, 1996, 2000a,b, 2006), although both Sinhala and Tamil would also be possible (Ansaldo 2005b, 2008). It furthermore assumes that the esoteric language was a diffuse variety.

---

[43] Actually, the membership of Sinhala in the Indo-Aryan language family (rather than Dravidian) has occasionally been disputed (Lassen 1847, Tennent 1859). See Geiger (1938) for a discussion.

[44] This does not preclude focusing of the language at a later point in time.

Mechanism   The mechanism leading to the present state of SLM is argued to be metatypy as described by Ross (1996, 1997, 2007). Bilingual speakers have a tendency to lessen the cognitive load which different grammars impose on them (Nadkarni 1975, Ross 2007). In the Lankan context, it seems plausible that the soldiers carried over some Lankan constructions into their Malay speech. Since everybody had some command of the local languages, the speaker could be sure that the hearer would be able to decode the message as intended, even if a Lankan construction was used instead of a Malay one. Furthermore, there was no fear of reprehension for the use of a foreign construction in Malay because of the diffuse status of the code, where no one felt the need to oversee linguistic usage. By and by, the use of Lankan constructions made inroads into the grammar up to the state of the language we find today.

Plausibility   The Lankan ecology very much resembles other settings where metatypy has been found, so that this hypothesis is plausible.

Sociological evidence   The esoteric nature of the Malay language can be derived from the confined dwellings in the barracks, with very tightly-knit networks, as described by (Hussainmiya 1990). The exoteric nature of Tamil as language of wider communication is obvious and agreed upon by all authors. The relevance of Sinhala can be derived from the high Sinhalese percentage in the population (80%). The diffuse nature of the immigrants' language is also uncontroversial (Smith & Paauw 2006:163).

Linguistic evidence   The linguistic evidence for this hypothesis is indirect. SLM exhibits both Tamil and Sinhalese structures, in about the same proportions (Ansaldo & Nordhoff 2006, submitted). This suggests that both languages had a comparable status in the genesis of SLM. The most obvious mechanism to accommodate this is to postulate that they were adstrates on a par with each other, which fits the metatypy hypothesis. Other explanations might be possible, but would require a more elaborate apparatus and require more coincidences (e.g. first Tamil influence, then subsequent Sinhala influence of the same importance in the same areas).

Counter-evidence   Evidence against metatypy could be found in influences which can only be attributed to one language, and not the other. This is the case for lexical influence from Tamil and phonological influence from Sinhala. There are many Tamil loanwords for basic vocabulary terms, like *kattil* 'bed', as well as many animal names (*vanaati* 'butterfly', *vavval* 'bat'). On the other hand, there are close to no integrated Sinhala loans. The Tamil loans are phonologically integrated in the SLM system, and also used by speakers who do not know Tamil. The Sinhalese borrowings seem to be nonce borrowings, and the speakers are aware of the code-switch. This difference in lexical influence suggests that at a certain point in time, which was decisive for the vocabulary, Tamil was more important than Sinhala. This point in time was most

likely very early, when the Malays had to find new words for unfamiliar concepts (although the concepts 'bed', 'butterfly' and 'bat' would have existed in South East Asia as well).

Sinhala seems to have had a greater influence on SLM phonology than Tamil, which might have to do with the typologically uncommon nature of Sri Lankan Tamil phonology (three coronal stops, retroflex nasal and lateral, no voicing distinction). While this shows that Tamil and Sinhala did not have precisely the same function in the genesis of SLM, it does not pose a problem for the metatypy hypothesis because metatypy is only concerned with grammatical structure, and not with the lexicon or phonology (Ross 2007).

Place in language contact theory   Metatypy is a well-described mechanism of language change in a contact setting. For the Sri Lankan context, it has recently (Ansaldo 2009a,b) been formalized along evolutionary ideas following the ideas of Croft (2000) or Mufwene (2001). It has furthermore been checked against empirical data (Ansaldo & Nordhoff submitted). Where the Sri Lankan context differs from more classical cases of metatypy is that the number of languages is greater than 2, a problem shared with the relexification approach (see above). But within the relexification approach, the necessity for two languages is a built-in feature, while for metatypy, it appears to be accidental. The shift in grammatical patterns, which is typical for metatypy, occurs towards the language of wider communication. There is nothing fundamental in the theory of metatypy to preclude change towards two typologically very similar languages of similar use in wider communication at the same time, in this case Sinhala and Tamil. The fact that this theoretical possibility has not been observed earlier is probably due to the very low chance of encountering a trilingual setting where the power relations are 'just right' to permit this kind of change.

Metatypy can be seen as the result of speakers adapting their linguistic behaviour in a bilingual setting in order to lessen their cognitive load (Nadkarni 1975, Ross 2007). In a 1:1 setting, this reduction of cognitive burden could go either way, but in a 2:1 setting, as in Sri Lanka, there is a clear preference for a 'winner-take-all' result, i.e. the development of a 3:0 pattern as the result of competition and selection among constructions (Ansaldo 2005b). The social and cognitive mechanisms at work are beyond the scope of this thesis, but are explored in detail in Ansaldo (2009b) and applied in Ansaldo & Nordhoff (submitted).

2.8.3   The linguistic ecology

In Section 2.8.1 we have seen the diverse influences from the South East Asian varieties in different periods of history. In this section, I will briefly sketch the present linguistic ecology of Sri Lanka to shed light on which languages the Sri Lankan Malays encounter in their daily life.

The majority language of the island is Sinhala. Sinhala is a diglossic language (Gair 1968, 1985b, Paolillo 1992, 1997), with important differences between the high

and the low variety. These differences cross-cut the spoken/written distinction, i.e. it is possible to combine high or low with spoken or written expression (De Silva 1974, Gair 1985b, Paolillo 1997). The high variety is described by Gunasekara (1891), Geiger (1938, 1973), and Gair & Karunatillake (1974), while the low variety is described by Chater (1815), Fairbanks et al. (1968), Matzel (1983), Gair & Paolillo (1997), Dissanayaka (2003), Jayawardena-Moser (2004) and Karunatillake (2004). Linguistic analyses of various topics of Colloquial Sinhala are compiled in Gair (1998) and Henadeerage (2002). The low variety is further divided between 'Upcountry' and 'Lowcountry' dialects, but the two are mutually comprehensible. The further dialectal subdivision is unclear and underresearched (Jayawardena-Moser 1996:xii). The descriptive material available for Colloquial Sinhala is good and covers the speech of the groups the Malays have contact with.

The next biggest language of Sri Lanka is Tamil. While there is abundant literature on Tamil, it very often does not address the varieties relevant in the Sri Lankan context. Like Sinhala, Tamil is a diglossic language, with Literary Tamil being distinguished from diverse forms of Colloquial Tamil (Annamalai & Steever 1998). Colloquial Tamil can be divided into Indian Tamil and Sri Lankan Tamil. Spoken Indian Tamil is well described (Pope 1867, Arden 1934, Beythan 1943, Asher 1982, Lehmann 1989, Schiffman 1999, Asher & Annamalai 2002, Annamalai & Steever 1998), but differs from Sri Lankan varieties. Indian Tamil is present in Sri Lanka via media (television, films), and via the Estate Tamils, who came during the British period to work on the tea plantations. 'Media Tamil' is quite homogeneous, but Estate Tamil is likely to still reflect the dialectal differences of the Indian immigrants from different regions of Southern India. The Malays do watch Tamil movies, and were in contact with Estate Tamils when they were overseers. This means that these two varieties of Indian Tamil can be considered as contact languages.

Dialectal variation in Tamil is underresearched (Shanmugam Pillai ca. 1986), especially in what concerns Sri Lanka. Zvelebil (1960, 1966) distinguishes Jaffna Tamil, Trincomalee Tamil, Batticaloa Tamil and General Ceylon Tamil on the basis of the phonology and morphology of a handful of speakers. Suseendirarajah (1967, 1973a) adds some more description of the Tamil of Jaffna. The dialects of Jaffna and Trincomalee are not important for the analysis of SLM since the Malays had little contact with those towns.[45] 'General Ceylon Tamil' is more promising, as is Batticaloa Tamil (Zvelebil 1966, Suseendirarajah 1973b, Selvarajagopal 1979), with an important percentage of Muslim speakers. With the exception of Batticaloa, all these descriptions cover Hindu/Christian varieties. As for dialects on the other side of the island, Bonta (2004, 2008) argues that varieties of Negombo Tamil have converged towards Sinhala and have lost verb agreement.

The Moors have to be divided on sociological grounds into 'Coast Moors' and 'Ceylon Moors', of which the Ceylon Moors speak a Lankan variety while the Coast Moors, hailing from India, speak an Indian form of Tamil (Nuhman 2007:25), also

---

[45]There are some Malays in Trincomalee, but it is unclear whether they still speak Malay (Hussein 2007:426).

discussed in Hussein (2007). Nuhman (2007:71ff) distinguishes Sri Lanka Muslim Tamil from the Hindu/Christian varieties. He further divides North-Eastern Muslim Tamil from South-Western Muslim Tamil (also see Hussein 2007:45) and argues that verb agreement was lost in the South-Western variety as well. He furthermore gives an overview of some additional phonological, morphological and lexical characteristics of Muslim Tamil. To sum up, overall description of Tamil is very good, description of Sri Lankan Hindu/Christian Tamil is fair, and description of Muslim varieties is seriously wanting, as is a description of Estate Tamil varieties. Unfortunately, it is the Muslim and the Estate varieties which are especially interesting as contact languages for the Malays.

Another important language is, of course, English, in its Sri Lankan shape (Kandiah 1996), while the oldest inhabitants of Sri Lanka, the Veddah, and their language (De Silva 1972), have not been in extensive contact with the Malays.

2.8.4   Heydays

Different languages had different impact in different periods on the formation of SLM, as is common in contact language formation (Arends 1989, Roberts 2004, Lim 2007). To go by chronological order, Portuguese was the first language to have the possibility to make an impact, although there is little evidence for that. The only item which could point to a Portuguese influence specific to Sri Lanka is *salba* 'escape', which does not seem to exist in the SEAA. There are numerous other Portuguese words in SLM, but these are either also found in the SEAA (e.g. *greeja/gereja* 'church') or in Sri Lanka (*kusini* 'kitchen', Sinh. *kussiya,* Tam. *kusini*), so that it cannot be established that they were adopted in Sri Lanka during Portuguese rule. They might have entered the language later, through new immigrants bringing words like *greeja* from the SEAA, or through adstrate influence bringing words like *kusini*. After the Portuguese were ousted, their influence ceased. Next came the Dutch, who also failed to have a huge impact on the language. No Dutch word seems to be exclusive to SLM; they are all shared with other varieties of Malay and the local languages (like *kaamar* 'room', Std. Malay *kamar*, Sinh. *kaamaraya*). During Dutch rule, many immigrants came from Java and other islands of the SEAA. There is a strong trace of them in SLM, with 11,7% of the vocabulary being traceable to them exclusively (Paauw 2004). The influx from island immigrants ceased in 1819, and the immigration from peninsular Malays grew stronger. This new wave of immigrants, however, did not have an important impact. Only 0,6% of SLM vocabulary can be traced to them (Paauw 2004). During the period of the regiment, Literary Malay was taught in the regiment schools, so this language could have exerted influence on the spoken variety, although it is unclear whether this indeed materialized. When the regiment schools were closed, Malays switched to English education, and would cling to this medium of instruction until independence. English was probably already a major factor, but schooling in English increased the impact of this language from the late 19th century onwards. After Sinhala became the medium of instruction in government schools

following independence, Sinhala increased its influence on SLM, and continues to do so up to today, due to the economic necessity to command this language.

As for Tamil, it is likely that it had an influence in the early days because many basic terms like *kattil* 'bed' are borrowed from this language, as are many names for animals. It is likely that the variety of Tamil was Muslim Tamil due to the shared religion; there is close to no evidence for prolonged contact with Hindu or Christian Tamils. The relative strength of Muslim Tamil influence over time is difficult to assess. In the beginning, when there were no Malay mosques, it was probably stronger, but as soon as Malay mosques were established and Malay life evolved around the military cantonments, Moor influence is likely to have decreased. The growing association with the Burghers and the adoption of a more intellectual and Westernized lifestyle in the 19th century suggest that the cultural ties between Malays and Moors became weaker in that period than what they had been before. But at the same time, the beginning of the estate economy, which employed many Malays as overseers over South Indian plantation workers, meant that the Malays had once again a good motivation to be conversant in Tamil. While the prestige of Estate Tamil was of course not very high, the relative size of labourers against overseers might have made up for that. As of today, the influence of Estate Tamil on Sri Lanka Malay (at least the Upcountry variety) remains understudied. With the orientation of Malays towards service economy, the contact with Estate Tamil decreased.

To sum up, most linguistic influence came, peaked and possibly disappeared. This is the case for the colonial languages as well as the languages of the immigrants, Estate Tamil and Sinhala (which came to stay). An exception to this 'peak' pattern is Muslim Tamil, which was probably present during the whole period in unclear proportions, possibly with a dip during the period of the regiment.

2.8.5 Envoi: Against reductionism

From the above descriptions, it should be clear that the linguistic history and ecology of Sri Lanka Malay are very complex. Sri Lanka Malay is definitely not a simple combination of Malay and Tamil and/or Sinhala. While this characterization could be used as a shorthand, it oversimplifies the diachronic complexities. To start with the immigrants, they surely spoke more than one language, and 'Malay' is probably a misnomer to designate most of their varieties, if used without a qualification (Moluccan Malay, Jakartan Malay, Tangsi Malay etc.). It is not clear whether we can speak of 'the immigrants' language' at all, given that the soldiers hailed from very diverse parts of the South East Asian Archipelago. Maybe there was a language which all of them shared, but it is more probable that that was not the case.[46] It is likely that different kinds of accommodations were used on the part of a speaker to facilitate communication with addressees from different ethnolinguistic backgrounds, and that the

---

[46]It should be borne in mind that Bahasa Indonesia, the national language of Indonesia, is a 20th century creation and is not spoken natively by a majority of the population (Steinhauer 2005).

'Malay lingua franca' was perpetually reinvented, very much like its Mediterranean counterpart (Selbach 2007, in preparation).

In Sri Lanka, Sinhala can be seen as a relatively homogeneous language. This is not the case for Tamil, where Hindu/Christian varieties and Muslim varieties differ considerably (Nuhman 2007), and the especially interesting Muslim varieties show diatopic variation. Estate Tamil also differs from estate to estate depending on the origin of the workers. Just as we must ask 'Which Malay?' when talking about the immigrants' language, we must ask 'Which Tamil?' and 'Which Sinhala?' when speaking about the language of the people they met in Sri Lanka. Reconstruction of historical contact settings is notoriously difficult (Renfrew et al. 2000, Lim 2007, Ansaldo 2009a), but it is probable that the answers to these questions will not be 'that Tamil' but rather 'these Tamil varieties, those Malay varieties, those other Sinhala varieties'. Only when we have a full and precise picture of the input varieties, and the end product, will we be able to make solid analyses of the evolution of Sri Lanka Malay.

Diachronic language contact studies investigate the development of a new variety out of the contact between two or more prior varieties. The more information we have on the historical languages and the present day language, the more precisely we can trace the development of the language. As for Sri Lanka Malay, we have little information about the historical varieties, and not too much information about the current variety. While the Sinhala part seems OK, we need a better description of historical Trade Malay varieties,[47] historical Muslim Tamil varieties, and present day Sri Lanka Malay to make solid analyses of the development of Sri Lanka Malay. I will not be able to shed much light on the historical varieties in this thesis, but in the following chapters, I hope to advance our knowledge of the current shape of Sri Lanka Malay and add the contemporary piece to the diachronic jigsaw puzzle.

---

[47]After writing this chapter, I received a copy of Paauw (2008), which is a big step forward into that direction.

Chapter 3

Methodology

3.1 Theoretical assumptions

The presentation in this grammar is intended to be as theory-neutral as possible. No knowledge of any particular theoretical framework is assumed in order to keep the book accessible for scholars and laymen alike.

Being theory-neutral does not mean that this grammar is atheoretical. There are many decisions in the presentation of the data that were guided by theoretical considerations. Structuring data is impossible without a solid theory that lays the foundation on which the grammatical findings can be grounded. But in a house, it is not necessary to see plans of the foundation in every room, and in a grammar book it is not necessary to cast everything into a powerful yet complicated formalism when it can also be said in simple, theory-neutral words.

One fundamental design principle of this grammatical description is the separation of formal and functional descriptions. In the formal part (part II of this book), I describe linguistic forms (morphemes, constructions) and indicate the functions they can fulfil (indicate recipient, situate event in the past). In the functional part (part III of this book), I take the converse approach and look at functions and how they can be encoded by morphemes and constructions.

The formal part of the grammar is inspired by Radical Construction Grammar (Croft 2001) and its separation of linguistic forms along the axes of atomic vs. complex and specific vs. schematic. Atomic schematic constructions are word classes; atomic specific constructions are morphemes; complex schematic constructions are NPs, VPs, clauses, among others; and complex specific constructions are idioms. These different types are discussed in the individual chapters of the formal part. It should be noted that phonology is separate from this approach, and is treated as a precondition for it.

The functional part of the grammar is structured along the levels found in Functional Discourse Grammar (Hengeveld & Mackenzie 2008), although some levels have

been collapsed for easier navigation. Propositional content is separated from information structure and speech acts. These three levels are discussed in separate chapters in the functional part.

Two chapters are in an intermediate position between a formal and a functional description. These are the chapters on grammatical relations and valency, which are inspired by Role and Reference Grammar (Van Valin & Foley 1997).

Care has been taken to avoid theory-specific jargon and formalisms as much as possible. Normally, standard descriptive nomenclature has been used. Every nomenclature in itself relies on theoretical considerations; the nomenclature here is based on Basic Linguistic Theory (Dixon 1997, Dryer 2006), which should be accessible to every reader with a basic knowledge of typology.

3.2   Data collection

The data for this grammar was collected in three field trips to Sri Lanka in 2005, 2006 and 2007. Additionally, two consultants visited the Netherlands in 2008. The first two trips took two and a half months each, while the 2007 trip was two weeks only. The 2008 visit was three weeks. Between the trips, I was in contact with the consultants vie email and telephone. During the trips, the focus was on data gathering, while the periods between them where devoted to analysis. The findings were cross-checked during the following field-trip, or by email. Focus was on natural discourse, monologues as well as conversation. In the first field trip, I did quite a lot of elicitation to get a feeling for the possible constructions in the language so that I could identify them when I heard/transcribed them. The phonological system had already been analyzed in its basics (Bichsel-Stettler 1989), so it was possible to devise a practical orthography quite quickly and proceed to morphology and syntax. It turned out that my main consultants had complementary strengths: Mr Cassiere provided a lot of narratives and cultural information, while the Hamit and Salim families helped me with grammatical analysis. Text production was stimulated using my basic command of Sri Lanka Malay as far as possible. This was done in order to reduce the possibility of interference from English through code switches. Grammatical analysis was done with English as a meta-language. Occasionally, Sinhala sentences were used to get at meanings which are difficult to render in English, e.g. evidentials. The Malays all have a good command of English, so English was also used in conversation, not only for difficult topics, but also when the speakers just felt like it. In order to maintain a relaxed and casual atmosphere, I did not interfere with that. Maybe some day in the future a researcher will want to do research on code-switching and will be happy to have this material as it is.

Between the first and the second trips, I transcribed the data in the Netherlands and collected interesting grammatical phenomena to check during the second trip. During my second trip, I taught transcription to Mr Izvan Salim, who then transcribed all new recordings, and also the recordings I had transcribed before myself. The great majority

`nar`	narrative
`cvs`	conversation
`sng`	song
`rec/rcp`	recipe
`wrt`	written text
`prs`	presentation
`eml`	email
`eli`	elicitation
`mix`	mix

Table 3.1: Genre codes used in text examples

of the example sentences in this book are taken from the phonemic transcriptions he made, with minor adjustments for typos and for more complicated aspects of the orthography, especially the use of ⟨nny⟩ and ⟨nng⟩, and spacing. The third, short, trip had as the main goal to restart the transcription process, which had come to a standstill in the meantime. This worked out fine. Due to the worsening security situation in Sri Lanka, the field trip in 2008 was canceled, and two of the main consultants came to the Netherlands instead, where the translations of the texts were perfected and remaining detailed grammatical questions were dealt with.

Besides elicitation, transcription and translation, I visited many Malay households in Sri Lanka, mainly in Kandy, but also in Gampola, Mawanella, Ulapane, Nawalapitiya, Badulla, Colombo, and Galle, and conducted interviews, mainly about family situation and history. The total number of Malays interviewed should be somewhere around 50. Impressionistically, the dialectal differences are slim, and I had no problems understanding the speakers of other towns. Note that I have not been to the South (Hambantota and Kirinda), where the dialect seems to be a bit different (Peter Slomanson, p.c.).

The examples in this book are mainly taken from natural discourse. For every example, its source is indicated. The source file follows certain conventions that permit to identify the text genre. The first letter indicates the town and the following numbers indicate the date (YYMMDD). The three following letters indicate the genre, for instance `nar` for narrative or `eli` for elicitation (see Table 3.1 for a list). The transcriptions as well as the media files are available on the Internet.[1] The primacy of natural discourse over elicitation means that there are not many examples that stem from an `eli`-file. These are mainly for negative evidence, which never occurs in normal discourse, and marginal phenomena that could not be found in natural discourse but are generally considered of linguistic interest. A third reason to use elicited examples was the desire to keep the context constant, e.g. in the elicitation of paradigms.

The recordings are transcribed phonemically. This means that phonological per-

---

[1] `http://corpus1.mpi.nl/ds/imdi_browser/`

formance factors (e.g. allegro forms) are eliminated. For instance, the dative marker is transcribed as =*nang* most of the time, even if phonetically it might be closer to [nə̃]. In contradistinction to phonology, the transcriber Mr Salim did not alter morphology or syntax of the recordings. He would transcribe every morpheme he heard, even if he thought that it was incorrect or misplaced. Likewise, he would not add morphemes which were not present in the recording, even if he thought that they would have been necessary. In case of doubt, it was decided to stick with the recorded data, even if it might contain performance errors. Performance errors are unavoidable in natural speech and metalinguistic elicitation alike, as Birdsong (1989:72) reminds us:[2]

> The hypocrisy of rejecting linguistic performance data as too noisy to study, while embracing metalinguistic performance data as proper input to theory, should be apparent to any thoughtful linguist. (Birdsong 1989:72)

3.3 Data gathering techniques

3.3.1 Elicitation

The Sri Lankan context has some advantages and some drawbacks for elicitation. The advantage is the familiarity of the speakers with the English language, so that the choice of metalanguage is easy, and very often subtle differences can be made clear. Another advantage is the multilingualism of the speakers, which allows switching to another language to exemplify the intended meaning. The major drawback is again multilingualism, which leads, at least in the case of Sri Lanka Malay, to a 'diffuse' grammar, which can accommodate a lot of variation. It is common that consultants repeat the word order of the sentence in the metalanguage in their translation. If asked in English, the answer will be SVO, if asked in Sinhala, the answer will be SOV. This has happened more than once, and was in fact used to test informants for consistency. Another drawback (for this task at least) is the social climate, something which could be approximated by a term like 'politeness'. Consultants, like any other person you meet, want the meeting to be enjoyable, and one way of making it enjoyable is giving the linguist the things he wants to hear, or at least not refusing him the things he longs for that much. The elicitor's wits, charm, subterfuge and creativity (or lack thereof) in making up contexts determine a great deal more the grammaticality judgment of a given sentence than does the consultant's intuition.

One thing that surely triggers politeness and accommodation is the famous sentence *Can you say X?* and its variations like *And you can't say Y, can you?* and *So you can also say Y?* In a diffuse multilingual setting like Sri Lanka, and under the absence of prescriptive schooling in Sri Lanka Malay, speakers are used to a lot of variation in their input and will only reject a sentence if it seriously hinders comprehension.

---

[2]See Schütze (1996) for a critical discussion of the reliability of different data gathering procedures in linguistics.

The approach taken to data collection in this grammar was to elicit grammaticality judgements in the role of a second language learner. These judgments helped shape my idea of the language and permitted me to understand more and more of the transcribed texts. As it happens, most of the elicited phenomena could also be found in the texts, and these naturally occurring examples are then used as illustration rather than the elicited sentences. In the remaining cases, the methods involving less target language production by the linguist were preferred over others. The normal way of prompting informants to produce a certain sentence was to establish a certain context that would logically lead to an utterance with the desired meaning. When checking for identificational constructions, for instance, the following context was presented:

(1) Suppose you come home and you see that someone has broken in and stolen the TV, some money and so on. You do not know who that was. But later you learn that Farook was the thief. How would you then tell a friend that Farook was the thief?

Normally consultants would not just give the last sentence, but they would actually repeat the story with the final climax that Farook was found to be the thief. This sentence is *Farook=jo maaling*, with the emphatic clitic *=jo* indicating the identification. Given that there was a whole story, the consultants would not really know what I was looking for, and accommodation effects should be reduced.

Additionally, the existence of other possibilities was also asked for (*Could you also say something else?*). If the construction I was looking for still did not show up, I would try to modify the sentence more and more, trying to hide my intention as much as possible.

(2) a. Can you start with another word?
     b. Can you leave out something, and does that change the meaning?
     c. Could you also start with X?

This involves an increasing amount of production by the linguist, but it is very limited, and the consultant still has to utter the sentence, instead of just nodding. Consultants easily accept doubtful sentences, but they rarely utter them themselves. The last level would be

(3) And if I say X Y Z, how does that sound?

This last sentence is already pushing the linguist's production quite far and is susceptible to being accepted as correct even if no native speaker would ever utter it. The second part of the sentence is crucial because the consultant cannot just say *yes*. He has to give a qualitative interpretation between something like *very good* and *You can't say that. That does not mean anything!* If the sentence X Y Z was accepted only at this late stage, I did two things. First, I asked the consultant to repeat the sentence. Many people can say that incomplete sentences are fine, for instance, but the dropped

material will invariably show up when they are asked to repeat. The second thing is
to ask which one of the sentences is 'better', the one that the consultant gave first, or
the one she finally accepted after direct prompt. The answer *Yours is better* clearly
indicates politeness effects. Fortunately, I did not come across this. The answer *both
same* is OK and indicates that there seems to be no great difference in acceptability be-
tween the two sentences. Still, it has to be noted down which sentence came first. The
answer *The second one is a bit odd, no?* or *You can say that in fast speech* indicates
that the second sentence would never be said in that way by the consultant and is thus
of dubious grammaticality. Another technique employed was to start the sentence in a
doubtful tone and leaving it suspended in the middle, like looking for a word. Often,
the consultant will join in then and finish the sentence.

3.3.2   Conversation and narratives

While I tried not to speak SLM during elicitation sessions in order to not distort the
data, I tried to avoid speaking English in conversation and when collecting narratives.
People tend to give answers in the languages they are asked in, so asking in English
would give stories in English, which was not desired. A very convenient paralinguistic
feature is the 'Indian Head Wobble', which indicates (roughly) that the hearer agrees
with what the speaker is saying. This gesture can then be used to stimulate further
linguistic production and normally does not lead to a code-switch to English or other
accommodation in the form of foreigner talk.

The inhabitants of Sri Lanka are generally very sociable and love to chat. The
Malays are no exception to this. They integrate foreigners very well, and so there
is some linguistic production of mine to be found in the cvs-files. As far as I can
see, this has not altered the speech of the Malays in general. I do not think that they
accommodated their speech beyond reducing speed and articulating better (and even
that they did not always do). I thus feel that examples from these conversations can
be used to illustrate language use as well as any other file. In fact, the analysis of
intonation is based on a conversation among several Malays and myself where they
inquire about my place of living etc.

Sometimes my consultants would become involved in a conversation and forget
about my presence. In that case, I often withdrew from the conversation and let them
speak among themselves. These data are about the most naturalistic as one can get.
Unfortunately, this is not only true for morphosyntax, intonation, discourse structure
etc., but also for what is generally known as common ground. Given that I did not
participate in the conversations, no need was felt to explain concepts and relations I
was not aware of, and I must say that very often I had not the slightest clue what a
conversation was about. This is partly due to specialized vocabulary, fast speech and
so on, but also to the fact that I did not know the people, places and events that were
interesting. Add this ignorance of the common ground to the general South Asian
tendency to drop known material and you get a completely lost linguist.

3.3.3 Translation

The 'elicitation' sessions were mostly spontaneous. I had noted down some phenomena I was interested in and would improvise contexts for variations as we went along. However, sometimes I would use prefabricated English or Sinhala sentences to systematically check some phenomenon. Normally, I wrote up sentences for three or four different topics which I asked in random order. This was done to prevent entrenchment of the constructions (Schütze 1996:134). When people are asked the very same construction twenty times in a row with only marginal variation, they tend to rate the 21st construction like the others. Mixing in some completely different constructions can prevent this feeling of 'assembly chain elicitation'.

I also used examples from a learner's grammar of Spoken Sinhala,[3] Karunatillake (2004). This was done to see whether some categories that are grammaticalized in Sinhala are grammaticalized in SLM, too. For instance, I had the suspicion that SLM would have a grammaticalized evidential marker because this is an areal feature. However, asking to translate English sentences with some evidential reading like *It is said/it seems that X* would always yield *aràbiilang ...* , which is the lexical word "to say". It was impossible to construct a context in English where speakers were led to mark evidentiality. Using Sinhala primes proved an instant success, though. The Sinhala evidential clitic *=lu* in some sentences was translated as the particle *kiyang*. *Kiyang* is the evidential marker that was impossible to elicit through English. Given this success, I copied all sentences from the learner's grammar and read them to the Malays, a source of big fun (*How do you say 'Oh no, the uncle fell in the well!'* will always be remembered for that ...). From their translations, I could see whether there was some more grammaticalized material in SLM that was very difficult to elicit through English. Of course, this technique cannot be used as evidence for semantics or word order because of interference effects. Where it can be used successfully is to establish the existence of elusive grammatical morphemes.

3.3.4 Word lists

Word collection was largely done in a spontaneous and associative way. Swadesh lists and extended Swadesh lists for South Asia taken from Abbi (2001) were also used, as were pictures from the MPI Nijmegen field manual (Levinson & Enfield 2001). Abbi also provides questionnaires for some typical areal phenomena, but it was found that they were too schematic and would not yield good responses.

Every now and then, I checked parts of my dictionary with consultants to see whether I had correctly written down what they had told me. This was especially necessary for the different coronal stops. I found it very hard to hear the difference between dental and retroflex stops, so cross-checking was necessary. The Sinhala alphabet has different graphemes for these stops, so I could ask the consultants to

---

[3]At that point in time I was not aware of Gair et al. (1979), a textbook for Spoken Sri Lankan Tamil.

write down the Malay word in Sinhala script and see which grapheme they would use.[4] This method relies on writing because, when asked which letter they would write, they would either say [ʈa̠jənə] or [ʈa̠jənə], the Sinhala names of the letters. This leaves the initial problem of whether the stop was dental or alveolar. But [ʈa̠jənə] is represented in Latin script by ⟨th⟩, so I found out that I could ask the informants whether a certain word should be spelled with ⟨t⟩ or ⟨th⟩. The answer would indicate whether the stop was dental (th) or alveolar (t). Other contexts where this asking to write down a word was useful was vowel quantity, the presence or absence of a labial approximant before [u], and the distinction between prenasalized consonants and combinations of nasal+homorganic stop.

3.4 Recording

Most sessions were recorded on audio with permission of the consultants. Sensitive material was deleted on request or not recorded. Recording devices are a bit obtrusive, and speakers were a bit concerned at first. I have the impression that after some time the consultants got used to the recording device and continued to talk like normal.

Some sessions were also recorded on video with a very small camcorder. This is more obtrusive than audio recording, and the data are less naturalistic. Some people were perfectly fine with being taped, while others seemed to be a bit uneasy. In the latter case, the video camera was not used much. I found that using a tripod improves the technical quality, but decreases the linguistic quality of the data because people tend to stare at the camera. This is why I did not use it much. I would normally sit with the camera on my lap, the LCD display turned 45° so that I could see it. This permitted me to keep eye contact with the consultant and to use mimics and gesture to signal that I was keeping track of the story. At the same time, I could occasionally shed a glance to the LCD to see how the camera was doing, whether the tape was full or the battery empty etc.

All audio data were recorded as linear pcm at 44,1kHz on flash memory cards using a Mayah Flashman and a Sony electret condenser microphone ECM-MS907. Video was taken with a Sony DCR-HC90E camcorder and a Sony ECM-HGZ1 gun zoom microphone. The videos seen in the archive have their audio track replaced by the Flashman audio track because the latter's quality is superior.

Every evening, the audio data were copied to the laptop, and video was captured with the program Premiere. Video was converted to mpeg1 and mpeg2 with the program Tsunami,[5] while audio was left as pcm/wav. The audio files were segmented into thematic chunks with the program `transcriber` and cut into smaller session

---

[4]Sri Lanka has a literacy rate of 95%, and the Malays are among the best educated ethnic groups so this technique is culturally adequate. Also, the Sinhala alphabet is phonemic and the pronunciation of a given grapheme is always predictable. My informants happened to be more familiar with the Sinhala script, but the Tamil script could also have been used.

[5]It is a somewhat awkward feeling to use a program called Tsunami in Sri Lanka on an everyday basis.

files with the export function of the same program. Mpeg1, mpeg2, master audio files and session files were stored in a folder that was burned twice on DVD when it reached 4 GB. One DVD was sent by snail mail to Amsterdam while the other DVD served as a local backup.

The data collected during one trip amount to about 50 GB, which is too much for my laptop's hard disk. An external USB-disk was used for extra storing capacity. This disk was also used as an additional backup place as long as space permitted.

The very first narrow phonetic transcriptions were done using `transcriber`. I wanted to use the program Keyman, which permits easy insertion of IPA symbols into any program. But I found that Keyman and transcriber would not work together. I finally used transcriber as if Keyman worked. This means, for instance, typing e= to get a schwa. In any other program, the conversion is done instantly, but not in transcriber. I stored the file with e= and later search&replaced the e= by ə. After I had worked out an orthography, this was no longer necessary because the orthography only draws on the standard Latin alphabet.

Transcriptions were also backed up on the same DVDs as the media files. I wrote a small perl script to convert the transcriber files into Toolbox files, keeping the time code information in a way such that ELAN would later be able to import my Toolbox files. Toolbox was used to interlinearize the transcriptions. As a final step, the Toolbox file was imported into ELAN and aligned with video if that was available (the audio file was already aligned through the time code I had kept).

3.5 Corpus

The corpus consists of about 150 texts of diverging nature, complemented by some elicitation sessions. They include narratives, conversations, recipes, emails, and written texts (see Table 3.1). Texts are mainly from Kandy, but can also stem from Badulla, Nawalapitiya, and Gampola. The shortest text is only 10 words, while the longest counts 998. The youngest speaker is around 12, while the oldest speaker is over 70. Table 3.2 gives an overview of the texts used to a greater or lesser extent to illustrate the phenomena discussed in this grammar. Examples are commonly used to illustrate more than one phenomenon. When my analyses are discussed in future work, reference should be made to this grammar. When only the examples are used, reference should be made to the individual text as available in the online corpus. Five representative texts are included in the appendix.

The data presented in examples follow a number of conventions which are illustrated below in example (4).

141	K070000wrt04	11	N060113nar05	6	K060103rec01	2	K061123sng01
99	K060108nar02	11	N060113nar03	6	B060115prs21	2	K061030mix01
84	K081105eli02	11	K070000wrt03	6	B060115cvs04	2	K061019nar01
80	K081106eli01	11	K060116nar11	5	N060113nar04	2	K060116nar15
59	K081104eli06	11	K060116nar06	5	K081111eli01	2	K060116nar08
54	K070000wrt01	11	K060116nar04	5	K081104eli03	2	K051222nar07
48	K051220nar01	11	K051201nar02	5	K061122nar03	2	K051222nar01
47	K051206nar02	11	B060115prs10	5	K060116nar05	2	K051213nar08
40	K081103eli02	10	K061120nar01	5	K051213nar07	2	K051206nar13
40	K051213nar06	10	K060116nar07	5	K051206nar18	2	K051206nar03
39	B060115nar04	10	K060103rec02	5	K051206nar17	2	B060115prs20
38	B060115cvs01	10	K051222nar03	5	B060115cvs16	2	B060115prs13
34	K081114eli01	10	K051201nar01	4	N060113nar02	2	B060115prs04
34	K081103eli04	10	B060115cvs08	4	K071011eml01	1	K090327eml01
34	K051205nar05	9	K061127nar03	4	K061019sng01	1	K081201eml01
33	N061031nar01	9	K061026rcp01	4	K060116nar23	1	K081105eli01
33	K051222nar06	9	K061019prs01	4	K060116nar09	1	K071203eml01
31	G051222nar01	9	K051205nar04	4	K060116nar03	1	K071113eml01
30	K051206nar07	9	B060115cvs13	4	K060116nar01	1	K071029eml01
29	B060115nar05	8	N061124sng01	4	K060103cvs01	1	K060116nar21
28	K061026prs01	8	K061019nar02	4	K051206nar14	1	K060116nar13
27	K070000wrt05	8	K060116nar10	4	K051206nar10	1	K051222nar02
27	K060103nar01	8	K060116nar02	4	K051206nar05	1	K051220nar02
27	K051213nar01	8	K060112nar01	4	B060115prs14	1	K051213nar04
26	B060115nar02	8	K051206nar20	4	B060115prs05	1	B060115prs18
24	N060113nar01	8	K051206nar11	4	B060115nar03	1	B060115prs12
24	K061026rcp04	8	K051205nar02	4	B060115nar01	1	B060115prs07
23	K070000wrt02	8	G051222nar02	4	B060115cvs02	1	B060115prs06
23	K051222nar04	8	B060115cvs03	3	K081118eli01	1	B060115cvs06
20	K051206nar12	7	K051222nar05	3	K060103cvs02	1	B060115cvs05
19	K060108nar01	7	K051206nar16	3	K051222nar08		
18	K081104eli05	7	G051222nar04	3	K051213nar05		
18	B060115rcp02	7	G051222nar03	3	K051206nar15		
16	K051213nar03	7	B060115rcp01	3	K051206nar08		
15	K061122nar01	7	B060115prs03	3	K051206nar06		
15	K051213nar02	7	B060115prs01	3	B060115prs17		
14	K081103eli03	6	K081104eli01	3	B060115cvs17		
14	K051206nar04	6	K061125nar01	3	B060115cvs09		
14	B060115prs15	6	K061026rcp02	3	B060115cvs07		
12	K051206nar19	6	K060116sng01	2	K061123sng03		

Table 3.2: Number of examples taken from different texts in the corpus and elicitation sessions. Many examples are used at more than one place.

(4) *[Aashik=nang hathu* soldier *mà-jaadi   suuka=si   katha] arà-caanya.*
Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT   NON.PAST-ask.
'(He) asks if you want to become a soldier, Ashik [because your father is one, too].' (B060115prs10)

The conventions are:

- Sri Lanka Malay words are in italics (*suuka* 'like').

- Words from another language (code-switching or borrowings which are not phonologically integrated), mostly English, are in roman (soldier).

- Clitics are marked with the equation mark (=*nang* 'DAT').

- Affixes are marked with the hyphen (*mà-* 'INF').

- Square brackets [] are sometimes used to indicate constituent structure where this facilitates the parsing of the sentence (e.g. heavy NPs, headless relative clauses, subordinates).

- Parentheses () are sometimes used for material needed in English but not present in the source, e.g *He* in (4)

- Background information can be given in brackets [] in the translation ([because your father is one, too]).

- The source from which the example is taken is given in parenthesis after the translation (B060115prs10). Most of the sources can be accessed in text and audio in the Dobes archive
`http://corpus1.mpi.nl/ds/imdi_browser/` . I would like to thank the Volkswagen foundation and its Dobes project for providing funds for the storage and long-term access of SLM data, next to funding the gathering and analysis of the data in the first place.

Part II

Phonology

Chapter 4

Phonology

Phonological research on Sri Lanka Malay started out with a Bachelor's thesis (Tapovanaye 1986) and two Master's theses (Bichsel-Stettler 1989, Tapovanaye 1995). It thus predates research on other areas, but then the domain lay dormant and most subsequent research concentrated on morphosyntax, although Smith et al. (2004) contains some analyses of phonological influence from Muslim Tamil.[1] According to Ansaldo (2005b) "phonology of [SLM] varieties is the least understood aspect." This chapter will first present the segmental inventory consisting of 6 vowels and 24 consonants (including 2 glides), and then proceed to the suprasegmental domains of syllable structure, word structure, and sentence intonation.

4.1 Segments

This section introduces the segments of SLM phonology.[2] Until the presentation of the practical orthography in Section 4.7, all examples are in broad phonetic transcription following the IPA alphabet.

4.1.1 Vowels

The SLM vowel system consists of five full vowels and schwa. It is thus the same as that of colloquial Indonesian (Ewing 2005:229) and many Malay varieties in general (Adelaar 1985, 2005).

---

[1] Bichsel-Stettler worked with informants from the Colombo-Gampaha area during about 14 months. Tapovanaye is Sri Lankan, and worked with Malays from the Colombo area as well, mainly in elicitation settings. The data in Smith et al. (2004) come from a transcript of about 500 sentences taken from a recording made in the Southern village of Kirinda in the late 1970s.

[2] I would like to thank Wolgang Kehrein and Diana Apoussidou for help with this section. They have corrected many of my oversights and set right my misconceptions about certain parts of phonology. All remaining errors are entirely my own.

<div align="center">

i   u

e   ə   o

a

</div>

Table 4.1: SLM vowel phonemes

The near minimal pairs in (1) illustrate the distinctions between the full vowels.

(1) paṉṯas preṉṯa biṉṯan oṉṯa buṉṯur
     hot   law   star   camel knock

Schwa   There is one centralized vowel in SLM, which can be realized as [ə], [ɪ], [ʊ], [i] or [u], depending on context. These contexts will be dealt with in more detail below. I will refer to this phoneme as 'schwa', without committing myself to any particular realization. The following examples show schwa in its realization as [ə] contrasting with [a] in (2) and schwa in its realization as [ɪ] contrasting with [i] in (3).

(2) kaṯːəm   : kiṯam   (3) mɪṉṯa : miṉṯa
     almsgiving : 1PL   vomit : ask

There are some speakers who do not have reduced vowels on the surface. These speakers will pronounce both words in (3) [miṉṯa] and consider them homophonous.[3] Other phonological effects triggered by schwa, like gemination (see Section 4.5) are not affected by this homophony. Speakers who do make the distinction have clearly different formants for realizations of /i/ as [i] and realizations of /ə/ as [ɪ], as shown in Figure 4.1

4.1.1.1   Major allophones of the full vowels

The vowels /a/ /i/ and /u/ are always realized as full vowels. /e/ and /o/ can be realized as lax [ɛ,ɔ] or tense [e,o]. Tense realizations are more often found in open syllables.

In some words, the pronunciation of /e/ approximates the pronunciation of /i/ in the same contexts, e.g. /ṯekus/, 'mouse',[4] is pronounced as [ṯːiːkus].

---

[3]SLM behaves like many other Austronesian languages in this regard, where the status of schwa is often difficult to determine (Himmelmann 2005:116).

[4]Vowel length is not phonemic, as will be discussed in Section 4.1.1.3 and therefore not indicated in the phonemic notation.
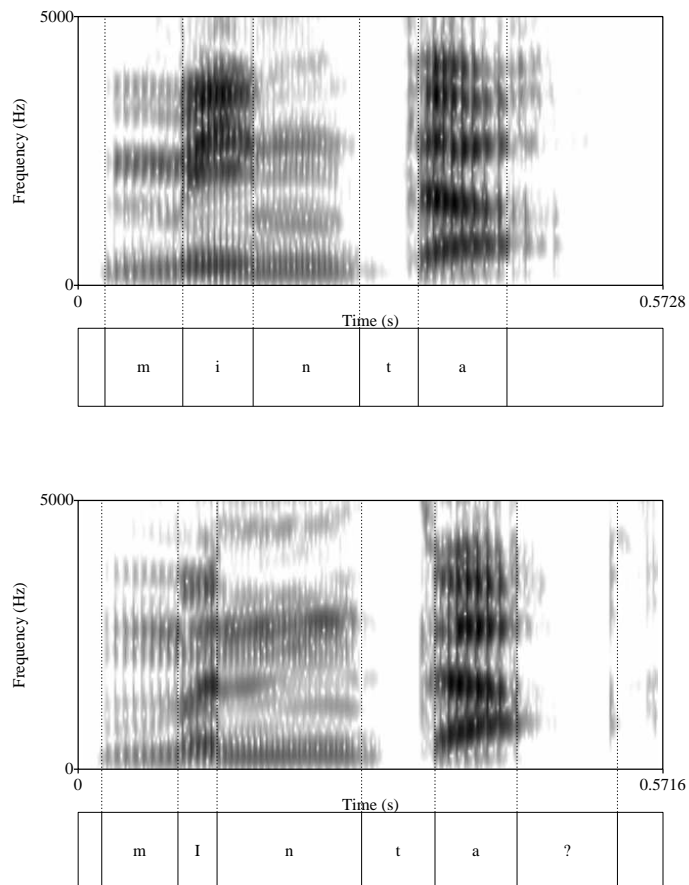
Figure 4.1: The formant structures of the first vowels in [minta] 'ask' and [mɪnta] 'raw' are clearly different, indicating that /i/ and /ə/ are different phonemes, which distinguish the two meanings. The second spectrogram indicates a glottal stop after the final vowel. This is not phonemic, as will be discussed in more detail below (Section 4.1.2).

4.1.1.2 Major allophones of schwa

The realization of schwa depends on its position within the stem and on the individual speaker. Schwa can be realized as [ɪ,ə,ʊ], and sometimes even as [i,u] (cf. Adelaar (1991:27f), Smith et al. (2004)).

Variations in height   In the final syllable, schwa is realized as [ə] or [a], depending on the speaker. In the penultimate syllable, schwa is raised by all speakers (e.g. [bɪsːaɾ] 'big'). The amount of raising depends on the speakers. Some speakers raise to [ɪ] or [ʊ], which makes the identification of the sound as a centralized vowel easy. Other speakers raise the vowel further, to [i] or [u]. In that case, the question whether the sound is underlyingly schwa or a full vowel can only be answered by looking at word structure (see Section 4.5). Especially a following geminate is indicative of the sound being schwa on the phonological level. In the antepenultimate and before, schwa is always realized as [ə]. An example is [cəcaʊak] 'wash'.[5]

No trisyllable can have [ɪ,ʊ] in the initial syllable. A first set of trisyllables bans raising altogether in the initial syllable, while a second set requires raising to [i,u]. Neither set admits alternation between schwa and a high vowel. Examples for the first set are [cəcaːʊak] 'body wash', [bərnaːma] 'famous', [m(ə)laːju] 'Malay', while an example for the second set is /kəmareŋ/ → [kumaːɾeŋ] 'yesterday'.

Since no reduced vowel is possible in the second set, contrary to the disyllables like [t̪{ʊ/u}lːɔr] 'egg', the question arises why a schwa should be present in the underlying form. The reasons for this have to do with syllable structure and the lengthening of the vowel and will be discussed in more detail in Section 4.5.

The verbal prefixes [aɾə-] 'NON.PAST',[6] [anə-] 'PAST' and [mə-] 'INF-'[7] have an unraised schwa [ə] irrespective of the syllable position. Given that most roots are disyllabic, schwa is normally in antepenultimate position as in [aɾə-maːkaŋ] 'is eating', [anə-miːnuŋ] 'drank', [mə-ɟaːlaŋ] 'to go', but it can also be found farther away from the end, as in [aɾə-cəcaːʊak] 'is washing'. There is only one monosyllabic verbal root [piː] 'go'. The prefixes keep the schwa unraised when combined with this root, even if this results in [ə] being in penultimate position (where normally only [ɪ] or [ʊ] are found): [aɾə-piː, anə-piː, mə-piː,*aɾɪ-piː, *anɪ-piː, mɪ-piː]. Evidence from vowel lengthening suggests that these prefixes do not form one phonological word with the verb. This means that the realization of schwa in these prefixes is independent of the structure of the verb the prefixes attach to.

---

[5]The normal word for 'to wash' is [cuːci]; 'to bathe' is [man̪d̪i]. [cəcaʊak] refers to an activity known in Sri Lanka as 'body-wash'. This reflects the etymology of *cuci+awak* 'body+wash'. I will use 'wash' in the gloss as a shorthand.

[6]The semantics of this prefix are more complicated than what can be capture in the gloss here, the reader is referred to Section 6.2.2.1 for a more thorough discussion of this prefix.

[7]And some other much less common prefixes, see Section 6.2.1.

Fronting and retracting   Whether schwa is raised to a front vowel [ɪ,i] or to a back vowel [ʊ,u] seems to be lexically determined, but there are some phonological cues. Some lexemes only admit a front vowel, some only a back vowel, some accept both. The following list gives examples.

(4)   /əmpat̪/ 'four' →[ʊmpat̪]. Further raise: [umpat̪]

(5)   /gəlap/ 'dark' →[gɪlːap]. Further raise: [gilːap]

(6)   /t̪əlɔr/ 'egg' →[t̪ɪlːɔɾ] or [t̪ʊlːɔɾ]. Further raise: [t̪ilːɔɾ] or [t̪ulːɔɾ]

(7)   /pəlaŋ/ 'slow' →[pɪlːaŋ] or [pʊlːaŋ]. Further raise: [pilːaŋ] or [pulːaŋ]

From this list and additional material it appears that two factors have an influence: the following consonant and vowel harmony. A back vowel is most often found before a labial consonant /p,b,m/ [dʊpːaŋ] 'front', [kʊbːoŋ] 'garden', [ʊmːa] 'mother' (cf. Adelaar 1991:27). Dorsal consonants seem to favor a front vowel: [sɪgːaɾ] 'healthy', [dɪkːat̪] 'vicinity'.[8] For the other consonants, a clear pattern could not be found.

Bichsel-Stettler (1989) found vowel harmony in words like [libːi] 'more' and [gʊmːuk] 'fat' (also cf. Adelaar 1991:27). The schwa in the initial syllable is raised and agrees in the feature [±back] with the following vowel.

All this being said, there seems to be a lot of variation between and even within dialects. In the text K060103rec02 for example, the word for 'raw' was transcribed by a native speaker once as ⟨mintha⟩ and two sentences later as ⟨muntha⟩. Notwithstanding the intended and actual realization by the speaker, this proves that for the hearer/transcriber in this case, both the rendering of the first vowel as ⟨i⟩ and as ⟨u⟩ were acceptable, suggesting that the phonological knowledge of the hearer has no strict distinction between the fronted and the retracted realization.

Deletion   In the initial syllable of di- and trisyllables, schwa is often dropped if the resulting sequence is phonotactically acceptable.[9] An example for this is [m(ə)laːju] 'Malay', which is rarely heard with the schwa pronounced; it is much more common to have the sequence *ml* as the onset of the first syllable ([mlaː.ju]). Schwa is not dropped in [cəcaːwak] 'wash' and [bərnaːma] 'famous' because the sequences *#cc(cː)* and *#brn* are not acceptable phonotactically. Deletion of schwa is also found in disyllables, such as /bəras/'rice', which can be realized as [brːas], next to [bɪrːas].[10] This reduction of schwa can also take place if no onset is present: /əma/'mother' can be realized as [ʊmːa], [umːa], or [m̩ma] with a syllabic nasal.

Realizations like [b(ɪ)rːas]'rice' or [p(ɪ)rːaŋ]'war' raise the question whether schwa is present in the phonological form, and then deleted, or whether it is absent in the

---

[8]For theoretical discussions of this, see Clements (1990) and Botma (2004:99).

[9]Jakarta varieties of Indonesian show frequent deletion of schwa in initial position (Uri Tadmor p.c., Ewing (2005:229)) Many of the Sri Lankan Malays came from/through Jakarta, so that this could very well be an inherited feature.

[10]The realization with [i] is never found for this lexeme.

phonological form and then epenthetically inserted.[11] There are two reasons to prefer the deletion account over the epenthesis account. First, there is a phoneme schwa anyway, so the presence of schwa in this position is no additional burden. Second, schwa varies with Ø only in some contexts, in lexemes like [cəcaːʋak] 'wash', Ø is not possible. An epenthetic account can thus not explain the whole story, while the deletion account can. Additionally, schwa is the historical vowel in these words, although the speakers are of course not necessarily aware of this.

4.1.1.3   Vowel length

Full vowels are predictably lengthened in open penultimate syllables. Long vowels are not found in other positions. Vowel length is thus not phonemic (*contra* Bichsel-Stettler (1989), Smith et al. (2004)), but conditioned by nature and position of the syllable (cf. Tapovanaye 1995). Example (8) shows this lengthening process for all full vowels.

(8)   maːɾa : meːɾa : d̪iːɾi : soːɾe   : muːɾa
      anger : red   : body : evening : cheap

Example (8) contains only disyllabic words. In trisyllabic words with an open penultimate syllable, the syllable is lengthened if the initial syllable contains a schwa (9) in any of its realizations ([ə,ɪ,i,ʊ,u]). It is not lengthened otherwise (10).[12]

(9)   cəcaːʋak : məlaːju : kumaːɾeŋ
      wash   : Malay   : yesterday

(10) niɡiɾi   : kut̪umuŋ : makanan
      country : see   : food

No surface form of an underlying schwa ([ə,ɪ,i,ʊ,u]) is ever lengthened. Instead, often another phonological process, consonant gemination, takes place, see Section 4.5.

This general pattern is very regular throughout the grammar, but a certain number of exceptions apply:

---

[11]Tamil for instance has epenthetical insertion of [ɪ] in onset clusters like *viyaaparam* 'business'<Skrt. *vyaabhara* (Annamalai & Steever 1998:103).

[12]There are close to no underived stems without schwa in the initial syllable. I analyze [niɡiɾi] 'country' and [kut̪umuŋ] 'see' as having no schwa synchronically in spite of the diachronic origins *nəgəri* and *kətəmu* because the vowels are always pronounced as full vowels, and none of the other phonological effects of schwa are found. As far as trisyllabic words derived with affixes go (like *makan-an*), all have only short vowels.

- function words never have long vowels: [ki(*ː)t̪am] '1PL', [ka(*ː)t̪a] 'QUOT', with the exception of interrogative pronouns.

- the sequence V*ᵢ*hV*ᵢ* never has a long vowel: [le(*ː)heɾ] 'neck', but [la*(ː)heɾ] 'be.born'

- some sequences VGV, where G is a glide, have a long vowel, others do not. [t̪u*(ː)ʋa] 'old', [t̪u(*ː)ʋan] 'Sir'. See Section 4.1.2.1 for detailed discussion.

- the sequence *uru* has no long vowel in the words [t̪uɾus] 'straight' and [guɾu] 'teacher' (but there is a long vowel in [kuːɾus] 'thin').

- two words are lexical exceptions [kubuːɾ] 'bury' and [kaːɾt̪u] 'quarter'. The former has a long vowel in the final, not the penultimate, syllable. The latter has a long vowel in the penultimate syllable, but the syllable is not open.

- Loanwords from English and Arabic tend to violate this pattern.

4.1.1.4 Distribution of the vowels

For the discussion of the distribution of vowels, we distinguish the root level from the stem level. In the context of this discussion, roots are defined as monomorphemic lexemes, while stems consist of at least one root and optional additional morphemes, lexical or derivational. To illustrate the distinction with English examples, *farm* would be a root,[13] whereas *dairy farm* or *farmer* would be stems, because they comprise additional morphemes, *dairy* and *-er* in these cases.

In SLM roots, there are no restrictions for the final and the penultimate syllable. Here, all vowels can be found. Examples are given in Table 4.2. In the antepenultimate position, only [i], [u] and [ə] are found in native roots. Any vowel can be found in the antepenultimate position in loanwords, which are given in parentheses in Table 4.2.

As for stems, no such restrictions apply. When adding the nominalizing suffix *-an* to a verbal root, the penultimate becomes the antepenultimate and keeps its vowel. In this way, [e], [o] and [a] can be found in the antepenultimate, e.g. [pake-yan] 'dress-NMLZR',[pot̪oŋ-an] 'cut-NMLZR', [pegan- an] 'catch-NMLZR'. Similar things can be said about compounding, which also shifts syllables farther away from the right edge. This does not affect the occurrence of [e], [o] or [a] either. Examples are [oɾaŋ ikːaŋ] 'man+fish=fisherman' and [babi uːt̪aŋ] 'pig+forest=wild boar'.

While schwa can occur anywhere, it is most often found in the penultimate and the antepenultimate. There are only two occurrences of schwa in the final syllable of two words, which are given below

---

[13]And also a stem.

	antepenultimate	penultimate	final
a	(vaʈːakːaː'pumpkin', Tamil)	ʈaksiɾ 'think'	kapːal 'ship'
e	(selinḍi'spider', Tamil)	seksa 'problem'	slampe 'handkerchief'
i	nigiɾi'country'	min̪ʈa 'ask'	sampi 'cow'
o	(goɾakaː'a condiment', Sinh.)	pompaŋ 'female'	nan̪ʈok 'sleepy'
u	kuʈumuŋ'see'	kumbaŋ 'flower'	campuɾ 'mix'
ə	cəcaʋak 'wash'	mɪn̪ʈa 'vomit'	kaʈːəm 'almsgiving'

Table 4.2: Vowels in antepenultimate, penultimate and final position of roots. In antepenultimate position, a, e and o are only found in loanwords, indicated by parentheses. The word *bahasa* 'language' could be an exception, but this seems to be a recent borrowing from Standard Malay, the native word is [baːsa].

(11) ʈaːnəm[14]   (12) kaʈːəm
      'plant(V)'   'almsgiving'

4.1.2   Consonants

SLM numbers 24 native consonants and glides plus 3 consonants only used for loanwords. 5 basic places of articulation are distinguished for stops: labial, dental, apical (with postalveolar and retroflex allophones),[15] palatal, and velar. These places have a series of voiceless stops, voiced stops and prenasalized voiced stops.[16] The palatal stops are realized as affricates (For the general treatment of affricates as strident stops see Jakobson et al. (1952), Clements (1999), and Kehrein (2002).

All stops are matched by homorganic nasals, with the exception of the laminodental nasal, which does occur phonetically, but lacks phonemic status.

The fricative series is less populated, with /s/ being the only fricative frequently encountered. There are some words with /h/, but overall /h/ is much less frequent than /s/. /f/, /z/ and /ʃ/ are finally only found in loanwords from Arabic ([ziharaʈ] 'shrine') or English ([femili] 'family', [bɾɪtɪʃ] 'British'). /f/ and /z/ are also found in many Islamic names like *Farzan* or *Fayizal*.

/l/ is an alveolar lateral approximant. /r/ is a rhotic realized as a tap or a trill.

The picture is completed by a labiodental approximant ʋ and a palatal approximant j (cf. Table 4.3).

---

[14]While Std. Malay does not permit schwa in the final syllable, Jakartanese does (Adelaar & Prentice 1996), and *tanəm* is also found in that variety, with the same meaning (Adelaar 1985:200).

[15]I use 'apical' as a cover term for postalveolar and retroflex realizations. This term as it is used here never includes the dentals, no matter how they are realized.

[16]The dental prenasalized stop is missing.

	labial	dental	apical	palatal	velar	glottal
stops						
voiceless	p	t̪	t	c	k	
voiced	b	d̪	d	ɟ	g	
prenasalized	ᵐb		ⁿd	ⁿɟ	ᵑg	
nasals	m		n	ɲ	ŋ	
fricatives	(f)		s (z) (ʃ)			h
approximants	ʋ			j		
liquids			r l			

Table 4.3: SLM consonant phonemes, including approximants. Palatal stops are phonetically affricated. Parentheses indicate phonemes only found in loanwords.

Bichsel-Stettler (1989) and Adelaar (1991) concur with this inventory, but do not list the prenasalized series, while Smith et al. (2004) express doubts about the apical series, and do not list the prenasalized series either. Minimal pairs establishing both apical as a place of articulation and prenasalized stop as a manner are given below.

Tapovanaye (1995) finds glottal stops before initial vowels and after final vowels, but does not grant them phoneme status. Bichsel-Stettler (1989) found a glottal stop before every initial vowel. For reasons of syllable structure, she granted phoneme status to it. She wanted to eliminate syllables without onset. In my data, the glottal stop is clearly optional. Figure 4.1 shows a glottal stop. While it is tempting to treat this as a reflex of historical [*məntah] 'raw' as opposed to historical [*minta] 'beg', the occurrence of the glottal stop is too irregular between speakers and lexemes to establish a clear correlation pattern between historical /h#/ and the occurrence of the glottal stop in the modern language. There is variation as to whether the glottal stop occurs word-initially, [ʔaːnak] or [aːnak] 'child'. Word-medial but syllable-initial vowels do not get a glottal stop preposed ([ɟaːu]'far', not *[ɟaːʔu]). The differences between the data of the aforementioned authors and my data might be due to the fact that they worked with elicited lexemes, while the data this analysis is based on naturalistic speech. In the context of an elicitation session (like the one which yielded the data for the spectrogram in Figure 4.1), a glottal stop is often inserted in citation forms of words before an initial vowel and after a final vowel. This glottal stop insertion is not very regular in non-citation contexts. Bichsel-Stettler (1989) and Tapovanaye (1995) treat it as a regular process, but this is not supported by the data in my corpus. Vowels in initial position are frequently found without glottal stop. As an example, the word [oːbat̪] 'medicine', pronounced after a pause of about 0.7 seconds (not represented completely in the graphic), does not receive a glottal stop, as can be seen from Figure 4.2.

In this description, I therefore do not assume a phonemic glottal stop.
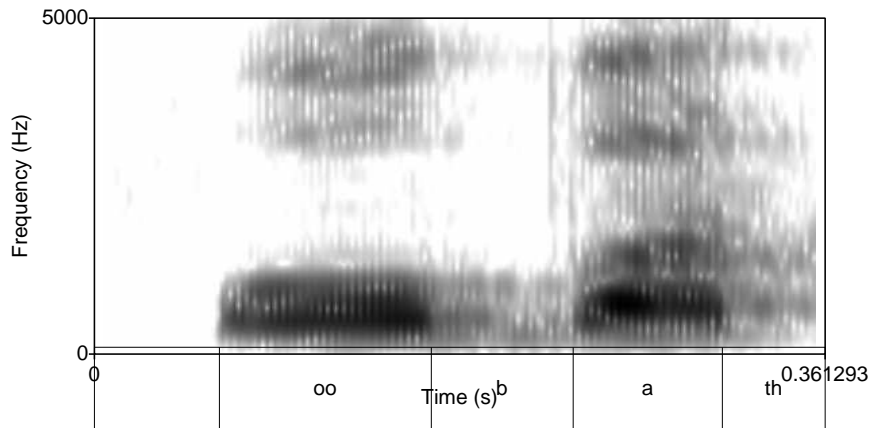
Figure 4.2: No glottal stop is found before the word initial [oː] in [oːbaṯ] 'medicine'

4.1.2.1   Distinctions in place of articulation

This section provides minimal pairs for all possible pairs of two neighboring places of articulation.

Stops   Distinctive places of articulation for stops are: labial, dental, apical, palatal, and velar. Apical comprises postalveolar and retroflex allophones. Apical as it is used in this description never comprises dental stops. In phonemic representation, /t/ will be used for apical stops, while the allophones will be represented as [ṯ,ḏ] and [ʈ,ɖ].

/p/ **vs.** /ṯ/		/b/ **vs.** /ḏ/		/ᵐb/ **vs.** /ⁿ̪ḏ/
				n/a
(13) baːpa : maṯːa		(14) baːṯu : ḏaːṯang		
father : eye		stone : come		
/ṯ/ **vs.** /t/		/ḏ/ **vs.** /d/		/ⁿ̪ḏ/ **vs.** /ⁿɖ/
				n/a
(15) kɔːʈɔɾ : kɔːṯɔɾ		(16) ḏuːa : ɖuːʋa		
dirt : fool		praise : two		

/t/ **vs.** /c/	/d/ **vs.** /ɟ/	/ⁿd/ **vs.** /ⁿɟ/

(17) kɔːʈɔɾ : cɔːcɔ   (18) aːɖa : aːɟaɾ   (19) ṭaːⁿɖak : paːⁿɟaŋ
      fool   : kind of   exist : teach   dance   : long
             fruit

/c/ **vs.** /k/	/ɟ/ **vs.** /g/	/ⁿɟ/ **vs.** /ⁿg/

(20) cuːci : kuːciŋ   (21) ɟaːga : gaːɟa   (22) paːⁿɟaŋ : maːⁿga
      wash : cat   protect : elephant   long   : mango

Fricatives   There are only two native fricatives, /s/ and /h/, for which example (23) provides a near minimal pair. In Tamil [s] and [c] are allophones of /c/ (ċ) (Suseendirarajah 1973a:174). Therefore it is necessary to establish that these phonemes are indeed different in SLM. Examples (24) shows this. For the sake of completeness, example (25) proves that /s/ is also different from /ɟ/. /ʃ/ and /z/ are only found in loanwords.

/s/ **vs.** /h/	/s/ **vs.** /c/	/s/ **vs.** /ɟ/

(23) kɔːsɔŋ : pohoŋ   (24) baːsa : baːca   (25) baːsa : blaːɟaɾ
      empty : tree   wet   : read   wet   : learn

   /f/ is marginally phonemic and is found in a few words or Arabic origin, like [fiṭena] 'slander' and [kafan] 'shroud', besides proper names and religious terms.

Nasals   The following list provides minimal pairs for all pairs of neighbouring nasals. Note that there is no dental nasal phoneme,[17] hence /m/ and /n/ are neighbours.

/m/ **vs.** /n/ **vs.** /ɲ/ **vs.** /ŋ/

(26) ṭaːma : ṭaːna : ṭaːɲa : ṭaːŋaŋ
      before : soil   : ask   : hand

Glides   SLM has two glides, /ʋ/ and /j/. These are distinctive as the following examples show.

---

[17][ṇ] does occur before dental stops, but this is always the result of assimilation.

/ʋ/ **vs.** /j/

(27)   saːʋa : saːjaŋ
       fields : love

Phonemic status of the glides   Some occurrences of the glides, like those in (27),
are clearly phonemic and present in the underlying form. Other occurrences could be
the result of a phonological constraint banning hiatus, especially when a long vowel
is involved. This is schematized in (28) (cf. Tapovanaye 1995:23).

(28) $CV_i.V_j \rightarrow CV_iːV_j \rightarrow CV_iGV_j$

   or, in a more formal way

(29)   [+*vocalic*][+*vocalic*] → [+*vocalic*][−*vocalic*]/_[+*vocalic*]

   The second vowel in a two-vowel sequence is not realized as vocalic in front of
a third vowel. Evidence for this comes from the absence of a long vowel in envi-
ronments which would normally require one (Section 4.1.1.3, p. 88). An example
is [lijat̪] 'see', where we would normally expect [liːat̪] or [liːjat̪] with a long vowel.
This problem of the missing long vowel can be solved if we assume that the sequence
VG is a special realization of a long vowel. It would then appear that the second part
of the vowel is turned into a glide, leaving only the first part as vocalic and thus result-
ing in a short vowel. Next to /i/, this is found for /u/, for example in [buʋaŋ] 'throw',
where the right part of the long /u/ is turned into a glide.

   If the first vowel is neither /i/ nor /u/, the glide formation is not found, resulting
in hiatus as in [ɟaːu] 'far'.[18] In these cases, a long vowel is always found. Also note
that a short vowel in the penultimate syllable of disyllabic words is only possible if
the onset of the final syllable is a glide. This distribution strongly suggests that [ij]
and [uʋ] are alternative realizations of a lengthened /i/ and a lengthened /u/.

   The analysis of the creation of a sequence VG out of a long vowel is complicated
by the fact that there are some words with a sequence VːG, like those given in (27).
While in (27), this is not problematic because /a/ does not have a corresponding glide,
there are also some words where a long [+high] vowel is followed by a glide, e.g.
[buːʋa] 'fruit'. This can be analyzed as an underlying form /bu.ʋa/, which under-
goes the regular vowel lengthening process of open penultimate syllables in disyllabic
words. There are thus words of the type /CV.V(C)/ like /bu.aŋ/ which have to be
distinguished from words of the type /CV.GV(C)/ like /bu.ʋa/.[19]

   The only other instance where a short vowel is found in the open penultimate of
a disyllabic lexical word is in the sequence V_ihV_i (e.g. [leheɾ] 'neck'). There might

---

[18]The phonological conditions of a preceding high vowel in order to form a glide are actually a common
feature of Austronesian languages (Himmelmann 2005:116).

[19]For an argument for the theoretical distinction between phonemic and non-phonemic glides see Levi
(2008).

be the possibility to explain this sequence in a manner akin to the occurrence of non-phonemic glides, similar to the solution proposed by Tapovanaye (see Section 4.5.1.2), but this needs more theoretical research.

4.1.2.2 Manner and voicing distinctions

Having discussed the different places of articulation, we now discuss the different manners of articulation available for the different places of articulation. Voicing contrasts are also discussed here for convenience.

Distinctions for labial consonants

/p/ **vs.** /b/ **vs.** /m/ **vs.** /ʋ/

(30) saːpa : ɾaːba : saːma : saːʋa
     who? : caress : together : fields

Distinctions for coronal consonants
/t̪/ **vs.** /s/ **vs.** /n/ **vs.** /r/ **vs.** /l/

(31) gaːt̪al : ɾaːsa : t̪aːna : d̪aːɾa : saːla
     itch   : tasty : soil   : blood : wrong

/t̪/ **vs.** /t/ **vs.** /d/ **vs.** /s/ **vs.** /n/ **vs.** /r/ **vs.** /l/

(32) kuːt̪u : luːʈuṯ : d̪uːd̪uk : suːsu : guːnuŋ   : t̪uːruŋ   : spuːlu
     louse : knee : sit   : milk : mountain : get.down : ten

/d̪/ **vs.** /d/ **vs.** /t̪/

(33) d̪uːa : d̪uːʋa : t̪uːʋa   apical /t/ is not possible initial position
     recital : two   : old

/t̪/ **vs.** /t/ **vs.** /d/

(34) kɔːt̪ɔɾ : kɔːʈɔɾ : kɔːd̪ɔk
     dirt   : fool   : frog

/d̪/ **vs.** /d/

(35) pad̪eɾi   : aːd̪e
     struggling : younger.sibling

[ɽ] (allophone of /d/) **vs.** [ɾ] (allophone of /r/)

  (36) ɖaːɽa : ɖaːɾa
        chest : blood

Distinctions for palatal consonants

/c/ **vs.** /ɟ/ **vs.** /ɲ/

  (37) caːri : ɟaːɖi   : ɲaːɲi
        find : become : sing

Distinctions for velar consonants

/k/ **vs.** /g/ **vs.** /ŋ/

  (38) klaːki : ɖaːgiŋ : aːŋiŋ
        boy   : beef : air

4.1.2.3   The prenasalized stops

SLM has four prenasalized stops matching the points of articulation of the non-pre-
nasalized stops. The dental prenasalized stop is not found. Prenasalized stops contrast
with N+stop clusters both phonetically and phonologically. Phonetically, the pre-
nasalized consonants have a much shorter nasal phase (Figure 4.3). Phonologically,
vowels preceding prenasalized consonants are lengthened when in penultimate posi-
tion of disyllabic stems (for more discussion of vowel lengthening see Section 4.4.1).
This suggests that the prenasalized consonants do not close the preceding syllable.
N+stop clusters, on the other hand, never occur after a long vowel, suggesting that
the nasal part closes the preceding syllable. This absence of vowel lengthening in het-
erosyllabic clusters can also be seen in Figure 4.3, where the vowel is shorter before
the N+C cluster of [ʈuɳɖu] 'piece' (above), and longer before the prenasalized ⁿd in
[ʈuːⁿɖuk] 'bend' below.

    Both heterosyllabic and tautosyllabic NC sequences are listed in Table 4.4, to-
gether with some other combinations of voiced stop and/or nasal.

    Additional support for the phonemic status of the prenasalized consonants comes
from acceptability judgments. Speakers were asked whether *sambal* and *ga:ᵐbar*
could be written with the sequence ⟨m+b⟩ (ⓜⓑ) or the prenasalized consonant ⟨ᵐb⟩
(ⓜ) in Sinhala script. Informants generally accepted the writing ⟨m+b⟩ for both forms,
but the writing ⟨ᵐb⟩ only for *ga:ᵐbar*. Analogous results were obtained for words with
other prenasalized consonants. This shows that these sounds are cognitively different
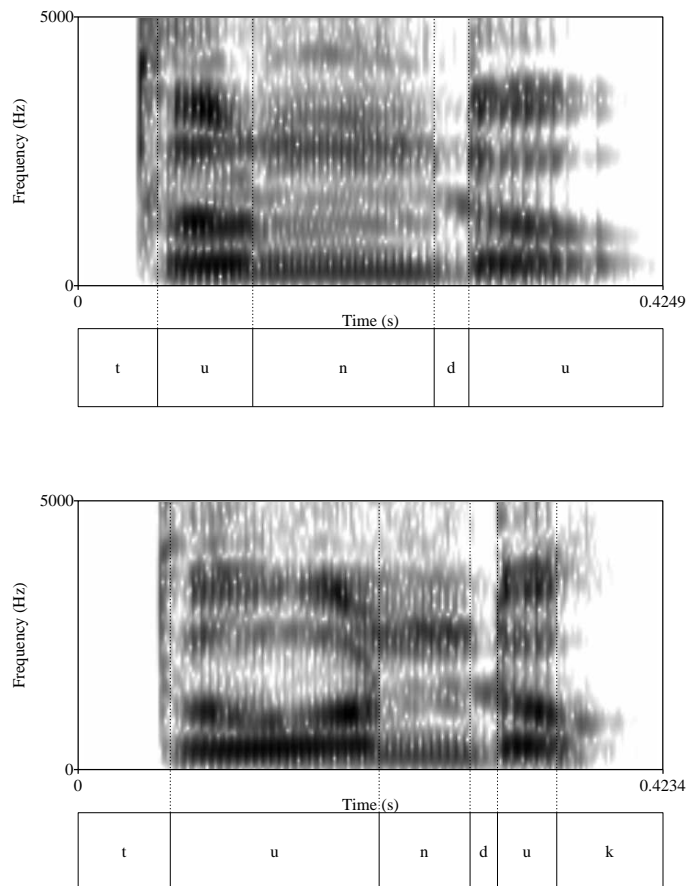
Figure 4.3: The heterosyllabic sequence /n.d/ in the Tamil loan [ʈuɳ.ɖu] 'piece' has a short vowel (69ms) and a comparably longer nasal (132ms) than the tautosyllabic sequence /.nd/ in the native word [ʈuː.ⁿɖuk] 'bend' below (151ms+66ms). The latter sequence is analyzed as a phonemic prenasalized consonant in this grammar.

	V:.CV	VC.CV	V:.NV	VN.NV	VN.CV	V:.ᴺCV
labial	baːɟi 'pig'	habːaɾ 'news'	t̪aːma 'earlier'	samːa 'every'	sambal 'spicy food'	gaːᵐbar 'picture'
dental	n/a	n/a	n/a	n/a	seliɳ̊ɖ̥i 'spider'	n/a
apical	liɖa 'tongue'	piɖːaŋ 'sword'	piːɳaŋ 'areca nut'	bmːaŋ 'thread'	t̪un̥ɖ̥u 'piece'	t̪uːⁿɖ̥uk 'bent'
palatal	gaːɟi 'salary'	haɟːi 'Haj'	ɟaːɲi 'sing'	baɲːak 'much'	aɲɟiŋ 'dog'	baːⁿɟir 'flood'
velar	ʈiːga 'three'	siɡːaɾ 'healthy'	iːɳaʈ̪ 'think'	t̪iɲːa 'middle'	siŋga 'lion'	mniːᵍgal 'die'

Table 4.4: Prenasalized consonants in opposition to plain and geminated stops, plain and geminated nasals, and clusters of nasal+stop. Dental stops normal do not occur word-medially. [seliɳ̊ɖ̥i] 'spider' is an exception, a loanword from Tamil.

for the speakers, otherwise they should have been treated alike. This is even more significant given that the vowel preceding a prenasalized stop tends to be reduced in Sinhala, but lengthened in SLM. Artefacts in the choice of grapheme resulting from vowel quality can therefore be excluded.

4.1.2.4 The palatal obstruents

SLM numbers three palatal obstruents /c,ɟ,ⁿɟ/, which, besides their realization as stops, have affricate allophones [tɕ,dʑ,ⁿdʑ]. Given that the other stops do not have the option of affrication, the question arises whether these phonemes pattern with the stops, or whether they are in a natural class of their own, the class of affricates (cf. Bichsel-Stettler (1989:63), Noonan (2006:357)). First, we have to establish whether they consist of one phoneme or two. In case they consist of only one phoneme, we have to answer the question whether this phoneme is internally complex.

Trubetzkoy (1939) contains a list of rules how to decide between the 'monophonematic' and the 'polyphonematic' analysis of a certain part of the sound sequence (*Schallstrom*). Particularly relevant for SLM are the rules IV and VI, which are given below.

> Regel IV. Eine potentiell monophonematische [...] Lautverbindung muß als Realisation eines einzigen Phonems gewertet wetden, wenn sie als Einzelphonem behandelt wird, d.h. wenn sie in solchen Lautstellungen vorkommt, wo in der betreffenden Sprache Phonemverbindungen nicht zugelassen werden.[20] (Trubetzkoy 1939:53)

> Regel VI. Wenn ein Bestandteil einer potentiell monophonematischen Lautverbindung nicht als kombinatorische Variante irgendeines Phonems derselben Sprache gedeutet werden kann, so muß die ganze Lautverbindung als Realisation eines Eigenphonems gewertet werden.[21] (Trubetzkoy 1939:54)

Rule IV states that a sequence must be analyzed as one phoneme if it occurs in a context where the language demands that only one phoneme be present, not more. In SLM, complex onsets are such a context. A complex onset can be formed by one consonant followed by any of /r,l,ʋ,j/ (see 4.3 for more discussion). This can be schematized as CXV, where X is one of the phonemes just listed. The word [criːʈa] 'story' has an /r/ in the second position. Preceding the /r/, only one phoneme is permitted by SLM syllable structure. Therefore, the sound sequence preceding /r/ must be analyzed as one phoneme /c/, not as a sequence of several phonemes, like /tɕ/. The argumentation for words like [ɟleːna] 'window' is analogous.

---

[20]A potentially monophonematic sound combination must be seen as the realization of exactly one phoneme if it is treated as a single phoneme, i.e. if it occurs in environments where the language under discussion does not admit combinations of phonemes.

[21]If a part of a potentially monophonematic sound sequence cannot be analyzed as a combinatorial variant of any phoneme of the same language, then the whole sound sequence must be seen as a realization of one phoneme.

Rule VI states that a sound sequence should not be analyzed as consisting of several phonemes if this would lead to the postulation of phonemes otherwise not found in the language. If the palatal obstruents are analyzed as polyphonematic, consisting of a stop and a following fricative, this yields to the postulation of palatal fricatives. Palatal fricatives are not found elsewhere in SLM phonology. Therefore, the palatal obstruents must be analyzed as monophonematic.

We now have firm evidence that the palatal obstruents are monophonematic. The following question is whether we have to assume a complex internal structure within the phoneme. In German, for instance, *Tal* 'valley' and *[*ts*]ahl* 'number' are distinguished by the former word having a simple monophonematic onset, while the latter has a complex monophonematic onset. There other features like voicing and place of articulation being alike, we need to recur to the internal structure of the phoneme to capture the difference between /t/ and /ts/. In SLM, such distinctions are not necessary. The palatal obstruents /c,ɟ,ⁿɟ/ are not in opposition to other segments which would require the distinction between simple and complex internal structure.

The last question is whether the palatal obstruents form a natural class with the other stops, or whether they are in a class of their own. The palatal obstruents behave in most ways like the other stops, and can occur in all places the stops can occur, affording the same phonological generalizations like vowel lengthening (Section 4.1.1.3) or gemination after schwa (Section 4.1.2.6). One exception to this rule is word-final position, where palatal obstruents are never found. This, however, need not be a result of /c,ɟ,ⁿɟ/ being distinct from the other stops, it rather seems to be a constraint banning palatals from this position since ɲ is also disallowed there.[22][23]

To conclude, the palatal obstruents can safely be treated as stops, there is no need them any different from other stops in phonology, but their phonetic realization can contain a fricative phase.[24]

4.1.2.5 Major consonantal allophones

This section discusses the free and complementary allophones of consonantal phonemes.

Stops /p, t̪, t, c, k/ are unaspirated voiceless stops. /b, d̪, d, ɟ, g/ are voiced stops. SLM voiceless consonants have a VOT close to zero (0-10ms), voiced consonants have a negative VOT and tend to be fully voiced.[25] The appendix (p. 789) contains

---

[22]The palatal glide [j] does occur in some words in final position, but this is uncommon.

[23]Standard Malay and Jakartanese also have a constraint against palatals in final position (Adelaar 1985:12,32).

[24]See Clements (1999) and Kehrein (2002) for more theoretical discussion of phonetic affrication of palatal stops.

[25] Voice onset time (VOT) is a measure used for consonants. It indicates the relative timing between the release of the articulatory stricture and the beginning of voicing, i.e. the vibration of the vocal chords (Ladefoged 1975:19f). A positive VOT is indicative of aspiration, unaspirated voiceless consonants have a VOT around 0, while a negative VOT is indicative of voiced consonants.

spectrograms illustrating this for a number of environments, namely the onset, intervocalic position, and following /s-/.

Phonologically apical /t, d/ are realized as retroflex [ʈ,ɖ] before /a,o,u/ and as postalveolar [t̠, d̠] before /e,i/.

The voiced apical stop /d/ can be realized between a vowel and /a,o,u/ as either a voiced retroflex stop [ɖ] or a postalveolar tap [ɽ].[26] This seems to depend on speech tempo but also on the individual lexeme. In the frequent lexeme /pada/, 'PLURAL', this is almost always the case, whereas the lexeme /duduk/ 'sit,stay,be', which is about as frequent as /pada/, the tap is found less often. [ɽ] as a realization of /d/ must not be confounded with [ɾ] as a realization of /r/. The minimal pair [ɖaːɽa] 'chest' and [ɖaːɾa] 'blood' attests to this.

The palatal stops are realized as stops [c, ɟ] or as affricates [tɕ,dʑ]. Affrication seems to be more common for the voiceless palatal stop

/p, b, t̠, d̠, g, k/ and the prenasalized stops have no noteworthy allophones.

All mentioned stops with the exception of /ɖ/ occur simple or as geminates, but geminate /t/ only occurs in loanwords from Tamil, such as [kat̠ːil] 'bed'.

Fricatives   /s/ is realized as [s] and, rarely, /s/ can be realized as [h], as in the following examples, where the normal pronunciation would be [samːa] (cf. Bichsel-Stettler (1989:65), Smith et al. (2004)).

(39) kaɾaŋ inːi hamːa s-abis
      now   PROX all   PAST-finish
      'Now all was finished' (K060116nar11)

(40) siːni klaːki prompaŋ hamːa an̠t̠i-ʈaːⁿɖak
      here boy   girl   all   IRR=dance
      'Here, boys and girls will all dance (together).' (K060116nar10)

The alternation between /s/ and /h/ is also found in colloquial Sinhala (Matzel 1983:18), where it is completely generalized. In SLM, it is a very marginal phenomenon (cf. Bichsel-Stettler 1989:65).

/h/ is a rare phoneme. While other varieties of Malay make more use of /h/, historically present initial h- has been lost in many words in SLM everyday speech (Paauw 2004:ch.4.3). It is sometimes pronounced in a handful of words for reasons of prestige, [(h)aːɾi] 'day' being by far the most frequent one. Words which retain initial /h/ are [habːaɾ] 'news' and [hat̠ːu] 'INDEF', furthermore words of Arabic origin like [haɟːi] 'Hajji' or [halal] 'permitted for Muslims'.

---

[26]See Schiffman (1999:14) for a similar process in Spoken Indian Tamil.

The distinction between /ʃ/ (occurring only in loanwords) and /s/ is frequently ignored.

Nasals /m/ is labialized before long [oː], as in [mʷoːɲɛt̪] 'monkey'.

Within a morpheme, nasals are not distinctive in preconsonantal position. Their place of articulation is always homorganic to the following consonant ([mp, n̪t̪, n̪d̪, ɲc, ŋk, mb] etc.). Coincidentally, this is the only way to produce a dental nasal. Between morpheme boundaries, non-homorganic nasals can be found, e.g. [oːraŋpaɖa] 'man PL', where a velar nasal meets a labial stop. But even in this position, assimilation is often found (/oraŋ pada/→[oːrampaɖa]).

Word-finally, nasals are often rendered as velar ([seːlon~seːloŋ] 'Ceylon'), but this is not obligatory (Bichsel-Stettler (1989), Saldin (2001), Smith et al. (2004), *contra* Adelaar (1991)). The following words illustrate that the place of articulation is distinctive in word-final position:

(41)   aːram   : makanan : baːraŋ
       charcoal : food   : goods

The former two could be heard with a velar nasal, but do also occur with the indicated consonant, whereas the final word can never occur with any other nasal than the velar one. This shows that speakers can neutralize the distinction in word final position, but this is facultative.

Liquids Consonant length (see next section) entails a difference in manner of articulation for /r/. Short /r/ is an alveolar tap [ɾ]; long /r/ is an alveolar trill [r] (see Figure 4.4).

/l/ has no noteworthy allophones.

4.1.2.6   Consonant length

All consonants besides the prenasalized stops, ɖ and ʋ can occur lengthened in intervocalic position. Long vowels and long consonants are mutually exclusive. (42) shows a near minimal pair with a) a long vowel and a short consonant and b) a short vowel and a long consonant (see also Figure 4.5).

(42)   a.   soːpi CV:CV
            liquor

       b.   t̪opːi CVC:V
            hat
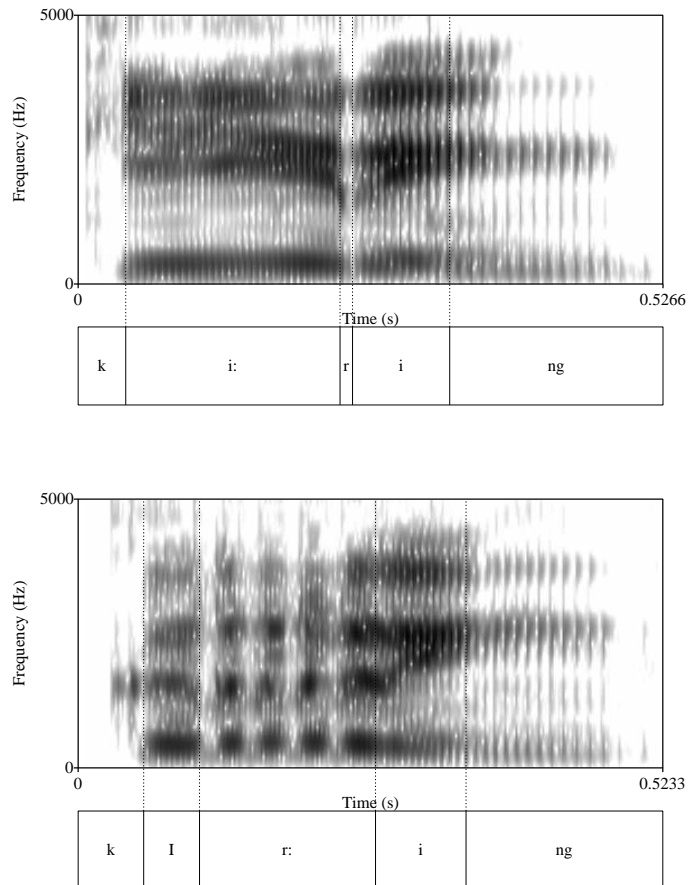
Figure 4.4: The words [kiːɾiŋ] 'send' (above) and [kɪrːiŋ] 'dry' (below) show a number of things. First, the formant structure of the first vowel is different, showing that [i] as a realization of /i/ must be distinguished from [ɪ] as a realization of /ə/. Second, they show the complementary distribution of long vowels and long consonants. The word above has a long vowel (206ms) and a short consonant (11ms), while the word below has a short vowel (50ms) and a long consonant (158ms). The spectrograms furthermore show that short /r/ is realized as a tap with minimal duration, while phonemically long/geminated r is realized as a trill.
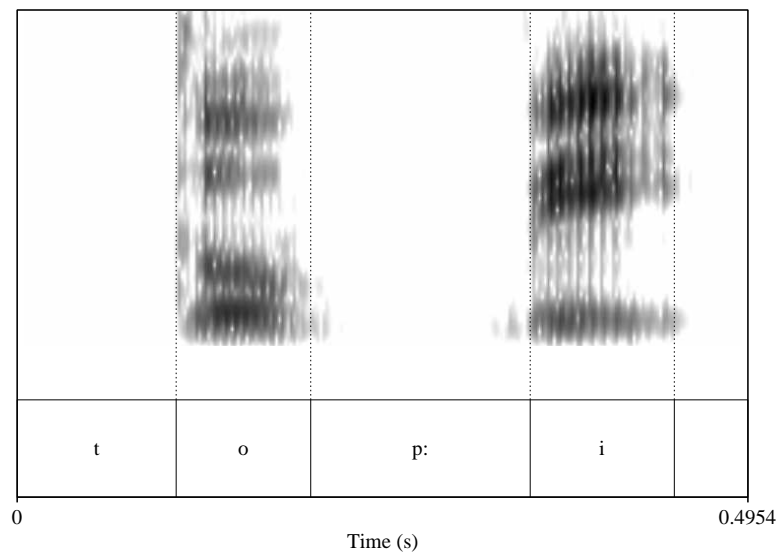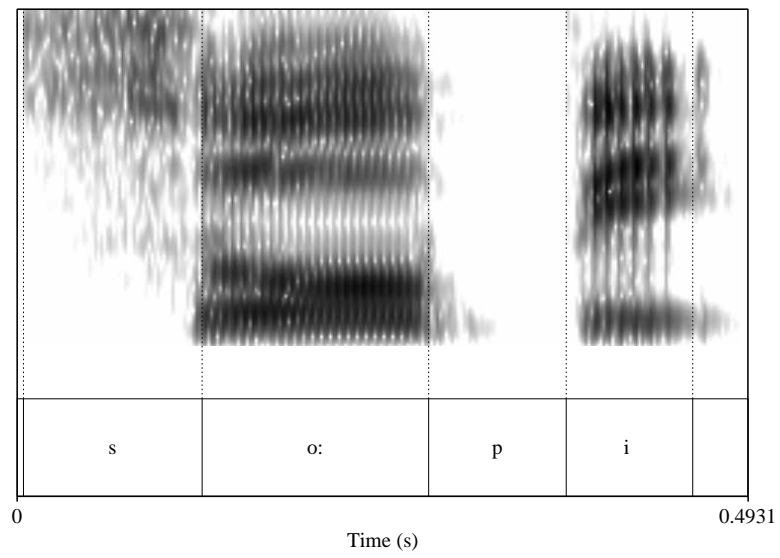
Figure 4.5: The word [soːpi] 'liquor' has a long vowel (153ms) and a short consonant (93ms), while the word [t̪opːi] 'hat' has a short vowel (91ms) and a long consonant (149ms).

Two types of long consonants can be distinguished on phonological grounds: Underlyingly long consonants, as in [t̪opːi] 'hat' or [ikːaŋ] 'fish', and lengthened consonants such as in [sɪgːaɾ] 'healthy'. The former always have a long consonant,[27] while the latter only have a long consonant if the consonant is located between penultimate and final syllable. If, through affixation, the consonant is found at another position, it is not geminated. Lengthened consonants can only occur after schwa, while underlyingly long consonants can occur after /a/ [at̪ːas] 'top', /e/ [pet̪ːa] 'parrot', /i/ [ikːaŋ] 'fish' and /o/ [t̪opːi] 'hat'. No long consonant has been found after /u/ or schwa. Tables 4.5 and 4.6 contain a column of words with long consonants.

There is only one instance of a long palatal glide, [maɟːet̪] 'corpse'.

Phonetically, long consonants are of about 1.5 times the duration of shorter consonants, as can be seen from Figure 4.5.

4.1.2.7 Distribution of consonants

All consonants may occur in intervocalic position, but d̪ is very rare there, one instance is [kɔd̪aɾat̪] 'miracle'. All consonants but ŋ and the prenasalized stops may occur in word-initial position, but the palatal glide is very rare there. There is only one word with initial ʈ, [ʈakːat̪a], which emphasizes the rapid nature of an event and might very well be onomatopoetic, so that it would not have to obey the general phonological rules.

There are more constraints word-finally. Voiced stops, apical stops, h, ɲ[28] and ʋ cannot occur in final position in native words, although some of those are frequently found in English loan words (*bag, pitch, bat*).

Few consonants ever occur in the coda of syllables inside a morpheme. Those that do occur are the nasals, the liquids, /k/ and /s/ (see Tables 4.5 and 4.6). If roots combine with an affix or a clitic with an onset, like [-kiŋ] 'CAUS' or [=pe] 'POSS', any consonant which can occur morpheme-finally can occur in a word-medial coda.

Tables 4.5 and 4.6 show examples for every consonant in initial and final position and also give examples of the consonant in intervocalic position (short and long) and morpheme-internal coda position.

4.1.3 Do the glides pattern with vowels or with consonants?

The labiodental approximant ʋ and the palatal approximant j are treated as phonological consonants in this grammar. Given that the approximants are on an intermediate position between consonants and vowels, it would equally be possible to treat them as vowels.[29] This fails to capture some important facts. First, tautosyllabic sequences

---

[27]In certain types of compound, like [kapːal+ʈɪɾbaŋ=kapalʈɪɾbaŋ] 'ship+fly=plane' the long consonant optionally disappears. These are rare.

[28]These four cannot occur word-finally in Sinhala, Tamil or Jakartanese either. -h# can occur in Std. Malay, but the other three cannot.

[29]Himmelmann (2005:116) reports that this problem is common in Austronesian languages and that, in most cases, the 'consonantal' analysis can be shown to be superior.

	initial	intervocalic short	intervocalic long	morpheme-internal coda	final
p	paːke 'dress'	baːpa 'father'	lʊpːas 'leave'	☐	ɡɪlːap 'dark'
ṯ	ṯaːna 'soil'	uːṯan 'forest'	aṯːas 'up'	☐	bansaṯ 'bedbug'
t	☐	baːʈok 'coconut.shell'	(kaṯːil) 'bed' (Tam.)	☐	☐
c	campuɾ 'mix'	kuːciŋ 'cat'	kicːil 'small'	☐	☐
k	kaːki 'leg'	baːkaɾ 'burn'	ikːaŋ 'fish'	ṯaksiɾ 'think'	saːnak 'relatives'
b	baːpa 'father'	ɾaːba 'stroke'	kʊbːɔŋ 'estate'	(ib.liːs) 'demon' (Ar.)	☐
ḏ	ḏaːṯang 'come'	(kɔḏaɾaṯ) 'miracle' (Ar.)	☐	☐	☐
d	daːɡiŋ 'meat'	aːɖa 'exist'	pʊdːas 'spicy'	☐	☐
ɟ	ɟaːlang 'walk'	blaːɟaɾ 'learn'	kiɟːa 'do'	☐	(haɟ) 'Hajj' (Ar.)
g	gaːɟa 'elephant'	gaːgak 'crow'	sɪgːaɾ 'healthy'	☐	☐

Table 4.5: Positions and gemination of stops. Boxed fields show positions not used in native words. If an example is given, it is from a loanword.

	initial	intervocalic short	intervocalic long	morpheme-internal coda	final
m	maːkaŋ 'eat'	saːma 'together'	samːa 'every'	lompa̪t 'jump'	ɡaːram 'salt'
n	naːsi 'rice'	aːnak 'child'	bɯnːaŋ 'thread'	ban̪tu 'help'	bin̪tan 'star'
ɲ	ɲaːri 'today'	mʷoːɲet 'monkey'	baɲːak 'much'	kɔɲci 'key'	□
ŋ	□	aːŋiŋ 'air'	t̪ɯŋːa 'middle'	maŋkok 'mug'	aːjaŋ 'chicken'
s	saːma 'with'	kaːsi 'give'	asːam 'sour'	miskin 'poor'	lʊpːas 'leave'
h	haːrum 'bad.smell'	laːher 'be.born'	ɹahːa̪t 'wicked'	□	□
r	raːba 'stroke'	ɡaːram 'salt'	bɪrːas 'rice'	ar̪tiyan 'meaning'	blaːɹar 'learn'
l	laːɖa 'pepper'	saːla 'wrong'	pʊlːam 'slow'	salba 'escape'	kapːal 'ship'
ʋ	ʋat̪ːu 'time'	kaːʋiŋ 'marry'	□	□	□
j	(jat̪iːm) 'orphan' (Ar.)	saːjaŋ 'love'	majːe̪t 'corpse'	□	bukulaj 'fight'

Table 4.6: Positions and gemination of other consonants. Boxed fields show positions not used in native words. If an example is given, it is from a loanword.

VG,[30] where G is one of the two approximants, only occur word-finally, like in [buku-laɟ] 'fight'. Sequences of VG do not occur before a consonant. Treating the approximants as vowels would lead to the postulation of complex nuclei VV only for final syllables, whereas treating them as consonants avoids this and is in line with general syllable structure in SLM, which permits the structure VC (Section 4.3).

Complex nuclei could be postulated for words like [bjaːsa] 'habit' or [pʋaːsa] 'fasting period', but this is not necessary either, since complex onsets are needed in this position anyway, illustrated by examples like [blaːɟaɾ] 'learn' or [cɾiːt̪a] 'story'.

Analyzing these approximants as vowels would needlessly complicate syllable structure, while analyzing them as consonants is fine with overall syllable structure, which permits final codas and complex onsets for the problematic positions where the approximants meet a vowel.

4.2   Suprasegmentals

The only suprasegmental which seems to of (albeit limited) relevance for SLM is vowel harmony. Vowel harmony might have or have had an influence on the realization of raised schwa as a front or back high vowel (see Section 4.1.1.2).

The following suprasegmental features have not been found to be relevant for SLM phonology: *stress, tone,* and *nasal harmony*. Nasal harmony is not present in many languages and its absence will not be discussed here. While tone is more common in the world's languages, its absence in SLM is not surprising either, given that neither the ancestor varieties nor the adstrates have lexical tone. Tone will not be discussed any further either, but see Section 4.8 for intonation patterns. A more intriguing feature is the absence of stress, which cannot be determined phonetically or phonologically. There are no cues of pitch or intensity, and vowel quality is inconclusive as well. This will be discussed in more detail in Section 4.5.4, after having established the general properties of the syllable.

4.3   Syllable structure

SLM syllables may have a complex onset and a simple coda. Both onset and coda are optional. Word-initially, /s-/ may precede nost of the consonants. Long vowels and a coda may not co-occur. The general syllable structure is then

(43)   (s)(C)(L)V(X)

where L is either /l,r,ʋ,j/ and X is either a consonant or the additional length of the vowel.

The following examples show

---

[30]Excluding geminates like in [maj̄ːet̥] 'corpse'.

- presence and absence of the onset (44),

- presence of a complex onset (45),

- presence of initial /s-/ with simple and complex onset (46),

- presence and absence of the coda (44), and

- mutual exclusion of long vowel and coda consonant (47).

(44) aːbu : caːbut̪   (45) ciːt̪ak : criːt̪a
      ash : remove   draw   : story

(46) kr̥iːɲat̪ : skr̥iːɲat̪   (47) soːpi : t̪opːi (t̪op.pi)
      sweat   : having.sweat   liquor : hat

Let us now turn to more detailed discussion of onset, nucleus and coda.

4.3.1 Onset

There are close to no restrictions on the occurrence of individual consonants in simple onsets of syllables. Any consonant or glide may occur in the onset of a syllable (cf. Tables 4.5 and 4.6), but /ŋ/ may not occur in the absolute onset of a word. Word internally, /ŋ/ is permitted in the onset, as [aːˑŋiŋ] 'air' shows.

Complex onsets are mainly found morpheme-initially. Exceptions I am aware of are instances of ə dropped between b and a liquid, like [sub(ə)la] 'side' or [lab(ə)ɾak] 'whack'. It is unclear as of now whether the b in these words is at the coda of the first syllable, at the onset of the last syllable, or whether these words are underlyingly still trisyllables. Other exceptions are [put̪r̥i] 'queen' and [kɔmplɔk] 'bush'.

Complex onsets are limited to combinations of stop or s combined with liquids or glides, and m+l and m+j, as shown in Table 4.7. Apical stops are not found in complex onsets, with the exception of [=d(ə)ɽiŋ] 'ABL' and [=d̪(e)ɾampad̪a] 'they'.[31] It is interesting to note that while all coronal voiced stops in native words are apical in intervocalic position, dental articulation can also be found in initial position, as shown by the pair [d̪uːʋa] 'two' and [d̪uːa] 'prayers'. Apical articulation is still more common overall, but if a liquid follows, dental articulation is the norm, as in [d̪laːpan] 'eight', [d̪raːpa] 'how much?', and [d̪raːka] 'hell'.

Complex onsets starting with /m-/ are also found. There are a number of /ml-/ onsets, but no /mr-/ onsets. Some informants claim that there is an onset /mɲ/ in [mɲaːlak] 'catch fire', but it was found that this word is is normally pronounced as [mjaːlak].

---

[31]It is unclear why the rhotic is retroflex in the former case and postalveolar in the latter.

[32]I would like to thank Uri Tadmor for suggesting this word.

	r	l	ʋ	j
p	[prompaŋ] 'girl'	[pliːʈa] 'lamp'	[puaːsa] 'fasting'	[pjaːra] 'adopt'
ʈ̪	[ʈ̪raː] 'NEG'	[t̪laːᵖjaŋ] 'naked'		
t				
c	[criʈa] 'story'	[claːna] 'trousers'		
k	[kreːʈa] 'cart'	[klaːpa] 'coconut'	[kuaːli] 'pot'	
b	[braːmak] 'birth'	[blaːkaŋ] 'after'	[buaːja] 'crocodile'	[bjaːsa] 'habit'
ɖ	[ɖraːpa] 'how.many'	[ɖlaːpan] 'eight'		
d	([d(e)ɽampaɖa] 'they')			
ɟ	[ɟraʊaʈ] 'pimple'32	[ɟleːna] 'window'		
g	[greːja] 'church'	[gluːɖup] 'thunder'		
s	[sraːʈus] 'one.hundred'	[slampe] 'handkerchief'	[suaːra] 'noise'	[sjaːmu] 'this person'
m		[mlaːraʈ] 'difficult'		[mjaːlak] 'catch fire'

Table 4.7: Complex onsets

Bichsel-Stettler (1989) states that syllables may not start with a vowel and that a preposed glottal stop is used to prevent this. This analysis is shared by Tapovanaye (1995). My data do not support this (See Section 4.1.2).

4.3.2 Nucleus

The nucleus normally consists of one vowel, which is long in open penultimate syllables of disyllables and trisyllables with initial schwa. Some words could be seen as having a rising diphthong [bjaːsa] 'habit', [sʋaːɾa] 'sound', but these are analyzed as complex onsets here.

Nasals left over after schwa deletion become syllabic, such as /əmpat̪/'four', realized as [ʋmpat̪] or [m̩pat̪] after deletion of schwa (see Section 4.1.1.2).

4.3.3 Coda

Syllables may have a simple coda. Morpheme-medially, the coda is most often identical to the following onset ([ik.kaŋ] 'fish'), or a homorganic nasal ([ban̪t̪u] 'help'), but /s, l, r, k/ are also found ([mis.kin] 'poor', [kal.t̪ɾa] 'unless', [kaɾ.cel] 'problem', [t̪ak.siɾ] 'think'). The velar nasal /ŋ/ is also found before /s/, as in [baŋsa] 'ethnic group', alongside the expected homorganic nasal found for instance in [bansat̪] 'bedbug'. Apical or palatal stops are never found in the coda, neither are voiced stops or /h/. Morpheme-finally, /p/ and /t̪/ are allowed in addition to the consonants named above, but /ɲ/ is impossible. Examples of coda position for the remaining consonants can be found in Tables 4.5 and 4.6 on page 106.

Given that voiced consonants never occur in coda position, the theoretical problem arises whether the coda stops are phonologically always voiceless, or whether some might be phonologically voiced and undergo final devoicing. The nominalizer *-an* can be used to test this because it has no onset. The root-final coda consonant is then resyllabified and found in onset position, where final devoicing is ruled out. It turns out that all coda consonants remain voiceless in this position, which shows that the lack of voicing is present in the underlying form and not caused by final devoicing. An example would be [t̪uːt̪up] 'close' and [t̪ut̪u{p/*b}an] 'closing'.

4.3.4 Initial /s-/

The phoneme /s/ can be found preceding an onset consonant. Sometimes this can be analyzed as reduction of the nucleus as in [s(ə).laː.mat̪], 'greetings' or [(a)s.maː.kaŋ] 'having eaten'. In these cases, the non-reduced form also exists. In other cases however, the non-reduced form does not exist. Cases in point are [spaːru] 'some' and [sbiːlan] 'nine'. Table 4.8 gives an overview of the possibilities.

In line with general findings about the syllabification of elements which do not show lower sonority than the element they precede, /s-/ before stops would be extrasyllabic (Halle & Vergnaud 1980, Wiese 1986), and would form a complex onset before

	labial	dental	apical	palatal	velar
voiceless stop	spaːɻu 'some'	stɨɲːa 'half'	n/a	(scuːci) 'cp-wash'	skaːɻaŋ 'now'
voiced stop	sɓiːlan 'nine'	(sɖaːʈaŋ) 'cp-come'	sɖiːkit̪ 'few'	(sɟaːlaŋ) 'cp-walk'	(sgaːli) 'cp-dig'
nasal	(smaːʈi) 'cp-dead'	n/a	snaːpan 'gun'	(sɲaːɲi) 'cp-sing'	n/a
lateral	n/a	n/a	slaːlu 'sad'	n/a	n/a
rhotic	n/a	n/a	sɾiːbu '1,000'	n/a	n/a
approximant	suaːɾa 'sound'	n/a	n/a	sjaːɾu 'this.person'	n/a

Table 4.8: S-clusters. S-clusters are possible in many combinations. These are indicated here. S-clusters cannot be formed with fricatives in SLM. Phonemes which are either not found at all in SLM, or not found in initial position, cannot form S-clusters. This is indicated by n/a. Some S-clusters could not be found in monomorphemic words, but only in combination with the conjunctive participle prefix [(a)s(ə)-]. These are indicated by parentheses.

sonorants, according to the Core Syllabification Principle (Clements 1990:299). A representation of extrasyllabic /s-/ is given in (48) for [sbantu] 'having helped', while a representation for the complex onset is given in (49) for [slampe] 'handkerchief'.[33]

(48) (49)
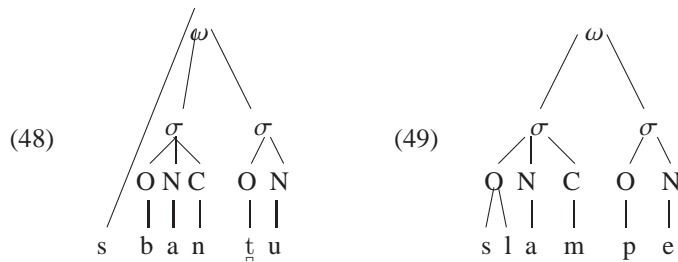
4.4 The internal structure of the word

Phonological words have an internal structure, which is based on feet, syllables and moras in SLM.[34] Furthermore, the structure of the morphological word has an influence on the structure of the phonological word. On the morphological level, we distinguish between roots, stems and words. Roots consist of one lexeme and no additional material, stems consists of one or more lexemes plus optional additional derivational morphemes, while the word consists of a stem and inflectional material. This is illustrated in (50). Outside of the domain of the word, there can be clitics attached, which do not form part of the word proper, but are interesting for delimiting purposes. This morphological structure of the word can thus be represented as follows:

(50) (CLT) (INFL) [LEXEME ROOT (LEXEME) (DERIV) STEM (INFL)] WORD
     (CLT)
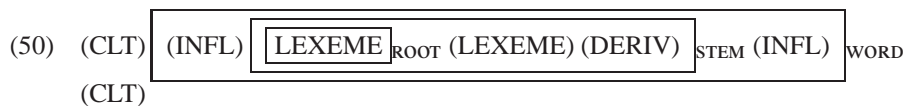
Generally speaking, the difference between roots and stems is not very important in SLM. Both are normally parsed into one phonological word. Material outside of the stem, like inflectional affixes and clitics, is not parsed into a phonological word with the stem.

---

[33]Unnecessary detail is removed from these representations. More detailed representations can be found below.

[34]This section would never have been as detailed as it is without the help of Diana Apoussidou, for which I am extremely grateful. Most of the theoretical insights presented here are hers. All misrepresentations are my own responsibility.

4.4.1 Roots

Sri Lanka Malay has roots consisting of two and three syllables, plus a handful of monosyllables. The polysyllables come in different types, listed in (51) (We disregard the difference between simple and complex onsets for this list). In the column for syllable structure, ə represents a phonemic schwa, which can be realized as [ə,ɪ,ʊ,i,u] phonetically. We can distinguish 5 groups: monosyllables, disyllables without schwa, disyllables with schwa, problematic disyllables, and trisyllables. At the syllable break, we can distinguish a short consonant (V.C), a long consonant ($C_i.C_iV$), and a hetero-syllabic cluster ($C_i.C_jV$).

	1)	CVː	e.g.	[piː] 'go', [t̪eː] 'tea', [caː] 'tea', [baː] 'bring'
	2)	CVC	e.g.	[pon] 'bride'[t̪aj] 'excrement'
	3)	(C)V.CV	e.g.	[kat̪a] 'QUOT', [ini] 'DIST'[35]
	4)	(C)Vː.CV(C)	e.g.	[t̪iːga] 'three', [uːt̪aŋ] 'forest'
	5)	(C)VC$_i$.C$_j$V(C)	e.g.	[kumpul] 'collect', [ont̪a] 'camel'
	6)	(C)VC$_i$.C$_i$V(C)	e.g.	[t̪opːi] 'hat', [at̪ːas] 'top'
(51)	7)	CəC$_i$.C$_j$V	e.g.	[mɪn̪t̪a] 'vomit'
	8)	(C)əC$_j$.C$_j$V(C)	e.g.	[sɪgːaɾ] 'healthy', [t̪ɪŋːa] 'middle'
	9)	(ə)N.CVC	e.g.	[(ʊ)mpat̪] 'four'
	10)	(ə)C$_i$.C$_i$V(C)	e.g.	[(ʊ)mːa] 'mother', [(ɪ)nːam] 'six'
	11)	CV$_i$.h$_i$C	e.g.	[pohoŋ] 'tree'
	12)	CV.GVC	e.g.	[lijat̪] 'see', [t̪uʋan] 'gentleman'
	13)	CVː.GVC	e.g.	[t̪uːʋa] 'old', [ɟuːʋal] 'sell'
	14)	CV.CV.CV(C)	e.g.	[kut̪umuŋ] 'see', [nigiɾi] 'country'
	15)	Cə.CVː.CV(C)	e.g.	[məd̪eːra] 'flag', [cəcaːʋak] 'wash'
	16)	CVC.CVC.CVC	e.g.	[kakːaɾlat̪] 'cockroach' (only example)

The following generalizations emerge from this list:

- there is maximum of one long vowel per word

- long vowels are always found in the penultimate, with the exception of a few monosyllabic words

- /ə/ in penultimate syllables is never realized as [ə], but always raised

- no allophone of /ə/ is ever long

- penultimate syllables with /ə/ are always closed

- long vowels never occur in closed syllables

- penultimate syllables of disyllabic words are always heavy ((C)VC or (C)Vː) with the exception of function words and words with a glide or h as onset of the final syllable

- long vowels are never preceded by a syllable with [a,e,o].

These generalizations point to an important role of the weight of the penultimate syllable. An analysis of these generalizations will be given in Section 4.5.

4.4.2 Stems

Derivational suffixes like the nominalizer *-an* add to the size of a root. A disyllabic word can become trisyllabic through suffixation. The resulting trisyllable conforms to the same word types listed above in (51) for roots. An example of this is [baːlas] 'answer(V)'+[-an] 'ɴᴍʟᴢʀ'=[balasan] 'answer(N)' with a short vowel (Figure 4.6). [*balaːsan] or [*baːlasan] are not structures listed in (51), and are disallowed.

There is one additional structure for stems which is never found in roots, namely

- Cə.CV.CVC e.g. [sɪgaɾ-an] 'healthy'+'ɴᴍʟᴢʀ'='health', where the schwa is always realized as [ɪ].

This structure has no long vowel, but a raised schwa in the first syllable. This stem structure contrasts with the structure of trisyllabic roots, where a schwa in the initial syllable triggers a long vowel in the penultimate. Additionally, schwa in trisyllabic roots is always realized as [ə], while here it is never realized as [ə], but always as [ɪ].

The second category-changing suffix next to the nominalizer *-an* mentioned above is the causativizer *-kiŋ*, which can behave like *-an* in making a word trisyllabic, resulting in a short vowel. The behaviour of *-kiŋ* is a bit more complex though and is treated in more detail on page 130.

Similar behaviour is found for the numeral derivational suffixes [-blas] '-teen' and [-pulu] '-ty', which disallow long vowels.

The derivational prefixes, [ka-] 'ᴏʀᴅ' and [kənə-] 'ɪɴᴠᴏʟ', do not have any influence on the vowel length of their host.

Compound stems like [kaca maːt̪a] 'mirror+eye=spectacles' can only have long vowels in the last part of the compound. This is true for nominal compounds like [kaca maːt̪a], but also for verbal compounds like [kasi kaːʋiŋ] 'give+marry=marry off'.

4.4.3 Words

The structure of stems is not changed by adding additional material on the word level, such as inflectional affixes or clitics. For instance, the quantity of the vowels remains the same if a clitic like [=pe] 'ᴘᴏss' is attached to the stem, e.g. [baːpa] 'father', [baːpa=pe] 'father's' with a long vowel in both, or [makanan=pe] 'of the food', with a short vowel as in [makanan]#. An inflectional suffix like the imperative *-la* would not change the vowel quantity of the stem it attaches to, either.
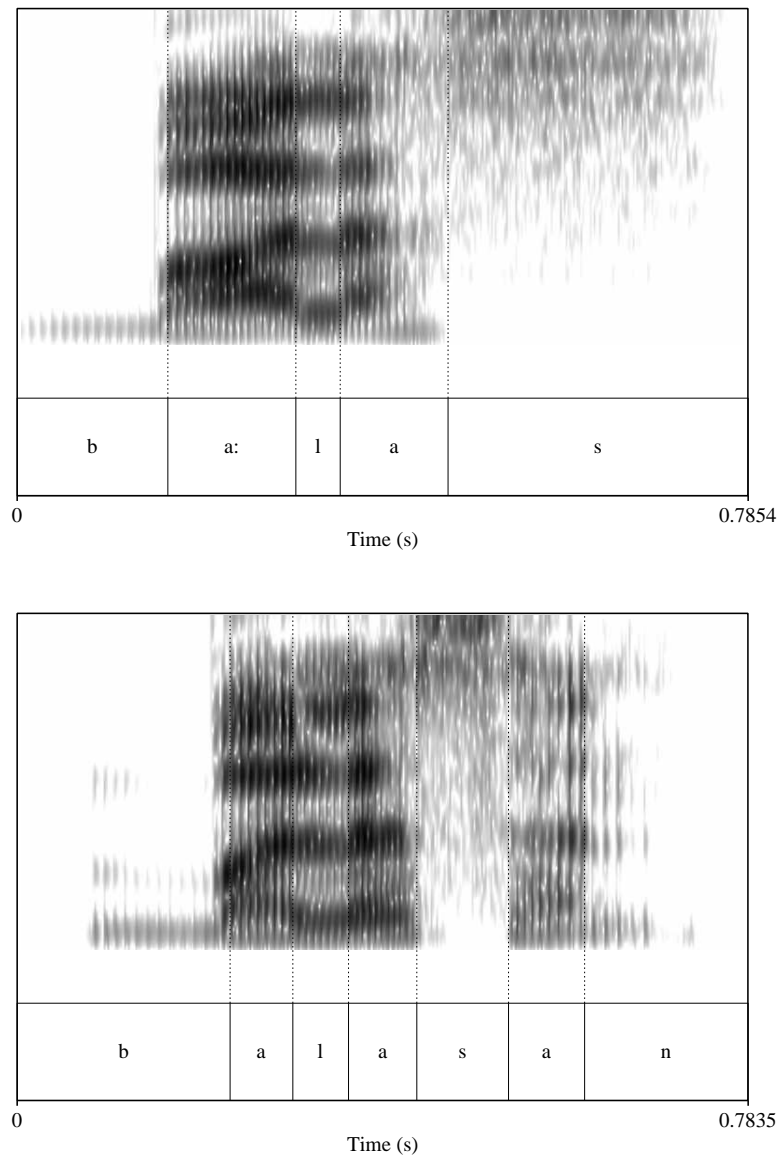
Figure 4.6: The word [baːlas] 'answer(v)' has a long first vowel (138ms). This same vowel is realized as short (67ms) upon the addition of the nominalizing suffix -*an* .

4.4.4 Word length

Disyllabic roots are most common (ca. 90%), while trisyllabic roots are more rare (ca. 10%). Only a handful of lexical roots consists of one syllable: [pii] 'go', [pon] 'bride', [t̪eː] 'tea', and [caː] 'tea', besides the function words [seː] '1s.polite', [goː] '1s'familiar, [luː] '2s.familiar', and [deː] '3s.familiar'. More than three syllables are found in the words [kenijaːja] 'in vain' and [bukaɾijɔk] 'cock-a-doodle-do'.

By affixation and compounding, longer words can be formed, but more than five syllables are rarely found. The longest *bona fide* word (i.e. without clitics) found up to now is [kə-sbilan-pulu] 'nintieth'. If clitics are included, there is technically no longest word because clitics can be stacked infinitely (see Section 6.4, p. 317).

4.4.5 Syllable prominence

Sri Lanka Malay words do not show different levels of prominence in their syllables, neither phonologically nor phonetically. In other words, there is no word accent or stress. Smith et al. (2004) claim that "stress falls on the long vowel of a word, or on the initial syllable if the word only contains only short vowels." They do not state by which acoustic cue they established the location of stress, so that we are probably dealing with an impressionistic statement here. Bakker (2006) reports accent on the first syllable, drawing on Robuchon (2003). These statements again seem to be impressionistic. Some instrumental analyses were undertaken to check these claims, but no cues for a saliently prominent (stressed) syllable could be found. 49 words with different syllable structures were checked in 6 different controlled environments with two male native speakers of SLM for pitch contour and intensity. The compiled measurements taken together do not suggest that a meaningful generalization about the correlation between word structure and pitch or intensity can be made (Apoussidou & Nordhoff 2008).

Figure 4.7 shows that neither pitch nor intensity distinguish the first two syllables from each other in the trisyllabic word [kut̪umuŋ] 'see'(The effects on the final syllable are due to the utterance boundary).[36] Instrumental analysis of a sample of words with different syllable structures in varying positions within the clause showed that *kut̪umuŋ* is no exception and that pitch and intensity do not provide cues of prominence (Apoussidou & Nordhoff 2008). This lack of word accent is actually not unheard of in the Malay world, as will be discussed in more detail in Section 4.5.4.

---

[36]I am aware of the fact that the historical source for this word had two schwas in the first two syllables, so that the lower intensity could be attributed to that. I do not think that this word has a synchronic schwa, though. The reason for this is that it behaves exactly like trisyllabic words without schwa like [makanan] 'food'. This word was chosen because it is a monomorphemic trisyllable, of which there are not many. Polymorphemic trisyllables are easier to come by, e.g. nominalizations with *-an*. These polymorphemic words were not retained for the recording session in order to avoid influences that the affixes might have on stress placement. In hindsight, this was a mistake. Future research should include some of these affixes trisyllables as well. I am confident that future research on pitch and intensity will provide no more stress cues for these trisyllables than they did for disyllables or the kut̪umuŋ-type words.
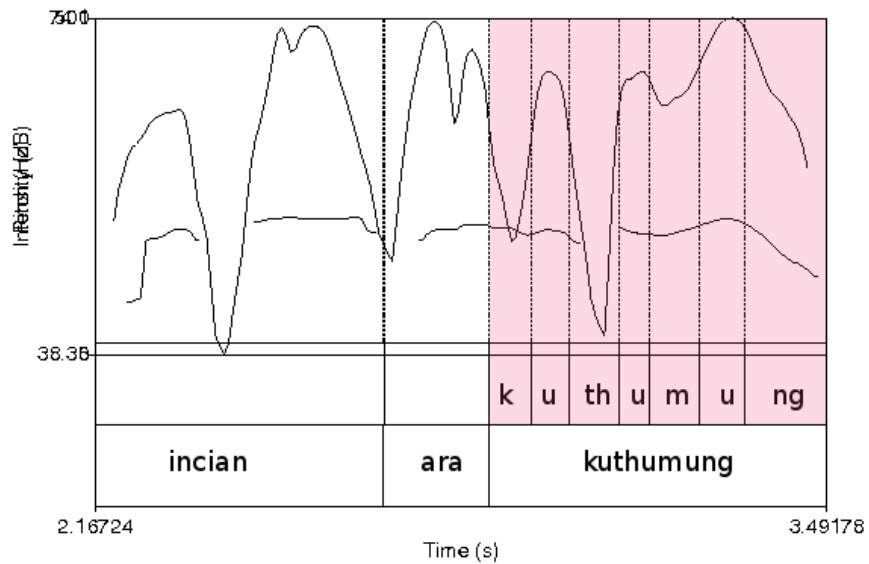
Figure 4.7: Pitch and intensity of an utterance involving the word [kuṯumuŋ] 'see' show that the first two syllables are not very different in intensity and pitch. Intensity and pitch cannot be used to determine which of the two should be the head. The final syllable has a higher loudness and a raise-fall pitch movement, both of which are the usual way of indicating utterance boundaries of declaratives.
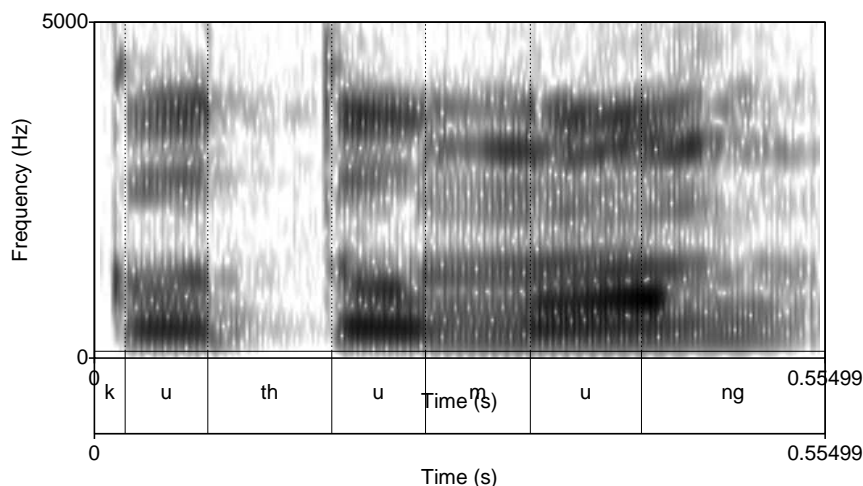
Figure 4.8: The spectrogram of [kut̪umuŋ] 'see' shows that all three u's are of about equal duration. Duration is therefore not a cue to determine syllable prominence.

Another stress cue which some languages use to indicate prominence of a syllable is vowel duration. This is not the case in SLM either. The spectrogram of *kut̪umuŋ* (Figure 4.8) shows that the vowels are of comparable length here (also see Figure 4.6 for a derived trisyllable). True, there are long and short vowels in SLM, as shown in Section 4.1.1.3, but if this was a stress cue, we would expect a difference in vowel duration (or at least moraic weight) to show up in every word, namely on the prominent syllable. This is not the case, as words like *kut̪umuŋ* show. The hypothesis makes wrong predictions, and must be amended or abandoned. A possible amendment would be to state *Trisyllables of the* kut̪umuŋ *type do not show stress.* This amendment would have to rely on lexical specification. The fact that there are differences in vowel duration needs explanation nevertheless. In Section 4.5, I will argue that vowel duration can better be explained by foot structure, and that this analysis does not need specification in the lexicon and is therefore superior to the "stress analysis".

The last stress cue which some languages use is vowel quality. Prominent syllables typically show a greater range of possible vowels, whereas non-prominent syllables have a smaller array and might only admit reduced vowels. In SLM, the antepenultimate syllable of a root seems to be a less prominent syllable, since it only admits [i], [u] and schwa (the latter often dropped altogether), but not [a], [e] or [o], the most sonorant vowels (Jespersen 1904, Selkirk 1984, Clements 1990).[37] While this is an

---

[37]Antepenultimate syllables are often restricted to schwa in Malay varieties, see Adelaar (1985:12,31,59) for examples of Standard Malay and Jakartanese.

indication of lack of prominence of this syllable, it does not indicate which of the remaining syllables is most prominent. Furthermore, this lack of prominence of the antepenultimate only applies on the root level; on the stem level, any vowel can be found in antepenultimate position (see Section 4.4.2).

To conclude, pitch and intensity do not distinguish syllable prominence in SLM. Differences in vowel duration are not found in all words and are therefore a bad stress cue. Vowel quality finally could indicate lack of prominence of one syllable on the root level. It does not provide positive prominence cues on the stem level, and does not provide any cues on the stem level.

4.5   Analysis of word structure

I will argue in this section that the different types of roots we find in SLM and the occurrence of long vowels can be explained by extrametricality of the final syllable (Kiparsky 1985) and a bimoraic foot requirement[38] (McCarthy & Prince 1993). The list in (51) is repeated and reordered here for convenience.

(52)

1)	$(C)VC_i.C_jV(C)$	e.g.	[kumpul] 'collect'
2)	$(C)VC_i.C_iV(C)$	e.g.	[ṭopːi] 'hat'
3)	CV.CV.CV(C)	e.g.	[kuṭumuŋ] 'see'
4)	$(C)Vː.CV(C)$	e.g.	[ṭiːga] 'three'
5)	CVː.GVC	e.g.	[ṭuːʋa] 'old' (G=glide)
6)	Cə.CVː.CV(C)	e.g.	[cəcaːʋak] 'wash'
7)	$CəC_i.C_jV$	e.g.	[mɪṃṭa] 'vomit'
8)	$(C)əC_j.C_jV(C)$	e.g.	[sɪgːaɾ] 'healthy'
9)	CVː	e.g.	[piː] 'go', [ṭeː] 'tea'
10)	CVC	e.g.	[pon] 'bride'
11)	(C)V.CV	e.g.	[kaṭa] 'ǫuoɪ'
12)	CVC.CVC.CVC	e.g.	[kakːaɾlaṭ] 'cockroach' (only example)
13)	CV.GVC	e.g.	[lijaṭ] 'see'
14)	$CV_i.h_iC$	e.g.	[pohoŋ] 'tree'
15)	(ə)N.CVC	e.g.	[(ʋ)mpaṭ] 'four'
16)	$(ə)C_i.C_iV(C)$	e.g.	[(ʋ)mːas] 'gold'

4.5.1   Roots

The following analytic generalizations can be drawn from this list: Within stems, a penultimate syllable is heavy iff it is not preceded by another syllable containing a full vowel, where heavy syllables are CVː and CVC. I propose that this distribution can be captured by the following principles:

- treat final syllables as extrametrical

---

[38]It is interesting to note that Sinhala also parses words into bimoraic feet (Letterman 1997).

- build bimoraic feet at right edge of stem

- ə cannot project a mora (Kager 1989)

- ɪ and ʊ can project a mora

- reduced vowels [ə,ɪ,ʊ] cannot lengthen

From these principles, it follows that disyllabic stems lengthen the vowel in an open penultimate syllable to build a bimoraic foot. The same is true for trisyllables with initial schwa. This will be developed below for the different types.

4.5.1.1 Canonical cases

A straightforward example of the internal structure of the root is /kumpul/'collect'. This is parsed as /(ku$_\mu$m$_\mu$) <pul>/, whose representation is given in (53).

(53)
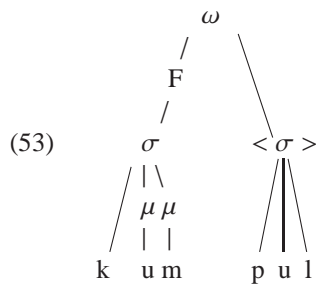
We exclude the final syllable *pul* and parse the remaining syllable *kum* in a bimoraic foot, where the vowel and the coda consonant contribute one mora each.

For disyllables with an underlying geminate, such as [ʈopːi] 'hat', the representation is similar, but the coda consonant of the first syllable serves at the same time as onset of the final syllable.

(54)
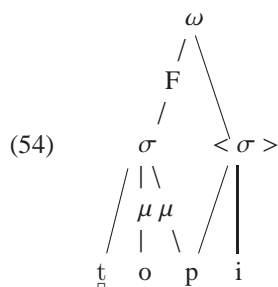
Trisyllabic words like [kuʈumuŋ] 'see' have the foot stretching over two syllables.

(55)

$$
\begin{array}{c}
\omega \\
/ \quad \backslash \\
\mathrm{F} \\
/ \quad \backslash \\
\sigma \quad \sigma \quad <\sigma> \\
| \quad | \\
\mu \quad \mu \\
| \quad | \\
\text{k u} \quad \underset{\sim}{\mathrm{t}}\,\text{u} \quad \text{m u ŋ}
\end{array}
$$
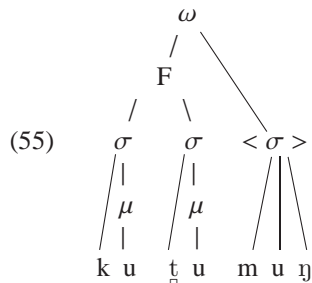
In disyllables without coda consonant in the penultimate, like [ṯiːga] 'three' (56), the vowel is lengthened to contribute the second mora for the bimoraic foot.

(56)

$$
\begin{array}{c}
\omega \\
/ \quad \backslash \\
\mathrm{F} \\
/ \\
\sigma \quad <\sigma> \\
| \, \backslash \\
\mu \, \mu \\
\backslash \, / \\
\underset{\sim}{\mathrm{t}} \quad \text{i} \qquad \text{g a}
\end{array}
$$
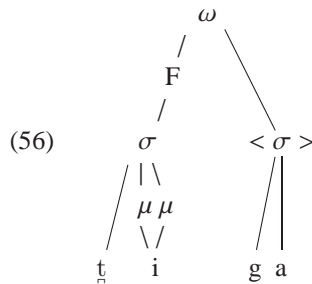
For one set of lexemes with a glide as the onset of the final syllable, the representation is the same as for [ṯiːga], just above. The word [ṯuːʋa] 'old' has an identical representation (57) to [ṯiːga] above.[39]

(57)

$$
\begin{array}{c}
\omega \\
/ \quad \backslash \\
\mathrm{F} \\
/ \\
\sigma \quad <\sigma> \\
| \, \backslash \\
\mu \, \mu \\
\backslash \, / \\
\underset{\sim}{\mathrm{t}} \quad \text{u} \qquad \text{ʋ a}
\end{array}
$$
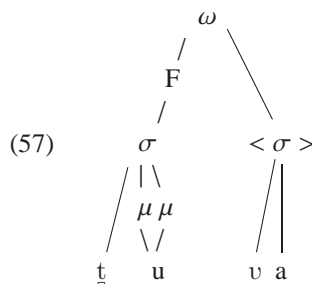
In trisyllables with an initial schwa and an open penultimate like [cəcaːʋak] 'wash', the vowel in the penultimate is also lengthened. This indicates that ə does not project a mora (cf. Kager 1989). The foot is then constructed only with the material of the penultimate syllable, which means that the vowel of this syllable needs to be lengthened.

---

[39]This contrasts with a more problematic case for [ṯuʋan] 'gentleman', shown below in (69).

(58)

$$
\begin{array}{c}
\omega \\
| \\
F \\
| \\
\sigma \quad \sigma \quad <\sigma> \\
\mu\,\mu \\
c\ \vartheta \quad c\ a \quad \upsilon\ a\ k
\end{array}
$$
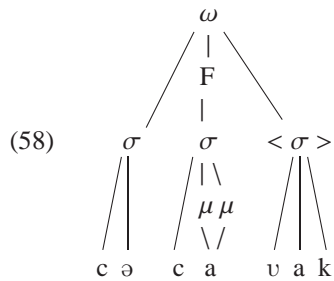
If schwa occurs in a disyllabic like [mɪn̪t̪a] 'ask', schwa is realized as ɪ, which, unlike ə, can project a mora. This raising of schwa is then a result of the need for an additional mora, since the coda consonant (in this case /n/) only provides one, not enough to construct a bimoraic foot.

(59)

$$
\begin{array}{c}
\omega \\
F \\
\sigma \quad <\sigma> \\
\mu\,\mu \\
|\ \ | \\
m\ \ \ \imath\ \ \underset{.}{n} \quad \underset{.}{t}\ a
\end{array}
$$
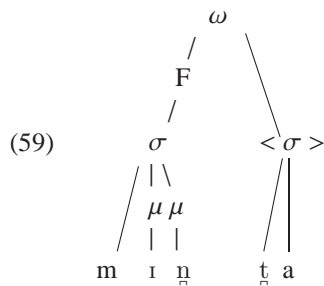
If the initial syllable of a disyllabic word contains schwa in an open syllable, moraic content is neither present in the vowel, since ə does not have moraic content, nor in the coda, which is absent. In this case, two things happen: ə is raised to ɪ, which does project one mora, and the onset of the following syllable is geminated, yielding a coda, and thereby an additional mora.

(60)

$$
\begin{array}{c}
\omega \\
F \\
\sigma \quad <\sigma> \\
\mu\,\mu \\
|\ \ \backslash \\
s\ \ \ \imath\ \ g \quad a\ r
\end{array}
$$
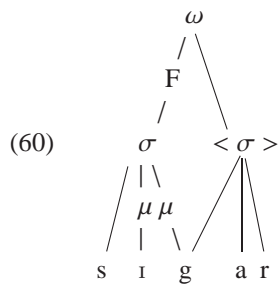
There are only a handful of monosyllabic lexical words such as [piː] 'go', [pon] 'bride' and [t̪aj] 'excrement'. Additionally, there are the pronouns [seː] '1s.ᴘᴏʟɪᴛᴇ', [goː] '1s.ꜰᴀᴍɪʟɪᴀʀ', [luː] '2s.ꜰᴀᴍɪʟɪᴀʀ', [deː] '3s.ꜰᴀᴍɪʟɪᴀʀ'. The fact that all monosyllables have a heavy syllable, especially the long vowel in piː suggests that the final

syllable is not treated as extrametrical for these words. Instead, the bimoraic foot is
constructed over the whole word.

(61)   (62)
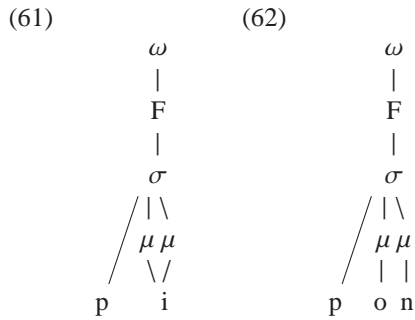
No function word has a long vowel, even not in penultimate open syllables. This
suggests that function words are not subject to the bimoraic foot requirement, and are
probably not parsed into a foot nor a word. (63) shows this for the quotative *kaṭa*.
In this representation, the levels of word, foot and mora are left out, since they are
irrelevant for the word *kaṭa*.

(63)
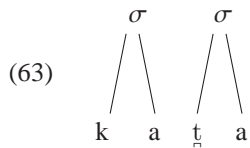
Trisyllabic roots with closed syllables are rare. Only one could be found, [kakːaɾlaṭ]
'cockroach'. It is not clear as of now whether this root is parsed into one bimoraic foot,
disregarding the first syllable (64), or into two feet, as in (65).

(64)   (65)
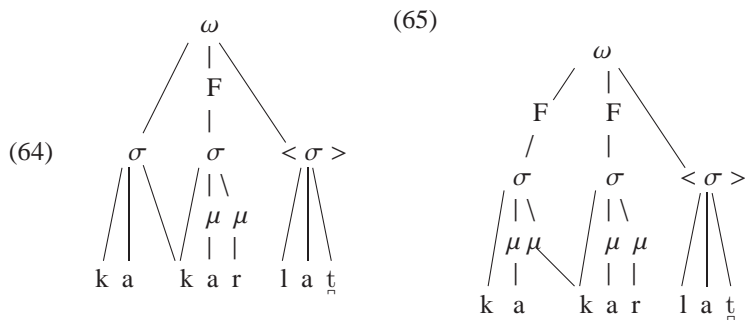
4.5.1.2 Problematic cases

A more problematic case for this analysis are co-occurrence of empty coda and empty onset. Our analysis would predict that for words of the form /CV$_i$V$_j$C/, we get [CV$_i$:V$_j$C] (66). However, what we find is [CV$_i$GV$_j$C] where G is a glide. This newly formed glide is resyllabified to the onset of the following syllable, yielding [CV$_i$.GV$_j$C]. This resyllabification means that the glide can no longer be the coda, unless it is geminated, which is not the case here. With the coda disappearing, the mora also goes, leaving insufficient material to construct a foot. This is illustrated in (66) and (68).

(66)   *   (67)

(68)   *   (69)
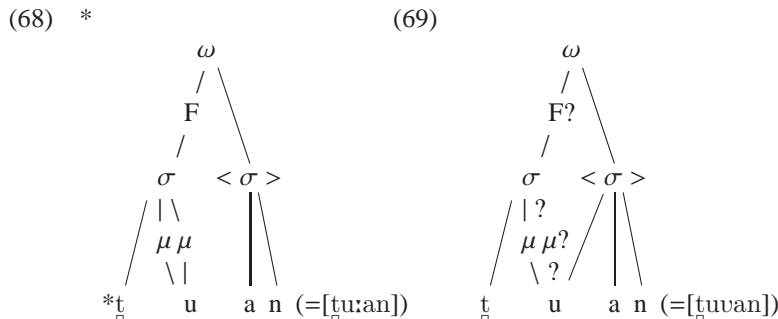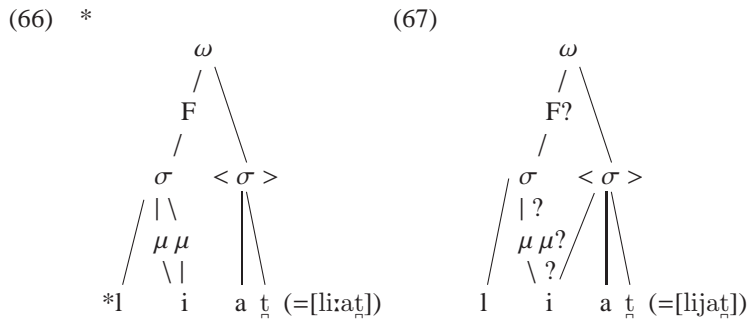
Under derivational accounts, it is easy to formulate a rule like

(70)   [+*vocalic*]:→ [+*vocalic*][−*vocalic*] / _[+*vocalic*]

In other words, a long vowel is split into a vowel and a glide before another vowel. It is unclear as of now how this could be made to work with the general foot structure.

The problem with words of the structure CV$_i$.hV$_i$C like [pohoŋ] 'tree' is similar. A vowel preceding the same vowel, separated by an /h/ is never lengthened. Tapovanaye (1995) assumes that these words have the underlying form /CV$_i$V$_i$C/ and that an h is inserted to prevent hiatus. Crucially, this takes place after vowel lengthening. In such a derivational account, this can be formulated as

(71) $V_i : \rightarrow V_i h/\_V_i$

This phenomenon cannot be captured with the analysis presented there. The bimoraic foot cannot be constructed for these words. On the other hand, the derivational account presented by Tapovanaye is not very plausible either, since the insertion of an h instead of a glide is not very well motivated, and historically, an h is present in these words anyway.

Another problem for the foot structure assumed in this analysis is the deletion of schwa between empty onset and nasals as in /əmpat̪/, which can be pronounced [ʊmpat̪], but also [m̩pat̪]. For the first pronunciation, there is no problem (72), but if the raised schwa is deleted, as is the case in the second pronunciation, a mora disappears and no bimoraic foot can be constructed. Whether the final syllable should be extrametrical (73) or not (74) is unclear as of now.

(72)   (73)   (74)
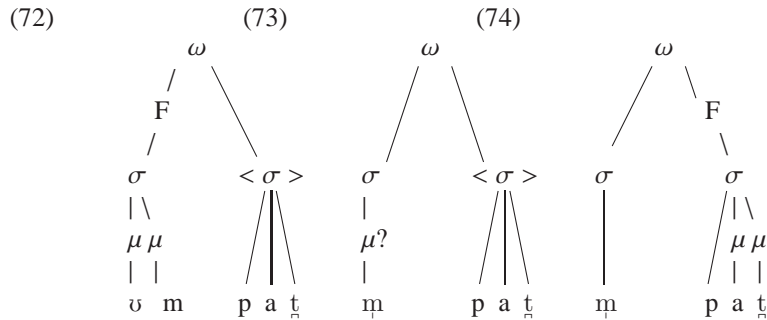
For the very last type in list (52), the (ə)$C_i C_i$V(C) type represented by [ʊmːas] 'gold', the discussion is analogous to [ʊmpat̪] 'four'.

(75)   (76)   (77)
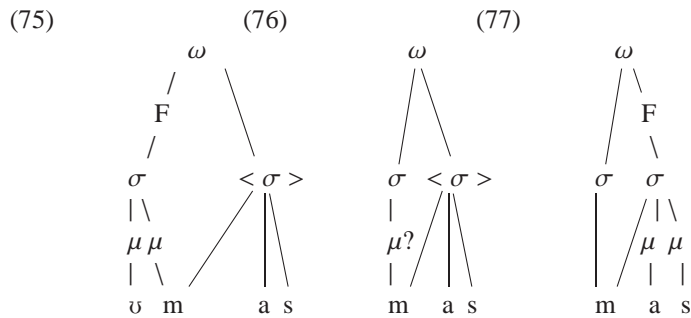
4.5.2   Stems

Stems consist of an obligatory root and optionally additional roots or derivational morphemes.

There are only a handful of monosyllabic roots, which do not combine well with each other nor with the derivational morphemes. This means that all bimorphemic stems have at least one disyllabic root, which makes the complex stems at least trisyllabic. These polysyllabic stems respond to the same criteria as roots, i.e. extrametricality of the final syllable and construction of a bimoraic foot.

4.5.2.1 Canonical cases

The additional material provided by the derivational affix or the second root in the compound means that more moraic content is available, and that strategies to provide missing moraic content, like vowel lengthening, gemination, or raising of schwa are less necessary with complex stems than with roots. If, through affixation or compounding, other possibilities to construct bimoraic feet arise, the vowel of the penultimate is not lengthened. The syllable containing the suffix becomes extrametrical, and the formerly extrametric final syllable now is penultimate and can provide moraic content. $[(ma_\mu ka_\mu)<nan>]$ 'food' has two moraic vowels in the first two syllable, canceling the need for an additional vowel for the second mora, which is necessary in $[(ma_\mu:_\mu)<kaŋ>]$ 'eat'. Both representations are given in (78) and (79)
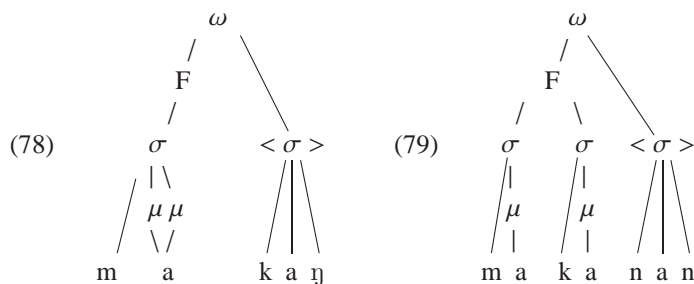
For disyllables with schwa in open syllables, material provided by affixes cancels the need for one extra mora. This means that gemination is no longer necessary, while realizing schwa as [ɪ] is. The representations in (80) show the difference in gemination between the disyllable and the nominalized trisyllable.
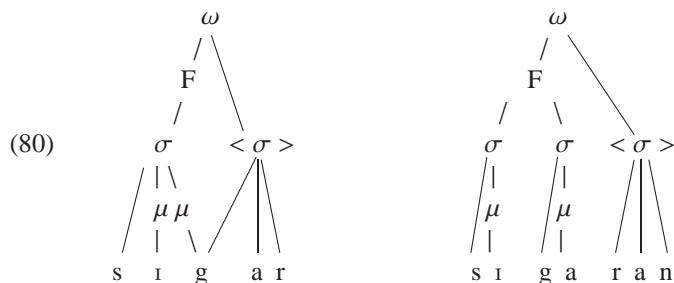
Another logical possibility would be to realize schwa as [ə] and geminate the onset of the final syllable. This would also provide an extra mora, but this solution is not employed. All geminated roots lose their gemination when nominalized.

The derivational suffix *-blas* is used to derive the numbers from thirteen to nineteen. It behaves like *-an* in the sense that it is parsed into the same phonological word as its host, which allows the formerly extrametrical final syllable to contribute moraic weight to the bimoraic foot, abolishing the need for vowel lengthening.

(81)
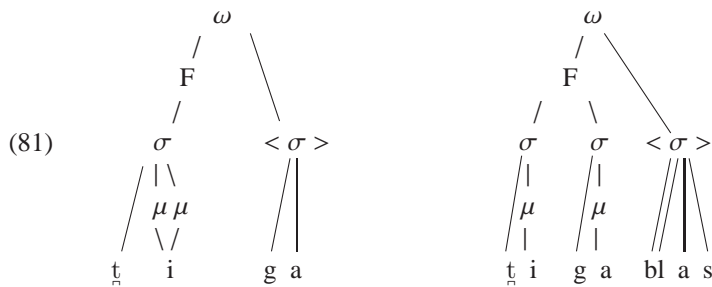
Compounding, just like affixation, provides extra material and influences the construction of feet and the need or not for additional moras. In (82), the words [kaːca] 'mirror' and [maːʈa] 'eye' are compounded. Both orders are possible, and the resulting meaning is 'spectacles' in both cases. The resulting word cannot be parsed into feet exhaustively, which is why the last possible vowel is lengthened, in this case the vowel in the penultimate syllable.

(82)
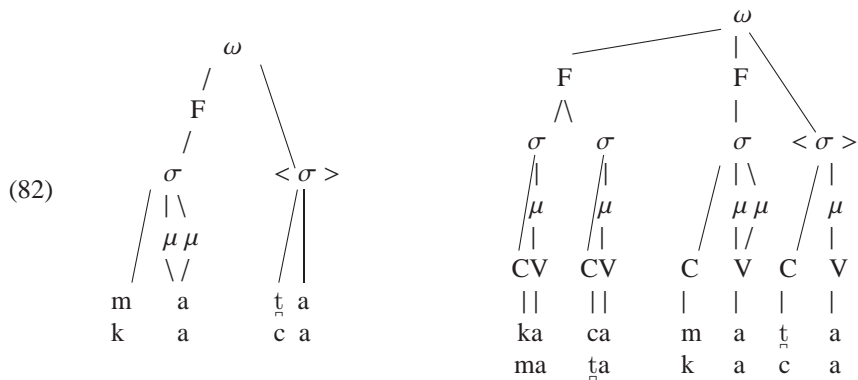
4.5.2.2 Problematic case

The derivational suffix *-pulu* is used to derive the numbers from twenty to ninety. The behaviour of this suffix is difficult to accommodate, since it precludes vowel lengthening of the stem, but the resulting word fails parse exhaustively.

(83)   a. $ṭi_\mu.(ga_\mu.pu_\mu). < lu_\mu >$

b. $(\underset{\sim}{t}i_\mu.ga_\mu).pu_\mu. < lu_\mu >$

4.5.3 Words

Words consist of a stem and optionally inflectional affixes or clitics. This extra material has no influence on the parsing of the stem. Inflectional material is not parsed into the same phonological word as the stem. This means that the representation of the word does not differ much from the representation of the stem. In SLM, there are two main types of extra material for words: verbal prefixes (see Section 6.2.1) or proclitics (see Sections 6.2.2 and 6.3), and enclitics, which are used for a variety of functions, like case marking, emphasis, coordination, etc. (see Section 6.4).

Prefixes are not parsed into the same phonological word as the stem. Both monosyllabic and disyllabic prefixes behave alike in this regard (84). The additional material provided by [su-] 'PAST' or [aɾə-] 'NON.PAST' does not take away the need to lengthen the penultimate syllable to construct the bimoraic foot.

(84)
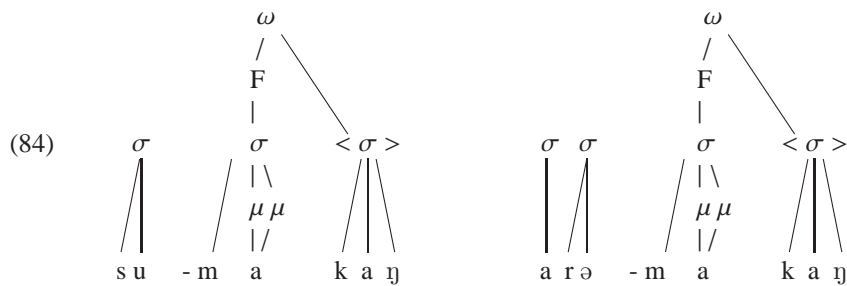
The same reasoning applies if a verbal proclitic like [mas$\underset{\sim}{t}$i=] 'must' is used (85).

(85)
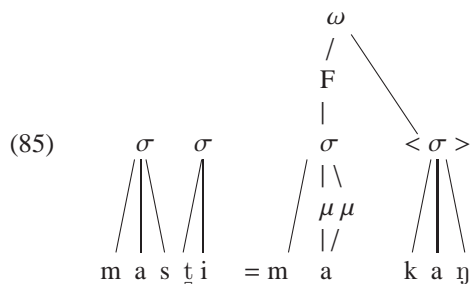
Enclitics, like inflectional prefixes and proclitics, are not parsed into the same phonological word. These include case endings, postpositions, focus clitics or coordinating clitics in SLM (see Section 6.4). Representation (86) shows a number of enclitics on */kake/* 'grandfather'. Crucially, neither of these enclitics has any influence on the vowel quantity in */kake/*.[40]

---

[40]Not all of the enclitics have a coda. For convenience and ease of exposition, the same structure is used here, but the missing coda is indicated by ∅.

(86)

$$
\omega \\
\Big/ \quad \Big\backslash \\
F \\
\Big/ \\
\sigma \quad <\sigma> \qquad\qquad \sigma \\
|\backslash \\
\mu\,\mu \\
$$

$$
k \quad a \qquad k\,e \quad = \left\{ \begin{array}{ccc} n & a & \eta & (\text{DAT}) \\ j & a & \eta & (\text{ACC}) \\ dr & i & \eta & (\text{ABL}) \\ l & e & \emptyset & (\text{ADDIT}) \\ j & o & \emptyset & (\text{EMPH}) \\ \cdots & \cdots & \cdots \end{array} \right\}
$$
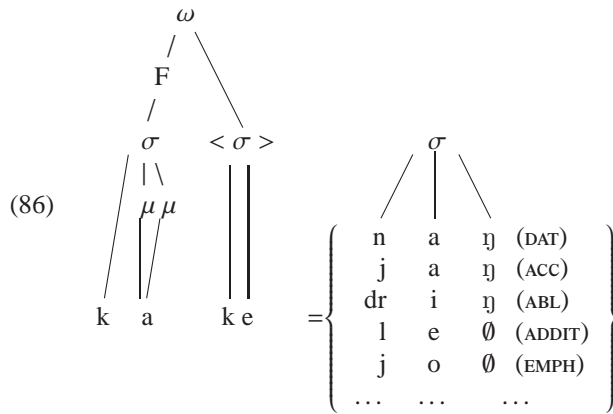
The causativizer *kiŋ* is sometimes parsed into the same phonological word as the stem (thus behaving like *-an*), sometimes not (thus behaving like an enclitic). This seems to vary with the speakers. Bichsel-Stettler (1989) states that *-kiŋ* does not take away lengthening, while Tapovanaye (1995) states that *-kiŋ* does indeed cancel lengthening. The speakers I asked about this answered that both possibilities are fine. This means that *-kiŋ* may optionally be parsed into the same phonological word as the stem. Example (88) shows both possibilities. One of the characteristics of grammaticalization is tighter integration of the grammaticalizing morpheme with its host (Hopper & Traugott 2003:140ff). In this sense, *-kiŋ* might be grammaticalizing from a more clitic-like status, not parsed into the same phonological word as its host, to a more affix-like status, integrated into the phonological word of the stem. This is shown in the final examples for the word [t̺iːt̺i] 'feed' (87) and the causative [t̺i(ː)t̺ikiŋ] 'make feed' (88).

(87)

$$
\omega \\
\Big/ \quad \Big\backslash \\
F \\
\Big/ \\
\sigma \qquad <\sigma> \\
|\backslash \\
\mu\,\mu \\
\backslash\,/ \\
t̺ \quad i \qquad t̺\,i \\
$$

(88)

```
          ω                              ω
         /                              /
        F                              F
       / \                            /
      σ   σ   < σ >          σ   < σ >   σ
      |   |                 /|\          
      μ   μ                μ μ           
      |   |                | /           
      t̪ i  t̪ i  k i ŋ      t̪  i   t̪ i   k i ŋ
```
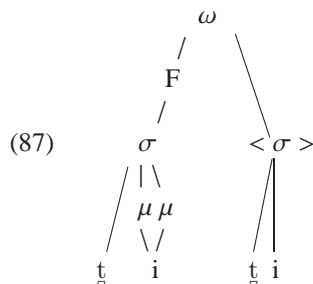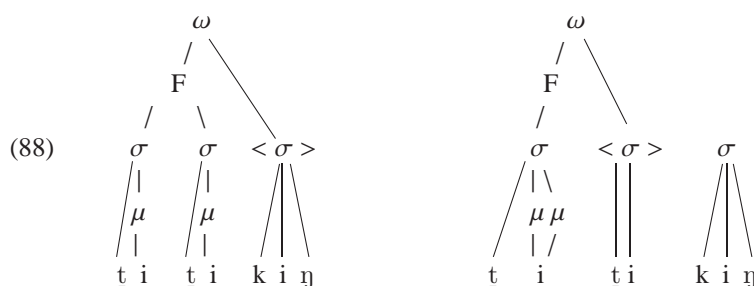
4.5.4 SLM as a stressless language

While the absence of stress as a relevant linguistic category could come as a surprise, it should be noted that stress in Malayic varieties such as (modern) Betawi, Javanese and Bahasa Indonesia is very elusive. Stress location in Indonesian has been proposed to be penultimate (Teeuw 1984, Alieva et al. 1991), final (Samsuri 1971) or arbitrary (Zubkova 1966, Halim 1974). See Odé (1994) for a discussion. For Javanese, stress has been claimed to fall on the penultimate (Ras 1982) and on the final syllable (Poedjoesoedarmo 1982). This variance is already indicative of the elusive nature of stress. Van Zanten et al. (2003)[41] conducted acoustic experiments for modern Indonesian and conclude:

> The Javanese and Jakartanese listeners, however, do no differentiate between the different stress locations [. . .] We can simply say that stress has no meaning to them, which is all the more reason to assume that word stress is neither a feature of Javanese and Jakartan, nor of the variant of Indonesian they speak. (van Zanten et al. 2003:169)

Betawi Malay[42] and to a lesser extent Javanese were among the main languages spoken by the immigrants in the 17*th* and 18*th* century (see Sections 2.2.2 and 2.8.1). In this light, absence of stress as a meaningful category in SLM seems less remarkable.

Furthermore, Tamil is also a language where the existence of stress has been questioned. Pope (1867), Arden (1934) and Arokianathan (1981) deny its existence, Balasubramanian (1980) states that no acoustic correlate of non-emphatic stress could be found, but Keane (2001) could find subtle stress cues and provides a good overview of the different positions.

Stress in Sinhala seems to be more straightforward than in the languages discussed above. Letterman (1997) summarizes research on Sinhala stress and concludes that there is broad consensus that Sinhala has a weight-sensitive system with stress falling on the first syllable unless the second syllable is heavier than the first, in which case the second syllable is stressed.

---

[41]I am grateful to Diana Apoussidou for pointing out this reference to me.

[42]Actually its predecessor, cf. Grijns (1991) for the history of Betawi Malay and Jakarta Malay.

With stress being dubious in the majority of the relevant input languages, the non-development of stress in SLM seems less extraordinary, and should indeed be expected.

4.5.5   Word structure of Arabic loans

Many loan words from Arabic violate the word structure as presented here in that they have a long vowel in the final syllable (Bichsel-Stettler 1989:54,69). Cases in point are [haɾaːm] 'prohibited for Muslims' or [imaːm] 'imam'.

4.5.6   Word structure in former research

The difference between CVːCV and CVCːV words such as [kiːɾiŋ] 'send' and [kɪrːiŋ] 'dry' has been a constant topic in discussion of SLM word structure and alternative accounts to the one here presented are available. Bichsel-Stettler (1989) refutes Tapovanaye's (1986) analysis which assumes vowel lengthening in open syllables. She shows some words of the structure CVːNCV(C) and argues that these words have a long vowel in a closed syllable. One pair she cites is [ṭʊmbu] 'coir fibre' vs. [ṭuːmbʊk] 'pound, pummel'. The long vowel in [ṭuːmbʊk] could not be explained by lengthening of open syllable vowels because it is in a closed syllable. However, all the examples she cites show a prenasalized consonant at the crucial position, so that the syllabification is CV.ŇCV(C) and not CVN.CV(C). This prenasalized consonant entails that the syllable is indeed open and the examples she cites are actually no counterexamples, but support the hypothesis advanced by Tapovanaye (1986) and further developed in Tapovanaye (1995).[43] Bichsel further shows that some trisyllables lengthen while others do not, which she takes as evidence for phonemic vowel length. All the examples she cites can be explained by the presence or absence of schwa in the initial syllable (see Section 4.5.1) and non-phonemic glide formation (see 4.1.2.1).

Bichsel analyzes long consonants as not phonemic as she assumes that they predictably occur after a short vowel in disyllabic words. The requirement for a short vowel assures that *CVːCːV(C) syllables cannot be obtained. Bichsel can rely on vowel quantity for her analysis since she assumes that it is phonemic, an analysis which is not shared here. Instead, I assume that lengthened consonants occur only after schwa, while for the remainder of the cases I assume lexical specification of a long consonant, in other words, long consonants are phonemic. It is clear that for the triple [dɪkːaṭ] 'close', [ikːaŋ] 'fish', [ṭiːkam] 'stab' one needs to assume either consonant length or vowel length as present underlyingly. The analyses are not very different in their applicability, but phonemic vowel length runs into problems with trisyllables like *makanan*, which are no problem for an analysis assuming phonemic consonant length for lexemes without a schwa.

---

[43]Furthermore, [ṭʊmbu] 'coir' is a Tamil loanword, corroborating the distribution of tautosyllabic and heterosyllabic NC-clusters on native words and loan words presented in Section 4.6.2.2.

Tapovanaye (1995) analyzes vowel lengthening and consonant gemination as a result of a minimal word requirement of four mora. A word like [ma$_\mu$ka$_\mu$ŋ$_\mu$] 'eat' has three moras, lacking one mora to fulfill the requirement. This mora is obtained through vowel lengthening: [ma$_\mu$ː$_\mu$ka$_\mu$ŋ$_\mu$]. This analysis inspired the analysis adopted here, but some modifications were necessary. One empirical problem is that Tapovanaye has to assume that all final syllables are obligatorily closed. Words which lack a final coda would be closed by a paragogic glottal stop. This glottal stop is necessary to contribute the essential fourth mora in words like [t̪i$_\mu$ː$_\mu$ga$_\mu$ʔ$_\mu$] 'three', otherwise the minimal word requirement would not be met. Unfortunately, this paragogic glottal stop is not always present, so that this analysis has empirical problems. The problematic part of this analysis, however, can be remedied by declaring the final syllable extrametrical, which reduces the requirement to two moras in the non-final syllables. This reduced requirement is then captured in the analysis presented here by the bimoraic foot requirement.

4.6 Some notes on historical phonology

4.6.1 Vowels

The Malay varieties the first immigrants spoke probably had a very similar 5+1 vowel system, which is very common in the Malay world (Adelaar 1985). The most interesting aspect is the development of schwa. Schwa generally retains its phonemic status, but is realized as ɪ or ʊ in the penultimate, or as i or u in the penultimate or the antepenultimate. Whether a front or back vowel is chosen depends on the following consonant and on vowel harmony (see Section 4.1.1.2).

Schwa in the first syllable of historic trisyllables is often dropped, as in [spuːlu] 'ten' (Std. Malay *səpuluh*), [criːt̪a] 'story'(Std. Malay *cərita*), [sbiːlan] 'nine' (Std. Malay *səmbilan*), [klaːpa] 'coconut' (Std. Malay *kəlapa*). This cluster formation can also occur word-medially as in [subla] 'side' (Std. Malay *səbəlah*).

An /a/ in the initial syllable of disyllabic words tends to lower historic /i,u/ in the final syllable to /e,o/ [ambel] 'take' (Std. Malay *ambil*), [t̪aːro] 'put' (Std. Malay *taruh*) (Bichsel-Stettler 1989:75). This is also found in Peninsular Malay (Paauw 2004), and might be a feature brought by the recruits from Malaysia. Paauw (2008) reports that this feature is also found in Eastern contact varieties of Malay.

Vowel lengthening was probably already present in the 19th century, there are records in (unvocalized) Arabic script where the word [peːgaŋ] 'catch' is written *p-y-g-ng*. The extra ⟨y⟩ probably indicated the long vowel (Hussainmiya 1987:118).

4.6.2 Consonants

4.6.2.1 Development of two additional coronal stops

The ancestor variety of SLM probably had two coronal stops, a voiceless superdental stop /t̪/ and a voiceless alveolar stop /d/ (Adelaar (1985:10), Adelaar (1991:26), Smith et al. (2004)). When SLM developed a fourfold distinction between voiced and voiceless dental and apical stops, only few lexemes underwent phonological change leading to instances of the new phonemes /d̪/ and /t/. There are few dental voiced stops, and few voiceless apical stops. In the dictionary of a non-representative text sample, the following distribution of coronal stops was found:

	dental		apical			
voiceless	186	(64%)	18	(6%)	182	(70%)
voiced	11	(4%)	77	(26%)	77	(30%)
	197	(68%)	95	(32%)	292	(100%)

We see that the distribution is heavily skewed towards t̪ and d, only about 10% of words have t or d̪. The only position where d̪ can occur in native words is initially, like [d̪aːt̪aŋ] 'come'. In loanwords, it can also occur in other positions, like [kɔd̪aɾat̪] 'power of god' (Arabic, probably through Tamil),[44] or [kald̪e] 'donkey' (Tamil).[45]

Smith et al. (2004) argue that language contact with Tamil caused the initial apical stops to be realized as dental, since Tamil disallows initial apical/retroflex stops. This account can explain why we find dental voiced stops in initial position, but apical voiced stops elsewhere. Still, there are a good deal of apical voiced stops also found in initial position,[46] so that the Tamil influence was limited to a subset of words. Furthermore, Tamil bans not only apicals, but also voiced stops as onsets altogether. This means that we would rather expect a *voiceless* dental stop for words where other Malay varieties have a voiced apical stop. This is not the case. There is no instance of diachronic initial devoicing.

Next to the development of a dental voiced stop, the other new phoneme is the apical voiceless stop. This is found mainly in loanwords from Tamil, like [kat̪il] 'bed', [t̪at̪ːi] 'hammer', [mut̪ːi] 'pot', where it always is geminated. Non-geminated apical stops are also found, but these are never in Tamil loanwords. They occur exclusively in native words.[47] Cases in point are [luːt̪ut̪] 'knee', [baːʈok] 'coconut.shell', [boːʈok]

---

[44]There is of course the cognate *kudrat* with the same meaning in Std. Malay (Uri Tadmor p.c.), but this should have given [kuːdrat̪] in SLM by regular sound changes. The phonologically deviant nature of [kɔd̪aɾat̪] suggests it is a loanword.

[45]The latter word is actually a Tamil loanword also in Nusantara Malay (*keldai*), but it is probable that the presence of Tamil as a contact language in Sri Lanka influenced the articulation. Instead of the apical realization of the voiced coronal stop, which is found in all other native phonemes word-internally, the Tamil pronunciation of the corresponding word has changed the articulation to dental.

[46]Actually a majority, Saldin's dictionary (Saldin 2007) contains 42 words with initial ⟨d⟩ and only 9 with initial ⟨d̪⟩.

[47]There could be Sinhala loanwords with a single apical voiceless stop, but these have not been found, and lexical influence from Sinhala is quite limited anyway.

'bald' or [paːʈok] 'hiss'. It appears that single apical voiceless stops are only found before [+back] consonants.

Manuscripts in Arabic script from the 19ᵗʰ century seem to distinguish a ⟨d⟩ from a ⟨d⟩ with a dot underneath, although the precise nature of this distinction is unclear (Hussainmiya 1987:120f). It is likely that one of these graphemes indicates the dental phoneme and the other one the apical phoneme, but the precise relationships still have to be established.

4.6.2.2 Prenasalized consonants

Another new feature due to language contact are the prenasalized consonants. Interestingly, *NC sequences which are heterosyllabic in Std. Malay today are tautosyllabic in SLM, while loanwords from the adstrates show the heterosyllabic pattern. It appears that historically heterosyllabic NC clusters in the following list have become tautosyllabic:

- [ɡaː.ᵐbar] 'picture'<*gambar

- [ṯaː.ⁿḓak] 'dance'<*tandak

- [baː.ⁿɟir] 'flood'<*banjir etc.

The heterosyllabic pattern is found in the following loanwords:

- [kaɲ.ɟi] 'groats'< Tamil *kanji*[48]

- [baɳ.ḓu] 'insect'< Tamil *baɳḓu*

- [ṯuɳ.ḓu] 'piece'< Tamil *ṯuɳḓu*

[ṯuɳ.ḓu] 'piece' can be contrasted with the inherited [ṯuː.ⁿḓuk] 'bend, bow', which has a tautosyllabic cluster, and a long vowel.

In some cases, there are cognates in Sinhala and Tamil, which caused a heterosyllabic pronunciation. An example is the spicy side dish, which is [sambol] (Sinh.), [sambal] (Tam.) and [sam.bal] (SLM), respectively. Compare this to [saː.ᵐbal], which would be expected in analogy to the cases mentioned above.

Some words with vocalic onsets are exception to this rule, like [a*(ː)m.bɛl] 'take' or [a(*ː)ɲ.ɟiŋ] 'dog'. Another exception is the word [pa(*ː)ŋgɛl] 'call'.

4.6.2.3 Consonant gemination

Lengthened consonants often occur after a synchronic schwa (see Section 4.1.2.6).[49] However, some consonants have to be assumed to be long underlyingly (see 4.1.2.6).

---

[48]Tamil stops following a nasal are allophonically voiced. This is represented in the lexemes given here.

[49]A similar phenomenon is found in Berau Malay (Collins 1992, Adelaar 2005) and other dialects of Malay (Adelaar 2005:206).

Examples for this are [apːi] 'fire', [kapːal] 'ship', [ikːaŋ] 'fish', [at̪ːas] 'top', [t̪opːi]
'hat', which are the same in Std. Malay, but without the geminate. At least [kapːal]
and [t̪opːi] can be explained historically, since these words have cognates in Sinhala
(t̪oppiya) and Tamil (kappal, t̪oppi), and were actually borrowed from Pali and Tamil
before the Malays arrived in Sri Lanka.[50] For the other words, a good explanation for
their having consonant length instead of vowel length does not seem at hand.

4.6.2.4   Loss of h in initial and final position

SLM has dropped h in initial and final position (Adelaar 1991). ([aːɾi] 'day' (Std.
Malay *hari*), [buːʋa] 'fruit' (Std. Malay *buah*)). This is already found in manuscripts
from the 19th century (Hussainmiya 1987:119). In intervocalic position, h is some-
times retained ([laːheɾ] 'be.born', [leheɾ] 'neck'), and sometimes dropped, with pos-
sible subsequent insertion of a glide ([lijat̪] 'see' (Std. Malay *lihat*), see Section
4.1.2.1). Initial /h-/ enjoys prestige in the SLM community and is used on special
occasions like speeches, mainly in the word [(h)aːɾi] 'day'. It is not used when pres-
tige is less important, and not used for words where the speakers are unaware of the
Standard Malay form having an h.[51]

4.6.2.5   Final velarization of nasals

Final velarization is found in many places, already in manuscripts from the 19th cen-
tury (Hussainmiya 1987:120). This could be an influence from Sinhala, where we find
the same process, but also Malay varieties of Moluccas and Manado (which could have
been among the ancestor varieties, see Section 2.8.1) show this phenomenon (Adelaar
1991, Paauw 2004, Smith et al. 2004)), so that it cannot be decided whether the causes
for this are internal or external. Furthermore, this process is common in other lan-
guages, e.g varieties of American Spanish (Hualde 2005:10f) and well-motivated on
theoretical grounds (Botma 2004:322ff), so that an independent development would
also be plausible.

   An intriguing phenomenon is the emergence of final velar nasals on some words
which do not have a nasal in that position in other varieties of Malay. Cases in point
are [kut̪umuŋ] 'see' (cf. *kət̪əmu*), [buːnuŋ] 'kill' (cf. *bunuh*), [deɾiŋ] 'from' (cf.
dari).[52]

---

[50]I would like to thank Uri Tadmor for pointing this out.

[51]Paauw (2008:99) notes that the sporadic retention of /h-/ in some words is also found in other Malay
varieties which have lost /h/ otherwise. The words he gives are also among the words with this feature in
SLM.

[52]Examples found in other authors are [kuluŋ], [ɖuːɖuŋ] and [ɲaːɲiŋ] .The Southern pronunciation of
the conditional marker, [kuluŋ] (Peter Slomanson p.c., cf. *kalau*). This marker is normally realized as
[kalu] in the Upcountry. A more distantly related case might be the Southern pronunciation of the animate
existential as [ɖuːɖuŋ] (Slomanson 2008b), which might be a reduction of the normal form [ɖuːɖuk] with
subsequent addition of a velar nasal. Tapovanaye has [ɲaːɲiŋ] 'song', where in the Upcountry [ɲaːɲi] is
heard.

4.6.2.6 Labial approximant

On the phonetic level, the labial approximant is bilabial in other Malay varieties, but labiodental in SLM (Bichsel-Stettler 1989, Smith et al. 2004).

4.6.2.7 #ṭ>c

In some words, SLM has a voiceless palatal stop alternating with the voiceless dental stop found in the Std. Malay cognate. This is the case for [ṭ/caːɲa] 'ask', [ṭ/cunɟikang] 'show' and [ṭ/ciːᵑgal] 'stay overnight'. Note that for the first two words, the palatal stop might be explained by assimilation to the palatal nasal later in the word. In the last example, we find a velar nasal, not a palatal nasal, so that this explanation is not possible there.

4.6.3 Word structure

Contemporary Sri Lanka Malay parses words into bimoraic feet. The historical variety did probably not have such a metrical structure of the word. Another change at the word level is that some historically trisyllabic words have lost the antepenultimate, or are in the process of losing it. Examples are [(ku)paːla] 'head' or [(ku)maːreng] 'yesterday' (cf. Hussein 2007:419). This is in line with the bimoraic foot analysis. Under this account, the antepenultimate syllable in these words is not parsed into a foot and actually against optimal word structure. As a consequence, it is likely to be dropped.

4.6.4 Loanwords from Tamil

A curious phenomenon is found in loanwords from Tamil with the phoneme /ɟ̇/. This phoneme is pronounced [cː] in Tamil when geminated, and [s] in the other instances that matter here (cf. Schiffman 1999:9,13).[53] SLM loans have [c] in initial position, where modern Tamil has [s-]. Cases in point are [ciŋgala] 'Sinhala'<Tamil *siŋgala*, [cərapːu] 'slipper'<Tamil *serappu* (Sinh. *serappuva*), [cipːi] 'shell'<Tamil *sippi* (Sinh. *sippiya*). We might either have to do with a spelling pronunciation, or the Tamil word was borrowed in a time when the pronunciation with [c] was still common.

4.7 Orthography

The exposition so far has been in broad IPA. The rest of this grammar will show examples in practical orthography developed on the basis of the phonological analysis presented above. The orthography presented here is phonetic in the sense that some

---

[53]However, Suseendirarajah (1973a:174) states that in Sri Lanka, the realization as an affricate and not s- is more common. This description does not seem to align with Tamil usage in Kandy, but this needs more research.

phonetic detail of surface structure is represented, e.g. vowel length or raising of schwa.

The mapping of phonemes on graphemes follows the orthography of Bahasa Malaysia/Bahasa Indonesia ('Bahasa' in the following) where possbile in order to facilitate exchange between the communities. Phonemes unknown to Bahasa will be rendered according to Lankan conventions.

No problems at all are found in the phonemes /p c k b g m n f z s h l r a e i o u/, which simply use the grapheme that inspired the IPA sign. The phonemes /ɟ, j/ are represented in the same way as in Bahasa, i.e. ⟨j, y⟩. /ɲ/ and /ŋ/ are written as ⟨ny, ng⟩ when occurring before a vowel. When occurring before a homorganic stop (/ɟ, c/; /g,k/), they can also be written as ⟨n⟩ because there is no risk of confusion, and the word for 'key' /koɲci/ could be mispronounced as [konici] when written ⟨konyci⟩. There is no risk for mispronunciation with ⟨konci⟩, since the nasal assimilates anyway. As for the velar nasal, the word for 'mug' can be written ⟨mangkok⟩ or ⟨mankok⟩. While the former is technically more correct, it was found that speakers do not see the need for the extra ⟨g⟩ and frequently omit it. In order not to burden the orthography with rules which make no sense for the speakers, it was decided to declare ⟨g⟩ optional in the phoneme sequence /ŋk/. ⟨g⟩ must be present in the sequence /ŋg/ on the other hand in order to disambiguate between /ŋ/⟨ng⟩ and /ŋg/⟨ngg⟩.[54] Another instance of accommodation of the orthography is the rendering of /ʋ/, which can either be ⟨v⟩ or ⟨w⟩, according to the preference of the speaker.

/ʃ/ is represented as ⟨sh⟩. Schwa in normal position is represented by ⟨à⟩. Schwa raised to [ɪ,ʊ] (see p. 123) is represented by ⟨ì,ù⟩. Schwa raised to [i,u] is represented by ⟨i,u⟩. The gravis is thus a sign of reduced/centralized articulation as compared to the vowel sign with which it combines.

The common rendering of Sinhala and Tamil words in Sri Lanka employs ⟨t,d⟩ for the apical sounds /t,d/ and ⟨th, dh⟩ for the dental sounds /t̪, d̪/. The same will be followed here. ⟨_h⟩ is thus not a sign of aspiration but of dental articulation.[55]

Prenasalized stops may be represented with a brevis over the nasal ⟨m̆b, n̆d, n̆j, n̆g⟩ or ⟨n̆gg⟩. Technically speaking, this is not be necessary because the preceding long vowels suffice to distinguish prenasalized stops from plain stops.

When consonants represented by digraphs are geminated, only the first part is doubled (*ddh, tth, nny, nng*, cf. Table 4.9).

While vowel length is predictable, the exact circumstances are quite intricate, which is why vowel length is represented in the orthography as a service to the reader unfamiliar to the language. This also makes the use of the brevis for the prenasalized series optional, adding to the usability of the orthography on typewriters and computer keyboards. Vowel length is indicated by doubling the vowel.

---

[54]Another contrast exists with /ŋː/, written ⟨nng⟩, see below.

[55]The rendering of dental and apical/retroflex sounds in orthography is therefore the exact mirror image of the orthography of Javanese (Robson 1992:10). This is a bit unfortunate for Austronesianists, but SLM speakers are more used to the Lankan spelling than to Javanese, which was taken to be a more important factor.

phoneme	grapheme	phoneme	grapheme	phoneme	grapheme
t̪	th	d̪	dh	ɲ	ny
t̪:	tth	d̪:	ddh	ɲː	nny
t	t	d	d	ŋ	ng
c	c	ɟ	j	j	y
ʃ	sh	ə	à,ì,ù	ʋ	v/w

Table 4.9: Mapping of phonemes to graphemes (Non-trivial cases)

Both phonemic and non-phonemic glides are represented, by ⟨y,v(w)⟩.

As for capitalization, the rules of the English languages are used (Begin of sentence, proper nouns, etc.).

Given the vast array of clitics, the problem of what shall be written together or apart arises. The speakers intuitively write most of the clitics and affixes as separate words, while linguists would prefer an indication of the bound nature of these morphemes. In order to keep the length of words within bounds, it was decided to not join all clitics to the host. The following list shows which items are joined to the host, or separated. This list is based on the relative length of the clitics, and on the importance of the host for the clitic.

- all affixes are joined (see Section 6.2)

- all proclitics are joined (see Section 6.2.2)

- all postpositions are joined (see Section 6.4.4)

- all Coordinating Clitics are joined (see Section 6.4.6)

- =*jo* and =*jona* are joined (see Sections 6.4.7.1 and 6.4.7.2)

- the remaining items are separated, most notably the indefiniteness marker *hatthu* and the plural marker *pada*.

It is also possible to write SLM in Sinhala script. Sinhala script will be less used by the scientific community but might be important for the speakers. In order to write SLM, the *śuddha* set is used. Since Sinhala and SLM have a very similar phoneme inventory, the biggest part of this is straightforward. The problematic cases are /ŋ/ in the onset, schwa and short /a/.

The śuddha set of Sinhala only has a grapheme for /ŋ/ in final position, ⟨◌⟩. This grapheme cannot combine with a vocalic diacritic. In order to render sequences of ŋ+vowel like [aːŋiŋ] 'air', the independent vowel sign (in this case ⟨ඉ⟩) can be employed, so that [aːŋiŋ] would be written ⟨ආ◌ඉ◌⟩.

In Sinhala, short [a] and [ə] are allophones of the same phoneme, /a/. This leads to confusion when writing SLM, where these sounds are not allophones. Furthermore, [a] and [aː] are complementary allophones of one phoneme in both Sinhala and

SLM, but the distribution is different. A mechanic application of Sinhala phoneme-to-grapheme rules is therefore not possible. When writing SLM, care must be taken to use the diacritic ⟨ɔ⟩ only for phonetically long [aː], to use the the inherent vowel (which receives no special marking in this script) only for phonetically short [a], and to use another sign for schwa. My transcriber found that the visarga ⟨�ః⟩ was well suited to represent schwa, as it is in the paradigm of vocalic diacritics, does not indicate a full vowel, and is not needed elsewhere in the system.

Writing SLM in Tamil script is more difficult since this script does not allow a differentiation between voiced and voiceless consonants. This is no problem for Tamil, where this distinction is not important, but it is a problem for SLM, which does distinguish voiced consonants from voiceless consonants. This means that the two words [ɡaːɡa] 'land' and [kaːka] 'elder brother' would be written alike in Tamil script ⟨காக⟩ and could not be differentiated.

Furthermore, Tamil script has no way to represent prenasalized vowels. There are some issues analogous to Sinhala in what concerns the velar nasal and schwa. Like in Sinhala, the velar nasal cannot combine with all vowels in Tamil. This can be solved by using the independent vowel sign, so that [aː.ŋiŋ] 'air' could be written ⟨ஆங்இங்⟩.[56] Schwa could be represented by the *āytam* ⟨∴⟩, analogous to ⟨�ః⟩ employed in Sinhala.

4.8 Intonation

We have now discussed the lower concepts of the prosodic hierarchy, i.e. mora, syllable, foot and word. The next higher levels are the phonological phrase and the intonational phrase, to which we will turn now.[57] The precise distinction between phonological phrase and intonational phrase is not easy to make in SLM, which is why the two concepts are discussed together here. There is certainly more to be said about SLM intonation than what can be alluded to here; what follows is merely a preliminary observation. The examples were drawn either from the case study (see Section 4.8.5) or from a recording of a senior female speaker in Kandy who was selected because of the quality of the recording (clear voice, little noise). The data set is thus not very large, nor balanced. However, it confirms my impressionistic findings about the relevant intonation patterns gathered earlier.

In SLM, the following three contours are found very frequently:

1. high tone until beginning of last lexical word, drop to the end: $^{H\longrightarrow H}$ ↘ $_L$:
   assertive contour

2. low tone at beginning, rising to final high tone target: $_L$ ↗ ^{H} **progredient contour**

---

[56]Depending on available fonts and purism on the part of the writer, a normally illicit combination involving ⟨ஞி⟩ (ɲi) could also be used.

[57]I would like to thank Cecilia Odé for help with this section. All remaining errors are my own.

3. high tone at beginning, downdrift to last syllable, steep rise to final high tone.
   $^{H}\searrow_{L}/^{H}$: **presuppositive contour** and **question contour**

In SLM, the distinction between presupposition and assertion is very important for intonation. The (syntactic) sentence consists of a number of (presupposed) NPs or subordinate clauses and a final assertion. Each of the presupposed elements receives the presupposed contour (HLH), and the assertion receives the assertive contour (HL) (89).

(89) a. H   L H   H   L
     [ADJ N] | *[*TAM- *V]* |
     PRESUP   ASSERTION

    b. H   L H   H   L
     [NP NP s-V] | *[*TAM- *V]* |
     PRESUP   ASSERTION

    c. H   L H   H   L H   H   L H   H   L
     SUBORD | NP | NP | *[*TAM- *V]* |
     PRESUP   PRESUP   PRESUP   ASSERTION

The above schema illustrates that NPs and subordinate clauses receive the same contour even if one is a nominal constituent as in a) and the other one a clause as in b). Both clauses and NPs can be seen in c). The fact is that both are presupposed, and therefore both receive the presupposed contour, which consists of a slow fall to L and a steep rise to a final high tone target. This will be discussed in more detail below. In distinction to that, the final verb is not presupposed, it forms part of the assertion. Thus, it receives an assertive contour, which means a high tone more or less at the beginning of the last lexical word and a steady fall to the end. This will also be discussed in more detail below.

In (89), we have illustrated the assertive contour in final position. If several assertions are chained, as in (90), the final one receives the same contour as above, but the preceding ones receive the progredient contour (LH).

(90) L   H H   L
     [NP TAM- *V]*$_{\text{CLAUSE}}$ | *[NP* TAM- *V]*$_{\text{CLAUSE}}$ |
     PROGREDIENT.ASSERTION   FINAL.ASSERTION

The two assertive contours and the presuppositive contour will now be discussed in more detail.

4.8.1   Assertive contour

The assertive contour is characterized by a fall from H to a low tone target. The exact position of H needs more research, but appears to be towards the beginning of the last lexical word of the assertion.

(91)
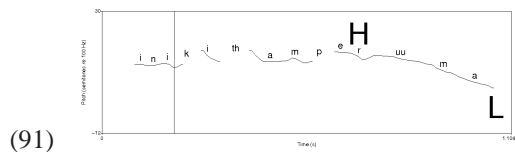
		H	L
ini	*kitham-pe ruuma*		
PROX	1PL=POSS	*house*	

'This is our house.'

In (91), there is no presupposition, only a deictic reference (whose contour we will ignore for now) and an assertion *kithampe ruuma* 'our house'. The assertion receives the HL contour, where H is linked to the beginning of the last lexical word *ruuma* 'house'[58] and the L is linked to the end of the utterance.

In (92), we have a verbal predicate *kitham=pe nigiri su-jaadi* 'became our country'. The assertive contour is applied to the predicate with the high tone on the prefix of the last lexical word, the verb. Then follows a drop to the end of the utterance. The constituents *kaarang* 'now' and *ini* 'PROX' receive a presuppositive contour, discussed further below.

(92)
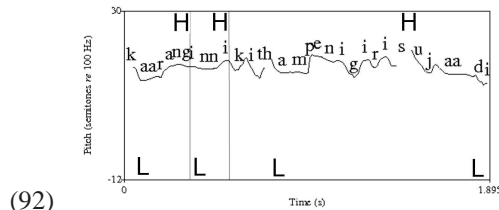
L	H L	H		H	L
kaarang |	*inni* |	*kitham=pe nigiri*	*su-jaadi*	|	
now	PROX	1PL=POSS *country*	PAST-*become*		

'Now this has become our country.'

---

[58]The pitch peaks after the k and the th are caused by aspiration and thus artefacts and not part of the intonation proper.

Above we have seen that the last lexical word (noun, verb) receives the high tone of the assertive contour. In the event that the assertion only contains functional elements, the description formulated above cannot apply. In this case, the high tone is linked to the last word nevertheless, even if it is not lexical. An example is *kitham* '1pl' in (93).
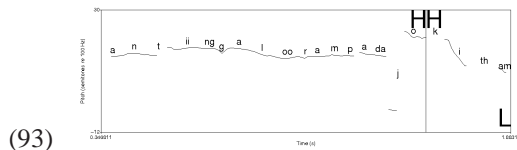
(93)

		H	H	L
an-tiinggal	*ooram*	*pada=jo*	|	*kitham*	|
PAST-*settle*	*man*	PL=FOC		1PL	

'The people who settled down here are we.'

In narrow focus, the high tone is linked to the element in focus even if this is not the last lexical element in the assertion. In (94), the speaker emphasizes the fact that the Malays are no longer East Asians, but Ceylonese. The two presuppositions *kaarang* 'now' and *kithang* 'we' are followed by the assertion *SEELON ooram pada* 'Ceylon people'. The presuppositions conform to the presuppositive contour, which will be discussed in more detail below. The assertion *SEELON ooram pada* receives the assertive HL contour, but H is linked to the emphasized element *SEELON* instead of the last lexical element *ooram*. This is marked in the graphic by a crossed out H where the normal tone would be expected.
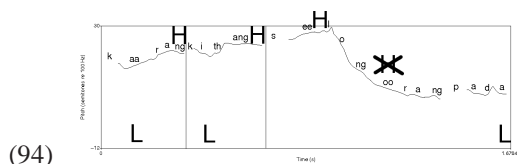
(94)

L	H	L	H	H		L
kaarang	|	*kithang*	|	*Seelong*	*ooram pada*	|
now		1PL		*Ceylon*	*man*	PL

'Now we are Ceylonese.'

Imperatives also take the assertive contour, as the following example, taken from cooking instructions, shows. Note that the rice is presupposed and has been talked about earlier, which can be seen from the accusative marker =*yang*, which normally

only attaches to inanimate referents when they are topical (see Section 6.4.4.1).
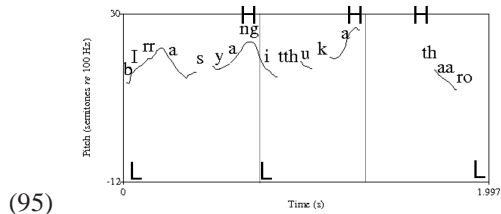
(95)

L		H	L		H	H		L
bìrras=yang		*| itthu=ka*			*| thaaro*	*|*		
raw.rice=ACC		DEM.DIST=LOC			*put*			

'Add the rice to it!'

4.8.2   Progredient contour

The progredient contour is characterized by a slow steady rise to a high tone at the end. It is used for the non-final elements in a chain of assertions. The enumeration of ethnic groups in (96) shows slow rise from start to end of each of the first three phonological phrases, which are progredient. The last phonological phrase has the "final" contour, but the intensity decreases so that the graphic cannot represent the very low final pitch. The movement is clear when listening, though.
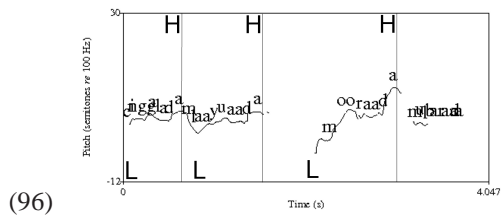
(96)

L		H	L		H	L		H		(L)
cinggala aada	*|*		*mlaayu aada*		*|*	moor *aada*	*|*	*mùlbar aada*	*|*	
Sinhala		*exist*	*Malay*	*exist*		*Moor exist*		*Tamil*	*exist*	

'There are Sinhalese, there are Malays, there are Moors, there are Hindus.'

4.8.3 Presuppositive contour

The presuppositive contour is characterized by a steady fall to the last syllable and a following steep rise to a high tone target: LH. It is used for presupposed elements, which can be NPs, adjuncts or subordinate clauses.

Subordinate clauses always precede the main clause in SLM and are normally presupposed. Being presupposed, they are realized with a final H tone, normally preceded by a low tone. (97) illustrates this pattern. The high tone is on the right edge of *kaaving* 'marry', while the low tone is immediately preceding. There is a second presupposition in this utterance *ini mlaayu* 'these Malays'. Just like the subordinate clause, this NP receives an LH contour. The assertion has a high tone on the past tense prefix *su-* of the last lexical word *braambath* 'spread'.
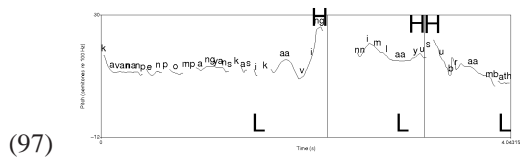
(97)

		L	H	L	H
ini	*kavanan=pe prompang=yang as-kasi kaaving*	|	*nni*	*mlaayu*	|
PROX	*group=*POSS *female=*ACC CP-*give marry*		PROX	*Malay*	

H		L
su-braambath	|
PAST-*spread*

'Having married their daughters, these Malays spread.'

In (98), there are two presupposed elements, and one assertion. The presupposed elements are *kithampe aanak pada=jo* 'our children' and *skaarang* 'now'. These receive a low tone on the nucleus of the last syllable (excluding the clitic =*jo*) and a high final tone. The assertion *baenang cinggala sablaajar* 'learnt Sinhala well' receives a high tone on the prefix of the last lexical word *blaajar* 'learn' and a low tone target assigned at the end.

(98)
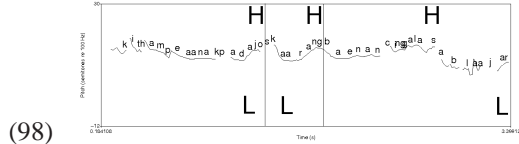
	L	H	L	H		H		L
kitham=pe aanak pada=jo			*skaarang*		*bae=na*	*cinggala sa-blaajar*		
1PL=POSS *child* PL=EMPH			*now*		*good*=DAT *Sinhala* PAST-*learn*			

'Our children have learnt Sinhala well.'

Things are similar in (99), where the presupposed constituents *laskalli* 'again' and *kithampe nigirinang* 'to our country' receive a LH tone at the right edge. The assertion *thàràpii* 'did not go' receives high tone on the prefix of the last lexical element and a low final tone.

(99)
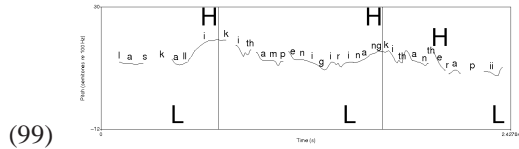
L	H		L	H		H	L
laskalli	|	*kitham=pe nigiri=nang*		|	*kithang thàrà-pii*		|
again		1PL=POSS *country*=DAT			1PL	NEG.PAST-*go*	

'We did not go back to our country.'

(100) shows a more complex sentence. The presupposed constituent *andaathang ooram pada* 'the men who had come' receives a LH tone on the right edge. So do the subordinates *spìrrang* 'having waged war' and *deranna asbanthu* 'having helped him'. The predicate *siinijo suciinggal* 'settled down right here' receives a high tone on the prefix of last lexical element *ciinggal* 'settle' and a low final tone.

(100)
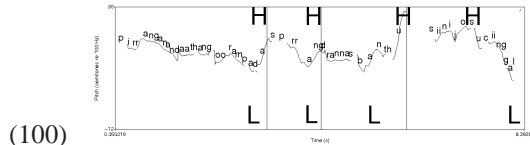
			L	H		L		H	
pìirang=na	*an-daathang*	*ooram*	*pada*	|	*s-pìrrang*	|	*deran=na*
war=DAT	PAST-*come*	*man*	PL		CP-*wage.war*		3=DAT

	L	H		H		L
as-banthu	|	*siini=jo*	*su-ciiňğgal*	|
CP-*help*		*here*=EMPH	PAST-*settle*

'The men who had come, after having waged war and helped him, settled down.'

4.8.4 Question contour

The question contour is identical to the presuppositive contour: LH on the last syllable. It might be possible to combine them into a "non-assertive" contour; for expository reasons this is not done here. The question contour is used for interrogative illocution. Given that there is also segmental material which signals questions (WH-words or the clitic =*si*), the question contour is not obligatory if either of these are present.

Example (101) shows the question contour on a WH-question introduced by *aapeyang* 'what'. The pitch rises to the right edge from an immediately preceding low tone linked to *kijja* 'do'. The preceding NP *dram pada* 'they' and the preceding subordinate *deram pada lae atthu pukuran mà-gijja thàrboole subbath* 'Since they could not do other work' have the presuppositive contour, which is emically identical to the question contour. All three contours in this utterance have a final high tone, and a preceding low tone, although the latter cannot be located exactly in the subordinate clause.
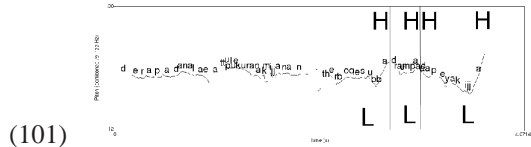
(101)

                                      L?   H   L
deram pada lae   atthu pukuran mà-gijja thàrboole subbath | dram pada
3PL   PL   *other* INDEF *work*   INF-*make cannot*   *because*   3PL   PL
H H   L   H
 | *aape=yang eng-kijja?* |
   what=ACC   PAST-*do*

'Since they could not do other work, what did they do?'

YN-questions without the clitic =*si* also have the question contour, as in (102)
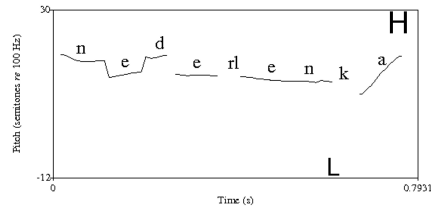
(102)

            L   H
Netherland-*ka* |
Netherlands-LOC

'Do they live in the Netherlands?'

Another example is (103), where the speakers ask a younger speaker for the word for 'ghee'. The pitch rises steadily to the end of the question.

(103)
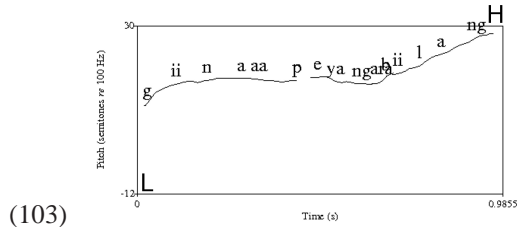

```
L                              H
ghee =na  aapeyang biilang |
      =DAT what        say
```

'How do you say for "Ghee"?'

When the interrogative clitic =*si* is present, the question contour can be present as in (104) or absent, as in (105). This seems to correlate with question type. In (104), the speaker requests information about the existence of siblings, whereas in (105) the speaker requests confirmation of her assumption that the addressee was not in Sri Lanka during the fasting period. Request for information is signalled both segmentally and suprasegmentally in (104), while request for confirmation is only marked segmentally in (105), and not suprasegmentally.

(104)
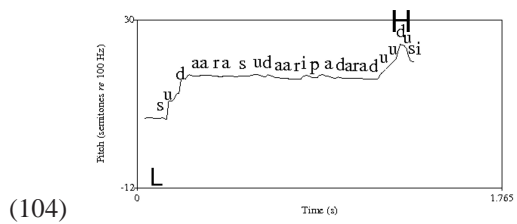

```
L                              H
sudaara sudaari pada arà-duuduk=si
brothers sisters  PL    NON.PAST-be=INTERR
```

'Do you have brothers or sisters?'
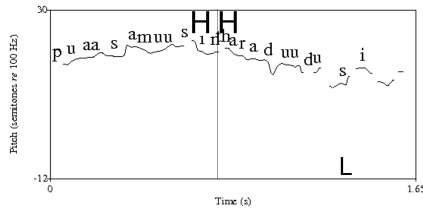
(105)


```
          H      H       L
puaasa muusing thàrà-duuduk=si
fasting  period   NEG.PAST-live=INTERR
```

'You were not here during the fasting period, were you?'

4.8.5   Case study

The intonational patterns for assertions, questions and commands will be exemplified by a short conversation recorded in Badulla on January 15, 2006. After some linguistic production stimulated by the researcher, the speakers have a little rest and chat about his family. The existence of the recording device is forgotten. This fragment was chosen because of the naturalistic setting and because various concise speech acts with distinct intonation contours follow each other in short sequence. Especially interesting are the different functions that interrogatives can have. They can be used to request information, or to request confirmation. This corresponds to different intonatory contours and to the presence or absence of the interrogative clitic =si.
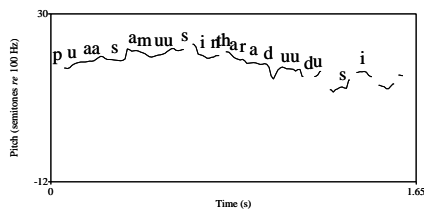
(106)


```
          H        H                    L
puaasa muusing | thàrà-duuduk=si      |
fasting  period   NEG.PAST-live=INTERR
```

'You were not here during the fasting period, were you?'

Comment: The negated predicate *thàràduuduk* 'were.not' together with =*si* indicate a presumption of absence, whose correctness is requested to be confirmed. The intonation contour is HL, like an assertion.
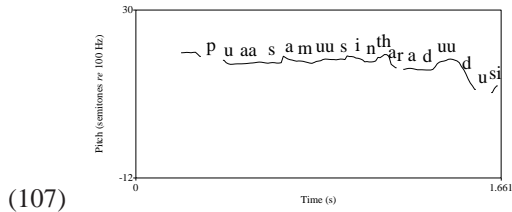
(107)

	H		L
puaasa muusing	*thàrà-duuduk=si*			
fasting period	NEG.PAST-*live*=INTERR		

'You were not here during the fasting period, were you?'

(108)
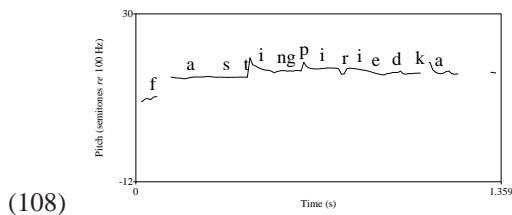
	H
fasting period=*ka*		
fasting period=LOC	

'In the fasting period'
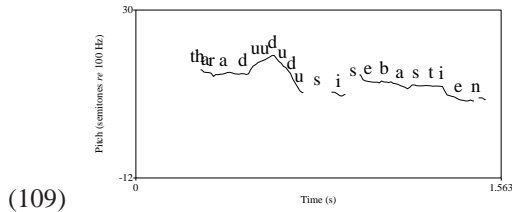
(109)

H   L   L

thàrà-duudu=si   *| Sebastian |*
NEG.PAST-*stay*=INTERR   *Sebastian*

'You(=Sebastian) were not here, were you?'

Comment: as above. The high tone of the assertional contour is linked to the beginning of the last lexical word, *duuduk* 'stay' in this case.   The afterthought/anti-topic (Chafe 1976) *Sebastian* receives a new assertional intonation contour, but the pitch is not reset.

(110)
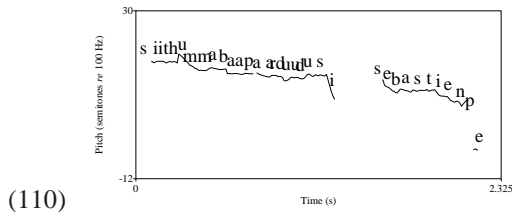
                                                                        L

SLM1: siithu mma-baapa   arà-duudu=si |
       there mother-father NON.PAST-*stay*=INTERR

                          L

Sebastian=pe |
Sebastian=POSS

'Do your parents live there (i.e. Europe), Sebastian?'

Comment: In contrast to the preceding question, the afterthought *Sebastian=pe* is located in the same contour. The rather strong fall compared to the previous questions indicates security that my parents do indeed live there and reinforces the reading as a request for confirmation, rather than information.

(111)   *SN: siithu?*
      there
      'What do you mean by "there"?'

Comment: No intonation is given for the speech of the non-native variety spoken by the German researcher in order to avoid false conclusions drawn from this learner's phonology.

(112)

               L   H
SLM1: Netherland=*ka* *|?*
      Netherlands=LOC
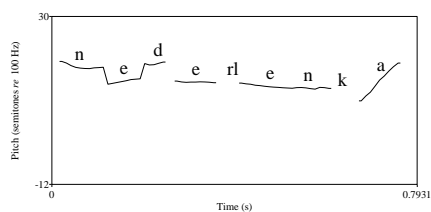
'Do they live in the Netherlands?'

Comment: Because the confirmation is declined, the question is restated, but with a clear contour of request for information.

(113)   *SN:* Germany=*ka.*
      Germany=LOC
      'They live in Germany.'

(114)
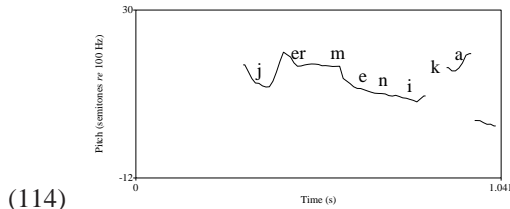
      H   L   H
SLM1: Germany=*ka* |
       Germany=LOC

'Oh, so they live in Germany?'

(115)   *SN:* Germany=*ka.*
         Germany=LOC
         'Yes, they live in Germany.'

(116)
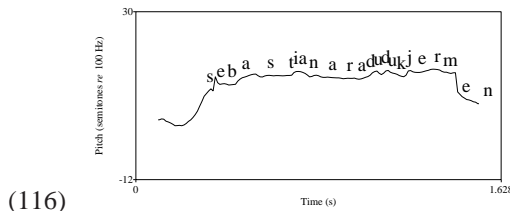
                  H   L
SLM1: Sebastian arà-duuduk German |.
     Sebastian NON.PAST-*stay Germany*

'So you live in Germany.'

(117)   *SN: no se se=jo*   Netherlands=*ka, se=ppe*
        (no) 1s 1s=EMPH Netherlands=LOC, 1s=POSS
        mma-baapa Germany=*ka*
        mother-father Germany=LOC
        'No, I live in the Netherlands, my parents live in Germany.'

Comment:Again a request for information, but less salient than the preceding one. It indicates the surprise that members of my family live in separate countries, and requests confirmation of the veracity of that information. =*si* is not present.

Comment: A tentative conclusion with a slight assertional contour. Listening to the example gives an impression of hesitation. The tentative conclusion seems to be 'If the parents live in Germany, so must the son.'

Comment: The solution of the mystery of parents and son living in different countries receives an assertional contour.
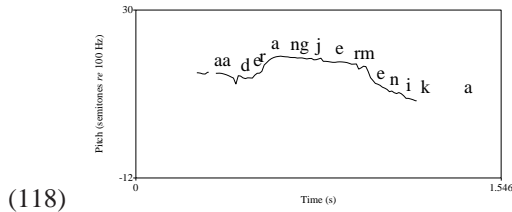
(118)


```
          H   H              L
SLM1: derang Germany=ka   |
      3p      Germany=LOC.
```

'So they are in Germany.'

Comment: A new topic: siblings. The contour is a request for information with a rising pitch since there are no presumptions available.
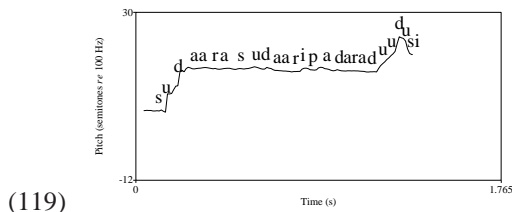
(119)


```
                                              H
SLM4: sudaara sudaari pada arà-duuduk=si   |
      brothers sisters  PL    NON.PAST-be=INTERR
```

'Do you have brothers or sisters?'

(120) *SN:* Germany=*ka   duuva aade   prompang*
           Germany=LOC two   younger.sibling woman
       aada
       exist
       'I have two younger sisters who live in Germany'

(121)
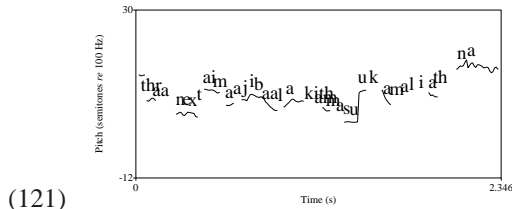
$$H \quad H \quad L \quad \quad H$$

SLM5: thraa. next time | aajibaa-la | kithan=na
      no   next time   bring-IMP   1PL=DAT

$$\qquad\qquad\qquad H$$

suuka mà-liiyath-na |
like   INF-*see*-DAT

'So bring them next time, we would like to see them
(won't you)'

Comment: This contour is difficult to interpret from pitch movement alone, but auditory impressions make it clear that we are dealing with a command. The rising pitch at the end might indicate mitigation, but this is unclear as of yet.

(122)
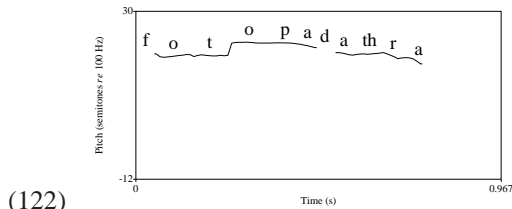
$$H \qquad\qquad\qquad L$$

SLM2: photo *pada thraa* |
      photo PL   NEG.*exist*

'You have no photos?'

Comment: Again a new topic, but with the underlying assumption that I have indeed not brought any photos: assertional contour. Note the absence of the clitic =*si* despite the question illocution.

(123)   *SLM1:* photo *pada thraa=si*
          photo PL   NEG.exist=INTERR
          'You have no photos, have you?'

Comment: No analysis due to problems with the sound file. (Multiple speakers at the same time, background noise)

(124) *SLM6: gaam̐bar   gaam̐bar.*
      "gaam̐bar" "gaam̐bar".
      'The word is "gaam̐bar".

Comment: No analysis due to problems with the sound file

(125) *SLM1: thraa,   thraa*
      NEG.exist, NEG.exist
      'No, he has no photos.'

Comment: No analysis due to problems with the sound file

(126) *SLM3: Sebastian s-kaaving?*
      Sebastian CP-marry
      'Is Sebastian married.' (asking other women)

Comment: No analysis due to problems with the sound file

(127)
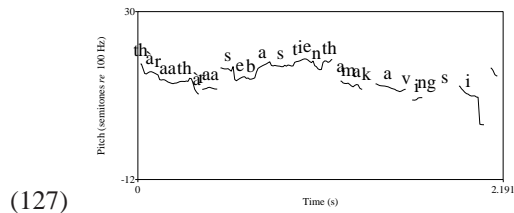
                          L   H
SLM1: thàra thàra | Sebastian |
        no   no   Sebastian
H   L
thama-kaaving=si |
NEG.NONPAST-*marry*=INTERR

'You are not married, are you?'

Comment: The presupposition is that I am not married. The overall contour is assertional, a request for confirmation. The very slight rise at the end indicates a remainder of doubt (cf. (122)).

(128)

	H	L
SLM1: Sri Lanka=ka=jo	|	*kaving* |
Sri Lanka=LOC=EMPH		*marry*
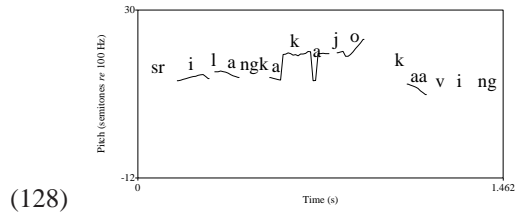
'Marry in Sri Lanka!'

Comment:A jocular imperative. The high tone is located at the end of the argument phrase/the beginning of the predicate phrase (The plateau around the second k is an artefact.)

Part III

Form

This third part of the book discusses the morphosyntactic forms of SLM, i.e. free morphemes, bound morphemes, phrases, predicates, clauses and superclausal constructions. I analyze the formal properties of every item and indicate the semantic or pragmatic functions it can be used for, which are treated in more detail in the fourth part of this description. References to other forms discussed in Part III are indicated by a boxed arrow ⊡→, while references to more detailed discussions of function in part IV are indicated by a circled arrow ⊙→.

Chapter 5

Major and minor word classes

This chapter gives an overview over SLM free morphemes. The chapter thereafter will treat bound morphology. Free morphemes comprise both lexical and functional morphemes. These fall into a number of classes, which will be discussed in turn.

Four major word classes are distinguished in SLM: verbs like *maakan* 'eat' ⊡→5.1, p. 163, nouns like *aanak* 'child' ⊡→5.2, p. 196, adjectives like *kiccil* 'small' ⊡→5.3, p. 210, and adverbs like *kumaareng* 'yesterday' ⊡→5.4, p. 214. These classes are all fairly large and can accept new members. The members of a class all share certain characteristics, by which class membership can be checked.

There are also a number of minor classes such as personal pronouns, which are rather small and which cannot accept new members. Of these classes, a full extensional list of members will be given, along with some characteristic traits.

5.1 Verbs

SLM Verbs can be subdivided into canonical verbs, which are the largest subclass ⊡→5.1.1, p. 164, two existentials ⊡→5.1.2, p. 164, and a number of defective verbs, which cannot take the whole array of morphology available for canonical verbs ⊡→5.1.3, p. 169. Verbs can combine with other verbs in multi-verb constructions, of which an overview is given in Section ⊡→5.1.4, p. 171. Some canonical verbs have also acquired a grammatical meaning when combined with other verbs. These are the vector verbs, which will be discussed in Section ⊡→5.1.5, p. 173. Finally, verbs can participate in a number of special constructions, like the perfect construction, which will be discussed in Section ⊡→5.1.6, p. 192.

5.1.1 Canonical verbs

Sri Lanka Malay has a large class of verbs, which most often denote events like *laari* 'run', but states like *thiidor* 'sleep' can also be found in this class. All verbs share the following characteristics:

1. They can take the following affixes: non-past tense *arà-*, past tense *anà-*, past tense *su-*, anterior tense *asà-*, irrealis *anthi-*, infinitive *mà-*, as well as some additional minor TAM-affixes.

2. They are negated by preposed *thàrà-* for past tense, postposed *thraa* for perfect, preposed *thama-* for present and future and *jamà-* for nonfinite contexts.

3. They are never negated by *bukang*.

4. They must be reduplicated in order to modify other predications

   There is a risk to mistake adjectives for verbs because adjectives can easily convert to verbs and take verbal morphology. Therefore, test for adjectivehood should also be done for candidate verbs to exclude this. See further discussion in the section on adjectives ⊡→5.3, p. 210. An example of a verb would be *laari* 'run', which could be inflected as *aràlaari, anàlaari, anthilaari,* etc. 'run, ran, will run, etc.'. This verb would be negated as *thàràlaari* for past tense, as *thamalaari* for non-past tense, and as *laari thraa* in the perfect. *Laari* can never be negated by *bukang*, and would have to be used as *laari˜laari* in order to modify another predication.

5.1.2 Existential verbs

There are two existentials, *aada* ⊡→5.1.2.1, p. 165 and *duuduk* ⊡→5.1.2.2, p. 168. The latter can only be used with animate referents, while the former can be used with both animate and inanimate referents. *Duuduk* thus carries a positive specification for animacy, while *aada* is underspecified for animacy.

   Example (1) shows that the animate existential *duuduk* cannot combine with the inanimate referent *ruuma* 'house'. Here, *aada* has to be used. Example (2) shows that both *aada* and *duuduk* can combine with the animate referent *mlaayu pada* 'Malays'.

(1) *Kluŭm̆bu=ka   ruuma pada **aada**/*arà-duuduk.*
     Colombo=ʟᴏᴄ house   ᴘʟ   exist/ɴᴏɴ.ᴘᴀsᴛ-exist.ᴀɴɪᴍ.
     'There are houses in Colombo.' (K081103eli02)

(2) *Kluŭm̆bu=ka   mlaayu pada **aada/arà-duuduk**.*
     Colombo=ʟᴏᴄ house   ᴘʟ   exist/ɴᴏɴ.ᴘᴀsᴛ-exist.ᴀɴɪᴍ.
     'There are Malays in Colombo.' (K081103eli02)

5.1.2.1 *aada*

Aada is one of the most frequent morphemes in SLM. It is normally pronounced [aaɖa] and rarely [aaɽa]. Most frequently, *aada* is used to refer to present situations. In distinction to other verbs, it does not carry the prefix *arà-* in those cases (cf. Smith & Paauw 2006:169).[1] In other temporal contexts, *aada* can take the relevant verbal morphology, just like any other verb, Example (3) shows the standard use of *arà-,* while example (4) shows the impossibility to combine *arà-* with *aada.*

(3) *Spaaru mlaayu pada **arà**-oomong.*
    some Malay PL NON.PAST-speak.
    '(Only) some Malays speak (it).' (G051222nar04)

(4) *Avuliya pada (***arà-**)aada.*
    saint PL NON.PAST-exist.
    'There are saints.' (B060115nar05)

In past contexts, *aada* can take the relevant morphology, i.e. the past tense prefixes *su-* (5)(6) and *anà-* (7)(8).

(5) *Hathu muusing=ka [...] hathu kiccil ruuma **su-aada***
    INDEF time=LOC [...] INDEF small house PAST-exist.
    'Once upon a time, there was a small house.' (K07000wrt04)

(6) *Se=dang bannyak creeveth pada **su-aada**.*
    1s.DAT much trouble PL PAST-exist
    'I had a lot of troubles' (K051213nar01)

(7) Talisman *hatthu **anà-aada** kiyang.*
    talisman INDEF PAST-exist EVID.
    'Apparently, there was a talisman.' (K051206nar02)

(8) *Itthukang **anà-aada** [Mr Janson katha hathu oorang].*
    then PAST-exist Mr Janson QUOT one man.
    'Then there was a certain Mr Janson.' (K051206nar04)

Examples (9) and (10) show that there is no difference in the use of the prefix between *aada* and other verbal predications as far as the prefixes *su-* and *anà-* are concerned.

---

[1]This defective status of the existential verb is a common feature of Austronesian languages (Himmelmann 2005:138).

(9) *Baapa=le   aanak=le guula **su-maakang**.*
    father=ADDIT son=ADDIT sugar PAST-eat
    'Father and son ate sugar.' (K070000wrt02)

(10) *Incayang=pe plajaran=nang incayang Kandi=nang **anà-dhaathang**.*
    3S.POLITE   education=DAT 3S.POLITE Kandy=DAT   PAST-come
    'For his education, he came to Kandy.' (K060108nar02)

The following examples show the inflectional potential of *aada* in other domains, namely irrealis mood (11), negation (12) and deontic modality (13). Example (14) finally shows the use of *aada* in a relative clause.

(11) *[Sebastian=ka su-meet-king]=nang derang=dang baae=dang suuka*
    Sebastian=LOC PAST-meet-CAUS=DAT 3PL=DAT   good=DAT   like
    ***thi-aada**.*
    IRR=exist
    'They would have very much liked to meet you.' (B060115cvs07)

(12) Confrontation *pada **thàrà-aada**   Tigers=samma.*
    Confrontation PL   NEG.PAST-exist Tigers=COMIT
    'There   were   no   confrontations   with   the   Tigers.'
    (K051206nar20,K081105eli02)

(13) *Prompang klaaki samma oorang inni=ka   **masà-aada**.*
    girl   boy   all   man   PROX=LOC must=exist
    'Boys and girls, everybody must be in this.' (K060116nar05)

(14) *[Incayang=pe kàpaala=ka **anà-aada** thoppi]=dering moonyeth pada=nang*
    3S.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
    su-buvang puukul.
    PAST-throw hit
    'He took the hat from his head and violently threw it at the monkeys.'
    (K070000wrt01)

The negation of *aada* if formed with *thàrà-* for past contexts, as seen in example (12), but for present tense, the negative particle *thraa* normally substitutes *aada* (15).

(15) *Se=ppe umma* **thraa**.
　　1s=POSS mother NEG
　　'My mother is dead.' (lit. 'is not here') (B060115prs03)

For negating events in the future, *thama-aada* can be used, but this is rarely ever found in the corpus. An elicited examples is given in (16).

(16) Iceland=*ka mlaayu thama-aada.*
　　Iceland=LOC Malay　NEG.IRR-exist.
　　'There will never be Malays in Iceland.' (K081103eli02)

Aada plays a key role in a number of constructions, namely the existential construction (17), the locational construction (18) and the possessive construction (19). Furthermore, *aada* is also used in a periphrastic construction with the bare verb or the conjunctive participle to form the perfect tense (20), and in another construction with a verb in the infinitive to form a periphrastic construction that conveys obligation (21). A more detailed discussion of these constructions can be found in Sections ⊡→5.1.6.1, p. 192 and ⊡→5.1.6.2, p. 194.

(17) *Hatthu kumpulan* **aada**.
　　INDEF　association exist
　　'There is an association.' (K060116nar23)

(18) *Mlaayu=ka=jo　bannyak avuliya Seelon=**ka** **aada**.*
　　Malay=LOC=EMPH many　　saint　Ceylon=LOC exist.
　　'Among the Malays there are many saints in Sri Lanka.' (K060108nar02)

(19) *Lorang=**ka** duvith* **aada**=si?
　　2PL=LOC　　money exist=INTERR
　　'Is the money with you?' (K060116nar21)

(20) *Uumur=nang kuurang, sdiikith oorang pada* **kaving**=le　**aada**.
　　age=DAT　　few　　few　　man　PL　marry=ADDIT exist
　　'Below that age, there are few people who are married, too.' (K061122nar01)

(21) *Lai　aapa, lai　aapa* **mà**-*biilang* **aada**?
　　other what　other what INF-say　　exist.
　　'What else (do I) have to say?' (K060108nar02)

Etymologically *aada* stems from the old TM existential **ada* and is thus related to

the present tense prefix *arà-* ⊡→6.2.2.1, p. 287. This might also be a reason why they cannot co-occur.

5.1.2.2   *duuduk*

The second existential verb, *duuduk,* is always inflected as a full verb. The final consonant is eroding, and pronunciations like [ɖuːɖuʔ] and [ɖuːɖu] can also be heard (in the South as well [ɖuːɖuŋ] Slomanson (2008b)). Just like *aada,* it can be used in existential (22), locational (23) and possessive contexts (24), provided that the referent existing, located or possessed is animate.

(22)   *Miskin oorang arà-**duuduk**.   Derang pada=nang samma bole=kaasi*
       poor   man   NON.PAST-exist.ANIM 3PL   PL=DAT   all   can-give
       'Poor people exist. (I) can give everything to them.' (B060115nar04)

(23)   *Sithu=**ka**, hathu bìssar beecek caaya Buruan su-**duuduk**.*
       there=LOC INDEF big   mud   colour bear   PAST-exist
       'There was a big brown bear there.' (K070000wrt04)

(24)   *Thiiga klaaki aade=le   hatthu pompang aade=le*
       three   male younger.sibling=ADDIT one   female   younger.sibling
       *se=**dang** arà-**duuduk**.*
       1s.DAT   NON.PAST-exist.ANIM.
       'I have three younger brothers and one younger sister.' (K060108nar01)

Duuduk cannot be used in either the perfect construction or the debitive construction. It is possible to combine the conjunctive participle of a verb with *duuduk*, but this is semantically different from the perfect construction with *aada*. In example (20), there is semantically only one predication with *aada* only contributing the aspectual value. This is different from *duuduk* in examples (25), which is a full lexical verb. Example (25) must be interpreted as two predications which follow each other in time.

(25)   *Baarang **as**-ambel arà-**duuduk**.*
       goods   CP-taken NON.PAST-stay
       'Having taken the goods, he stayed/sat/was there.' (not:'He has taken the goods.') (K081103eli02)collect items and wait

Duuduk has a second meaning, which is 'to sit', already mentioned in the example above. Additionally, it can mean 'stay'. Sinhala also has two different existential verbs, *tiyenavaa* (inanimate) and *innavaa* (animate). Interestingly, the Sinhala animate

verb historically also means 'to sit', so that the grammaticalization path of *duuduk* can be explained by this adstrate influence.[2]

While *duuduk* has as its etymological origin 'to sit', today it can be used as an existential even for animate beings that are unable to perform the act of sitting, for instance bats in (26). *Duuduk* has thus grammaticalized from a full verb 'to sit' to an existential.

(26)   *Kiccil vavvaal pada daalang=ka arà-duuduk.*
      small bat   PL   inside=LOC NON.PAST-exist.ANIM.
      'There are small bats inside.'

Duuduk is also present on a second grammaticalization path, namely the ablative. When used in its conjunctive participle form, it is used to denote source of motion ⊙→15.1.2.6, p. 570. The use of the perfective form of an existential for SOURCE closely parallels the Sinhala form *innavaa, iňdalaa* and the Tamil *iru(kkiradu), irundu*. Example (27) shows this construction.

(27)   *Suda see* Trinco=*ka*   ***asàduuduk*** *Kluum̌bu=nang su-dhaathang.*
      So   1s Trincomalee=LOC from   Colombo=DAT PAST-come
      'So I went from Trincomalee to Colombo.' (K051206nar20)

Furthermore, *duuduk* can be used as a vector verb to indicate continuous aspect ⊡→5.1.5.7, p. 187.

5.1.3 Defective verbs

A number of verbs are defective in the sense that they cannot combine with the prefix *arà-* 'non-past'. In addition to *aada* discussed above, these are most notably *suuka* 'like' and *thaau* 'know'. In affirmative sentences in the present, they do not take a prefix. (28) and (29) give examples for this. These verbs also rarely mark past tense overtly, which is obligatory for canonical verbs.[3]

(28)   *Go=dang*   *baae=nang* ∅ ***thaau.***
      1s.familiar.DAT good=DAT   know
      'I know very well.' (B060115nar04)

---

[2]Tamil on the other hand does not code animacy in the choice of the existential verb, and Tamil existential verbs do not stem from a historic verb 'to sit'.

[3]The temporal underspecification of 'know' is also found in Sinhala and Tamil.

(29) *Kithang=nang ∅ **suuka** mà-diyath=nang.*
     1PL=DAT   like   INF-see=DAT
     'We would like to meet them.' (B060115cvs03)

(30) *Se=dang baapa=ke* soldier *mà-jaadi* ∅ **suuka**.
     1s.DAT   father=SIMIL soldier INF-become like.
     'I want to become a soldier like daddy.' (B060115prs10)

Anà- can be used on *thaau* when it is combined with the vector verb *ambel* 'take' to yield the meaning of *got to know.*In example (31), the allomorph *nya-* is used instead of *anà-*.

(31) *Itthukapang=jo derang **nya-thaau ambel** derang pada politic=nang suuka*
     then=EMPH   3PL   PAST-know take 3PL   PL   politic=DAT like
     katha.
     QUOT
     'Only   then   will   they   come   to   know   that   they   like   politics'
     (K051206nar12,K081105eli02)

As for negation, the defective verbs have a particular pattern. They are negated by *thàrà-* regardless of reference time, which distinguishes them from other verbs, which take *thama-* in non-past tenses (32)(33). In the cited examples, the prefix *thàrà-* is used despite of the time reference clearly being to the present. It should be noted, however, that for *suuka*, an alternative negation is also available. This negation patterns with adjectival negation. The latter pattern can refer to present (34) or past (35).

(32) *Incayang=pe baapa=pe naama see **thàrà-thaau**.*
     3s.POLITE=POSS father=POSS name   1s NEG-know
     'I do not know his father's name' (K060108nar02)

(33) *Luvar nigiri   kithang=nang mà-pii **thàrà-suuka**.*
     outside country 1PL=DAT   INF-go NEG-like
     'We do not want to go abroad.' (K051222nar04)

(34) *Se **suuka thraa**.*
     1s like   NEG
     'I don't like (it).' (K060116nar08)

	phonological words	full verbs	morphological words	events
verbal compound	1	2	1	1
full verb + vector verb	2	1	1	1
full verb + full verb	2	2	1	1
clause chain	2+	2+	2+	2+

Table 5.1: Different types of multi-verb constructions in SLM and their characteristics

(35) *Bìssar pukurjan see **suuka thraa**.*
      big   work   1s like   NEG
      'I did not want a high post.' (K060116nar08)

5.1.4 Multi-verb constructions

SLM verbs can combine in a number of ways to yield multi-verb constructions. If two verbs are compounded and parsed into one phonological word, we are dealing with verbal compounds ⊡→5.1.4.1, p. 171. If the verbs are parsed into at least two phonological words, there are several possibilities. If one of the verbs is bleached and only provides aspectual or other grammatical information, we are dealing with a vector verb construction ⊡→5.1.4.2, p. 172. If both of the verbs contribute with their literal meaning, we are dealing with a serial verb construction ⊡→5.1.4.3, p. 172. As a last possibility, which is a multi-clause construction rather than a multi-verb construction, clause chains with the conjunctive participle *asà-* also comprise several verbs ⊡→5.1.4.4, p. 173. These four types of multi-verb constructions can then be differentiated by looking at the number of phonological words, the number of full verbs, the number of morphological words, and the number of events (Aikhenvald 2006b, Dixon 2006). Table 5.1 gives an overview.

The following sections will discuss prototypical instances of the four types.

5.1.4.1 Verbal compounds

The string *kasithaau* in (36) is parsed into one phonological word, which can be seen from the absence of a long vowel in *kasi*. It includes two full verbs, *ka(a)si* 'give' and *thaau* 'know', which both contribute to the meaning with their literal reading. We are dealing with one morphological word, which is only inflected once for TAM, with the debitive prefix *masà-*. Finally, the act of informing is only one event, not two separate events of first giving and then knowing.

(36)   *Badulla* Kandy *Matale samma* association=*nang* **masà-kasi-thaau**.
        Badulla Kandy Matale samma association=DAT   must-give-know
        'Badulla, Kandy, Matale, we must inform all other associations.'
        (K060116nar06)

Verbal compounds are discussed in more detail in ⊡→7.1.5, p. 403.

5.1.4.2   Vector verbs

The string *naangis ambel* in (37) is parsed into two phonological words, which can
be seen from the long vowel in *naangis*. This string includes one full verb, *naangis*
'weep', which contributes it literal meaning, and one vector verb, *ambel*, which only
contributes inchoative aspectual information, and not the literal meaning of 'take'.
There is no action of taking taking place. We are dealing with one morphological word
naangis ambel, which can be seen from the TAM-inflection, which is only present for
naangis and not for *ambel*. Finally, the crying of the child is only one event, not one
event of crying and a separate event of taking.

(37)   *Suda Andare=pe   aanak=le   baapa anà-biilang=kee=jo **asà-naangis***
        thus Andare=POSS child=ADDIT father PAST-say=SIMIL=EMPH CP-cry

        ***ambel** su-dhaathang.*
        take   PAST-come
        'So Andare's son started to cry and came (running) just like the father had told
        him.' (K070000wrt02)

Vector verbs are discussed in more detail below in ⊡→5.1.5, p. 173.

5.1.4.3   Full verb serialization

The string *salba laari* in (38) constitutes two phonological words. There are two full
verbs in it, *salba* 'escape' and *laari* 'run', which both equally contribute with their
literal reading. We are dealing with one morphological word, indicated by one TAM
marking, perfect *aada*. We are also only dealing with one event, escaping in a running
way, not with one event of escaping and a further event of running.

(38)   *Incayang theeṁbak abbis,   **salba laari** aada thumpath=nang.*
        3s.POLITE shoot   finished escape run   exist place=DAT.
        'After the shooting he escaped to the (aforementioned) place.'
        (K051206nar02)

Full verb serialization is discussed in more detail in ⊡→10.1.3, p. 466.

5.1.4.4 Clause chains

Example (39) contains two subordinate clauses where the verb is inflected with the conjunctive participle prefix *(a)s(à)-*, and one main clause in the end. All the verbs are full verbs which contribute their literal meaning. We are dealing with much more than one phonological word, and every verb is inflected individually, indicating that we are dealing with three morphological words as far as the verbs are concerned. Finally, we are not dealing with one event, but with three, first fighting, then helping, then settling down.

(39)   a.   *Oorang pada **asà**-pirrang*
            man   PL   CP-wage.war.
            'After having waged war'

       b.   *derang=nang **asà**-banthu*
            3PL=DAT   CP-help
            'and after having helped them'

       c.   *siini=jo   se-ciiňǧgal.*
            there=EMPH PAST-settle
            'the people settled down right here.' (K051222nar03)

Clause chains are discussed in more detail in ⊡→13.3, p. 529.

5.1.5   Vector verbs

Vector verbs (Pray 1970, Hook 1974, Masica 1991, Kachru 1993) (called 'auxiliaries' by Smith & Paauw (2006) for SLM) are verbs which are used after full verbs to highlight a certain semantic aspect of the verb, like perfectivity, intensity or beneficiary.[4] Abbi & Gopalakrishnan (1991) defines these verbs as "a sequence of at least two verbs where the first member is the main or predicating verb and the second member, although homophonous with an independent verb in the language, does not appear in its primary lexical meaning; V2 only occurs in the sequence to mark the main verb V1 for certain 'grammatical' features."

Thus, the verb *aajar* means 'teach', but complemented with the vector verb *kaasi* 'give', it still means 'teach', but highlights the profitable aspect of the teaching. *Bu-*

---

[4]Similar verbs in languages of South Asia and beyond have been given a variety of names. 'Aspectual verbs', 'intensifying verbs', 'verbal auxiliaries', 'explicator verbs', 'light verbs'. For SLM, some of these terms are not appropriate: not all of the verbs convey aspectual or intensifying meaning, so that those terms would be misnomers. 'Auxiliary' implies that the verb is used to carry tense or agreement, neither is the case in SLM, so this is a bad term as well. 'Light verb' and 'Explicator verb' seem to be viable alternatives to the term 'Vector verb' used in this grammar. 'Explicator Compound Verb' is actually also found very often (Abbi & Gopalakrishnan 1991, Abbi 1994), but, in this grammar, 'Compound Verbs'⊡→7.1.5, p. 403 is used for a different construction, so that this name might lead to confusion. 'Light verb' is very often used in combination with a noun (Butt 2003), which is not what we find in SLM, hence this term is not used either.

vang means to throw, and combined with the vector verb *puukul* 'hit' it means 'throw violently'. A characteristic of vector verb is that they always occur after another verbs and that their meaning can be bleached as compared to the full verb. The full verb can carry a TAM-prefix, while the vector verb never does (Smith & Paauw 2006:171).[5]

All vector verbs have a corresponding full verb. The number of vector verbs is limited, the following eight vector verbs have been retrieved:[6]

- *ambel* 'take' for (self)-benefactive and ingressive

- *kaasi* 'give' for benefactive

- *abbis* 'finish' for completive

- *thaaro* 'put' for affectedness

- *simpang* 'keep' for continuative

- *puukul* 'hit' for intensity

- *duuduk* 'sit' for progressive

- *kìnna* 'strike' for adversative

Abbi & Gopalakrishnan (1991) divide vector verbs into three groups: ASPECTUAL, ADVERBIAL and ATTITUDINAL. It appears that the SLM vector verbs can carry aspectual information (inchoative, completive etc.), adverbial information (intensity, affectedness) and attitudinal information (adversative).

Examples of non-aspectual information are found in the following examples: Intensity of the action is expressed by *puukul* 'hit' in (40); motivation of the action is expressed by *kaasi* 'give' in (41).

(40) *Incayang=pe   kàpaala=ka anà-aada thoppi=dering moonyeth pada=nang*
      3s.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
      *su-**buvang puukul**.*
      PAST-throw hit
      'He took the hat from his head and violently threw it violently at the monkeys.'
      (K070000wrt01)

---

[5]This is actually a striking difference to many other South Asian languages, where the main verb is typically in the (non-finite) conjunctive participle form, and the vector verb carries tense and agreement (Masica 1976:141).

[6]This list fits well with the list compiled by Abbi & Gopalakrishnan (1991:173f) for common vector verbs in India, which includes GO, COME, GIVE, TAKE, KEEP, PUT, SIT and FALL. FINISH, HIT and STRIKE are additions found in SLM.

(41) *Kithang=pe ini*　younger generation=*nang=jo konnyong masà-**biilang***
　　　1PL=POSS　PROX younger generation=DAT=EMPH few　　must-say

　　　***kaasi**, masà-aajar.*
　　　give　must-teach

　　　'It is to the younger generation that we must explain it, must teach it.'
　　　(B060115cvs01)

Note that *biilang* 'say' and *kaasi* 'give' do not differ in their basic valency as both can be used with both a theme and a goal. *Kaasi* thus does not serve to introduce a new participant here, but rather highlights the beneficial nature of the action. This is rephrased at the end of the sentence as *aajar* 'teach'. One could be tempted to see *biilang+kaasi* 'say+give' as equivalent to *aajar*, but this does not capture the entirety of the facts, as *aajar+kaasi* 'teach+give' can also be found (42).

(42) *Itthu muusing Islam igaama nya-**aajar kaasi** 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)

As in the preceding example, *kaasi* is not used to introduce a new participant here, but rather to highlight the beneficial nature of the action. To drive the point home, if *aajar* is equivalent to the trivalent construction *biilang+kaasi* in example (41), then there is certainly no need to use *kaasi* again on *aajar* in example (42) to introduce yet another participant.

TAM-prefixes can only attach to the first verb, while postverbal material can only be used after the last verb. No material can intervene between the two verbs. This distinguishes vector verbs from the similar clause chaining construction ⊡→13.3, p. 529, which permits intervening material, as in the following example, where a second TAM-prefix (*su-*) separates the two verbs.

(43) *Spaaru oorang pada biini pada=yang **asà-ambel su-pii**.*
　　　some　man　PL　wife PL=ACC　CP-take　PAST-go.

　　　'Some men took wives and left/Some men left with wives.' (K051206nar07)

It is possible to use vector verbs in clause chains, as the following example shows.

(44)   a. *Itthu=nang blaakang [inni oorang lìkkas˜lìkkas thoppi pada=yang*
           dist   after   PROX man   fast˜RED   hat   PL=ACC

           *asà-[**kumpul ambel**]vector]chainedclause*
           CP-collect   take.

           'After that the man quickly picked up his hats and' (K070000wrt01)

       b. *[sithu=ka=dering su-pii]mainclause.*
           there=LOC=ABL   PAST-go

           'left from there.' (K070000wrt01)

In this example, the vector verb combination *kumpul ambel* 'collect for oneself' is
used within a clause chain *asàkumpul ambel ... supi* 'having collected, he left'. The
conjunctive participle *asà-* is used only once, on *kumpul*, so that we are dealing with
two clauses, comprising a total of three verbs. Two are in the dependent clause and
one is in the matrix clause. The two verbs in the dependent clause count as a unity for
matters of TAM-marking and so the prefix only attaches to the first one.

5.1.5.1   *ambel* 'take'

Ambel has 'take' as its meaning when used as a full verb. When used as a vector verb,
it can be used to highlight the benefactive aspect of a verb. Most often this benefactive
aspect applies to the speaker, but it can occasionally also be used for other people
profiting from the action. The second use of *ambel* as a vector verb is to indicate
ingressive aspect.[7]

Ambel as a vector verb has to be distinguished from both the use as a full verb and
from the homonymous postposition found by Slomanson (2008b).

The first use of *ambel* is to indicate the profitable nature of the action denoted by
the full verb ⊙→15.1.2.10, p. 573.[8] In example (45), the speaker inquires whether he
should ask money from an elder member of the community. Since this money would
profit him, the full verb *mintha* 'ask' is complemented by *ambel* as a vector verb.

(45)   *[Tony Hassan* uncle=*nang anà-kaasi duvith] athi-mintha **ambel**=si?*
           Tony Hassan uncle=DAT   CP-give   money IRR-ask   take=INTERR

           'Shall I ask for the money you gave to uncle Tony Hassan?' (K071011eml01)

---

[7]Smith et al. (2004) note that *ambel* denotes progressive aspect (also see Ansaldo 2009b). Slomanson
(2008b) shows that this reading of *ambe(l)* is not related to the verb 'take', but a development from the
historical adposition **sambil* 'while', which is reanalyzed as progressive aspect marker. This marker seems
to be a Southern morpheme and is not found in the Upcountry.

[8]The use of the verb TAKE in this function is typical of Indo-Aryan languages (Abbi & Gopalakrishnan
1991:175), found in Sinhala (Jayawardena-Moser (2004:44), Matzel (1983:81)), but also in Tamil (Arden
(1934:193), Schiffman (1999:95), Fedson (1993)).

The use of *mintha* without *ambel* is perfectly possible, as the following example shows. *Ambel* is thus optional.

(46) *Derang pada arà-mintha   Ø nigiri.*
     3PL   PL   NON.PAST-ask country
     'They are asking for a country of their own.' (K051206nar10)

Like in example (45), example (47) shows a use of *ambel* for an action which profits the agent, namely picking up hats that a group of monkeys had thrown to the ground. The man is a hat-seller and has thus an interest in gathering his hats.

(47) *[Itthu=nang blaakang inni oorang lìkkas˜lìkkas thoppi pada=yang*
     dist   after   PROX man   fast˜RED   hat   PL=ACC
     *asà-**kumpul ambel**], sithu=ka=dering su-pii.*
     CP-collect   take   there=LOC=ABL PAST-go
     'After that, the man quickly picked up his hats and left that place.'
     (K070000wrt01,K081104eli06)

Slomanson (2008b) postulates a rule that aspect morphemes are always realized pre-verbally in non-finite clauses. Example (47) shows that this rule does not hold. The subordinate clause, indicated by brackets, contains the verb *kumpul* 'collect', which is marked with the non-finite prefix *asà*. The vector verb *ambel* contributes aspectual information, but is found *post*-verbally in the subordinate clause, invalidating the rule stated by Slomanson.

In example (48), the British captured countries, to their profit. The verb *peegang* 'catch' is used for this, but the additional use of *ambel* makes clear that the British profited from the action.

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

This contrast with the use of *peegang* in (49), where the action is not beneficial.

(49) Heart *attack asà-**peegang** baapa=le   su-niiňǧgal.*
     heart attack CP-catch   father=ADDIT PAST-die
     'My   father   got   a   heart   attack   and   died   as   well.'
     (K051205nar05,K081104eli06)

Still the use of *peegang* is optional even in beneficial context, as the following example shows.

(50)   Singapore=*jona anà-**peegang** ∅.
       Singapore=PHAT PAST-catch
       'They captured Singapore.' (K051206nar07,K081104eli06)

The benefit need not be realized at the time of speaking, as shown by (51), where the wish is expressed that in the future the addressee will be able to enquire from other people the things he is interested in.

(51)   *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=**caanya ambel***
       Hopefully other know brother sister   PL=LOC   can-ask   take
       [nya-gijja lai   saapa=kee aada]=si   katha.
       PAST-make other who=SIMIL exist=INTERR QUOT
       'Hopefully, you can enquire from another person you know whether there is someone else who did something.' (N061031nar01,K081104eli06)

Normally, the beneficiary of the action is the agent, but it is also possible that the action denoted by the full verb followed by *ambel* is beneficial for another participant. This is the case in the following example, where the family keep the bear in their home. This is profitable for the bear because it protects him from the coldl it is not so profitable for the family.

(52)   *See=yang lorang=susamma diinging muusing sangke-habbis anà-**simpang***
       1s=ACC   2PL=COMIT   cold   season until-finish PAST-keep
       ambel.
       take
       'You have kept me together with you until the cold season was over.'
       (K070000wrt04,K081104eli06)

The second meaning that can be conveyed by *ambel* as a vector verb is ingressive, especially with the verb *thaau* 'know'. In the first example, (53), two fooled women learn about their being fooled. Before they were not aware of this, but now they are, so that there is a change of state taking place, expressed by *ambel*.

(53) *Kanabisan=ka=jo duva oorang=le anà-**thaau ambel** [Andare duva*
last=LOC=EMPH   two man=ADDIT PAST-know take   Andare two
oorang=yang=le asà-enco-kang aada] katha.
man=ACC=ADDIT CP-fool-CAUS   exist EMPH
'At the very end, both women understood that Andare had fooled both of
them.' (K070000wrt05,K081104eli06)

A similar situation obtains in (54), where the young people will come to know
about their interest in politics only when a certain condition is met. Only then will the
state change from ignorance to knowledge, again conveyed by *ambel*.

(54) *Itthukapang=jo derang **nya-thaau ambel**, derang pada* politic=*nang suuka*
then=EMPH   3PL PAST-know take 3PL   PL politic=DAT like
katha.
QUOT
'Only   then   will   they   come   to   know   that   they   like   politics'
(K051206nar12,K081105eli02)

This ingressive meaning can be combined with expression of modality. In (55),
the change of state is put into facultative modality.

(55) a. *Malaysia samma oorang=pe naama pada Maas.*
Malaysia all   man=POSS name PL   Maas
'People from Malaysia are all called "Maas".'

  b. *Suda itthu=dering=jo kithang=nang ini   Indonesia=pe   oorang=si*
thus DIST=ABL=EMPH 1PL=DAT   PROX Indonesia=POSS man=disj
*giithu   kalthraa Malaysian oorang=si katha **bàrà=thaau ambel**.*
that.way if-NEG   Malaysian man=disj QUOT can-know   take
'So with that we can come to know whether someone is Indonesian or
otherwise Malaysian' (K060108nar02)

In some cases, it is difficult to decide whether *ambel* is used in the beneficiary
sense or the ingressive sense because both are possible. In (56), a student writes down
the correct answers in an exam after she had seen the answers in dream. *Ambel* in this
case can both mean that she started to write down, or that the action of writing down
was beneficial to her.

(56) *Baaye=nang ingath-an   tak tak katha su-**thuulis ambel**.*
good=DAT   think-NMLZR tak tak QUOT PAST-write take
'She wrote the thoughts down correctly like tak tak tak.' (K051220nar01)

Another example is taking up again well-paid work, which is both a change of state and beneficial.

.

(57)   a. *Bannyak* experience *se=dang* engineering branch=*ka anà-daapath.*
        much   experience 1s.DAT   engineering branch=LOC PAST-get.
        'I got a lot of experience in the engineering branch.'

      b. *Suda itthusubbath=jo see laile* modern engineering *asà-**gijja ambel***
        thus therefore=EMPH 1s again modern engineering CP-make take
        arà-pii.
        NON.PAST go
        'So therefore I still take up modern engineering and continue.'
        (K051206nar20,K081104eli06)

Finally, in a story where Andare fools the king with the help of his son who starts to cry, *ambel* could be interpreted as inceptive, or as beneficial to the cryer, since it furthers the cause of fooling the king.

(58)   *Suda Andare=pe aanak=le   baapa anà-biilang=kee=jo **asà-naangis***
       thus Andare=POSS child=ADDIT father PAST-say=SIMIL=EMPH CP-cry
       ***ambel** su-dhaathang.*
       take   PAST-come
       'So Andare's son started to cry and came (running) just like the father had told him.' (K070000wrt02)

Smith & Paauw (2006:171f) give the meaning of this vector verb as progressive, the relevant example is repeated in (59)

(59)   *Deraŋ daataŋ   dupaŋ n-duuduk siini*_(,) *Islam ayn-duuduk ambel*
       3PL   come-NOMZ-DAT before past-be   here Islam past-be   pres
       jo aaɖa, tuʋan.
       foc be   sir
       '[Islam] was [here] before they came, Islam was [present] here [all along], sir [Lit: was being here, was staying here].' (Smith & Paauw (2006:171(29)), original orthography, comma added)

My informants found it very hard to make sense out of this example. They offered the following correction:

(60) *Duppang, derang anà-dhaathang duuduk siini, Islam anà-duuduk ambel=jo*
earlier 3PL PAST-come stay here, Islam PAST-stay take=EMPH
aada
exist
'Earlier, they have come here, at that time also they were Muslims.'
(K081104eli06)

For the English meaning given by Smith & Paauw (2006), they give the following translation:

(61) *Derang siini mà-dhaathang=nang duppang, siini Islam su-aada*
3PL here INF-come=DAT before here Islam PAST-exist
'Before they came here, Islam was (already) here.' (K081104eli06)

The second example which Smith & Paauw (2006) use to show the progressive reading of *ambel* is given below in the original orthography.

(62) *Kaaram suɖa kubaali le ini reepot atu su baaʋuŋ ambel aada.*
now link again emph this trouble one past arise prog be
'Now, see, this trouble is arising again.' (Smith & Paauw (2006:171(30)),
original orthography)

My informants reacted to this sentence by stating: "Anybody who has talked to you would have tried to convince you that he knows a bit of Malaysian also." While this statement is to be taken with a grain of salt, it illustrates the puzzlement which the sentence caused. There are some reasons for this:

- Given the past tense marker *su-*, the gloss as present progressive is surprising. On the other hand, the gloss 'was arising' seems problematic as well.

- The presence of both a deictic and an indefiniteness marker around *reepot* 'trouble' makes little sense.

- Why is there marking for past tense (*su-*) and perfect tense (*aada*) at the same time (similar to English *have arose* instead of past *arose* or perfect *have arisen*)?

- How is the cooccurrence of *kaaram* 'now' and the past inflection on the predicate to be explained?

When asked how they would render the sentence, the informants gave the following:

(63) *Karang suda laskalli ini   reepoth hatthu arà-bavung (ambel) dhaathang*
      now   thus again   PROX trouble INDEF NON.PAST-rise take   come
      'Now, once again, another problem is arising' (K081104eli06)

The use of *ambel* is optional. In this sentence, *ambel* could have the meaning of
progressive, as suggested by Smith & Paauw (2006), or the meaning of inchoative, as
argued for in the rest of the preceding section.

5.1.5.2   *kaasi* 'give'

Kaasi has as its literal meaning 'give' and is used to highlight the beneficial nature
of an action for another person than the agent (termed 'alterbenefaction' by Lehmann
(1989:227) for Tamil[9]) ⊙→15.1.2.10, p. 573. This is the case in the following two
examples, which deal with imparting knowledge. Note that *aajar* is used as a synonym
for *biilang kaasi* in (64), but as *aajar kaasi* in (65), so that a valency changing function
of *kaasi* is unlikely.

(64) *Kithang=pe ini   younger generation=nang=jo konnyong masà-**biilang***
      1PL=POSS   PROX younger generation=DAT=EMPH few   must-say
      ***kaasi**, masà-aajar.*
      give   must-teach
      'It is to the younger generation that we must explain it, must teach it.'
      (B060115cvs01)

(65) *Itthu muusing [Islam igaama nya-**aajar kaasi** ∅] 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)

Another example of the beneficial nature of an action highlighted by *kaasi* is (66),
where needy people receive soup. *Kaasi* is optional, but used to emphasize the benefit.

(66) *Bannyak=le soup=jo   nya-baapi **kaasi**.*
      much   soup=EMPH PAST-bring give
      'We brought (them) a lot of soup.' (B060115nar02)

Looking up things for another person is also marked by *kaasi* in (67).

---

[9]Also see Chater (1815:93) for the analogous Sinhala construction.

(67) *Se=dang aathi-ka   asà-kluuling bole=caari **kaasi** Malays pada=pe atthas.*
1s.DAT   heart=LOC CP-roam   can-search give Malays PL=POSS about
'I am going to go through my mind, and I will forward it to you what I find
about the Malays.' (G051222nar01)

5.1.5.3   *abbis* 'finish'

This vector verb is used to indicate completive aspect ⊙→15.5.4.4, p. 638 (Smith &
Paauw (2006:171), Ansaldo (2009b)). It can occur with a simple or a geminate stop.
It is related to the full verb *abbis*, which means 'finished'.

Examples (68)-(70) give the use of *abbis* as a full verb, the only verb in the re-
spective clauses.

(68) Water *boothol samma **abbis**.*
water bottle   all   finish
'All water bottles were finished.' (K051206nar16)

(69) a. *[Ruuma pukurjan **abbis**]=nang blaakang,*
house   work   finish=DAT   after.
'After the house work was finished,'

   b. *Derang [dìkkath=ka aada laapang]=nang mà-maayeng=nang su-pii.*
3PL   vicinity=LOC exist ground=DAT   INF-play=DAT   PAST-go.
'they went to play to the nearby ground.' (K070000wrt04)

(70) a. *Ini   mlaayu pada siini pìrrang samma **abbis**=nang blaakang,*
PROX Malay PL   here war   all   finish=DAT after
'After these Malays had finished the war,' (K051206nar07)

   b. *derang pada=pe duty samma **abbis**=nang blaakang, ...*
3PL   PL=POSS duty all   finish=DAT after
'after they had completed their duty, ...' (K051206nar07)

The above examples do not show tense marking on the full verb *abbis*. Tense
marking is often suppressed for *abbis*, but the following example indicates that tense
marking is possible.

(71) *Karang inni hamma **s-abis**,   bukang.*
now   PROX all   PAST-finish, TAG
'Now all was finished, wasn't it?' (K060116nar11)

The use as a full verb exemplified above contrasts with the use as a vector verb given in (72), where *abbis* indicates the completive aspect of the full verb *rubbus* 'boil'.

(72) *Baaye=nang rubbus **abbis**.*
      good=ᴅᴀᴛ   boil   finish
      'When it has boiled well.' (K061026rcp01)

Other examples of *abbis* as a vector verb are (73), where it indicates completive aspect of the converted adjective *maasak* 'cooked', and (74), where it indicates the extent of bending the back which is necessary to enter the cave.

(73) *Vattakka maasak **abbis**, ...*
      pumpkin cooked compl
      'When the pumpkin was (completely) cooked ...' (K051206nar16)

(74) *Giithu=jo   thuuňduk **abbis**=jo   masà-pii.*
      like.that=ᴇᴍᴘʜ bent   finish=ᴇᴍᴘʜ must-go
      'You must enter there completely crouched.' (K051206nar02)

When combined with *dhaathang* 'come', the meaning is lexicalized as 'arrive' as seen in the following examples:

(75) ***Dhaathang abbis** [...] Seelon=ka   arà-duuduk*
      come   finish [...] Ceylon=ʟᴏᴄ ɴᴏɴ.ᴘᴀsᴛ-stay
      'When they had finished coming (=arrived), they stayed in Ceylon.'
      (K051206nar07)

(76) *Suda derang pada **dhaathang abbis**=jo   derang pada itthu=jo*
      thus 3ᴘʟ   ᴘʟ   come   finish=ᴇᴍᴘʜ 3ᴘʟ   ᴘʟ   ᴅɪsᴛ=ᴇᴍᴘʜ
      arà-biilang.
      ɴᴏɴ.ᴘᴀsᴛ-say
      'So when they had arrived, they said the following:' (K051206nar05)

The use of *abbis* as a vector verb has to be distinguished from its use as a full verb, given above. Furthermore, there is a less common use of *abbis-* as a completive verbal prefix (Slomanson 2008a). An example of this is (77), where *abbis* is found in preverbal position.

(77) *Itthu **abbis maakang** kalu kithang=nang bole=duuduk hatthu=le*
DIST finish eat   if   1PL=DAT   can-stay   INDEF=ADDIT
jamà-maakang=nang two duuva two o'clock=ke=sangke bole=duuduk.
NEG.NONFIN-eat=DAT   two two   two o'clock=SIMIL=until can-stay.
'If we eat it up, we can stay up until two o'clock without eating anything.'
(K061026rcp04)

A similar example is (78). It is probable that preverbal *abbis-* is a variant (and possibly the etymological source) of the conjunctive participle, normally pronounced *asà-* ⊡→6.2.1.4, p. 276.

(78) *Thuuju thaaun luvar nigiri=ka   asà-duuduk karang **abbis**-dhaathang*
seven year   outside country=LOC CP-stay   now   finish-come
aada.
exist.
'Having stayed seven years abroad, I have now returned.' (B060115prs13)

5.1.5.4 *thaaro* 'put'

Thaaro has 'put' as its literal meaning and highlights the affectedness of the patient ⊙→15.1.2.2, p. 563. Smith & Paauw (2006:171) give its meaning as 'wilful action with irreversible change of state or condition', which captures the facts presented below very well.[10]

The following two examples show that *thaaro* can be used when the change of state is important, detrimental and permanent, such as tearing (79), whereas *thaaro* cannot be used when no such permanent detrimental change of state takes place, as with writing, which leaves the book basically unaffected (80).

(79) *Se ini   buk arà-soovek-kang   thaaro.*
1s PROX book NON.PAST-torn-caus put
'I am tearing the book to pieces.' (K081104eli06)

(80) *Se ini   buk arà-thuulis   (*thaaro).*
1s PROX book NON.PAST-writing put
'I am writing this book.' (K081104eli06)

A naturalistic example of the use of *thaaro* as a vector verb is (81), where the detrimental nature of killing is highlighted by *thaaro*.

---

[10]Also see the discussion of the semantically related Tamil verb *pooḍu* 'put' in Schiffman (1999:89f).

(81)   *Incayang=yang **siaanu** asà-buunung thaaro.*
       3s.POLITE=ACC   3s.PROX CP-kill   put
       'This one has killed him.' (K051220nar01)(K081103eli02)

5.1.5.5   *simpang* 'keep'

Simpang indicates that the state-of-affairs is continuous and will persist beyond the
reference time frame. An example for this is (82), where the state of sugar lying on
a mat to dry will persist for some more time. The example is a bit involved in that it
consists of a vector verb construction in a purposive clause (*mà-kìrring simpang* 'to
keep to dry'), which in turn is a complement of the verbal predicate *siibar* 'spread' in
the pluperfect, marked by the conjunctive participle form (CP) *asà-* and the past tense
form of the existential, *su-aada*.

(82)   *Raaja hathu thiikar=ka guula [[asà-siibar [**mà-kìrring simpang**]*purpclause*]*cp**
       king   INDEF man=LOC   sugar CP-spread   INF-dry   keep
       *su-aada]*verbalpredicate*·*
       PAST-exist.
       'The King had sprinkled sugar on a mat and had left it to dry.'

5.1.5.6   *puukul* 'hit'

Puukul has as its literal meaning 'hit'. It is used to indicate that an action was very
violent. The only instances are related to throwing an item to (83) and fro (84).

(83)   *[Incayang=pe kàpaala=ka anà-aada] thoppi=dering moonyeth pada=nang*
       3s.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
       *su-buvang **puukul**.*
       PAST-throw hit
       'He took the hat from his head and violently threw it at the monkeys.'
       (K070000wrt01,K081104eli06)

(84)   *Ithu=kapang ithu moonyeth pada=le [anà-maayeng duuduk thoppi]*
       DIST=when   DIST monkey   PL=ADDIT PAST-play   sit   hat
       *pada=dering inni oorang=nang su-**bale-king**   **puukul**.*
       PL=ABL   PROX man=DAT   PAST-return-CAUS hit
       'Then the monkeys threw back the hats with which they had been playing.'
       (K070000wrt01)

5.1.5.7 *duuduk* 'sit'

This verb has been discussed extensively above ⊡→5.1.2.2, p. 168. When used as a vector verb, it indicates progressive aspect ⊙→15.5.4.2, p. 636. This is the case in (85), where the playing of the monkeys is ongoing (Smith et al. 2004, Ansaldo 2009b). Abbi & Gopalakrishnan (1991:175) note that the use of SIT as a vector verb is typical of Indo-Aryan and generally not found in Dravidian.

(85) *[Dee   arà-**sbuuni**   **duuduk** cave] asaraathang sini=ka   asàduuduk*
     3S.IMPOLITE NON.PAST-hide sit   cave] COPULA   here=LOC from
     hathu three miles cara jaau=ka.
     INDEF three miles way far=LOC
     'The cave where he remained hidden is about three miles from here.'
     (K051206nar02)

Another example is (86), where the continuous nature of the job search is highlighted by *duuduk*.

(86) *Incayang suda [aapa=ke   hathu pukurjan] mà-girja arà-**diyath**   **duuduk**.*
     3S.POLITE thus what=SIMIL INDEF word   INF-make NON.PAST-try sit
     'Now he is looking forward to do some kind or other of work.'
     (K051222nar08,K081104eli06)

Duuduk can be used in different temporal references, such as past in (87) of future in (88).

(87) *Loram pada anà-dhaathang vakthu=dika se spaathu **anà**-gijja   **duuduk**.*
     2PL   PL   PAST-come   time=LOC   1s shoe   PAST-make stay
     'When you came, I was making shoes.' (K081104eli06)

(88) *Loram pada arà-dhaathang vakthu=dika se spaathu **anthi**-gijja **duuduk**.*
     2PL   PL   NON.PAST-come time=LOC   1s shoe   IRR-make stay
     'When you will come, I will be making shoes.' (K081104eli06)

5.1.5.8 *kìnna* 'strike'

Kìnna is a vector verb used to highlight the surprising and adversative nature of the event for the undergoer. When used as a full verb, its meaning is 'to strike'. It differs from the other vector verbs by preceding the main verb, where other vector verbs follow.

First, let us discuss the use of *kìnna* as a full verb. In example (89), *kìnna* is the only verb, hence a full verb. This is also confirmed by the semantics, which involve a

clear instance of striking.

(89)   *Boola* wicket=*ka   su-kìnna.*
       ball   wicket=LOC PAST-strike
       'The ball struck the wicket.' (K081104eli06)

A naturalistic example of *kìnna* being used as a full verb is (90).

(90)   *Suda itthu=dering de=dang   thama-***kìnna***   kiyang*
       thus   DIST=ABL   3S.IMPOLITE=DAT NEG.NONPAST=strike EVID.
       'So, because of it [the talisman], (his persecutors' bullets) would not hit him,
       it seems.' (K051206nar02,K081105eli02)

Let us now turn to clauses where *kìnna* is accompanied by another verb.

(91)   *Sdiikith thaaun=nang duppang see ini   Aajuth=nang* **su-kìnna   daapath**.
       few   year=DAT   before   1s PROX dwarf=DAT   PAST-kìnna get
       'Some years before, I was captured by this dwarf.' (K070000wrt04)

In this example, the verb *daapath*, which normally means 'to get', is modified by
kìnna to highlight the surprising and adversative nature of this act of getting, which is
then best translated into English as 'to capture'. If the event of getting is not surprising
and adversative, as in (92), *kìnna* is not used.[11]   Importantly, the argument structure
remains the same (dative for the recipient, zero for the undergoer). It is therefore not
the case that *kìnna* changes the syntactic status of arguments, as a passive construction
would do. Rather, it contributes a semantic shade of meaning, very much in the way
other vector verbs do.

(92)   *Se=dang bannyak panthas atthu pon   su-daapath.*
       1s.DAT   much   beautiful INDEF bride PAST-get.
       'I got a very nice bride (from the matchmaker).' (K081103eli02))

This 'unfortunate' reading of *kìnna* is also found in the following example. The
dwarf is grasped and carried away by a big bird.

---

[11]This only holds if the selection of the speakers as a bridegroom does not come as a bad surprise for the
bride.

(93) *Derang su-kuthumung [ithu buurung=pe kuuku=ka Aajuth **asà-kìnna***
     3PL   PAST-see   DIST bird=POSS   claw=LOC dwarf CP-strike
     ***sìrrath** arà-duuduk].*
     stuck   SIMULT-stay.
     'They saw that the dwarf sat stuck in the claws of the bird.' (K070000wrt04)

This contrasts with (94), where the act of getting stuck is (although being unpleasant) not particularly adversative for the undergoer, the shoe.

(94) *Se=ppe spaathu beecek=ka asà-sìrrath.*
     1s=POSS shoe   mud=LOC CP-stuck.
     'My shoe got stuck in the mud.' (K081105eli02)

A further argument against an analysis as syntactic passive comes from the fact that *kìnna* can be used with intransitive verbs like *picca* 'break down'[12] (95). The transitive verb *picca-king* involves a causitivizer, which is absent here. Generally, intransitives do not lend themselves to passivization in the languages of the world, and if they do, we get an impersonal reading like German *Hier wird getanzt* 'Here is being danced=People dance here' (cf. Foley & Van Valin 1985:325). This is not the case in (95), where the argument *seppe ruuma* 'my house' is still present.

(95) *Buumi ginthar=subbath se=ppe ruuma su-kìnna   picca.*
     earth   shake=because   1s=POSS house PAST-strike break.down
     'My house broke down because of the earth quake.' (K081105eli02)

The second component of *kìnna* is surprise. An expected adversative result cannot be marked by *kìnna*. The following three statements about US presidential elections show this. In 2000, Al Gore lost to George W. Bush by a small margin, after having already declared his victory. In this context, *kìnna* is the perfect choice, since the event was surprising and adversative. Note that *kaala* 'lose' is intransitive in (96), proving again that *kìnna* cannot be analyzed as a syntactic passive.

(96) *Al Gore su-kìnna   kaala.*
     Al Gore PAST-strike lose.
     'Al Gore lost (with a small margin).' (K081105eli02)

If the event is not adversative, such as victory, *kìnna* cannot be used.

---

[12]In this case actually a converted adjective. In stative reading, it means 'be.broken', in dynamic reading, 'to break down'.

(97)   *George W Bush su-kìnna   bìnnang.*
        George   W. Bush PAST-strike win.
        (no meaning, unless the victory turned out to be adversative to
        GWB)(K081105eli02)

If the event is adversative, but not surprising, *kìnna* cannot be used either. This is
the case for the result for the 2008 presidential elections in the USA, where the defeat
of John McCain occurred as expected.

(98)   *John McCain su-kìnna   kaala.*
        John   McCain PAST-strike lose.
        (John McCain lost surprisingly)(K081105eli02)

This element of surprise can also be found in the following three examples, where
a person, probably a masochist, gives money to be beaten. If *kìnna* is used, the beater
cannot be the receiver, because that would not come as a surprise. Rather, the identity
of the beater must come as a surprise to the client, i.e. a third person must hit (100).
If, through the use of the emphatic clitic *=jo*, the interpretation of identity of receiver
and beater is forced, the sentence becomes ungrammatical (101).

(99)   *Se ini   oorang=nang duvith su-kaasi incayang se=dang mà-puukul=nang.*
        1s PROX man=DAT   money PAST-give 3s.POLITE 1s.DAT   INF-hit=DAT
        'I gave money to this man so that he would hit me.' (K081105eli02)

(100)   *Se ini   oorang=nang duvith su-kaasi incayang se=dang mà-**kìnna***
        1s PROX man=DAT   money PAST-give 3s.POLITE 1s.DAT   INF-strike
        puukul=nang.
        hit=DAT
        'I gave money to this man so that he would have me hit by a third person by
        surprise.' (K081105eli02)

(101)   **se ini   oorang=nang duvith su-kaasi incayang=**jo**   se=dang mà-**kìnna***
        1s PROX man=DAT   money PAST-give 3s.POLITE=EMPH 1s.DAT   INF-strike
        puukul=nang.
        hit=DAT
        '(I gave money to this man so that he himself would hit me.)' (K081105eli02)

Kìnna does not combine well with the expression of voluntary actors, or actors in
control. It is thus difficult to combine *kìnna* with wilful actions, such as assassinations.
(102) shows the normal way of expressing an assassination. If *kìnna* is to be used, as

in (103), the actor may not be overtly specified, but must be put in a periphrasis.

(102) *Oswald Kennedy=yang su-buunung.*
       Oswald Kennedy=ACC PAST-kill.
       'Oswald assassinated Kennedy.' (K081105eli02)

(103) *Oswald=pe   thaangang=dering Kennedy(=yang) se-kìnna   buunung.*
       Oswald=POSS hand=ABL   Kennedy=ACC   PAST-strike kill
       'Kennedy was killed by the hands of Oswald.' (K081105eli02)

If no agent is specified for verbs modified by *kìnna*, a non-volitional actor, or a non-human actor is automatically assumed. In example (104) and (105), we are dealing with an execution. If *kìnna* is used for this state-of-affairs, it automatically implies that the entity performing the execution was not volitional. Since executions cannot be done involuntarily, the lack of volition expressed by *kìnna* implies that some entity incapable of volition must intervene between the instigator of the execution and the end, such as a machine, the guillotine.

(104) *Saddam Hussein=yang su-buunung.*
       Saddam Hussein=ACC PAST-kill
       'Sassam Hussein was executed.' (K081105eli02)

(105) *Saddam Hussein=yang su-kìnna   buunung.*
       Saddam Hussein=ACC PAST-strike kill
       'Saddam Hussein was guillotined (or killed by another device, but not directly by a human).' (K081105eli02)

This lack of volition expressed by *kìnna* can also be seen in the following two sentences, where an explicit statement of volition *kamauvan* 'desire' cannot combine with *kìnna*.

(106) *Oswald=nang Kennedy=yang mà-buunung kamauvan su-aada.*
       Oswald=DAT Kennedy=ACC INF-kill   desire   PAST-exist.
       'Oswald wanted to kill Kennedy.' (K081105eli02)

(107) **Oswald=nang Kennedy=yang mà-kìnna buunung kamauvan su-aada.*
       Oswald=DAT   Kennedy=ACC INF-strike kill   desire   PAST-exist
       '(Oswald wanted to kill Kennedy involuntarily).' (K081105eli02)

Kìnna is of Malayic descent; today we still find the cognate *kena* in Std. Malay (Chung 2005) and Malay contact varieties (Ansaldo 2009b). *Kìnna* is probably also

related to the verbal derivational prefix *kanà-* ⊡→6.2.5.3, p. 312. This needs further research. The main difference between the two seems to be that *kanà-* does not have the adversative component; it only has the involuntary/surprise component. It might be possible to present a unified analysis of *kìnna* and *kanà-*. What speaks against this is that they can cooccur in the same predication (108), which shows that they are not in complementary distribution.

(108)   *Itthu haari=ka=jo   aanak pompang duuva=nang hathu duuri*
        DIST day=LOC=EMPH child female   two=DAT   INDEF thorn
        *pohong=nang jeeňǧgoth=yang anà-**kànà**-daapath **kìnna** hathu Aajuth*
        tree=DAT   beard=ACC   PAST-INVOL-get   strike INDEF dwarf
        hatthu=yang su-kuthumung.
        INDEF=ACC   PAST-see.
        'On the same day, the two girls saw a dwarf whose beard had got stuck in a thorn tree.' (K070000wrt04)

5.1.6   Special constructions involving verbal predicates

Verbs are used in a number of periphrastic constructions which yield TAM meanings. The perfect tenses are formed with a verb followed by the existential *aada* ⊡→5.1.6.1, p. 192. Obligational constructions can be formed by a verb in the infinitive followed by *aada* ⊡→5.1.6.2, p. 194 or *jaadi* 'become' ⊡→5.1.6.3, p. 195.

5.1.6.1   The perfect tenses

The perfect is formed by a verb followed by the existential *aada* in the affirmative and its negation *thraa* in the negative (cf. Slomanson (2006:143), Slomanson (2008a), Smith & Paauw (2006:169f), and Ansaldo (2009b)). The verb can either be in the stem form (110), or the conjunctive participle prefix *asà-* can be prefixed (111) (Slomanson (2006:143),Slomanson (2008a)). Normally, the existential is present, but this is not always the case. The perfect is used to refer to events located in the past. No particular aspectual reading is implied ⊙→15.5.2.1, p. 616.

(109) $\left\{ \begin{array}{c} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\}$ *(asà-)*V $\left\{ \begin{array}{c} \textit{(aada)} \\ \textit{thraa} \end{array} \right\}$

(110)   *Uumur=nang kuurang, sdiikith oorang pada ∅-kaving=le **aada**.*
        age=DAT   few   few   man   PL   marry=ADDIT exist
        'Below that age, there are few people who are married, too.' (K061122nar01)

(111) *Itthu=le kitham=pe mlaayu pada=jo itthu thumpath samma **asà**-kaasi **aada***.
but   1PL=POSS Malay PL=EMPH DIST place   all   CP-give exist
'Therefore it was our Malays who have all given away those lands.'
(B060115cvs04)

This construction is very frequent. Both Sinhala and Tamil have an analogous construction consisting of the conjunctive participle and an existential, but lack the construction with the bare verb instead of the conjunctive participle. The use of cognates of *aada* is also widespread in other Malay varieties (Adelaar & Prentice 1996, Bakker 2006).

The perfect tense is negated by replacing the existential with *thraa*, as in (112).

(112) *Hatthu dhaatha   **asà**-kaaving **thraa***.
one   elder.sister CP-marry   NEG
'One elder sister has not married.' (K061019prs01)

The pluperfect consists of the same element as the perfect construction, but with the existential in the past tense, instead of the present tense. Example (114) illustrates this for *su-*. No instance of *anà-aada* in this construction has been found.

(113) $\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\}$ *(asà-)*V $\left\{ \begin{array}{l} su- \\ anà-? \end{array} \right\}$ *aada*

(114) Letter=*ka thaaro **s-aada***.
letter=LOC put   PAST-exist
'(I) had put it on the letter.' (K060116nar10)

Note that the use of the past tense marker is not obligatory, the normal perfect tense can also be used where other languages, like English, might require the use of the pluperfect. An example for this is given in (115), where the action of giving their word is clearly anterior to the point in time the narrative is dealing with, but still the perfect construction is used.

(115) a. *Derang pada kathahan thama-thuukar*.
3PL   PL   word   NEG.NONPAST-change
'They would not change their word.'

b. *Karang* British *government=nang derang kathahan **asà**-kaasi **aada***.
now   British Government=DAT 3PL   word   CP-give exist
'Now, they had given their word to the British government.'
(K051213nar06)

The combination of the perfect construction with the irrealis markers *anthi* (affirmative) (117) ⊡→6.2.2.2, p. 290and *thama* (negative) (118)⊡→6.2.2.5, p. 298 yields the anterior irrealis construction.

(116) $\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\}$ (*asà-*)V $\left\{ \begin{array}{l} \textit{anthi–} \\ \textit{thama–} \end{array} \right\}$ *aada*

(117)   *Bìssar aanak **asà-dhaathang anthi-aada** ruuma=nang.*
         big   child CP-come   IRR=exist   house=DAT
         'My big child will have come home.' (B060115cvs08)

(118)   *Inni bedahaan **jaadi**=jo   **thuma=aada**.*
         PROX difference become=EMPH NEG.NONPAST=exist
         'These differences would not have arisen.' (K060116nar03)

The anterior irrealis is not found very frequently. Tamil has an analogous construction (Lehmann 1989:207), while in Sinhala, such a construction cannot be found because of the lack of a verbal irrealis marker.

5.1.6.2   Dative+infinitive+V+*aada*

Besides the formation of the perfect tense, the existential is also used in an obligational construction ⊙→15.7.2.2, p. 653, where it combines with a verb in the infinite. The entity for which the obligation holds is marked with the dative. The Kleene star * marks that an element can be present 0, 1, or more times.

(119) $\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\}$ * NP=*nang mà-*V (TAM)-*aada*

Example (120) illustrates this construction with the dative, the infinitive and the existential highlighted.

(120)   *Kithang=pe sudari pada=**nang** makanan **mà**-maasak **aada**.*
         1s=POSS   sister PL=DAT   food   INF-cook   exist
         'Our sisters all have to cook food.' (B060115cvs01)

The person under the obligation need not be expressed overtly if he or she in inferable from discourse. This is the case for the speaker in (121).

(121) *Lai aapa lai aapa ∅ **mà**-biilang **aada**.*
 other what other what INF-say exist.
 'What else (do I) have to say?' (K060108nar02)

Occasionally, one of the elements can be dislocated to a position to the right of *aada* . This is the case for the complement clause in (122). The underscore indicates the position where this clause would normally be expected.

(122) *Se=dang ___ aada [ini army pada=yang mà-salba-kang=nang]*CLS.
 1s.DAT exist PROX army PL=ACC INF-save-CAUS=DAT
 'I had to save these soldiers' (K051213nar01)

This construction with *aada* has a component of duty to it imposed by law or moral values, which distinguished it from general obligation by *masthi-* and from the following construction with *jaadi* which do not require (*masthi*) or allow (*jaadi*) the duty component.

5.1.6.3 Dative+infinitive+V+*jaadi*

This periphrastic construction can also be used to express obligation ⊙→15.7.2.2, p. 653. It is very similar to the construction with *aada*, but *aada* is replaced by *jaadi*. Its structure is as follows:

(123) $\left\{\begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array}\right\}$ * NP=*nang mà*-V TAM-*jaadi*

The three important elements of this construction, the dative, the infinitive and *jaadi* are highlighted in the following example.

(124) *Se=ppe profession=subbath se=**dang** siini **mà-pii** su-**jaadi**.*
 1s=POSS profession=because 1s=DAT here INF-go PAST-become
 'I had to come here because of my profession.' (G051222nar01)

This use of *jaadi* 'become' must not be confounded with the literal use, which can also occur with a dative. This is shown in (125).

(125) *Se=dang buthul seksa pada anà-jaadi.*
 1s.DAT correct problem PL PAST-become
 'I got a lot of problems.' (K051213nar01)

The person experiencing the obligation need not be mentioned if it can be established from discourse. This is the case for *kithang* 'we' in (126).

(126)   *Blaakang, ∅ cinggala **mà**-blaajar su-**jaadi**.*
       after   Sinhala INF-learn   PAST-become.
       'Then (we) had to learn Sinhala.' (K051222nar06)

Besides the past tenses illustrated above, this construction can also be used with irrealis inflection *athi-* on *jaadi*.

(127)   *[Saapa=pe=ke thaangang di   baava]=ka=jo   pukurjan **mà**-gijja*
       who=POSS=SIMIL hand   LINK under=LOC=EMPH work   INF-make
       *athi-**jaadi**.*
       IRR-become
       'You will always have to work under someone's command'
       (K051206nar07,K081104eli06)

The following two examples also show this construction and are given as additional illustration.

(128)   *Suda derang**=nang** hathyang muusing=sangke **mà**-duuduk su-**jaadi**.*
       thus 3PL=DAT   other   time=until   INF-stay   PAST-become
       'So they had to wait until another time' (K051220nar01)

(129)   *Ini   thaaun=ka derang sama oorang=**nang** England=nang **mà**-pii*
       PROX year=LOC 3PL   all   man=DAT   England=DAT INF-go
       *su-**jaadi**.*
       PAST-become.
       'All men had to go to England this year.' (K051222nar06)

The construction with *jaadi* conveys some kind of circumstantial obligation. Due to circumstances, a certain obligation arises. This is different from the construction with *aada* discussed above, where some kind of moral authority is responsible for the obligation, and it is also different from general obligation by *masthi*, which does not carry any specification as to whether the obligation arose because of circumstances or not.

5.2   Nouns

SLM nouns can be divided into common nouns ⊡→5.2.1, p. 197, relator nouns indicating figure-ground relations ⊡→5.2.2, p. 199, and proper nouns ⊡→5.2.3, p. 209.

5.2.1 Common nouns

Sri Lanka Malay has a large open class of nouns, which most often denote objects like *ruuma* 'house' or persons like *aanak* 'child', but also some abstract concepts like *vatthu* 'time'. Many abstract concepts are derived from another word class by means of the nominalizer *-an*, like *makanan* 'food' from *maakang* 'eat'. All nouns share the following characteristics:

1. They cannot take the following affixes: present tense *arà-*, past tense *anà-*, past tense *su-*, anterior tense *asà-*, irrealis *anthi-*, infinitive *mà-*.

2. They are always negated with *bukang*, never with *thàrà-* or *thama-*.

3. They must combine with *=nang* in order to modify a predication

4. Many nouns can also combine with *pada* 'PLURAL' and the deictics *ini* 'proximal' and *itthu* 'distal'.

These criteria will now be discussed in turn.

Example (130) shows the impossibility of nouns to combine with the present tense marker *arà-*. Combination with other TAM-markers is not possible, either.

(130)   *se arà-oorang.*
        1s   NON.PAST-man
        (word salad, no possible meaning)

Nouns in ascriptive position are negated by *bukang*.[13]

(131)   *Deram Islam oorang **bukang**.*
        3PL   Islam man   NEG.NONV
        'They were not Muslims.' (K051213nar03)

(132)   *See innam   blaakang, hatthu aanak **bukang**.*
        1s   PROX.DAT after   INDEF child NEG.NONV
        'I will never be a child again.' (K081106eli01)

The verbal negators *thàrà-* or *thama-* are not possible.

(133)   **Sindbad   thàrà-Islam   oorang.*
        Sindbad the sailor NEG.PAST-Moor man
        (Sindbad was no Muslim) (K081103eli02)

---

[13]Another negator often found in the vicinity of nouns is *thraa*, which negates the existence or the availability of the item denoted by the noun in question. But the scope of this negator is the existential predicate, not the nominal argument thereof, whereas *bukang* negates the nominal predicate.

When nouns are used as ascriptive predicates on arguments in the singular, they must obligatorily carry the indefiniteness marker *hatthu* (134). This differentiates them from adjectives (135), which can be used without *hatthu*.

(134)   *Farook *(atthu) maaling.*
        Farook INDEF   thief.
        'Farook is a thief.' (K081103eli02)

(135)   *Farook kiccil.*
        Farook small.
        'Farook is small.' (K081103eli02)

Nouns can be used to modify predications ('adverbial function'). For this, they have to take the dative marker =*nang* (136). This distinguishes them from adjectives, where the use of =*nang* is optional (137).

(136)   *Incayang svaara$_N$*(=nang) arà-oomong.*
        3s.POLITE noise=DAT   NON.PAST-speak.
        'He speaks loud.' (K081103eli02)

(137)   *Incayang pullang$_{ADJ}$(=nang) arà-oomong.*
        3s.POLITE slow=DAT   NON.PAST-speak.
        'He speaks softly.' (K081103eli02)

Many, but not all nouns can combine with the plural marker *pada* and the deictics *ini* and *itthu*. This distinguishes nouns from verbs, which cannot combine with these markers. It does not distinguish them from adjectives, some of which have the same possibilities. Example (138) shows the use of the noun *baarang* 'goods' combined with the proximal deictic *ini*. Example (139) shows the combination of the nouns *vatthu* 'time' and *nigiri* 'country' with the plural markers and the distal deictic *itthu*.

(138)   ***Ini baarang** pada=yang asà-baapi laayeng nigiri=ka   anà-juval.*
        PROX goods   PL=ACC   CP=bring other   country=LOC PAST-sell
        'He brought the goods and sold them in other countries.' (K060103nar01)

(139)   ***Itthu vatthu**=ka **itthu nigiri pada**=ka arà-duuduk.*
        DIST time=LOC DIST land PL=LOC   NON.PAST-stay
        'At that time, (they) lived in those countries.' (N060113nar01)

Nouns are also often found with postpositions marking case as in (140), but this does not distinguish them from the other word classes because in SLM, case markers

relator noun	referential meaning	relational meaning
blaakang	back	behind, after
duppang	front	in front of, before, ago
baa(wa)	bottom	below, under
atthas	top	on, above, about
thìnnga	middle	in
daalang	inside	in
luvar	outside	out
vatthu	time	during
dìkkath	vicinity	at
mosthor	manner	like

Table 5.2: Relator nouns

can combine with members from any word class ▢→6.4.4, p. 327.

(140) *Itthu baathu_N=**yang** incayang Seelong_N=**dering** laayeng nigiri_N=**nang***
       DIST stone=ACC   3s.POLITE Ceylon=ABL   other   country=DAT
       asà-baapi.
       CP-bring
       'These stones, he brought them from Ceylon to other countries.'
       (K060103nar01)

5.2.2 Relator nouns

A subset of nouns can be used as relator nouns (Starosta 1985, DeLancey 1997) to indicate (mostly spatial) relationship between two referents ⊙→15.4.3, p. 609 (Adelaar 1991:25). This is a typical South Asian feature (Masica 1976:23). These relator nouns can be used in postpositional phrases to indicate the relation between figure and ground more precisely than with a simple locative =*ka*.

Relator nouns and the NP denoting the ground are either simply juxtaposed as in (141)(142) or they are joined with the possessive =*pe* on the ground NP (143)-(145)(Saldin 2001:62).[14]

(141) *Ini   **pohong** ∅ **atthas**=ka moonyeth hathu kavanan su-aada.*
       PROX tree   top=LOC monkey INDEF group   PAST-exist
       'On top of this tree was a group of monkeys.' (K070000wrt01)

---

[14]It is possible for different relator nouns to require/permit/ban a possessive/genitive, as shown by De-Lancey (1997) for Tibetan and Tamang.

(142)   *Soore=ka,   Snow-white=le   Rose-red=le   derang=pe umma=samma*
        Evening=LOC Snow.white=ADDIT Rose.Red=ADDIT 3PL=POSS   mother=COMIT

        appi Ø ***dìkkath***=*ka   arà-duuduk ambel.*
        fire   vicinity=LOC NON.PAST-sit take
        'In the evening, Snow White and Rose Red used to sit down next to the fire
        with their mother.' (K070000wrt04)

(143)   *Andare hathu **pohong=pe baava**=ka   kapang-duuduk.*
        Andare INDEF tree=POSS   bottom=LOC when-sit.
        'When Andare sat down below a tree.' ((K070000wrt03))

(144)   *Bìssar hathu buurung **derang=pe atthas**=dering su-thìrbang.*
        big   INDEF bird   3PL=POSS top=ABL   PAST-fly
        'A big bird flew over them.' (K070000wrt04)

(145)   ***Derang pada=pe atthas*** *kithang=nang bannyak mà-biilang thàràboole.*
        3PL   PL=POSS top   1PL=DAT   much   INF-say   cannot
        'We cannot tell you much about them.' (K051206nar12)

Relator nouns are not simply postpositions, because they can occur without a host
as in (146), where *daalang* 'inside' does not refer to the interior of the immediately
preceding constituent *vavvaal pada* 'bats',[15] but rather to the interior of a cave men-
tioned earlier in discourse.

(146)   *Kiccil vavvaal pada* Ø ***daalang=ka*** *arà-duuduk.*
        small bat   PL   inside=LOC prog.exist.ANIM
        'There are small bats inside.' (K051206nar02)

Another example of a relator noun occurring without a host is (147), where the
ground to which *baava* 'bottom' relates is the house, which is present in discourse,
but not mentioned in the sentence.

(147)   *Goo   Ø baava=ka=jo   arà-duuduk, Ø baava=ka.*
        1s.FAMILIAR bottom=LOC=EMPH NON.PAST-live   bottom=LOC
        'I live downstairs.' (B060115cvs13)

Relator nouns can also attach to pronouns (148) and declausal NPs (149). This is
more often found with temporal and causal relator nouns, and less often with local re-

---

[15]This word is a loanword from Tamil, which explains its uncommon word structure with the long vowel
in the last syllable.

lator nouns. The reason for this is of course that clauses typically denote non-concrete entities, which cannot be located in space.

(148) ***Se=ppe blaakang*** *arà-raathang* *[. . .]* *se=ppe* *aanak klaaki pada.*
1s=POSS after   NON.PAST-come [. . .] 1s=POSS child boy   PL
'After me are coming my sons.' (K060108nar02)

(149) *[[Mlaayu pada anà-dhaathang]*CLS*]*NP*=pe* ***atthas*** *se=dang hatthu=le*
Malay   PL   PAST-come=POSS   about 1s=DAT INDEF=ADDIT
mà-biilang thàràboole.
INF-say   cannot
'I cannot tell you anything about the coming of the Malays.' (K081105eli02)

In the following, I will discuss some relator nouns in more detail. I will give occurrences with ∅-marking, occurrence with =*pe*, occurrence with =*nang*, relation to space and relation to time if examples for these cases could be found. I give the referential/free form of a relator noun as well, and the use as an adverb, if this exists.

5.2.2.1 *blaakang* 'back'

This relator noun is used in spatial context with =*pe* (150) and in temporal contexts with =*nang* (151). Additionally, it is also possible to use *blaakang* without a postposition. (152).

(150) *Kandi=ka* *Malay mosque*space***=pe blaakang***=ka incayang=pe   zihaarath*
Kandi=LOC Malay mosque=POSS   back=LOC   3s.POLITE=POSS shrine
aada.
exist.
'In Kandy behind the Malay Mosque, there is his shrine.'

(151) *Dr Draaman duuva thaaun*time*=**nang blaakang** incayang su-mniiňǧgal.*
Dr Draaman two   year=DAT   after   3s.POLITE PAST-die
'After two years, Dr Draman died.' (K051213nar08)

(152) *Itthu ∅ **blaakang** lai   [se=dang kalu] bannyak itthu=pe   atthas*
DIST   after   other 1s.DAT   if   much   DIST=POSS about
thàràthaau.
NEG-know
'After that, as for me, I cannot tell you much.' (K051205nar02)

When used referentially, *blaakang* means 'back'; for the human anatomical part,

this has to be complemented by *thìnnga* 'middle' (153).

(153) *Se=ppe* **thìnnga blaakang** *arà-saakith.*
       1s=POSS middle   back   NON.PAST-hurt
       'My back hurts.' (K081103eli02)

Example (154) shows the use of *blaakang* as an adverb to indicate the spatial orientation of *diyath* 'look'.

(154) **Blaakang**=*yang diyath!*
       back=ACC   look
       'Look behind!' (K081103eli02)

Another adverbial use of *blaakang* is to indicate sequence of events, like English *then* or *after that*. This use is shown in (155).

(155) a. *[Thàrà-dhaathang oorang pada]=nang nya-force-kang kiyang.*
           NEG.PAST-come   man   PL=DAT   PAST-force-CAUS EVID
           'They apparently forced the people who did not come.'
       b. **Blaakang**_{adv} *thàrà-pii   kiyang.*
           after   NEG.PAST-go EVID
           '(But) after (that) (they) apparently (still would) not go.' (K051206nar07)

5.2.2.2   *duppang* **'front'**

This relator noun can indicate spatial location in front of the ground. In this case, it is construed with =*pe* (156).

(156) *Ruuma=pe duppang pohong aada*
       house=POSS front   tree   exist
       'There are trees in front of the house.' (K081103eli02)

As for temporal relations, two constructions involving *duppang* have to be distinguished. When used with =*nang*, *duppang* means 'before, ago, earlier' (157) ⊙→15.5.3.1, p. 625. When construed with =*pe*, it means 'future' (158). In the former case, *duppang* thus refers to a past event, while in the latter case, it refers to the future.

(157) *Kithang=**nang duppang** lai   duuva bàrgaada asà-dhaathang aada.*
       1PL=DAT   before other two   family   CP-come   exist
       'Before us, there were two other families.' (K060108nar02)

(158) *Kithang=**pe duppang** muusing(=yang) arà-dhaathang.*
1PL=POSS future time(=ACC)   NON.PAST-come.
'Our future is approaching.' (N061031nar01, K081103eli02)

When used referentially, *duppang* is translated as 'front' (159).

(159) *Laile ithu ruuma pada aada, **duppang**=ka sajja sdiikith laayeng-kang*
still DIST house PL exist front=LOC only few different=CAUS
aada.
exist.
'Those houses still remain, only the front portion has been renovated.'
(K051206nar02)

(160) *Sithuka aada, duppang$_{adv}$=ka aada.*
DIST=LOC exist before=LOC exist
'There are some there, there are some in front.' (K051206nar02)

5.2.2.3 *atthas* 'top'

This relator noun has two meanings. The first, literal, is 'top', the second one, 'about, on' ⊙→15.1.2.3, p. 568. *Atthas* is very often found without =*pe* (161), but the occurrence of =*pe* is also possible (162).

(161) *Ini   **pohong** ∅ **atthas**=ka moonyeth hathu kavanan su-aada.*
PROX tree   top=LOC monkey INDEF group PAST-exist
'On top of this tree was a group of monkeys.' (K070000wrt01)

(162) *Ruuma=**pe atthas**=ka=jo derang asà-duuduk aada, laile itthu ruuma*
house=POSS top=LOC=EMPH 3PL   CP-exist.ANIM exist still DIST house
pada aada
PL   exist
'Above the house some others have stayed, the houses still remain.'

The relator noun *atthas* can in turn host lative information, perlative in (163).

(163) *Bìssar hathu buurung derang=pe **atthas**=**dering** su-thìrbang.*
big   INDEF bird 3PL=POSS top=ABL   PAST-fly
'A big bird flew over them.' (K070000wrt04)

Used referentially, *atthas* means 'top' or 'upstairs' (164).

(164)   ***Atthas**=ka cumma baarang pada=jo.*
        top=LOC   idle   goods   PL=EMPH
        'Upstairs, there is only rubbish.' (B060115cvs13)

The second, figurative, meaning of *atthas* is 'about', as given in (165) for a nomi-
nal host and in (166) for a clausal host.

(165)   *Kithang Islam=**pe**   **atthas** arà-oomong   ambel.*
        1PL   Islam=POSS about NON.PAST-speak take.
        'We talk about Islam.' (K061026prs01)

(166)   *[[Mlaayu pada anà-dhaathang]=_{CLS}]_{NP}=pe atthas se=dang hatthu=le*
        Malay   PL   PAST-come=POSS   about 1S=DAT   INDEF=ADDIT
        mà-biilang thàràboole.
        INF-say   cannot
        'I cannot tell you anything about the coming of the Malays.' (K081105eli02)

5.2.2.4 *baa(va)* **'bottom'**

This relator noun indicates position below/under another entity. It must not be con-
founded with the verb *baa* 'bring'. Like the other relator nouns, *baava* can be con-
strued with =*pe* (167) or without (168).

(167)   *Seelong su-aada   soojor=**pe**   **baava**=ka.*
        Ceylon PAST-exist European=POSS under=LOC
        'Sri Lanka was under European rule.' (K051222nar06)

(168)   *[Jaalang hathu piiňǧgir=ka anà-aada hathu pohong]_{space}* ∅ ***baava**=ka*
        street   INDEF border=LOC PAST-exist INDEF tree   bottom=LOC
        su-seeňder.
        PAST-rest
        '(He) rested under a tree which stood at the side of the street.' (K070000wrt01)

It can be used without a host to indicate location. This use can be seen either as
referential ('on the ground') (169) or adverbial ('below') (170), but the two uses are
difficult to tell apart.

(169) *Dee su-thiidor **baava**$_{adv}$=ka.*
       3s   PAST-sleep bottom=loc.
       'He slept downstairs.' (K051205nar05)

(170) ***Baava**=ka Kaasim katha hatthu* **family***.*
       below=LOC Kaasim QUOT INDEF family
       'Below, there is a family called "Kaasim".' (K051206nar06)

5.2.2.5 *daalang* '**inside**'

This relator noun refers to the inside of an entity. A use with =*pe* is given in (171), an example without =*pe* is given in (172).

(171) *Ruuma=**pe** **daalang**=ka gìllap*
       house=poss inside=loc   dark
       'Inside the house, it is dark.' (K081103eli02)

(172) *Aanak pompang duuva=le, derang=pe umma=le   Buruan=yang*
       child female two   3PL=POSS mother=ADDIT bear=ACC
       *ruuma=∅ **daalang**=nang su-panggel.*
       house   inside=DAT   PAST-call.
       'The two girls and their mother invited the bear to come inside the house.'
       (K070000wrt04)

   The referential use of *daalang* is given in (173).

(173) *Kiccil vavvaal pada ∅ **daalang=ka** arà-duuduk.*
       small bat   PL   inside=LOC prog.exist.ANIM
       'There are small bats inside' (K051206nar02)

5.2.2.6 *luvar* '**outside**'

This relator noun refers to the outside of an entity. It does not occur in the corpus with the relator function, but elicitation shows that it can be used with or without =*pe* (174).

(174) *Nigiri(=pe) luvar=ka   bannyak uuthan.*
       town=POSS outside=LOC much   jungle
       'Outside of the town is a lot of jungle.' (K081103eli02)

The referential use of *luvar* is given in (175).

(175) *Ruuma=pe luvar   panthas*
       house=poss outside beautiful
       'The outside of the house is beautiful.' (K081103eli02)

The adverbial use of *luvar* is given in (176) and (177).

(176) *Itthu vakthu=ka hathu bìssar beecek caaya Buruan mlaarath=ka*
       DIST time=LOC INDEF big   brown colour bear   difficulty=LOC
       *uuthang=dering **luvar=nang** su-dhaathang.*
       forest=ABL   outside=DAT PAST-come
       'Then, a big brown bear came out of the woods with difficulties.'
       (K070000wrt04)

(177) *Aavi **luvar=nang** kapang-dhaathang itthu bambu=yang giini*
       steam outside=DAT when-come   DIST bamboo=ACC like.this
       angkath=apa pullang arà-thoolak.
       lift=after   slow   NON.PAST-push
       'When the steam comes out, lift the bamboo like this and push it slowly.'
       (K061026rcp04)

Example (178) shows a rare instance of the adverbial use of a relator noun with a complement.

(178) *Duuva thiiga Kluum̌bu=deri, Kluum̌bu=dering luvar   duuva=le*
       ttwo   three Colombo=ABL, Colombo=ABL   outside two=ADDIT
       'Two or three (co-presidents) are from Colombo, two more are from outside
       Colombo.' (K051206nar13)

5.2.2.7 *dìkkath* 'vicinity'

This relator noun refers to the vicinity of an entity (Smith et al. 2004). The use with *=pe* is given in (179), the use without *=pe*, in (180).

(179) Kandy town=*ka,* Bank *of Ceylon=pe   dìkkath=ka   aada derang=pe*
       Kandy town=LOC Bank of Ceylon=POSS vicinity=LOC exist 3PL=POSS
       paasar.
       shop.
       'Their shop is in Kandy town, close to the Bank of Ceylon.' (K051220nar01)

(180) *Hatthu avuliya aada kitham=pe ruuma ∅ dìkkath.*
       INDEF saint exist 1PL=POSS house vicinity
       'There is a saint close to our house.' (K060108nar02)

The referential use of *dìkkath* is given in (181) and (182).

(181) *Kithang sama oorang dìkkath*_{adv}=*ka arà-duuduk.*
       1PL   all   man   vicinity=LOC NON.PAST-exist.ANIM
       'All of us live close together.' (B060115cvs08)

In (182), the emphatic clitic *=jo*, which intervenes between *see* '1s' and *dìkkath* shows that *see* is not the ground for *dìkkath* here. Rather, the ground is the patient, and the figure is the speaker. The speaker is in a constellation to the ground indicated by the relator noun, i.e. 'vicinity' in the case of *dìkkath*.

(182) *Mr. Yusuf thàràsìggar vakthu, see=jo   dìkkath*_{adv}=*ka asà-duuduk samma*
       Mr Yusuf sick   time 1s=EMPH vicinity=LOC CP-stay   all
       incayang=nang nursing *samma girja aada*
       3s.POLITE=DAT   nursing all   make exist
       'When Mr. Yusuf was sick, I stayed close to him and I cared for him.'
       (K060116nar07,K081104eli06)

5.2.2.8 *didaalam* **'period'**

This relator noun is quite rare and can only combine with temporal expressions, where it has a meaning of 'during' or 'within'. It is normally used without *=pe* as in (183)-(185), but one instance has been found where it is used with *=pe*, albeit with a temporal expression metonymically derived from a ruler, Queen Elizabeth (186).

(183) *Ini   vakthu **didaalam** se=dang Mrs Cunci   nya-biilang kithang baae*
       PROX time   during   1s.DAT   Mrs Chunchee PAST-say   1PL   good
       mlaayu baaru mlaayu hatthu marà-mulain.
       Malay new Malay INDEF ADHORT-start
       'Within this period, Mrs Chunchee had told me that we must start a new Malay
       one, a good Malay one (= association).' (K060116nar04,K081105eli02)

(184) *Duuva week **didaalang** cinggala mulbar reepoth.*
       two   week within   Sinhala Tamil problem.
       'Within two weeks, there was the Sinhala-Tamil problem.' (K051213nar01)

(185) *Hathu thaaun **didaalam** thiiga skali quran   nya-thamaam-king.*
       one   year   within   three time Qur'an PAST-complete-CAUS
       'I   completed   the   Qur'an   thrice   within   one   year.'
       (K051213nar02)(K081103eli02)

(186) a. *Independence anà-daapath=nang=apa*
          Independence PAST-get=DAT=after
          'After having obtained independence'

       b. *laile derang anaduuduk*   under the Commonwealth
          still 3PL   PAST-exist.ANIM under the Commonwealth
          'they still were under the Commonwealth'

       c. *Puthri* Queen Elizabeth=***pe   didaalang**=jo anà-duudu.*
          queen Queen Elizabeth=POSS during=EMPH PAST-exist-ANIM
          'during Queen Elizabth's reign.' (K051222nar06)

5.2.2.9   Spatial and temporal relator nouns

Some of the relator nouns can also be used with =*nang* on their complement, predom-
inantly when the item serving as ground is verbal (Smith et al. 2004).

  It appears that =*nang* is used with temporal contexts, whereas =*pe* is used in spatial
contexts. The following examples illustrate this.

(187) *Duuva ruuma=pe duppang.*
       two   house=POSS front
       'In front of the two houses.' (K081103eli02)

(188) *Duuva aari=nang duppang.*
     two   day=DAT   front
     'Two days before/ago.' (K081103eli02)

(189) *Duuva ruuma=pe blaakang.*
     two   house=POSS behind
     'Behind the two houses.' (K081103eli02)

(190) *Duuva aari=nang blaakang.*
     two   day=DAT   after
     'Two days later.' (K081103eli02)

The temporal readings indicated by *aari* 'day' in the examples above take the dative marking, whereas the spatial readings forced by *ruuma* 'house' trigger the possessive *=pe*.

In rapid speech however, it is frequent that *=nang* as well as *=pe* are dropped.

5.2.3 Proper nouns

Proper nouns are a special class of nouns in SLM. They are used to designate individual entities, but do not have semantic content. Proper nouns share many properties with common nouns, but there are some differences as well. Normally, proper nouns can only refer to exactly one entity, hence the use of the indefiniteness marker, the deictics and the plural marker is normally not found. These markers would give information about the referential status of the referent, but this is not necessary, since proper nouns designate only one entity. In rare cases is it possible to find a combination of a deictic with a proper noun as in (191) or (192).

(191) *See **ini** **Sri Lanka**=ka nya-blaajar.*
     1s   PROX Sri Lanka=LOC PAST-learn
     'I studied here in Sri Lanka.' (K061026prs01)

(192) *Thuan Shaaban=jo   anà-capture-king **inni** **Saradiyel**=yang.*
     Thuan Shaaban=EMPH PAST-capture-CAUS PROX Saradiyel=ACC
     'It was Thuan Shaaban who captured this Saradiyel.' (K051205nar02)

Proper nouns are normally not modified by adjectives either, but in some metonymical extensions as in (193) this is possible.

(193)   *Go=ppe*   *naama **Badulla buulath** thaau.*
        1s.FAMILIAR=POSS name   Badulla whole   know
        'Whole Badulla knows my name.' (B060115nar04)

They can take all the postpositions, though. Since proper nouns are very often human, the special constructions to use humans as LOCATION or GOAL must be taken in to account as shown in (194).

(194)   a.   *Tony=pe   dìkkath=nang mari.*
             Tony=POSS vicinity=DAT   come.IMP
             'Come close to Tony.'

        b.   **Tony=nang mari.*
             Tony=DAT   come.IMP (K081103eli02)

As for the frequency of proper names among the Malay population, first names are in the overwhelming majority of the cases of Islamic origin. The letters ⟨F⟩ and ⟨Z⟩ are very popular, and alliteration and assonance between siblings' names are very common (e.g. Fazlin, Falin, Fazeer, Faizel, Faiz; Zahira, Zarina, Farina; Imran, Izvan). Very often, the Sri Lanka Malays carry pet names which have nothing to do with their real name, e.g. Tony, Daisy or Joy. The use of this pet name is often so rampant that even very good friends do not know each other's official name. Another common pattern is the use of acronyms of the person's given names, a famous example being BDK Saldin [biɖiˈkeː].

There are few genuinely Malay place names (see Table 5.3). In most cases, the Malays simply use the Sinhala name for Sinhalese place names like *Singhapitiya*, which is clearly parseable as 'lion place' in Sinhala. There are few Tamil place names in the Upcountry, so the temptation to use Tamil place names is limited. A notable exception is Kandy itself, whose SLM name *Kandi* is taken over from Tamil *Kaɳɖi* (see Table 5.3). For places abroad, the English name is generally used, even if a Tamil or Sinhala word does exist, so [ʤəˈpæːn] 'Japan', not [ʤaˈpaːniya] (Sinhalese), [ʤappan] (Tamil) or [ʤepaŋ] (Std. Malay). Occasionally, Sinhala placenames also carry the Sinhala locative case in *-e,* e.g. *Badull-e* instead of *Badulla=ka*.

5.3   Adjectives

The third major word class in SLM are Adjectives. In distinction to nouns and verbs, it is difficult to give simple tests for adjectivehood because adjectives permit a lot but require little. They accept the whole of verbal morphology and most nominal morphology, but do not require it. In fact it is this permissiveness that singles out adjectives. In order to ascertain the status of an adjective, one has to test for the possibility to co-occur with a certain other item, and then for the possibility to occur

English	Sinhala	Tamil	SLM
Kandy	Nuvara	Kaṇḍi	Kandi
Colombo	Kolaṁba	Koḻumbu	Kluuṁbu
Negombo	Migamuva	Koḻumbu nir	Guuṁbu
Galle	Galla	Kalla	Gaali
Jaffna	Yapanaya	Yaḻppaṇam	Jaapna
Batticaloa	Maḍakalapuva	Maṭṭakkaḻappu	Batticaloa
Trinco(malee)	Tirikuuṇaamaḻaya	Tirukooṇamalai	Trinco

Table 5.3: Sri Lankan toponyms in English, Sinhala, Tamil and SLM

without said item. If both are possible, then it is probably an adjective, otherwise a noun or a verb, as the following list shows.

- They can be used as head of predicate without TAM morphology

- They can be used as head of predicate with TAM morphology

- They can take the superlative marker *anà-*

- They can be used as heads of predicates together with *atthu*

- They can be used as a pre- or postnominal modifier of a term without further modification

- They can modify predicates with =*nang*

- They can modify predicates without =*nang*

- They are negated by following *thraa* in all tenses in most idiolects. Occasionally also prefix *thàrà-* or following *bukang*.

These criteria will now be examined in turn. The adjective can be used as head of predicate without TAM-morphology, as given in (195). It is also possible to use TAM-morphology with an adjective, but then the predicate gets a dynamic meaning (196) instead of the static meaning in (195).[16]

(195) *Aanak thiiňǧgi.*
       child   tall
       'The child is tall.' (K081103eli02)

(196) *Aanak arà-thiiňǧgi.*
       child   NON.PAST-tall.
       'The child is growing tall.' (K081103eli02)

---

[16]Interestingly, the same is true in Mauritius Creole (Alleyne 2000:131).

Adjectives can combine with the superlative marker *anà-* (197). This marker is not very common. Furthermore, it is homophonous to the verbal past tense marker *anà-*, so that it is a less ideal test then what one could imagine.

(197) *Seelon=ka **anà**-bìssar pohong.*
       Ceylon=LOC SUPERL-big tree
       'The biggest tree in Sri Lanka.' (K081104eli06)

Adjectives can be used with or without the indefiniteness marker *hatthu*.

(198) incayang bìssar/iitham
       3s.POLITE big/black
       'He is big/black.' (K081103eli02)

(199) *Incayang hatthu bìssar.*
       3s.POLITE INDEF big
       'He is a big one/head.' (K081103eli02)

(200) *Incayang hatthu iitham*
       3s.POLITE INDEF black
       'He is a dark person.' (K081103eli02)

Adjectival lexemes can head referential phrases.

(201) *Iitham/bìssar su-dhaathang.*
       black/big   PAST-come.
       'The black one/big one/boss came.' (K081103eli02)

Adjectives can pre- or postmodify nouns, but postmodification is dispreferred for ad hoc formations (202). It is possible in lexicalized N+ADJ combination like *baavam puuthi* 'onion'+'white'='garlic'.

(202) a. *Bìssar ruuma hatthu aada.*
          big   house INDEF exist.
          'There is a big house.' (K081103eli02)
       b. *??ruuma bìssar hatthu aada.*
          big   house INDEF exist.
          'There is a big house.' (K081103eli02)

Adjectives can modify verbal predications with or without the dative marker =*nang*.

(pàr/kà)thaama 'earlier'	*siini* 'here'	*thapi* 'but'
kumareng 'yesterday'	*sana(ka)* 'there yonder'	*incalla* 'hopefully'
nyaari 'today'	*siithu* 'there'	*suda* 'thus'
(s)karang 'now'	*giini* 'this way'	*sajja* 'only'
beeso 'tomorrow'	*giithu* 'that way'	*sinderi* 'from here'
luuso 'later than tomorrow'	*sgiini* 'this much'	*sanderi* 'from there yonder'
subbang 'often'	*sgiithu* 'that much'	*sithari* 'from there'

Table 5.4: SLM Adverbs

(203) *Baae(=nang) arà-nyaanyi.*
       good=DAT   NON.PAST-sing.
       'He sings well.' (K081103eli02)

Adjectives are negated with *thraa* (204). They can also use verbal negation when used in a verbal frame (205), and nominal negation with *bukang* when used in a nominal frame (Slomanson 2006:141).

(204) *Se=ppe aanak **iitham thraa**.*
       1s=POSS child   black
       'My child is not black.' (K081103eli02)

(205) *Aanak thàrà-thiiňǧgi.*
       child   NEG.NONPAST-big.
       'The child did not grow tall.' (K081103eli02)

A small subclass of adjectives are always negated by *thàrà-*, without implying a verbal reading or reference to the past (206). Negation with *thraa* is not possible for the adjectives in this class.

(206) *Se=dang **thàrà-sìggar**.*
       1s=DAT   NEG[adj]-healthy
       'I am/was not healthy.' (K081104eli06)

Adjectives can take both nominal and verbal morphology. A variety of tests are thus necessary to determine whether a lexeme is an adjective, and to rule out the possibility that it could be a noun or a verb.

5.4 Adverbs

Adverbs are words that carry lexical meaning but are neither nouns nor verbs nor
adjectives and do not quantify. This class is quite small in SLM. Monomorphemic
adverbs are given in Table 5.4. Adverbs can combine with postpositions as shown in
the examples below.

(207)   ***Karam=pe*** *mosthor=nang, mpapulu aari=ka=jo   sunnath=le*
        now=POSS manner=DAT   forty   day=LOC=EMPH circumcision=ADDIT
        arà-kijja.
        NON.PAST-make.
        'For today's way of doing (it), it is on the fortieth day that they also do the
        circumcision.' (K061122nar01)

(208)   a. ***Dovulu=pe*** *oorang pada itthu anà-kirja hathu thundu bambu*
            before=POSS man   PL   DIST PAST-make INDEF piece bamboo
            asà-ambel=apa.
            CP-take=after
            'The people from former times to make this, they took a piece of bamboo
            and'
        b. *Itthu=yang baaye=nang arà-*wrap-*kang   athu kaayeng*
            DIST=ACC   good=DAT   NON.PAST-wrap-CAUS INDEF cloth
            thundu=dering.
            piece=ABL
            'wrapped it well in a piece of cloth.' (K061026rcp04)

(209)   *[Incayang cinggala asà-blaajar]=apa **sini=pe**   raaja=nang mà-banthu*
        3s.POLITE Sinhala CP-learn=after   here=POSS king=DAT INF-help
        anà-mulain.
        PAST-start
        'He learned Sinhala and began to help the local king.' (K060108nar02)

   SLM adverbs can be divided into temporal adverbs, adverbs with a spatial deictic
component, and a 'remainder' class of other adverbs.

5.4.1 Temporal adverbs

Temporal adverbs normally do not take postpositions, with the exception of *beeso* 'to-
morrow' and *luusa* 'some day after tomorrow', which often combine with the dative.
The following examples illustrate this.

(210) *Muula **pàrthaama** Badulla ruuma saakith=ka s-riibu sbiilan raathus*
before first Badulla house sick=LOC one-thousand nine hundred
lima-pulu dhlaapan=ka pukurjan arà-gijja vakthu.
five-ty eight=LOC work NON.PAST-make time
'Before, when I was working in Badulla in 1958.' (K051213nar01)

(211) *Inni railway department=ka head guard hatthu **kàthaama**.*
PROX railway department=LOC head guard INDEF before
'He was a head guard in the railway department before.'

(212) *Punnu mlaayu pada **kàthaama** English=jona anthi-oomong.*
many Malay PL earlier English=PHAT IRR=speak
'Many Malays would speak English in former times, wouldn't they?'
(K051222nar06)

(213) a. *Non-Muslims pada=subbath kithang muuka konnyong*
non-Muslims PL=because 1PL face little
arà-cunji-kang siini.
NON.PAST-show-CAUS here
'Because of the non-Muslims, we show our faces here.'

b. *Thapi **karang karang skarang skarang** Sri Lanka=pe=le Islam*
but now now now now Sri Lanka=POSS=ADDIT Islam
pada muuka arà-thuuthup karang dovulu abbis-dhaathang muuka
PL face NON.PAST-close now before COPULA face
thama-thuuthup.
NEG.NONPAST-close
'But now, Muslims in Sri Lanka cover their faces, now; before it was that
they would not cover their faces.' (K061026prs01)

(214) a. ***Beeso luusa** lubaarang arà-dhaathang.*
tomorrow some.day.after.tomorrow festival NON.PAST-come.
'The day after tomorrow is the festival.'

b. *Itthu lanthran kithang=pe ruuma see arà-cuuci.*
DIST reason 1PL=POSS house 1s NON.PAST-clean
'That's why I am cleaning the house.' (K061019prs01)

5.4.2 Adverbs with a deictic component

A number of adverbs have clear diachronic relations to the deictics *ini* 'proximal' and *itthu* 'distal' ⊡→5.8, p. 237. Sometimes, a third adverb with *san-* exists in the set, whose deictic value is unclear. As for local adverbs, *siini, siithu* and *sana* exemplify the pattern.

(215)   *[Non-Muslims pada=subbath] kithang muuka konnyong arà-cunji-kang*
       non-Muslims PL=because   1PL   face   little   NON.PAST-show-CAUS
       siini.
       here
       'Because of the non-Muslims, we show our faces here.' (K061026prs01)

(216)   ***Siithu*** *umma-baapa arà-duuduk=si.*
       there   mother-father NON.PAST-exist.ANIM=INTERR
       'Do your parents live over there?' (B060115cvs03)

(217)   *Itthu thoppi=yang **sana**=ka   simpang.*
       DIST   hat=ACC   there=LOC keep.
       'Keep that hat over there.' (K081103eli03)

Sana- seems to be mutually exclusive with *siini*; the precise difference in meaning between *siithu* and *sana(ka)* is unclear. *Siithu* is the normal adverb used and occurs frequently in the corpus, *sana(ka)* does not occur at all. Note that *siini* and *siithu* can occur on their own, while *sana* must take the locative postposition *=ka*. *Siini* and *siithu* can also combine with *=ka* (218)(219), and with the ablative postposition *=dering* (220)(221). Combinations with the allative postposition *=nang* are not possible.

(218)   Malay regiment *hatthu dhaathang=apa **sini=ka**   settle=apa.*
       Malay regiment INDEF come=after   here=LOC settle=after.
       'After the Malay regiment had come and settled here.' (G051222nar03)

(219)   *Derang anà-baalek sajja=jo,   **sithu=ka** panthas hathu Aanak raaja*
       3PL   PAST-return only=EMPH there=LOC beautiful INDEF child   prince
       su-aada!
       PAST-exist
       'And when they returned, there was beautiful prince!' (K070000wrt04)

(220) *Incayang **siini=dering** su-pii.*
      3s.POLITE here=ABL   PAST-go.
      'He went off from here.' (K081103eli03)

(221) *Incayang **siithu=dering** su-dhaathang.*
      3s.POLITE there=ABL   PAST-come.
      'He came back from there.' (K081103eli03)

The combination of the adverbs with =*dering* is also the origin of a related set of adverbs: *sindari* 'nearside', *sithari* 'farside' and *sandari*, whose meaning is unclear as of now. These adverbs require the dative postposition =*nang* when the ground to which they relate is expressed (a grandfather in (222) and (223)). In this respect, they resemble temporal relator nouns ⊡→5.2.2, p. 199, what distinguishes them is the absence of the lexical counterpart of relator nouns.

(222) *Incayang=nang **sithari** se=dang mà-biilang thàràthaau.*
      3s.POLITE=DAT that.side 1s.DAT   INF-say   cannot.
      'From this person onwards, I am unable to explain.' (K060108nar02)

(223) *Incayangnang **sindari**, se=dang mà-biilang bannyak thaau.*
      3s.POLITE=DAT this.side 1s.DAT   INF-say   much   can.
      'Up to this person, I call tell you a lot.' (K081104eli06)

Related adverbs are also found in the domains of manner (*giini* and *giithu*) and amount (*sgiini* and *sgiithu*), but there, only two degrees of distance are distinguished. The equivalent of *sana(ka)* is missing.[17]

(224) *Ithukang ithu bambu **giithu**=jo   luvar=nang arà-dhaathang.*
      then   DIST bamboo like.that=EMPH outside=DAT NON.PAST-come
      'Then, when the bamboo come out like that.' (K061026rcp04)

(225) *[**Sgiini**   lakuvan de sindari   arà-baa   katha] asà-thaau blaakang,*
      this.much wealth 3s from.here NON.PAST-take QUOT CP-know   after
      soojer   pada incayang=sàsaama Seelon=nang asà-dhaathang ...
      European PL   3s.POLITE=COMIT   Ceylon=DAT CP-come
      'After having learnt that he brings this much wealth from here, the Europeans came to Ceylon together with him (and ...)' (K060103nar01)

---

[17]This restricted use of *sana* seems to be the normal case in Malay varieties (cf. **Paauw** 2008:276).

(226)   *Derang pada* **sgiithu**   *braani*
   3PL PL that.much strong
   'They were so strong.' (K060108nar02)

The proximal and the distal deictic are ubiquitous, while *sanaka* is quite rare.

5.4.3   Other adverbs

Three other adverbs were found, which do not pattern with the ones discussed above: *suda* 'thus' (227) (Smith & Paauw 2006:167), *sajja* 'only' (228) and *incalla* 'hopefully' (229).

(227)   **Suda** *skaarang kitham=pe aanak pada laayeng pukurjan pada*
   thus now   1PL=POSS child PL other work PL
   arà-girja.
   NON.PAST-make.
   'So now our children work in other professions.' (K051222nar05)

(228)   *Aathi=yang* **sajja** *hatthu oorang=nang bole=ambel hathyang bagiyan*
   liver=ACC only one man=DAT can=take other part
   bole=baagi sama oorang=nang.
   can=divide all man=DAT.
   'The liver can only be taken by one person; the other parts can be divided among all men.' (K060112nar01)

(229)   **Incalla** *[lai thaau sudaara sudaari pada]=ka bole=caanya ambel*
   Hopefully other know brother sister PL=LOC can-ask take
   [nya-gijja lai saapa=kee aada=si katha].
   PAST-make other who=SIMIL exist=INTERR QUOT
   'Hopefully, you can enquire from another person you know whether there is someone else who did something.' (N061031nar01)

5.5   Copula

SLM has a copula which is derived from the verb *dhaathang* 'come' by either the conjunctive participle prefix *asà-* , or the postposition *=apa*, or both. Possible forms then include *as(à)dhaathang, dhaathang(a)pa, dhaathampa, as(à)dhaathampa* and *as(a)dhaathang(a)pa*. The choice of form depends on the idiolect. The three main patterns with *asà-*, *=apa* and both are illustrated below in sentences introducing names,

but there are more uses to be discussed below.

(230) a. *Ini* head master ***asà****dhaathang hathu Jayathilaka.*
           PROX head master COPULA   INDEF Jayathilaka
           'This head master was a certain Jayathilaka.'

       b. *Se=ppe* class master ***asà****dhaathang hathu Mr Senevirathna.*
           1s=POSS class master COPULA   INDEF Mr Senevirathna
           'My class master was a certain Mr Senevirathna.' (K051201nar02)

(231) *Karang se=ppe* father-in-law *dhaathang**apa** Mr Asali.*
       now   1s=POSS father-in-law COPULA   Mr Asali
       'Now my father-in-law is Mr Asali.' (G051222nar01)

(232) Owner ***as****dhaathang**apa** Sir Handy Kothalawela.*
       owner COPULA   Sir Handy Kothalawela
       'The owner was Sir Handy Kothalawela.' (K051213nar04)

In the above examples, the copula is combined with foreign names (Sinhala and English), but it can also be found with native words, as in (233).

(233) *Se=ppe baapa dhaathangapa **Jinaan Samath***.
       1s=POSS father COPULA   Jinaan Samath.
       'My father was Jinaan Samath.' (N060113nar03)

As the preceding examples show, the copula is underspecified for tense and allows for both a present and a past interpretation.

The use of *(asà)dhaathang(apa)* as a copula has to be distinguished from the use as a conjunctive participle in the perfect construction as in (234) or in clause chains as in (235).

(234) *See skarang Sri Lanka=nang **asà-dhaathang** aada.*
       1s now   Sri Lanka=DAT CP-come   exist.
       'I have now come to Sri Lanka.' (K051206nar17)

(235) *[Itthu nigiri=deri   **as-dhaathang** anà-thiiňǧgal oorang pada]=jo kithang.*
       DIST country=ABL CP-come   PAST-stay   man   PL=EMPH 1PL.
       'The people who have come from those countries and stayed (here) are we.'
       (K051222nar03)

In example (235), we are dealing with the literal meaning of *dhaathang* 'come' as used in a clause chain where it is a non-final element and thus marked with the

conjunctive participle prefix *asà-*. An interpretation as a copula is not possible here.

When used for naming purposes, the copula can combine with the person herself, as in the examples given above, or with the word *naama* 'name', as in the next example.

(236) *Se=ppe naama asàdhaathang Cintha Sinthani.*
       1s=poss name   copula   Chintha Sinthani.
       'My name is Chintha Sinthani.' (B060115prs04)

Besides for introducing names, the copula is mainly used for giving information on kinship status and profession. These two uses are illustrated below for all three forms, yielding six possibilities. Both the use of *dhaathangapa* with kin and the use of *asdhaathangapa* with professions is not attested. This is thought to be an accidental gap in the corpus.

(237) *Baapa=pe umma **asà**dhaathang kaake=pe   **aade**$_{kin}$.*
       father=poss mother copula   grandfather=poss younger.sibling.
       'My paternal grandmother was my grandfather's younger sister.'
       (K051205nar05)

(238) Estate=*pe*   field officer ***asà**dhaathang**apa** kithang=pe **kaake**$_{kin}$*
       estate=poss field officer copula   1PL=poss   grandfather.
       'The estate field officer was our grandfather.' (N060113nar03)

(239) *[Seelong=nang duppang duppang anà-dhaathang mlaayu] **asà**dhaathang*
       Ceylon=DAT   before before PAST-come   Malay] copula
       ***oorang ikkang**$_{profession}$.*
       man   fish
       'The Malays who came to Ceylon very early were fishermen.'
       (K060108nar02)

(240) *Ummape   baapa dhaathang**apa** hathu **inspector of police**$_{profession}$.*
       mother=poss father copula   INDEF inspector of police.
       'My mother's father was an inspector of police.' (N060113nar03)

Next to these central uses, which are very frequent, the copula can also be used for the following: age (241)(242), nationality (243), sex (244).

(241) *Suda se=ppe thuuva anak klaaki **asàdhaathang dhlapan-blas thaaun**ₐgₑ.*
     so   1s=ᴘᴏss old   child male ᴄᴏᴘᴜʟᴀ   eight-teen   year
     'So my eldest son is eighteen.' (K060108nar02)

    Just as with names, the predicate (a number in this case) can be attributed to either the person or the quality, in this case the age.

(242) *Go=ppe age **asàdhaathang 78**ₐgₑ.*
     1s=ᴘᴏss age ᴄᴏᴘᴜʟᴀ   78
     'My age is 78[18].' (B060115nar04)

(243) a. *Se=ppe daughter-in-law=pe   mother **asàdhaathang bìnggaali**ₑₜₕₙᵢ𝒸gᵣₒᵤₚ.*
         1s=ᴘᴏss daughter-in-law=ᴘᴏss mother ᴄᴏᴘᴜʟᴀ   Bengali

         'My daughter-in-law's mother is Bengali.'
    b. *Ithukapang daughter-in-law=pe father **asàdhaathang***
        then   daughter-in-law=ᴘᴏss father ᴄᴏᴘᴜʟᴀ
        ***mlaayu**ₑₜₕₙᵢ𝒸gᵣₒᵤₚ.*
        Malay
        'Then my daughter-in-law's father is Malay.' (K051206nar08)

(244) *Kàthama aanak dhaathangapa **klaaki**ₛₑₓ.*
     first   child ᴄᴏᴘᴜʟᴀ   male
     'My oldest child is a boy.' (G051222nar01)

    These are the central uses of the copula. In the following, I will give some more peripheral uses of the copula. Example (245) can be seen as a special case of naming. However, it does not apply to persons, but to a problem caused by an organization.

(245) *Itthu vakthu kithang=nang nya-aada **asàdhaathang** ini   **JVP katha** hathu*
     ᴅɪsᴛ time   1ᴘʟ=ᴅᴀᴛ   ᴘᴀsᴛ-exist ᴄᴏᴘᴜʟᴀ   ᴘʀᴏx JVP ǫᴜᴏᴛ ɪɴᴅᴇғ
     problem.
     problem.
     'What we had at that time was the so-called JVP-problem.' (K051206nar10)

    The copula can also be used for more complex relations, like the complement clause in the following sentence.

---

[18]This speaker has little command of English, so that a calque on the English sentence is unlikely.

(246) *Suda karang kithang=nang aada* problem ***dhaathangapa** [kithang=pe*
      SO   now   1PL=DAT   exist problem COPULA   1PL=POSS
      aanak pada mlaayu thama-oomong].
      child PL   Malay NEG.IRR-speak
      'So our problem is now that our children do not speak Malay.'
      (G051222nar01)

A complement clause is also introduced by the copula in the following sentence. Actually, it could be argued that the whole stretch of following utterances is predicated on *Sepakthakrawpe rules* 'the rules of Sepaktakraw'.

(247) a. *Sepakthakraw=pe* rules ***dhaathangapa.***
         Sepaktakraw=POSS rules COPULA
         'The rules of Sepaktakraw are as follows:'
     b. *Ini   hathu* badminton court=*ka   arà-maayeng.*
         PROX INDEF badminton court=LOC NON.PAST-play
         'You play it on a badminton court.'
     c. Game *hatthu itthe same* measurement, badminton=*pe   height=le*
         game INDEF other same measurement badminton=POSS height=ADDIT
         same=*jo.*
         same=foc.
         'In a game, you use the same measurements and the same height as in badminton.' (N060113nar05)

A property is introduced by the copula in the following sentence.

(248) *Gampola=ka bannyak mulbar. Giithu=le   thraa. Siini **dhaathangapa***
      Gampola=LOC many Tamil   like.that=ADDIT NEG here COPULA
      mixed
      mixed
      'There are many Hindus in Gampola. It's not that much here. Here, the population is mixed.' (G051222nar04)

An object is specified in the following example.

(249) *Itthu asàdhaathang [baaye=nang vaasil-kang   aada dagger] hatthu.*
      DIST COPULA   good=DAT   blessing-CAUS exist dagger INDEF
      'That one was a well-blessed dagger.' (K051206nar02)

Finally, a date is introduced by the copula in (250).

(250) *Ini   laaher dhaathangapa* 1940=*ka.*
      PROX birth   COPULA   1940=LOC
      'This birth took place in 1940.' (N060113nar03)

Missing arguments from preceding discourse can also be introduced by the copula.

(251) *Itthu asàdhaathang inni Raagala subala.*
      DIST COPULA   PROX Raagala side.
      'That was in the Raagala region.' (K051205nar05,K081105eli02)

(252) *Itthu abbisdhaathang* custard powder=*dering=jo arà-kirja.*
      DIST COPULA   custard powder=ABL=EMPH NON.PAST-make
      'This   one   is   such   that   it   is   made   with   custard   powder.'
      (K061026rcp02,K081105eli02)

Note that in the great majority of cases, an English loan word is present. This might trigger some patterns containing copulas in the speakers' minds. On the other hand, there are normally no English loan words in the naming use, which is the most frequent one. However, loanwords and names are treated alike in other areas of the grammar as well, namely in that they often co-occur with the quotative *katha* ⊡→6.4.9.1, p. 389. This connection might merit further research. To close this section, let's have a look at a short stretch of discourse with three occurrences of *asdhaathang*, one as conjunctive participle and two as a copula, for introducing a profession and a name, respectively.

(253) a. *Se=ppe kaake   **asàdhaathang*** estate tea factory officer.
           1s=POSS grandfather COPULA   estate tea factory officer.
           'My grandfather was estate tea factory officer.'

       b. Estate tea factory officer. *Itthusubbath=jo incayang=yang siithu*
           estate tea factory officer therefore=EMPH 3s.POLITE=ACC   there
           anà-braanak
           PAST-be.born.
           'Estate tea factory officer. Because of that, he was born here'

       c. *Siithu asà-braanak incayang=pe   plajaran=nang incayang*
           there CP-be.born   3s.POLITE=POSS education=DAT 3s.POLITE
           Kandi=nang anà-dhaathang.
           Kandy=DAT   PAST-go.
           'After his birth, he went to Kandy for his education.'

       d. *Kandi=nang **asà-dhaathang** Kandi=ka   asà-kaaving=apa itthu=nang*
           Kandy=DAT   CP-come   Kandy=LOC CP-marry=after   DIST=DAT
           blaakang=jo kithang pada anà-bìssar.
           after=EMPH 1PL   PL   PAST-big
           'After having come to Kandy and after having married, after that we grew
           up.'

       e. *Kaake=pe   naama **asàdhaathang** TN Salim.*
           grandfather=POSS name   COPULA   TN Salim
           'My grandfather's name was TN Salim.' (K060108nar02)

5.6   Personal pronouns

SLM pronouns distinguish three persons and two numbers. Additionally, there are
some additional distinction of politeness and animacy. Unlike in other Malay varieties
or in Tamil, there is no inclusive/exclusive distinction in the first person plural. Table
5.5 gives a list of all the words in the class of pronouns.

   The standard form to refer to the speaker is *se* in the Upcountry. *Go* is understood,
but its use is frowned upon, making it one of the most stigmatized sociolinguistic
features.[20] I have found it in actual use in two households in Badulla, and in Slave
Island (not in the Upcountry), but normally it would only be the topic of metalinguistic
commentary.

   The following two examples show the use of *se* and *go* with the possessive post-
position =*(p)pe* and the dative form with *dang*.

---

[19]Adelaar (1991:32) has *lurang*.

[20]The cognates of *goo* are also stigmatized in other Malay varieties (cf. Ansaldo 2009b).

pronoun	=yang	=nang	=pe	meaning	origin
goo	goyang	godang	goppe	1s.FAMILIAR (Southern dialect)	Hokkien 1991, 2005b (Adelaar Ansaldo)
see	seeyang	sedang	seppe	1s.POLITE.	
luu	lu(u)yang	ludang	luppe	2s.FAMILIAR	Hokkien 1991, 2005b (Adelaar Ansaldo)
dee	deeyang	dedang	deppe	3s.INANIMATE/ 3s.FAMILIAR	
diya	diyayang	diyanang	diyape	3s.PROXIMAL	
siaanu	siaanuyang	siaanunang	siaanupe		
incayang	incayangyang	incayangnang	incayangpe	3s.POLITE	*encik ia* + addition of velar nasal ☐→4.6.2.5, p.136 (Slomanson 2008b)
spaaman	spaamanyang	spaamannang	spaamanpe	3s.POLITE	uncle?? (Std, Malay, Javanese *paman* (Adelaar 1985:141))
kithang	kithangyang	kithannang	kithampe	1PL	<*kita orang
lorang[19]	lorangyang	lorangnang	lorampe	2PL/2s.POLITE	<*lu orang
derang	derangyang	derangnang	derampe	3PL	<*de orang, neutral
incayang pada	incayang padayang	incayang padanang	incayang padape	3PL.POLITE	

Table 5.5: Pronouns. The etymology of the plural pronouns includes a contraction of *orang* 'man' to *rang* (Adelaar & Prentice 1996, Adelaar 2005, 1991)

(254)   ***Se**=ppe* profession=*subbath **se=dang** siini mà-pii su-jaadi.*
        1s=POSS profession=because 1s.DAT   here INF-go PAST-become
        'I had to come here because of my profession.' (G051222nar01)

(255)   ***Go**=ppe*   daughter *pada **go=dang***   *makanan arà-kaasi.*
        1s.FAMILIAR daughter PL   1s.familiar.DAT food   NON.PAST-give.
        'My daughters give me food.' (B060115nar04)

One speaker can use both forms in one sentence. This is thought to be an avoidance
of *go* which is not applied to the whole string.

(256)   ***See**=le*   *pii aada dhraapa=so*   *duuva thiiga skali **go***   *pii*
        1s=ADDIT go exist how=many=UNDET two   three time 1s.FAMILIAR go
        aada.
        exist
        'I also went, several times, two or three times I went.' (B060115nar05)

This is different for the forms to address the hearer, which shows a clear distinction
of familiarity and politeness. *Lu* is almost exclusively used to address children, or very
intimate friends, whereas in other contexts, *lorang* (otherwise second person plural)
has to be used and *lu* would be considered offensive.

The following two sentences show the use of *luu* when singing a song to a child
and the use of *lorang* when addressing a stranger.

(257)   a. ***Luu**=le*   *asà-dhaathang kanaapa nya-laaher.*
           2s.FAMILIAR=ADDIT CP-come   why   PAST-be.born
           'Why did you come to this earth?'

           Luu   *abbis bùssar **lu**=ppe umma-baapa=nang kaasi thaangang.*
           2s.FAMILIAR finish big   2s=POSS mother-father=DAT give hand.
           'When you will have finished growing up, lend a hand to your parents.'
           (K060116sng01)

(258) a. *See=yang luppas, Thuan Buruan.*
    1s=ACC leaave sir   bear
    'Let me go, Mister Bear.'
    Lorang *se=dang mà-hiidop thumpath kala-kaasi,*
    2PL   1s.DAT   info-live place   if-give
    'If you spare my life,'

    b. *see **lorang**=nang **lorang**=pe samma duvith=le,   baarang pada=le*
    1s 2PL=DAT   2PL=POSS all   money=ADDIT goods   PL=ADDIT
    anthi-bale-king.
    IRR-return-caus.
    'I will return all your money and goods to you.' (K070000wrt04)

The third person finally shows a greater deal of variation. The most frequent way to refer to a third person topic is probably zero expression, i.e. not mentioning the person, leaving it to the hearer to infer what the the topic of the utterance is. The second most common way is *incian/incayang,* which is the most commonly heard overt form in the Upcountry. It seems to be less common in Colombo (Peter Slomanson p.c.). The form *spaaman* seems to carry greater prestige, but is not heard very often. *Dee* is an impolite form. The plural form *derang* can also sometimes be found with singular reference and does not appear to be impolite. *Siaanu* finally can only refer to a person that is visible. This morpheme had escaped my attention for a long time, indicating that is not very common.

The text K051205nar02 narrates the story of a thief being captured by a sergeant. The thief is referred to as impolite *dee*, whereas the policeman receives respectful marking, *incayang*.

(259) ***Dee***   *buthul jahhath.*
    3s.IMPOLITE very   wicked
    'He (the thief) was very wicked.' (K051205nar02)

(260) ***Incayang=pe*** *naama Thuan Shaaban.*
    3s.POLITE=POSS name Thuan Shaaban
    'His name (of the sergeant) was Thuan Shaaban.' (K051205nar02)

Spaaman is used in B060115nar05 to refer to a saint.

(261)   a. *Seelonka   avuliya pada aada.*
         eylon=LOC saint   PL   exist.
         'There are saints in Sri Lanka.'

         b. ***Spaaman**=pe   naama Sekiilan avuliya.*
         3s.POLITE=POSS name   Sekiilan saint.
         'His name was Avuliya Sekiilan.' (B060115nar05)

That saint is later referred to by *derang*, technically a plural pronoun, but of clear
singular reference here.

(262)   a. *Sithu=ka=jo   kuburan samma asà-gaali.*
         there=LOC=EMPH grave   all   CP-dig
         'They dug his grave right there and'

         b. *Karang itthu avuliya **derang**=pe zihaarath aada.*
         now   DIST saint   3=POSS   shrine   exist
         'now there is his shrine there.' (B060115nar05)

The only instance of *siaanu* in the corpus is given below.

(263)   *Incayang=yang **siaanu** asà-buunung thaaro.*
         3s.POLITE=ACC   3s.prox CP-kill   put
         'This one has killed him.' (K051220nar01)

The plural pronouns have less distinctions. One interesting observation is that de-
spite their being inherently plural, the plural marker *pada* can optionally be used on
them. Very often, the final nasal of *kithang, lorang, derang* then assimilates to the fol-
lowing labial stop in *pada*, yielding *kithampada, lorampada, derampada. Derampada*
can further be reduced to *drampada*.

(264)   *Basra=ka hathu duuva haari **kithang** ∅ anà-duuduk.*
         Basra=LOC INDEF two   day   1PL   PAST-stay
         'We stayed one or two days in Basra.' (K051206nar19)

(265)   ***Kithang pada** siini=jo   arà-duuduk.*
         1PL   PL   here=EMPH NON.PAST-live
         'We live here.' (K051206nar07)

(266) *Boole lìkkas=ka [see **lorang**=∅=yang mliige=nang anthi-panggel].*
can quick=LOC 1s 2PL=ACC palace=DAT IRR=call
'As soon as possible I will call (the two of) you (girls) to the palace.'
(K070000wrt04)

(267) ***Lorang pada** pukurjan arà-gijja.*
2PL PL work NON.PAST-make.
'(The two of) you (researchers) are working.' (N061031nar01)

Whereas there are politeness distinctions in the third person singular, this is not the case for the plural, where *derang (pada)* is used for any referent. As with the first and second person plural, the use of *pada* is optional.

(268) *Irish* nuns*, **derang**=∅=pe English baaye.*
Irish nuns 3PL=POSS English good.
'The English of the Irish nuns was very good.' (K051222nar06)

(269) Dutch period=*ka* ***derang pada** dhaathang aada.*
Dutch period=LOC 3PL PL come exist
'They came in the Dutch period.' (K051206nar05)

The same plural referent can be referred to in the same utterance once by *derang* and another time by *derang pada*.

(270) *Itthukapang=jo **derang** ∅ nya-thaau ambel **derang pada** politic=nang suuka*
then=EMPH 3PL PAST-know take 3PL PL politic=DAT like
katha.
QUOT
'Only then will they come to know that they like politics' (K051206nar12)

5.7 Interrogative pronouns

SLM has a small set of basic interrogative pronouns. These can be combined with postpositions to give a greater array of semantic possibilities. The interrogative pronouns are reduplicated to indicate that exhaustiveness of the answer is required (271), and only once to indicate that this is not so (272).

(271) *Aapa˜aapa anà-bìlli?*
       what˜RED   PAST-buy
       'What all did you buy?' (K081103eli03)

(272) *Aapa anà-bìlli?*
       what PAST-buy
       'What did you buy?' (K081103eli03)

Besides in interrogative clauses, interrogative pronouns are also used together with clitics to form indefinite expressions ⊡→8.4, p. 447. Examples are given in (273) for an affirmative sentence and in (274) for a negative sentence.

(273) *Thapi **aapacara=so** itthu samma asà-iilang   su-aada.*
       But   how=UNDET   DIST all   CP-disappeared PAST-exist.
       'But somehow it had all disappeared.' (K20070920eml01)

(274) *Bannyak haari=dering **saapa=yang=ke** thàrà-enco-kang   katha*
       much   day=ABL   who=PAT=SIMIL NEG.PAST-fool-CAUSE QUOT
       anà-iingath Andare.
       PAST-think   Andare
       'Andare thought of all the days that he had not taken anyone for a fool.'
       (K070000wrt05)

A similar construction, but then with reduplicated interrogative pronouns is used to form maximizing headless relative clauses ⊡→8.1.12, p. 424 like in (275).

(275) *Inni **saapa~saapa**=ka inni mlaayu pakeyan pada aada**=so**, lorang pada*
       prox who~who=LOC   PROX Malay dress   PL   exist=disj 2PL   PL
       ini   mlaayu pakeyan=samma ini   kaving=nang mà-dhaathang bannyak
       PROX Malay dress=with   PROX wedding=DAT INF-come   much
       uthaama.
       honour
       'Whoever owns such Malay dresses, your coming together with this Malay dress to the wedding will be greatly appreciated .' (K060116nar04)

Finally, interrogative pronouns can be used with the additive =*le* to yield a universal quantifier.

(276) *Skarang **maana** aari=le   atthu atthu oorang=yang arà-buunung.*
      now   which day=ADDIT one   one   man=ACC   PAST-kill.
      'Now people kill each other every day.' (K051206nar11)

The interrogative pronouns can also be used in subordinate clauses (277)(278)
⊡→13.5, p. 532.

(277) *Andare raaja=ka su-caanya [inni mà-kirring simpang aada **aapa**=yang]*
      Andare king=LOC PAST-ask   PROX INF-dry   keep   exist what=ACC
      katha.
      QUOT
      'Andare inquired from the King what was it that was left [on the mat] to dry.'
      (K070000wrt02)

(278) *[Laayeng oorang pada **apcara** kijja] se thàrà-thaau.*
      different man   PL   how   make 1s NEG-know
      'I don't know how other people do it.' (B060115rcp02)

5.7.1   *aapa* 'what'

Aapa is used to query inanimate entities. It can be combined with a number of post-
positions, e.g. *aapanam* 'why, what for' or *aapadring* 'with what' (280). *Aapa* must
not be confounded with the postposition =*apa* with a short vowel ⊡→6.4.4.10, p. 353.

(279) ***Aapa** n-jaadi   mlaayu pada?*
      what PAST-become Malay PL
      'What became of the Malays?' (K051213nar06)

(280) *Ini   daging aapa=dering arà-poothong?*
      PROX meat   what=ABL   NON.PAST-cut
      'With what should I cut this meat?' (K081105eli02)

Aapa is also often found with the accusative postposition =*yang* even if it does not
have the semantic role of patient.[21]

---

[21]This actually contrasts with a statement in Slomanson (2006:150): "=*nya* [=*yang*] cannot be suffixed
to subjects in SLM." In example (281), the NP =*yang* attaches to is one part of an equational clause (the
other part is a headless relative clause), and the parts of equational clauses should probably be analyzed as
subjects in the theory underlying Slomanson's analysis. However, Slomanson does not assume that =*nya*
and =*yang* are identical; but given his etymological argumentation, what is said about the development of
=*nya* should be true of the development of =*yang* as well.

(281)   *Andare raaja=ka su-caanya [inni mà-kirring simpang aada ∅]* **aapa=yang**
        Andare king=LOC PAST-ask   PROX INF-dry   keep   exist   what=ACC
        katha.
        QUOT
        'Andare inquired from the King what was it that was left [on the mat] to dry.'
        (K070000wrt02)

This is presumably because very often the items queried for by *aapa* were patients, so that this form (over)generalized.[22]

5.7.2   *saapa* 'who'

Saapa is used to query for persons. It can also be combined with a number of postpositions, like *saapanang* 'for whom', *saapayang* 'whom' or *saapasàsaama* 'with whom' (283). An additional realization is [θaːpa].

(282)   a. **Saapa** *anà-maathi?*
            who   PAST-dead
            'Who died?'

        b. *Samma mlaayu pada!*
            all   Malay PL
            'All Malays!' (K051213nar07)

(283)   **Saapa=sàsaama** *Kluum̌bu anthi-pii?*
        who=COMIT   Colombo IRR-go
        'With whom will you go to Colombo?' (K081105eli02)

Like the other interrogative pronouns, *saapa* can be used in clauses as above, or on its own as below.

(284)   **Saapa***? See=si?*
        who   1S=INTERR.
        'Who? Me?' (B060115prs18)

The combination of *saapa* with a postposition and its indefinite use in a subordinate are illustrated in (285).

---

[22]Note that this poses problems for the proposed etymology of *=yang*, which should be a patient marker which generalized from patient questions (Slomanson 2006:150). If *=yang* can be used in non-patient questions, this analysis loses credibility.

(285) *Lorang pada asà-dhaathang [**saapa=pe**=ke baa   thaangang=ka=jo*
     2PL   PL   COPULA   who=SIMIL   under hand=LOC=EMPH
     pukurjan mà-gijja athi-jaadi].
     work   INF-make IRR-become
     'You will always have to work under someone's command' (K051206nar07)

The use of *saapa* in reported questions is given in (286).

(286)   a. *Luu=nya   jadi-kang   rabbu saapa; lu=ppe nabi   pada*
           2s.FAMILIAR=ACC become-CAUS prophet who   2s=POSS prophet PL
           ***saapa** katha biilang.*
           who   QUOT say
           'Say who the prophet is who made you, who are your prophets.'

       b. *Lu=ppe rabbu **saapa** katha buthul balas-an   asà-biilang,*
           2s=POSS prophet who   QUOT correct answer-NMLZR CP-say
           'Having given the correct answer as to who your prophet is,'

       c. *lu=ppe nabi   **saapa** katha buthul balas-an   asà-biilang, ...*
           2s=POSS prophet who   QUOT correct answer-NMLZR CP-say
           'having given the correct as to who your prophets are, ...'

5.7.3   *ma(a)na* 'which, what'

This pronoun is polysemous. It can be used for querying a place (287), or for querying which members of a set have a positive truth value for the proposition (288).

(287) ***Mana** nigiri=ka   arà-duuduk?*
      which country=LOC NON.PAST-stay
      'Which country do you live in?' (B060115cvs16)

(288) *Itthu blaakang se=dang karang **maana**=ke   pii thàràboole.*
      DIST after   1s.DAT   now   where=SIMIL go cannot
      'Thereafter, I cannot go anywhere now.' (K061120nar01)

Just like the other interrogative pronouns, *mana* can combine with postpositions to increase the semantic range of possible queries, like *mana(de)ri* 'where from' or *manaka*, which highlights the locative reading. An example is given in (289), which also shows the use for indefinite reference.

(289) *Saudi=so,*   ***mana=ka=so;***   *athu nigiri=ka.*
       Saudi.Arabia=UNDET where=LOC=UNDET INDEF country=loc.
       'In Saudi Arabia or somewhere, in some country.' (B060115nar02)

5.7.4 *kaapang* 'when'

Kaapang is used to query for a point in time, like English *when*. The interrogative pronoun is nearly homonymous with the postposition =*kapang* meaning 'then', but the two differ in the length of the vowel. There happens to be no instance of *kaapang* used in a question in the corpus. The following example is elicited.

(290) *Kaapang loram pada siini arà-dhaathang?*
       when   2PL   PL   here NON.PAST-come.
       'When are you coming here?' (K081106eli01)

There are some instances of *kaapang* being used with clitics to yield a variety of indefinite (291) or universal readings (292)-(294).

(291) *See lorang=nang arà-simpa   **kaapang=ke** see lorang=nang ithu*
       1s 2PL=DAT   NON.PAST-promise when=UNDET 1s 2PL=DAT   DIST
       uuthang arà-baayar   katha.
       debt   NON.PAST-pay QUOT
       'I promise you that I will pay back that debt some day.' (K070000wrt04)

(292) *Suda itthu kithang=nang **kaapang=pon** thama-luupa.*
       thud DIST 1PL=DAT   when=any   NEG.IRR-forget
       'So, we will never forget this.' (B060115nar02)

(293) *Go   **kaapang=le** saala thamau-gijja.*
       1s.FAMILIAR when=ADDIT wrong NEG.NONPAST-make.
       'I never do any wrong.' (B060115nar04)

(294) Girls High School *Kandi=ka   se=dang **kaapang=le** udahan   hatthu*
       girls high school Kandy=LOC 1s.DAT   when=ADDIT invitation INDEF
       arà-kiiring.
       NON.PAST-send.
       'I am always invited to the Girls High School in Kandy.' (K061127nar03)

5.7.5   *càraapa* 'how'

Càraapa is used to query manner. Its meaning is composed of *aapa* 'what' and *caara* 'way'. The inverted form *a(a)p(a)caara* is also possible.

There is no example of this pronoun used in a question in the corpus, hence (295) gives an elicited example.

(295)   *Aapacara/caraapa ini   arà-gijja.*
       how   PROX pro-make.
       'How do you do that?' (K081103eli03)

The corpus does contain uses in subordinate clauses, as in (296)-(298).

(296)   *[Laayeng oorang pada **apcaara** kijja] se thàrà-thaau.*
       different man PL   how   make 1s NEG-know
       'I don't know how other people do it.' (B060115rcp02)

(297)   *Derang thàrà-thaau [ini   pàrhaal pada **aapacara** anà-jaadi=so].*
       3PL   NEG-know PROX problem PL   how   PAST-become=UNDET
       'They did not know how these problems had occurred.' (K051213nar01)

(298)   *[Cinggala=nang=le Dutch=nang=le   **aapcara** anà-banthu katha] thaau.*
       Sinhala=DAT=ADDIT Dutch=DAT=ADDIT how   PAST-help   QUOT know
       'He knows how (the Malays) helped the Dutch and the Sinhalese.'
       (K051206nar04)

Indefinite uses are given in (299) and (300).

(299)   *Thapi **aapacara=so** itthu samma asà-iilang   su-aada.*
       But   how=UNDET DIST all   CP-disappeared PAST-exist.
       'But somehow it had all disappeared.' (K20070920eml01)

(300)   ***Aapcara=ke** incayang ini   ciina oorang Islam=nang asà-dhaathang,*
       how=SIMIL 3s.POLITE PROX China man   Islam=DAT CP-come
       asà-kaaving=apa, karang màsiigith=nang arà-pii.
       CP-marry=after   now   mosque=DAT   NON.PAST-go.
       'Somehow he, this Chinaman converted to Islam and married and now goes to
       the mosque.' (K051220nar01)

The 'semi-reduplicated' from *caraapacara* also exists.

(301)   *Kithang caraapacara=kee mà-pii anà-aada thumpath=ka su-sampe.*
        1PL   how=SIMIL   INF-go PAST-exist place=LOC   PAST-reach
        'We arrived at our destination anyhow.' (K081103eli03)

5.7.6   *dhraapa* 'how much'

Dhraapa is used to query an amount or a quantity. As a special case, it can also be used to query the length of a period of time when combined with *laama* 'while', so *dhraapa laama* 'how long'.

Example (302) shows the use of *dhraapa* in a question.

(302)   **Dhraapa** *thaaun* **dhraapa** *buulang lu   arà-baapi   suusa.*
        how.many year   how.many month   2S.FAMILIAR NON.PAST-bring sad
        'How many years, how many months are you bringing sadness (into my life).'
        (K061123sng01)

Dhraapa can also be combined with *bannyak* to emphasize the quantity. This can be done in declaratives (303) but is mainly done in exclamatives (304)(305).

(303)   **Dhraapa=so**   **bannyak** doctors *pada=nang inni* native treatment *samma*
        how.many=UNDET many   doctors PL=DAT   PROX native treatment all
        anà-kirja.
        PAST-make.
        'Several doctors,   many   doctors   applied   these   native   treatments.'
        (K051205nar05)

(304)   *Spuulu aanak aada kalu* **dhraapa**   **bannyak** *pasiith athi-aada.*
        ten   child exist if   how.many much   trouble IRR-exist
        'To have ten children, how much trouble there must be!' (B060115nar04)

(305)   *Inni* two months=*ka* **dhraapa**   **bannyak** *blaajar aada!*
        PROX two months=LOC how.many much   learn   exist
        'How much he has learned in these two months!' (B060115cvs01)

The indefinite use of *dhraapa* with a clitic is shown in the following examples.

(306) *See=le   pii aada,* **dhraapa=so**,   *duuva thiiga skali go   pii*
      1S=ADDIT go exist   how=many=UNDET two   three time 1S.FAMILIAR go
      aada.
      exist
      'I also went, several times, two or three times I went.' (B060115nar05)

(307) *Kaake=nang*   **dhraapa=so**   *bannyak aanak pada, se=dang kalu*
      grandfather=DAT how.many=UNDET much   child PL   1S.DAT   if
      blaangang thàrà-thaau.
      amount   NEG-know
      'Grandfather had several children, many children, as for me, I do not know
      the number.' (K051205nar05)

An example of the universal quantifier is found in (308), where the additive clitic attaches to the verb.

(308) **Dhraapa**   *oorang theem̃bak=**le**, incayang=nang thama-kìnna.*
      how.many man   shoot=ADDIT 3S.POLITE=DAT   NEG.IRR-affect
      'How ever many people shot at him, none would touch him.' (K051206nar02)

5.7.7   *kànaapa* 'why'

This pronoun is used to query a reason and is an alternative to *aapa=nang* 'what=DAT'.

(309) *Sama oorang masà-thaksir [kithang=pe nigiri=ka   **kànaapa** kithang ini*
      all   man   must-think 1PL=POSS   country=LOC why   1PL   PROX
      bedahan   arà-simpang].
      difference NON.PAST-keep.
      'All men must think why we maintain these differences in our country.'

5.8   Deictics

This class comprises words that serve to encode different distances between the speaker and the entity talked about. The basic forms are proximal *inni* 'this' and distal *itthu* 'that' as given in the following two examples.

(310) *Suda **inni** moonyeth pada aapa thaau=si anà-gijja!*
       Thus PROX monkey   PL   what know=IRR PAST-make
       'So these monkeys do you know what they did !?' (K070000wrt01)

(311) ***Itthu** aanak pompang duuva=nang thriima thàrà-kaasi.*
       DIST child female   two=DAT   thank NEG.PAST-give
       'He did not thank those two girls.' (K070000wrt04)

   Both forms can also occur without gemination.

(312) *Sdiikith thaaun=nang duppang see **ini** Aajuth=nang su-kìnna   daapath.*
       few   year=DAT   before 1s PROX dwarf=DAT   PAST-strike get
       'Some years ago, I fell prey to this dwarf.' (K070000wrt04)

(313) *See lorang=nang arà-simpa   kaapang=ke see lorang=nang **ithu***
       1s 2PL=DAT   NON.PAST-promise when=UNDET 1s 2PL=DAT   DIST
       uuthang arà-baayar   katha.
       debt   NON.PAST-pay QUOT
       'I promise you that I will pay back that debt some day.' (K070000wrt04)

   Note that the four examples above are all taken from written texts by the same
speaker, so that the use of geminates in deictics can be both present and absent in the
same idiolect. In the corpus, *itthu* (452 instances) is much more common than *ithu*
(96). The difference is smaller between *inni* (224) and *ini* (147).
   The deictics can occur with all postpositions. As an example, (314) gives the use
of the proximal deictic combined with the possessive marker *=pe*.

(314) *Se=dang kalu, suda bannyak thàrà-thaau **inni=pe**   atthas*
       1s.DAT if   thus much   NEG-know   PROX=POSS about
       mà-biilang=nang.
       INF-say=DAT
       'So, as for me, I cannot tell you much about this.' (K051205nar04)

   Deictics can be used either to modify a noun as in the examples (310)-(313) above,
or referentially as in (314) or below.

(315) ***Ini**   kithang=pe ruuma.*
       PROX 1PL=POSS   house
       'This is our house.' (K051222nar02)

base form	place	location	source	manner	amount	distance
ini	siini	siinika	sindari	giini	sgiini	close to speaker
ithu	siithu	siithuka	sithari	gitthu	sgiithu	away from speaker

Table 5.6: Deictics

(316) *Andare=yang mà-enco-king=nang raaja su-biilang [**itthu** paasir] katha.*
Andare=ACC INF-fool-CAUS=DAT   king PAST-say DIST sand   QUOT
'To fool Andare, the king said that that was sand.' (K070000wrt02)

The normal function of the deictics is to indicate the spatial location of a referent ⊙→15.4.2, p. 607. In the following example, two women scream from different sides. One side is indicated with *itthu*, the other one with *inni*.

(317) a. *Biini **itthu** subla=dering arà-bitharak;*
wife DIST side=ABL   NON.PAST-scream
'(Andare's) wife screamed from that side;'

   b. *puthri **inni** subla=dering arà-bitharak.*
queen DIST side=ABL   NON.PAST-scream
'the queen screamed from this side.' (K070000wrt05)

Occasionally, proper nouns can also be found together with deictics. In that case, the deictics serve to emphasize the proper noun.

(318) *See **ini**   Sri Lanka=ka nya-blaajar.*
1s   PROX Sri Lanka=LOC PAST-learn
'I studied here in Sri Lanka.' (K061026prs01)

(319) *Suda kithang karang **inni** Gampola=nang anà-dhaatham=pa hatthu thuuju*
thus 1PL   now   PROX Gampola=DAT PAST-come=after INDEF seven
thaaun=ka s-jaadi.
year=LOC   PAST-become
'So, now, it has become about seven years after we moved here to Gampola.' (G051222nar01)

There are a number of regular correspondences between the lexemes in this class and a number of adverbs ⊡→5.4.2, p. 216, which are listed in Table 5.6.

There might be a possibility of a third deictic *anà*, which was found in one elicitation session but could not be reproduced. The meaning of this deictic is not proximal, and thus different from *ini*, but the difference between *anà* and *itthu* could not be

established.

While continental varieties of Tamil only have a twofold contrast, at least Jaffna Tamil has a threefold contrast in deictics (proximal *i*-, medial *u*-, distal *a*-), while Sinhala has a fourfold contrast, adding an anaphoric deictic (Gair 1982, 1991a).

5.9   Quantifiers

There are five quantifiers in SLM: *sdiikith* 'few', *konnyom* 'few', *spaaru* 'some', *bannyak* 'many' and *sa(a/m)ma* 'all'. Another common way to express universal quantification is the WH=*le*-construction discussed in ▱→8.1.12, p. 424.

Quantifiers are distinguished from verbs and adjectives by their inability to combine with TAM-morphology. From nouns, they are distinguished by the fact that they are able to modify a predication without having to be derived with =*nang* as can be seen from example (320).

(320)   *Incayang bannyak(*=nang) arà-nyaanyi.*
        3s.POLITE much(=DAT)   NON.PAST-sing
        'He sings a lot.' (K081103eli02)

Normally, quantifiers are used as modifiers as in (321) for a noun, (322) for an adjective or (323) for a verb.

(321)   *Itthu=nam blaakang **bannyak oorang** pada siini se-duuduk.*
        DIST=DAT after   many   man   PL   here PAST-stay.
        'After that, many people lived here.' (G051222nar03)

(322)   *Suda inni kaving   **bannya(k) panthas**.*
        thus PROX wedding much   beautiful
        'So this wedding was very beautiful.' (K060116nar04)

(323)   *Itthu kumpulan=dang derang=jo **bannyak** arà-**banthu**.*
        DIST association=DAT 3PL=EMPH much   NON.PAST-help
        'It is them who help that association a lot.' (B060115cvs01)

Less often, they are used referentially as in (324)-(327).

(324) a. ***Samma**=le atthi-maati.*
           all=ADDIT   IRR-die
           'All would die.'

    b. ***Konnyong** atthi-salba.*
        little   IRR-escape
        'Few would escape.' (K081105eli02)

(325) ***Bannyak** se=dang e-daapath.*
       much   1s.DAT   PAST-get
       'I earned a lot.' (N060113nar04)

(326) *Incayang=pe baa=ka   **spaaru** aada.*
       3.POLITE=POSS down=LOC some   exist
       'Below him, there are (only) few (Malay houses).' (K051213nar05)

(327) ***Spaaru** Indonesia=dering dhaathang aada.*
       some   Indonesia=ABL   come   exist
       'Some came from Indonesia.' (K060108nar02)

Even rarer is the use as a predicate as in (328).

(328) *Buthul **konnyong**=jo mulbar, hatthu period.*
       very   few=EMPH   Tamil   INDEF period
       'Tamil was very sparse at that time.' (K051222nar06,K081105eli02)

Both the referential and the modifying use can be seen in (329).

(329) ***Spaaru** pada **bannyak** suuka arà-blaajar.*
       some   PL   much   like   NON.PAST-learn
       '(Only) some like to study a lot.' (B060115cvs01)

Normally quantifiers precede the element they modify (330), but other material can intervene between the quantifier and the head word, such as *Muslim* in (331).

(330) *Inni sudaari=pe femili=ka   **bannyak oorang** tsunami=da spuukul su-pii.*
       PROX sister=POSS family=LOC many   people tsunami=DAT CP-hit PAST-go
       'In this sister's family, many people were swept away by the tsunami.'
       (B060115nar02)

1	(s/h)a(t)thu	11	subblas	100	sraathus
2	duuva	12	dooblas	200	duuva raathus
3	thiiga	13	thigablas	300	thiiga raathus
4	ùmpath	20	duvapulu	1000	sriibu
5	liima	21	duvapulsatthu	2000	duuva riibu
6	ìnnam	22	duvapulduuva	3000	thiiga riibu
7	thuuju	23	duvapulthiiga	26452	duvapulu ìnnam riibu
8	dhlaapan	30	thigapulu		ùmpath raathus limapulu
9	sbiilan	40	mpathpulu		duuva
10	spuulu	90	sbilanpulu		

Table 5.7: Numerals

(331) **Bannyak** *Muslim oorang pada arà-duuduk.*
       many   Muslim man   PL   NON.PAST-exist.ANIM
       'There are many Muslims.' (K061026prs01)

Occasionally, quantifiers float in the sentence. An example for this are given below, where the underscore marks the expected position of the quantifier.

(332) __ *cinggala su-aada* **sdiikith**.
          Sinhala PAST-exist few.
       'There were few Sinhalese.' (K051222nar06)

5.10 Numerals

SLM can use native numbers until 999,999 (Table 5.7). The words for the powers of 10 have a /s-/ which disappears in their multiples. This is a relic of historical *satthu* 'one'.

The following sentences show use of the smaller numbers in natural speech.

(333) *Kithang* **hathu** *week=nang* **hathu** *skaali* **duuva** *skaali=ke arà-maakang.*
       1PL   one   week=DAT one   time   two   time=SIMIL NON.PAST-eat.
       'We might eat it one or twice a week.' (K061026rcp04)

(334) *Pon=pe   ruuma=ka=jo   thaama* **duuva thiiga** *aari atthi-duuduk.*
       bride=POSS house=LOC=EMPH earlier two   three days IRR-stay.
       'They used to stay two or three days in the bride's house.' (K061122nar01)

(335)   a.   *Thapi ithu **thiiga**=nang arà-baagi.*
           but   DIST three=DAT   NON.PAST-divide.
           'But you divide it into three (parts).

     b.   ***Thiiga** asà-baagi.*
           three   CP-divide
           'Having divided it,'

     c.   ***hatthu** kithang=pe mà-use-king=nang;*
           one   1PL=POSS   INF-use-CAUS=DAT
           'one (part) is ours to use;'

     d.   ***hatthu** saanak sudaara;*
           one   relative brother
           'one is for the relatives;'

     e.   *hathyang miskiin=pada=nang.*
           other   poor=PL=DAT
           'and one is for the poor.' (K060112nar01)

(336)   *Incayang **thuju-pul-liima** thaaun=sangke incayang anà-iidop.*
        3s.POLITE seven-ty-five   year=until   3s.POLITE PAST-live
        'He lived up to his 75th year.' (K060108nar02,K081104eli06)

(337)   ***Duva-pulu ìnnam riibu**   **ùmpath raathus lima-pulu duuva*** votes
        two-ty   six   thousand four   hundred five-ty   two   votes
        incayang=nang anà-daapath.
        3s.POLITE=DAT   PAST-get
        'He got 26,452 votes.' (N061031nar01)

(338)   a.   *Kandi=ka   hathu **thiga-pulu riibu**=kee   mlaayu pada*
           Kandy=LOC INDEF three-ty   thousand=SIMIL Malay   PL
           arà-duuduk.
           NON.PAST-exist.ANIM.
           'There are 30,000 Malays in Kandy.'[23]

     b.   *Punnu=le Seelong=ka arà-duuduk.*
           full=ADDIT Ceylon=LOC NON.PAST-exist.ANIM
           'There are a lot in Sri Lanka.'

     c.   *Hathu **ìnnam-pulu liima riibu**=ke.*
           INDEF six-ty   five thousand=SIMIL
           'About 65,000.' (K060108nar02)

In everyday speech, compound numbers are normally in English and many people find it difficult to correctly construe the Malay numbers. The following example shows a complex number occurring in natural discourse, but this is quite rare, and normally, people would say *nineteen-fifty-two=ka*.

(339) **Sriibu   sbilang-rathus lima-pulu duuva=ka se nnam** class=*ka.*
       thousand nine-hundred five-ty   two=LOC 1s six   class=LOC
       'In 1952, I was in 6th grade.' (K051213nar02)

Mixing of small Malay numbers and bigger English number is common, an example is (340) (English numbers are written in digits in the examples, not in words).

(340) *Se=ppe nya-laaher*   date **duuva duuva** 1960.
       1s=POSS PAST-be.born date two   two   1960.
       'My birthday is 2-2-1960.' (K061019prs01)

Powers of ten above 100,000 are often expressed in *lakhs,*[24] a South Asian number meaning 100,000 (341).

(341) *Itthuka kithang=nang* **one lakh** *nya-daapath.*
       Then   1PL=DAT   one lakh PAST-get.
       'Then we received 100,000 Rupees.' (K060116nar06)

Fractions commonly used include *kaarthu* 'quarter' and *(s)thìnnga* '(a) half'. Percents are formed by using the dative on the numeral for '100', *sraathus*. The following example shows the use of both a fraction and a percentage in discussing the amount of ritual almsgiving Muslims have to perform.

(342) *Deram=pe cari-an=dering* **duuva sthìnnga**, **sraathus=na**, **duuva sthìnnga**
       3PL=POSS find-NMLZR=ABL two   half   hundred=DAT two   half
       itthu blaangang derang zakath masà-kaasi.
       DIST among   3PL   zakat must-give.
       'From their earnings, they must give 2 1/2%, 2 1/2, that amount they must give as *Zakat.*' (K061122nar01)

All cardinals are of Malay origin. The ordinals are derived by prefixing *ka-*, thus *kaduuva, kathiiga, ka-mpath* etc. The ordinal for *satthu* 'one' is an exception: it is *kàthaama*.

---

[23]This number is overestimated by order of magnitude. The actual number should be around 3,000.

[24]Sri Lankan English *lakh* < Indian English *lakh* < several Indian languages *la(a)kh* < Sanskrit *lakṣa*. Sinhala and Tamil do not use *lakh* , but *lakṣaya* and *ilaṭcam*, respectively.

(343) *Thahun baaru muusing=pe **kàthaama** aari=ka.*
      year   new   season=poss first   day=loc
      'On the first day of the new year.' (K070000wrt04)

5.11 Interjections

Interjections are morphemes which can form a whole utterance on their own, and are not morphologically integrated into the grammatical system of the language.

The most frequent interjection is *iiya* 'yes' This interjection is used to give affirmative answers.

(344) *SN: Butthul* sportsman*, bukang?*
      correct sportsman, TAG
      'They are god sportsmen, aren't they?'

(345) *SLM: **Iiya iiya**,* sportsman.
        yes yes sportsman.
      'Yes indeed, they are good sportsmen.'
      [. . .]

(346) *SN:* Football *aada?*
      Football exist.
      'Is there football being played in Badulla.'

(347) *SLM: **Iiya*** football *aada.*
        yes football exist
      'Yes, there is football.' (B060115cvs01)

The negative counterpart *thraa* is not treated as an interjection, but as a particle in this description ⊡→5.13.1, p. 254. This is because *thraa* has other meanings than 'negative answer'. It is integrated much more tightly into the grammar and participates in a number of constructions (verbal negation, adjectival negation, locational negation). Furthermore, it can serve as a base for causative derivation. This means that it fails the criteria for interjections, which are not integrated morphologically or syntactically.

The interjection *saya* can be used instead of *iiya* and is considered more polite (Saldin 2001:55). Saldin states that it corresponds to Malaysian *ia saya*.

*Ayyoo*n is an expression of surprise. It seems to be of Lankan origin. There is no instance of it in the corpus, but it was frequently used in conversations I overheard.

There are at least two interjections of Arabic origin, *yammuhaiyaddiin,* an expression of regret, and *alhamdulillah* 'praise to god'.

5.12 Modal particles

Modal particles are a small class of words that are used to indicate permission, capacity and volition. There are four of them, namely:

- *boole* 'can' (348) ⊡→5.12.1, p. 248

- *thàrboole* 'cannot' (349) ⊡→5.12.2, p. 249

- *maau, kamauvan* 'want' (350) ⊡→5.12.3, p. 250

- and *thussa* 'NEG.want' (351) ⊡→5.12.4, p. 253.

(348)   *Boole.*
        can
        'It is possible/(I/You/He/She/It/one) can.' (K081103eli03)

(349)   *Derang [thàràboole] katha su-biilang.*
        3PL   cannot   QUOT PAST-say
        'They said "It is not possible".' (K060116nar06)

(350)   *[Kamauvan] vakthu=nang, kithang=nang itthu mosthor=nang,* Malaysian
        want   time=DAT   1PL=DAT   DIST manner=DAT   Malaysian
        hathu mosthor=nang kithang=nang bole=duuduk.
        INDEF manner=DAT   1PL   can-exist.ANIM
        'When we want, we can be like this or in a Malaysian way' (K060108nar02)

(351)   *Thussa!*
        NEG.want.
        'Don't (you...)!!' (K081103eli03)

There are many similarities in behaviour, but also some slight differences, so that the members of this class are in a relation of family resemblance[25] (Wittgenstein 1953). Modal particles are distinguished from verbs by their total lack of verbal morphology, such as prefixes of tense, or causative derivation. Modal particles are distinguished from nouns by their inability to take any nominal modifier, such as adjectives, deictics, the indefiniteness marker, etc.

Modal particles are independent words, which can form an utterance by themselves (348), and they can be found in different position in the clause (352).

---

[25]The corresponding 'quasi-verbs' or 'modal adjectives' in Sinhala are in a similar relationship to each other (Gair 1967).

(352)   a.   *Boole=pii.*
        can=go.
        '(X) can go' (K081103eli03)

       b.   *Mà-pii boole.*
        INF-go   can
        '(X) can go' (K081103eli03)

Furthermore, they can be used without a verb, as in (353), where *thàrboole* is the predicate of the headless relative clause. The event of the relative clause is the same as the one denoted by the verb in the main clause, *ambel* 'take', but this is not overtly realized in the relative clause since it can be inferred.

(353)   *[Se=dang thàràboole] pada see thàrà-ambel.*
       1s.DAT   cannot   PL   1s   NEG.PAST-take
       'What I couldn't take, I didn't take.' (K051213nar01)

This distinguishes modal particles from both affixes and clitics expressing modality. Those have a fixed preverbal position and cannot be used by themselves to form a complete utterance. Modal particles are grammaticalizing to clitics and affixes, though, as can be seen in (352), where the modal procliticizes to the verb, and can undergo phonological erosion. In the preverbal position, [bɔl] and [bər] are also possible, which is not the case for the postverbal position. *Boole* is thus becoming very much like the other preverbal affixes, a process that has progressed to a lesser extent for the other modal particles. *Thàrboole* 'cannot' can only be used in clause final position, while *thussa* and to a certain extent *maau* can also procliticize to a verb.

All modal particles require the dative (354) (cf Slomanson 2008a). Since the dative is very often used to mark lack of volitionality (Ansaldo 2005b), the use of the dative with modal particles can be explained by the fact that the entity feeling possibility, urge, or desire is normally not wilfully doing so.

(354)   *Kithang=**nang** mà-pii thàràboole.*
       1PL=DAT   INF-go cannot
       'We can't go.' (K051206nar07)

These particles can take nominal (355) or clausal complements (356)(357). If they take a clausal complement, the verb of the complement clause is either in the bare form (356) or in the infinitive (357). If the modal particle is procliticized, only the bare form is possible (358).

(355)   ***Buula*** *maau=le*   ***paasir*** *konnyo maau.*
        flour   want=ADDIT sugar little   want
        'You need flour and you need some sugar.' (B060115rcp02)

(356)   *Ithukapang umma baapa su-biilang [[lorang=**nang** kaaving] **thàrboole**].*
        then   father mother PAST-say 2PL=DAT   marry   cannot
        'Then the parents said that they could not marry.' (K051220nar01)

(357)   *Se=**dang** karang jaau **mà**-pii thàrà-boole.*
        1s=DAT now   far   INF-go cannot.
        'I cannot go far.' (K061120nar01)

(358)   *Kithang=**nang** baaye=nang mulbar bole=∅=baaca.*
        1PL=DAT   good=DAT   Tamil can=read
        'We can read Tamil well.' (K051222nar06)

5.12.1   *boole* 'can'

This particle is used to indicate capacity ⊙→15.7.2.1, p. 652 as well as permission ⊙→15.7.2.2, p. 653 (Smith & Paauw 2006:174). When used as a free word, there are no allomorphs, but when procliticized, all *boole=, bolle=* and *bàr(à)=* can be found (cf. Slomanson 2006:139).

(359)   *[Deram pada=pe ini   sthri pada se-baava   katha] **boole**=biilang.*
        3PL   PL=POSS PROX wife PL   PAST-bring QUOT can=say.
        'It is possible to say that they brought their wives.' (K051206nar08)

Boole is normally used with clausal complements, but it can also be found with nominal complements. In that case, the canonical activity associated with the noun is implied. For instance, *Mlaayu boole* 'Malay can' then asserts that the referent is able to speak Malay, without the verb for speaking (*oomong*) actually being present. The non-realization of the verb is possible because speaking is the canonical activity one performs with a language (as compared to eating, killing, or hitting). There is thus no big chance of misinterpretation on the part of the hearer.

The functions of *boole* are to indicate capacity and permission.

(360) *Itthu=ka aathi=yang sajja hatthu oorang=nang **bole=ambel**.*
      DIST=LOC liver=ACC   only one   man=DAT   can-take
      'Then only one person can take the liver.'

In example (360), it is an ontological fact that only one person can take the intact liver, since animals only have one liver. Two different modal uses of *boole* can be found in the following example.

(361) *[[**Boole**] **oorang** pada]=nang siithu **boole=pii**.*
      can   man   PL=DAT   there can=go.
      'All men who are able to go may go.' (B060115cvs01)

Example (361) shows the use of *boole* without a verb in the relative clause, indicated by brackets. The (omitted) verb in the relative clause is the same as in the main clause *pii* 'go' and can therefore be inferred. At the same time, this sentence shows both the use of *boole* to indicate ability (in the relative clause) and permission (in the main clause). Finally, we can also see the dative marker =*nang* on the overt argument *boole oorang pada* 'the man who can'.

Another example of *boole* being used in a relative clause is (362). Note that there is no special marking on the relative clause. The relativizing meaning is indicated by position only ⊡→13.2, p. 516.

(362) *Deram pada [[baae=nang pìrrang mà-kijja **boole**] oorang].*
      3PL   PL   good=nang war   INF-make can   man
      'They were men who could well make war.' (K051213nar06)

Slomanson (2003) discusses the variable positions of *boole*. He states that the categorical status of *boole* is ambiguous between a verb and an adjective, but chooses to treat *boole* as an independent finite verb. He notes certain similarities to Sinhalese *puluvan*, which he chooses to call *modal adjective*.

5.12.2 *thàr(à)boole* **'cannot'**

This particle is the negation of *boole* and expresses incapacity ⊙→15.7.2.1, p. 652 and interdiction ⊙→15.7.2.2, p. 653. *Thàrboole* can only be used postverbally and requires the infinitive with clausal complements (363)(364). The infinitive is optionally accompanied by the dative marker =*nang* on the verb (365). Since the pronominal argument is also marked by =*nang*, this means that there are two instances of =*nang* in this sentence, one on the pronoun and one on the verb.

(363) *Bannyak **mà**-biilang **thàrboole**.*
    much   INF-say   cannot
    '(I) can't tell you much.' (K051206nar12)

(364) *Se=dang karang jaau **mà**-pii **thàrboole**.*
    1s.DAT   now   far   INF-go cannot.
    'I cannot go far.' (K061120nar01)

(365) *Derang pada=**nang** atthu=le   mà-kijja=**nang thàràboole**=subbath ...*
    3PL   PL=DAT   one=ADDIT INF-do=DAT   cannot=because   ...
    'Because they couldn't do anything, ...' (N060113nar01)

    Thàrboole is used to indicate impossibility, which can be either factual (366) or social (367).

(366) *Daalang=ka light=le   mà-ambel baapi thàràboole.*
    inside=LOC light=ADDIT INF-take   bring cannot.
    'Inside, you cannot take a light, either.' (K051206nar02)

(367) Cigarette *mà-miinong thàràboole.*
    cigarette INF-drink   cannot
    'It is forbidden to smoke.' (K060116nar04)

5.12.3 *maau/(ka)mau(van)* 'want'

This particle is used to indicate desire (Smith & Paauw 2006:172f) ⊙→15.7.2.3, p. 656. The three possible forms are *maau, mauvan* and *kamauvan*, which are derived from the first by the suffix *-an* and the historic circumfix *ka-...-an*. All can be used postverbally (368)-(370). There might be a possibility to use *mau=* preverbally, but elicitation about this topic was inconclusive.

(368) *Baapa=nang mosque=nang mà-pii **maau**.*
    father=DAT   mosque=DAT INF-go want
    'Father, you should go to the mosque.' (B060115nar04)

(369) *Cumma **mauvan** thraanang, incayang=nang fithena asà-thaarek=apa*
　　　idle　　need　　without　3s.POLITE=DAT quarrel cp-pull=after
　　　incayang=nang su-thiikam.
　　　3s.POLITE=DAT　PAST-stab
　　　'Unnecessarily, he was dragged into a quarrel and then stabbed.'
　　　(K051220nar01,K081105eli02)

(370) *Kithang=nang hathu* application *mà-sign **kamauvan** vakthu=nang=jo,*
　　　1PL=DAT　　　INDEF application INF-sign want　　　time=DAT=EMPH
　　　kithang arà-pii　inni politicians *pada dìkkath=nang.*
　　　1PL　　NON.PAST-go PROX politicians PL　vicinity=DAT
　　　'When we want to sign an application, we approach these politicians'
　　　(K051206nar12)

Besides the use as a modal particle, *kamauvan* also has a lexical meaning of 'desire' (371), 'need' (372) or 'valuable' (373).

(371) *[**Kamauvan** su-aada]　see=yang dhaathang* remand=*ka.*
　　　desire　　　PAST-exist 1s=ACC　come　　remand=LOC
　　　'They had the desire that I got remanded.' (K061122nar03)

(372) *Incayang thàràsìggar vakthu incayang=nang duvith **kamauvan** su-aada.*
　　　3s.POLITE sick　　　time 3s.POLITE=DAT money desire　　PAST-exist.
　　　'When he was sick, he had a need for money.' (K060116nar07)

(373) *[Ini　records bannyak **kamauvan** athi-jaadi　katha] kithang arà-iingath.*
　　　PROX records much　　valuable　IRR=become QUOT　1PL　　NON.PAST-think
　　　'We think that these records will become very valuable in the future.'
　　　(N061124sng01)

Like the other modal particles, these three particles govern the dative, and clausal complements are in the infinitive.[26] With nominal complements it means that the argument desires to possess the item or property under discussion (374). With clausal complements, it means that the subject wants to perform the action denoted by the clause (375).

---

[26]Smith & Paauw (2006:172f) report that the infinitive marker can be left out, and that the prefix *mas(th)i* can be used on the verb. Dropping the infinitive marker is sometimes done in the Upcountry, but the combination of *masthi* and *maau* is completely ungrammatical.

(374) *Deran=nang thumpath **maau**.*
       3PL=DAT   place   want.
       'They wanted land.' (N060113nar01)

(375) *Kithang=nang hathu* application *mà-sign **kamauvan** vakthu=nang=jo,*
       1PL=DAT   INDEF application INF-sign want   time=DAT=EMPH
       kithang arà-pii   inni politicians *pada dìkkath=nang.*
       1PL   NON.PAST-go PROX politicians PL   vicinity=DAT
       'When we want to sign an application, we approach these politicians'
       (K051206nar12)

If the complement clause has no subject, it is inferred that the entity mentioned in
the matrix clause wants to perform the action (376) (377), whereas if the person wants
some other body to perform an action, this second body must be overtly mentioned as
in (378).

(376) *Itthusubbath=jo incayang=nang **maau**, ini*   Sri Lankan Malay *mà-blaajar*
       therefore=EMPH 3p.POLITE=DAT   want   PROX Sri Lankan Malay INF-learn
       ***maau**.*
       want
       'This is why he wants to learn this Sri Lanka Malay.' (B060115prs15)

(377) *[Andare kanabisan=nang anà-mintha] [hathu raaja=ke   asà-paake=apa*
       Andare last=DAT   PAST-ask   INDEF king=SIMIL CP-dress=after
       *kampong=nang mà-pii **maau** katha].*
       village=DAT   INF-go want QUOT
       'What Andare wanted as a last wish, was to go to the village dressed up as a
       king.' (K070000wrt03)

(378) ***Kithang**=nang maau kitham=pe mlaayu **lorang** blaajar, lorang=pe mlaayu*
       1PL=DAT   want 1PL=POSS Malay 2PL   learn   2PL=POSS Malay
       ***kitham** blaajar.*
       1PL   learn
       'We want that you learn our [Sri Lankan] Malay, and we learn your
       [Malaysian] Malay.' (K060116nar02)

Note that in (378), the second person becomes topical, and that the first person
must therefore be reintroduced in the final part of the sentence to get the references
right, even if that is normally not necessary if the wanting person and the performing
person are identical. Also note that the infinitive marker *mà-* is missing on the verbs

in the complement clause, which is a particularity of this speaker.

Constructions with *maau* can also be used for mild commands (379) ⊙→17.1.5, p. 699. A particular use of *maau* is found in the genre of recipes, where it is used to indicate the needed ingredients (380).

(379) *Baapa=nang* mosque=*nang mà-pii **maau**.
        father=DAT   mosque=DAT   INF-go want
        'Father, you should go to the mosque.' (B060115nar04)

(380) *Manis-an=nang   mà-thaaro guula **maau**, gula   paasir konnyong **maau**.*
        sweet-NMLZR=DAT INF-put   sugar want sugar sand few   want
        'In order to make it sweet, you need sugar, you need a bit of crystal sugar.'
        (B060115rcp02)

These particles are reasonably common and seem to occur in about equal proportions.

The negation of *(ka)mauvan* is *thàr(ka)mauvan*. This modal particle can only be used postverbally. The negation of *maau* is *thussa*, to be discussed in the next section.

(381) *Go   size   mà-ambel **thàràmauvan**.*
        1s.FAMILIAR measure INF-take   NEG.want.
        'I did not (even) need not take measures [for tailoring].' (B060115nar04)

5.12.4 *thussa* 'want.not'

This particle is the negation of *maau*, which is also often used for negative imperatives. This particle can be used both before (382) and after the verb (383). As a difference to the other modal particles, the infinitive can also be used when *thussa* is in preverbal position (382). Some speakers also have a procliticized allomorph *thus=*.

(382) *'**Thussa** mà-thaakuth', Buruan su-biilang.*
        neg.IMP INF-fear   bear   PAST-say
        ' "Don't be afraid," said Bear./"I don't want you to be afraid," said Bear '
        (K070000wrt04)

Thussa can also be used with clauses in which the actor is different from the wanter (383).

(383) *Se=dang lorang ini   buuva mà-picca-kang=nang thussa.*
     1s.DAT   2PL   PROX fruit   INF-broken-CAUS=DAT   NEG.want
     'I do not want you to pluck this fruit.' (K081103eli03)

Just like *maau, thussa* can be used with verbal (382), nominal (384) and with clausal complements (383). Additionally, *thussa* can be used on its own as a command, similar to the English *Don't!*

(384) *Se=dang **paayong**=yang **thussa**.*
     1s=DAT   umbrella=DAT   NEG.want
     'I don't want an umbrella, thank you.' (K081103eli03)

The function of *thussa* is to express lack of desire ⊙→15.7.2.3, p. 657. It can also be used for interdictions ⊙→15.7.2.2, p. 653. Adelaar (1991) gives the etymon as **tra usa* (also see Saldin 2001:60).

5.13   Negative particles

There are two negative particles, *thraa* ⊡→5.13.1, p. 254 and *bukang* ⊡→5.13.2, p. 258. The latter is used for negating nominal predicates, and for constituent negation. The former is used for negating adjectival and locational predicates, and verbal predicates in the perfect tense.[27][28]

5.13.1   *thraa*

Thraa is the negation of *aada*, (< *tidak aada* (cf. Adelaar 1991:26)). There is only one form. This form can be used for giving a negative answer as in (385) ⊙→17.1.6, p. 701, where it is used to decline an offer, in this case in reported speech.

(385) *Ithu=kapang derang nya-biilang '**thraa**, kithang giithu   thama-pii.'*
     DIST-when   3PL   PAST-say   no   1PL   like.that NEG.NONPAST-go
     'Then they said "no, we cannot go like that." ' (K051213nar06)

The following examples show more instances of *thraa* in reported speech.

---

[27]Verbal negation is other tenses is not done by a particle, but by the prefixes *thàrà-* ⊡→6.2.2.4, p. 297 for past and *thama-* ⊡→6.2.2.5, p. 298 for nonpast.

[28]Sinhala also has two negators, with similar uses (Gair & Paolillo 1988). Continental Tamil is normally described as having only one (*illai*), but in Sri Lanka, an additional constituent negator *alla* can be found (Gair & Suseendirarajah 1981).

(386) *Oorang pada su-biilang: 'thraa, thraa, incayang* saint*, incayang avuliya*
      man   PL   PAST-say   NEG   NEG   3s.POLITE saint, 3s.POLITE saint
      hatthu jaadi   aada'
      INDEF become exist.
      'The men said "No, no, he is a saint, he has become a saint." ' (K051220nar01)

(387) *'Thraa thraa inni=yang masà-picca-kang' katha biilang.*
      NEG   NEG   PROX   must-broken-CAUS QUOT say
      ' "No, no," he said "You must pick these ones." ' (K051220nar01)

Additionally, *thraa* can be used in all the contexts where *aada* could be used in affirmative contexts ⊙→15.12, p. 671. It is always found in final position then. These contexts are

- existential. Negation of both concrete and abstract existence is done with *thraa*

- locationals. Since locationals are a subtype of existentials, it is little surprising that they are also negated by *thraa*.

- possession. Predicative possession is also negated by *thraa*.

- perfect tense of verbal predications. The negation of this tense is formed by changing *aada* to *thraa*.

The first function is the negation of the existential, i.e. predicating that a certain entity does not exist, like *bìrras* 'rice' in (388).

(388) *Bìrras thraa.*
      raw.rice NEG
      'There is no rice.' (K051206nar12)

Another example is (389), where the fact is conveyed that the mother does not live anymore and is thus not among the things of which one can say that they exist.

(389) *Se=ppe umma thraa.*
      1s=POSS mother NEG
      'My mother is dead.' (B060115prs03)

The following example also shows a use of *thraa* which does not predicate complete absence of *bedahan* 'difference', but rather the absence of a lot of differences. Some differences remain, but these are a minority.

(390) *Punnu bedahan **thraa.***
       many   difference NEG
       'There are not many differences.' (K060108nar02)

The entity whose non-existence is predicated need not be overtly present, but can be one which is inferable from context. In (391), the entity whose absence *thraa* predicates is not overtly realized.

(391) *Kàthaama su-aada, karang ∅ **thraa**.*
       before   PAST-exist now   NEG
       'Before (there) was [rugby], (but) not now.' (B060115cvs01)

Absence of an entity from a certain location, rather then general absence as discussed above, is also indicated by *thraa*, the second use indicated above. The semantic difference between complete absence and absence from a certain place is mirrored in grammar by the possibility to use *duuduk* with animate entities in the latter case but not in the former. In example (392), the fact that there are neither cattle nor goats found in Saudi-Arabia is indicated by *thraa*. This does not entail that these animals do not exist at all, it merely entails that they are not found in Saudi-Arabia.

(392) *Saudika   ontha, **samping kambing thraa**.*
       Saudi.Arabia camel, cattle   goat   NEG
       'In Saudi Arabia (they offer) camels, there are no cattle or goats.'
       (K060112nar01)

In the following examples, the absence of the Malay language in a certain place and the imputed absence of grammar in the language of the speaker are indicated by *thraa*. In both cases, the domain from where the entity is absent is indicated by the locative postposition =*ka*, attached to a deictic referring to a concrete place in (393) and to a noun referring to the abstract concept *bahasa* 'language' in (394).

(393) ***Itthu=ka mlaayu thraa**, bannyak=nang English=jo   aada.*
       dist=LOC Malay NEG   much=DAT   English=EMPH exist
       'There is no Malay over there, it is all English which is there.' (B060115prs15)

(394) *Kitham pada=pe **bahasa=ka**   ini **grammar thraa**.*
       1PL   PL=POSS language=LOC PROX grammar NEG
       'There is no grammar in our language.' (G051222nar02)

Thraa in these uses is underspecified for tense and can be used with past tense reference, as the following example shows, where *itthu muusingka* 'back in those

days' clearly indicates that we are dealing with reference to the past, but this is not reflected in the behaviour of *thraa*, there is no prefix or other device indicating past tense on the predicate.

(395) *Itthu muusing=ka cinggala **thraa**.*
       DIST time=LOC   Sinhalese NEG
       'At that time there was no Sinhala.' (K051222nar06)

The third use of *thraa* is for negating possession. Just as with affirmative possession with *aada* ⊡→5.1.2.1, p. 165, the possessor is indicated with a postposition. This is normally the dative =*nang* as in (396) or (397) or its allomorph =*dang* used on pronouns as in (398).

(396) *Ini   aanak pada=**nang** time hatthu **thraa**.*
       prox child PL=DAT   time INDEF NEG
       'These children do not have any time.' (G051222nar01)

(397) *Derang pada=nang [itthu mà-kumpul athu mosthor] **thraa**.*
       3PL   PL=DAT   DIST INF-add INDEF way   NEG
       'They have no way to colllect all this.' (G051222nar01)

(398) *Suda **go=dang**   hatthu kurang-an **thraa**.*
       SO   1S.FAMILIAR INDEF few=NMLZR NEG
       'So, I am not missing anything.' (B060115nar04)

The fourth use of *thraa* is in negated predicates in the perfect. In this construction like in the ones mentioned before, *thraa* takes the position of its affirmative counterpart *aada*.

(399) *Itthu=nang blaakang kithang=pe hatthu oorang=le minister **jaadi**   **thraa**.*
       dist=DAT   after   1PL=POSS INDEF man=ADDIT minister become NEG
       'After that, not one of our men has become minister again.' (N061124sng01)

(400) *[Kithang baaye mlaayu arà-oomong   katha incayang] **biilang thraa**.*
       1p   good Malay NON.PAST-speak QUOT 3s.POLITE say   NEG
       'He has not said that we speak good Malay.' (B060115prs15)

(401) Invitations *daapath **thraa**.*
       invitations get   NEG
       'The invitations had not arrived.' (K060116nar23)

Sometimes, *thraa* can also be found when negating adjectives (cf. Slomanson 2006:137).

(402)   *Itthu muusing gampang **thraa**.*
         DIST time   easy   NEG
         'It was not easy back then.' (B060115nar05)

In this function, *thraa* competes with *thàrà-*. The exact conditions which trigger one or the other adjectival negation are unclear. There seems to be some variation between speakers, and even within idiolects, so that this is difficult to test. This is discussed in more detail under adjectives ⊡→10.6, p. 477.

Thraa also has a a verbal reading meaning 'disappear, become unavailable' (403). When combined with =*nang* it means 'without' (404).

(403)   *Bannyak kithang=nang **su-thraa**ᵥ.*
         much   1PL=DAT   PAST-NEG
         'We lost many.' (B060115nar02)

(404)   *Cumma mauvan **thraanang**, incayang=nang fithena asà-thaarek=apa*
         idle   need   without   3s.POLITE=DAT quarrel CP-pull=after
         incayang=nang su-thiikam.
         3s.POLITE=DAT   PAST-stab
         'Without need, he was dragged into a quarrel and then stabbed.'
         (K051220nar01,K081105eli02)

Thraa can also be the source of a causative derivation, as shown in the following example.

(405)   *Itthu=nang kithang arà-**thraa-king**   kithang=pe Seelon=pe*
         DIST=DAT   1PL   NON.PAST-NEG-CAUS 1PL=POSS   Ceylon=POSS
         mosthor=nang.
         manner=DAT
         'Therefore we are losing our Ceylonese ways.' (B060115rcp01)

5.13.2   *bukang*

Bukang is used for the negation of nominal predicates ⊙→10.4, p. 472 and for the negation of constituents in a predication ⊙→15.12, p. 671. *bukang* occurs after the constituent it negates.

The following examples show the use of *bukang* as negation of a nominal ascriptive predicate. (406) shows the use on a simple predication, while in (407) we have

an affirmative ascription *hatthu Muslim*, followed by a negative ascription *mlaayu bukang*.

(406)   *Deram Islam oorang **bukang**.*
       3PL   Islam man   NEG.NONV
       'They were not Muslims.' (K051213nar03)

(407)   *Sindbad* the Sailor *hatthu Muslim, mlaayu **bukang**.*
       Sindbad the Sailor INDEF   Muslim, Malay NEG.NONV
       'Sindbad the sailor was a Moor, he was not a Malay.' (K060103nar01)

Negation of nominal predicates is the most frequent function *bukang* occurs in, but it can also be used for constituent negation, as example (408) shows.

(408)   *Thaangang=dering **bukang**   kaaki=dering masà-maayeng.*
       hand=instr   NEG.NONV foot=instr   must-play
       'You must not play with the hands, but with the feet.' (N060113nar05)

In this example, the predicate is *maayeng* 'play', but it is not the predicate which is being negated, but the adjunct giving information about the entitiy which participates in the act of playing as an instrument. This entity happens to be not the hand, as could be expected, but the foot. Hence it is the argument, realized as the constituent *kaakidering* 'with feet', which is negated, and *bukang* as the constituent negator must be used. If the verbal negative prefix *thàrà-* was used on *maayeng* 'play', it would be the verb which would be negated, which does not correspond to the information structure intended here. It is not the act of playing which is negated, but the instrument of playing, the hand.

A final use of *bukang* is to request confirmation, like a question tag as in (409). This example is a nice contrast to the use of *bukang* in (406), where almost the same content is present, but used in a negation context, whereas we are dealing with a tag question here.

(409)   *Deram pada* Moors*, **bukang**?*
       3PL   PL   Moors, TAG
       'The are Moors, aren't they?' (K051206nar12)

Bukang has a transparent etymology and is cognate to Std. Malay *bukan*. It is frequently found as a negator of nominal ascriptive predicates and as a question tag, but is very infrequently found in the constituent negation context, which might have to do with the fact that the communicative need for constituent negation arises less often.

5.14   Other particles

In this section I will discuss the remainder of SLM free morphemes, which do not fit into any other class.

5.14.1   *suuda* 'enough'

This is a particle which can combine with nouns and means that there is a sufficient quantity available (410). It can thus be glossed as 'enough'. It has to be distinguished from the adverb *suda* with a short vowel, which means 'thus, then, so', which is given for comparison in (411).

(410)   Kithang=*nang bìrras*   **suuda**.
        1PL=DAT   raw.rice enough.
        'We have enough rice.' (K081103eli04)

(411)   **Suda** *[puthri=le   biini=le   arà-caaňda aari]=le   su-dhaathang.*
        thus   queen=ADDIT wife=ADDIT SIMULT-meet day=ADDIT PAST-come
        'Then came also the day when the queen and the wife were to meet.'
        (K070000wrt05)

   The use of a particle for this function parallels Sinhala *æti* and Tamil *podum*.

5.14.2   *kalu* 'if'

This particle is used to mark conditionals ⊙→15.8, p. 660. It has a free form, which occurs after the predicate (412)(413). Additionally, there is a procliticized form *kal(a)=* which occurs only preverbally and blocks the use of other TAM-marking on the verb (414)-(417) (Slomanson 2006:153). Smith et al. (2004) also note the form *ka(ng)-*, which is not found in the corpus.

(412)   *Nnam thullor ar-**ambel**   **kalu**.*
        six   eggs   NON.PAST-take if
        'Suppose you take six eggs.' (B060115rcp02)

(413)   *Lai   aapa=ke   **aada kalu**, se=dang aathi-ka   asà-kluuling bole=caari*
        other what=SIMIL exist if   1s.DAT heart=LOC CP-roam   can-search
        kaasi Malays pada=pe atthas.
        give Malays PL=POSS about
        'If there should be something else, I am going to go through my mind, and I will forward it to you' (G051222nar01)

(414)   a. *Lorang se=dang mà-iidop thumpath **kala=kaasi**,*
            2PL   1s.DAT   INF-stay   place   if-give
            'If you give me a place to stay,'

        b. *see lorang=nang lorang=pe samma duvith=le   baarang pada=le*
            1s 2PL=DAT   2PL=POSS all   money=ADDIT goods   PL=ADDIT
            anthi-bale-king.
            IRR=turn-CAUS
            'I will return all your money and the goods to you.' (K070000wrt04)

(415)   a. *See lorang=nang thama=sakith-kang,*
            1s 2PL=DAT   NEG.NONPAST=pain-CAUS
            'I will do you no harm,'

        b. *lorang see=yang diinging=dering **kala=aapith**.*
            2PL   1s=ACC cold=ABL   if-protect
            'if you give me shelter from the cold.' (K070000wrt04)

(416)   *Go=ppe naama Badulla buulath thaau, lorang Mr. Mahamud katha*
        1s=POSS name Badulla whole know 2PL   Mr Mahamud QUOT
        ***kala=biilang**.*
        if=say
        'Whole Badulla knows my name if you say Mr Mahamud' (B060115nar04)

(417)   Important occasion *pada **kala=aada** aapa=ke*   festival *pada laayeng*
        important occasion PL   if-exist   what=SIMIL festival PL   other
        wedding ceremony *[. . .] ithu vakthu kithang arà-kirja*
        wedding ceremony   DIST time 1PL   NON.PAST-make.
        'If there is an important occasion, some festival or otherwise a wedding cere-
        mony, we will prepare it.' (K061026rcp01)

The utterances (414) and (415) also show the use of the irrealis markers *anthi-* (affirmative) and *thama-* (negative) in the apodosis (i.e. the sentence containing the consequence). Those two utterances also show that the order of protasis and apodosis is free.

In rare cases, *kal-* and another TAM-(quasi-)prefix can be found combined, as in (418)(419). The conditions for this are unclear as of now.

(418) Wicket=*yang* **kal**=*su-picca,*   out.
      wicket=ACC   if-PAST-fall-broken out.
      'When the wicket has fallen, you are out.' (K051201nar02)

(419) *Lorang=nang duvith kal=su-daapath, kitham=pe banthuan*
      2PL=DAT   money if-PAST-get   1PL=POSS help-NMLZR
      thàrà-kamauvan.
      NEG-need.
      'When you will have got money, you will not need our help (any more).'
      (K081103eli04)

One context in which the stacking of *kal-* with another verbal prefix is possible is the expression of negation (420).

(420) *Lorang=nang duvith kal=thàrà-daapath, kithang anthi-banthu.*
      2PL=DAT   money if-NEG.PAST-get   1PL   IRR-help.
      'Should you get no money, we would help.' (K081103eli04)

However, this combination of *kal-* with another verbal prefix is not always possible. *Anthi-* in (421) cannot be combined with *kal-*; the postverbal form *kalu* has to be used instead.

(421) *Lorang=nang duvith anthi-daapath kalu, kithang svaara thraa=nang*
      2PL=DAT   money IRR-get   if   1PL   noise NEG=DAT
      anthi-duuduk.
      IRR-stay.
      'If you get your money, we will be finally left in peace.' (K081103eli04)

The different combinatorial possibilities of various TAM prefixes and *kal-/kalu*, as well as the resulting semantics are in need of further research.

Kalu can not only be used with verbs, it can also combine with (proper) nouns (422)(423)(424) and pronouns (425) or deictics (426).

(422) *Hatthu samping hatthu oorang bolle=kaasi;* **kumpulang kalu***, thuuju*
      one   cow   one   man   can=give; party   cond seven
      oorang.
      men.
      'One man can sacrifice one cow; when people are in a party, seven men are
      required.' (K060112nar01,K081105eli02)

(423) *See, **pukurjan**=nang **kalu**, thama-pii   ruuma pukurjan asà-kirja ambel*
      1s   work=DAT   if   NEG.NONPAST-go house work   CP-make take
      ruuma=ka arà-duuduk.
      house=LOC NON.PAST-stay.
      'As for work, I don't go to work, I do housework and stay at home.'
      (B060115prs04)

(424) ***Galle**=ka **kalu**, se=ppe cousin brothers pada=samma see anà-jaalang*
      Galle=LOC if   1s=POSS cousin brothers PL=COMIT   see PAST-walk
      skuul=nang.
      school=DAT
      'As for when I was in Galle, I walked to school with my cousins.'
      (K051201nar02)

(425) ***Incayang kalu**, duuva skali asà-dhaathang baaye=nang asà-practice*
      3s.POLITE if   two   time CP-come   good=DAT   CP-practice
      baaye=nang arà-oomong.
      good=DAT   NON.PAST-speak
      'As for him, he has come twice and practiced well and (now) speaks well.'
      (K061030mix01)

(426) ***Itthu kalu**, [. . .] Dubai=ka asà-duuduk laama kar pada kitham*
      DIST if   [. . .] Dubai-LOC CP-stay   old   car PL   1PL
      arà-baapi   Iraq ports=dang.
      NON.PAST-take Iraq ports=DAT
      'In that case, we bring old cars from Dubai to the ports in Iraq.'
      (K051206nar19)

In these contexts, it is not a real conditional meaning that emerges, but rather a reading of 'given X, Y is likely/required'. Like this, the translations above can be paraphrased as 'Given a party, it will be seven people', 'Given work, I do not go there', 'Given my staying in Galle, I walked to school' and 'Given that state of affairs, we bring them from Dubai to Iraq.' The historically related Indonesian form *kalau* is used in similar contexts (Ewing 2005:242f).

This use is often found when attributing opinions to people. In example (427)(428), *kalu* refers to the speaker. This use cannot be conditional in the strict sense, since the speaker is not apt to being there or not. She is always there when making an utterance. *Kalu* in this case rather translates to something like 'as for me.'

(427) ***Se=dang kalu*** *suda bannyak thàrà-thaau inni=pe   atthas*
        1s=DAT   if   thus   much   NEG-know   PROX=POSS about
        mà-biilang=nang.
        INF-say=DAT.
        'So, as for me, I cannot tell you much about this.' (K051205nar04)

(428) ***Se=dang kalu*** *bannyak mà-biilang thàrà-thaau sdiikith see athi-biilang.*
        1s.DAT   if   much   INF-say   NEG-know   few   1s IRR-say
        'As for me, I cannot tell you a lot, I will tell you a little.' (K051205nar02)

This use indicating topicality can also be found with preverbal allomorphs of *kalu*, as in (429).

(429) *Karang se=ppe* *family=yang* ***kal**-ambel,* now *se=ppe* *naama Thuan Kabir*
        Now   1s=POSS family=ACC   if-take   now 1s=POSS name   Thuan Kabir
        Samath.
        Samath
        'Now, as far as my family is concerned, my name is Thuan Kabir Samath.'
        (N060113nar03)

Furthermore, *kalu* can also be combined with modal particles as in (430).

(430) *Lorang=nang* ***kala**-boole, kithang=nang hathu* camera *baa.*
        2s.POLITE=DAT if-can   1PL=DAT   INDEF camera bring
        'Bring us a camera if you can.' (Letter 26.06.2007)

The border between conditional and temporal meaning of *kalu* is often blurred, as the following example shows:

(431) *Laskalli*   ***ka**-dhaathang, mari Badulla=nang.*
        other.time if-come   come Badulla=DAT
        'Come to Badulla next time.' (B060115cvs17)

In this example, *ka-* could have a temporal reading referring to a point of time in the future when the addressee will return to Badulla, or it could have a conditional meaning making a conditional request which is only realized in the event that the addressee actually does return to Badulla.

The boundary is further weakened by the fact that one allomorph of the temporal prefix *kapang* 'when' is *kam-* ⊡→6.2.1.7, p. 283, which is very similar to *kal(u)*, so that this leads to further confusion (Slomanson 2008b). In the following example, the speaker refers to a past situation, which cannot be interpreted in terms of conditional

truth values, because all events have already occurred. Normally, *kam-* would be used in this context, but in this example we find *kal-*, a further indication that the border between temporal and conditional is fuzzy. This seems to be true especially of the preverbal use. Temporal use of postverbal *kalu* is not attested.

(432) *Sebastian mlaayu mà-oomong katha **kala**=biilang, butthul asà-suuka*
Sebastian Malay ɪɴF-oomon QUOT cond-say   very   PAST-like
su-jaadi.
PAST-become.
'When (you) told (us) that Sebastian spoke Malay, we became very happy.'
(B060115cvs01)

When combined with present tense predicates, *kalu* can have a counterfactual (433) or a realis meaning (434), when combined with perfect predicates, only the counterfactual reading seems possible (435).[29]

(433) *Se=dang saayap kala-aada, bole=thìrbang.*
1s.DAT   wing   if-exist   can=fly.
'If I had wings I could fly.' (K08114eli01)

(434) *Lorang asà-caape **kala=blaajar**, lorang=nang A/L bole=pass.*
2PL   CP-tired   if-learn   2PL=DAT   A/L can=pass
'If you study hard, you will be able to pass the A/L exams.' (K08114eli01)

(435) *Sampi! Luu   maalas! A/L thàrà-pass.   Lorang baae=nang*
cow   2s.FAMILIAR lazy   A/L ɴEG.PAST-pass 2PL   good=DAT
*asà-caape **asà**-blaajar **kala=aada**, mà-pass=nang su-aada*
CP-tired   CP-learn   if-exist   ɪɴF-pass=DAT PAST-exist
'Lazybones! If you had studied harder, you would have passed your A/L exams.' (K08114eli01)

De Silva Jayasuriya (2002) treats *kalu* as a suffix in her theoretical discussion, but in the examples she gives it is written as an independent word, so that the interpretation as particle seems intended.[30]

---

[29]Jayawardena-Moser (2004:46) states that in Sinhala, there is no difference between a conditional involving a verb in the perfect tense, and a conditional involving a present tense verb. It is unclear, whether in SLM, we are dealing with a similar situation or not.

[30]The preverbal and negative forms are not discussed.

5.14.3 *kalthra* 'if.not'

This particle is the negation of *kalu* used on NPs, and is used to construct negative conditionals of hypothesized unavailability (436) ☉→15.8, p. 660. It can only be used after the predicate.

(436) *Lorang=ka duvith (*aada) **kal=thra**, kithang anthi-banthu.*
      2PL=LOC   money exist   NEG=if,   1PL   IRR-help.
      'If you have no money available, we will help you.' (K081103eli04)

As given above in (436), *kalthraa* can be transparently segmented into *kal=thraa* "if=neg". This is less the case when used as a linker on discourse level, when it is used without a proposition. In this case, it is best glossed by 'otherwise'(437)(438). Also cooccurrence with *giithu* 'like.that' are common (439). This also has the meaning of *otherwise*. Given the lexicalized meaning of *kalthraa* in this function, it is treated as an particle in its own right here, instead of an analysis as *kal-thraa* "if-NEG".

(437) ***Kalthraa**, kithang arà-baapi*   vehicles*.*
      otherwise 1PL   NON.PAST-take vehicles
      'Otherwise, we take vehicles.' (K051206nar19)

(438) *Dee maana aari=le   asà-dhaathang, thìngaari vakthu=nang, **kalthraa***
      3   which day=ADDIT CP-come   noon   time=DAT   otherwise
      maalang vakthu=nang.
      night   time=DAT
      'He came every day, at noon or otherwise during the night and (attacked).'
      (K051206nar02)

(439) *Suda itthu=dering=jo kithang=nang ini   Indonesia=pe   oorang=si*
      thus DIST=ABL=EMPH 1PL=DAT   PROX Indonesia=POSS man=disj
      ***giithu   kalthraa** Malaysian oorang=si katha bàrà=thaau ambel.*
      that.way otherwise Malaysian man=disj QUOT can-know   take
      'So with that we can come to know whether someone is Indonesian or otherwise Malaysian' (K060108nar02,K081105eli02)

The use of *kalthra* standing on its own is reasonably common, while the use in negative conditionals is rare.

5.14.4   imperative *mari*

This particle is used for commands ☉→17.1.5, p. 699. When used with a verb, the command is to perform the action denoted by that verb (440)(441), when used on its

own, the action of coming is implied. The goal of coming can either be preverbal (442) or postverbal (443).

(440) *Mari kuthumu!*
come see
'Look!' (K051220nar02)

(441) ***Mari*** *maakang.*
come eat
'Eat!' (B060115rcp02)

(442) *Kithang=pe ruuma pada=nang **mari**-∅.*
1PL=POSS house PL=DAT come.IMP
'Come to our houses!' (B060115cvs08)

(443) *Laskalli   ka-dhaathang, **mari** ∅ Badulla=nang.*
other.time when-come   come Badulla=DAT
'Come to Badulla on your next visit!' (B060115cvs17)

Both uses are shown in (444). The first *mari* implies the meaning 'come', while the second one indicates imperative mood for the verb *oomong* 'talk'.

(444) *Mari   darling, mari oomong.*
come.IMP darling IMP   talk
'Come darling, come and speak' (K060108nar01)

It is possible to combine *mari* with the imperative suffix *-la* as in the following example.

(445) a. *Saayang se=ppe thuan **mari**   laari-**la**.*
love   1s=POSS sir   come.IMP run-IMP
'Come running my beloved gentleman.'
   b. *See=samma kumpul **mari**   thaaňdak-**la**.*
1s=COMIT   gather come.IMP dance-imp.
'Come and dance with me.' (N061124sng01)

The addressee who is asked to perform the action can be overtly specified, as in the following two examples.

(446) *Derang=nang nya-biilang raaja:* **lorang mari***.*
       3=DAT   PAST-say   king   2PL   come.IMP
       'The king said to them "you come!".' (K051213nar06)

(447) a. *Lorang pada baaye pìddang=dering, baaye=nang pìrrang mà-gijja*
           2PL   PL   good sword=ABL   good-DAT   war   INF-make
           paaňde oorang pada.
           brave   man   PL
           'You are good with the sword, you are brave men who know how to wage
           war.'
       b. **Lorang mari** *siini kithang=nang.*
           2PL   come here 1PL=DAT
           'Join us.' (K051213nar06)

5.15   Conjunctions

There are no conjunctions. The role that conjunctions fulfill in other languages is
taken over by postpositional clitics that attach to a declausal NP ⊡→6.4.4, p. 327.
Coordination is accomplished by clitics ⊡→6.4.6, p. 359.

5.16   Classifiers

There are no classifiers, contrary to many other western Austronesian languages (Him-
melmann 2005:173). The only remnant of historical classifiers might be the expres-
sion *anak buva* 'child-fruit=children'.

Chapter 6

Bound morphemes

Sri Lanka Malay is a language with comparatively little affixation,[1] but regular use of clitics.[2] The distinction between clitics and affixes on the one hand and free words on the other is not always trivial, so the criteria for treating a certain morpheme as an affix, or a clitic, or a free word shall be explicated briefly ⊡→6.1, p. 269.[3] Based on these definitions, affixes ⊡→6.2, p. 270, simple clitics ⊡→6.3, p. 316, and bound words ⊡→6.4, p. 317 will be discussed. The final section of this chapter ⊡→6.5, p. 396 squares the presentation based on phonological criteria with the traditional notions of nominal and verbal morphology.

6.1 Defitions of different types of bound morphemes

In SLM, five classes of items can be distinguished (Table 6.1):

Free words are phonologically independent, i.e. they can occur without any additional material and can form a phonological word ω (Nespor & Vogel 1986) on their own. In SLM, they can furthermore also form an utterance on their own, e.g. when answering a question. Some free words have variants which are phonologically not independent and have to attach to some other material. These are *simple clitics*. SLM modal particles like *boole* 'can' have simple clitic variants ([bər]). Some other phonologically not independent words lack a corresponding free form. These are *bound words*. In SLM, these are postpositions, Coordinating Clitics, the emphatic clitic, the indefiniteness marker and the plural marker. These attach to their host on the level of syntax. Morphological information like the word class of the host they attach to is invisible to them. This distinguishes them from *affixes*. Affixes are selective with regard

---

[1]Most of the inherited morphology has been lost, cf. Adelaar (1991).

[2]The SLM tendency towards bound morphology for TAM as compared to other Malayic varieties is noted by Ansaldo (2005b).

[3]Terminology and criteria are based on Zwicky (1977), Zwicky & Pullum (1983)

	phonologically independent	free variant	selectivity for hosts	X=*jo*	example
free word	+	n/a	n/a	+	*ruuma* 'house'
simple clitic	–	+	+	+	*boole* 'can'
bound word	–	–	–	+	*=pe* 'poss'
true affix	–	–	+	–	*mà-* 'inf'
quasi affix	–	–	+	+	*anthi÷* 'irr'

Table 6.1: Defining features of clitics

to the word class of the hosts they attach to. Affixes furthermore do not tolerate that non-affixal material intervene between them and their host. In SLM, there are a number of verbal inflectional affixes, one adjectival prefix, and some more derivational affixes. Finally, there are some morphemes which behave exactly like affixes, with the exception that there is one morpheme which may, under restricted circumstances, separate them from their host: the emphatic marker *jo*. These morphemes are hence called *quasi-prefixes*. They encode TAM and attach only to verbs (and converted adjectives). Affixes are indicated by a hyphen (-), simple clitics and bound words are indicated by two superposed hyphens (=). Quasi-affixes are indicated by a division sign (÷) in this chapter, and by a simple hyphen elsewhere. The reason for the avoidance of the division sign as a general marker of quasi-affixation is that it was found to interfere with readability. This must be accepted when the particular phonological status of these morphemes is discussed. When the phonological status is irrelevant, as is the case for all examples outside this section, there is no reason to forgo readability. For all practical purposes, affixes and quasi-affixes can be treated alike, which is why the hyphen was chosen as a general marker for these morphemes.

We will discuss these morphemes in order of decreasing boundedness, i.e. starting with affixes and reaching bound words via the middle stages of quasi-prefixes and simple clitics.

6.2   Affixes

Affixes are the grammatical elements integrated most closely with their host. It is impossible to separate them from their host or to move them around. They have no free variants. Furthermore, every affix is selective with regard to its host. There are affixes which only combine with verbs, others which only combine with numerals etc.

The most important category of affixes in SLM are verbal affixes expressing TAM ⊡→6.2.1, p. 271. A related category are the quasi-prefixes ⊡→6.2.2, p. 285. The mapping of different TAM categories on these morphemes is discussed in ⊡→6.2.3, p. 302. Besides that, two inflectional suffixes ⊡→6.2.4, p. 302 and four derivational affixes

⊞→6.2.5, p. 304 and some affixes used for compound numbers ⊞→6.2.6, p. 315.

6.2.1 Prefixes

There are seven inflectional prefixes in SLM, and two suffixes. Of the prefixes, *mà-* and *jamà-* can only occur in subordinate clauses while the others can also occur in main clauses.

The phonological status of the affixes is characterized by the absence of an independent form and the impossibility to make a pause between them and the stem they attach to. The affixes can be used with any verb or adjective, and cannot be used with any other word class. An exception is the superlative marker *anà-*, which can only attach to adjectives. All of them are used in any variety of the language with high frequency, with the exception of *kànà-* and *-de*. Stacking (i.e. adding more than one prefix to a verb) is not possible (cf. Slomanson 2006:144).

6.2.1.1   past tense *su-*

This prefix is used to mark events as occurring in the past (Smith & Paauw 2006:166) ⊙→15.5.2.1, p. 616. Alternative realizations are *sa-, sà-* or even *s-*, which is easily confused with a homonymous allomorph of the conjunctive participle *asà-* (see below) (cf. Slomanson 2006:137). The following examples show the different allomorphs of *su-*.[4]

(1) *Derang laayeng nigiri   pada   nang   s[ʊ]-pii.*
     3PL   other   country PL=DAT PAST-go
     'They went abroad.' (K051222nar06)

(2) *Umma=le   s[a]-mnniiňǧgal.*
     mother=ADDIT PAST-die
     'My mother also died.' (K061120nar01)

(3) *Baapa=le   s[ə]-mniiňǧgal.*
     father=ADDIT PAST-die
     'My father also died.' (K061120nar01)

---

[4] Adelaar (1991:31) lists *suda* as an allomorph ('full form' in his words) of *su-*, but this form is not used as a prefix in the variety described here. There are other uses of *suda*, which are described in ⊞→5.4, p. 214 and ⊞→5.14.1, p. 260. Ansaldo (2009b) has *si-*, which is not found with this function in the corpus either. Ansaldo furthermore conflates past and perfective (which is called conjunctive participle in this book); these two morphemes are clearly distinct in the Upcountry.

(4) *Karang inni hamma s-abis,   bukang.*
     now   PROX all   PAST-finish, TAG
     'Now all was finished, wasn't it?' (K060116nar11)

The examples above show the use of *su-* on verbs. The following two examples
show the use of *su-* on a converted adjective.

(5) *Kumpulan su-mampus_{ADJ}.*
     association PAST-dead
     'The association became defunct.' (K060116nar01)

(6) *Kumpulan pada bannyak se-kuurang_{ADJ}.*
     association PL   much   PAST-few
     'The associations shrank heavily.' (K060116nar01)

Su- is a general past marker. It can refer to any point before the speech situation,
be it distant as in (7) or close as in (8). There is thus no semantic specialization like
in the English, where the past tense is used for remote events and the perfect tense for
events having a bearing on the speech situation. This is reflected in the translation of
these examples, which give both English forms for the sake of illustration.

(7) *Suda **1994=ka** se=ppe husband **su-niiňǧgal**.*
     thus 1994=LOC 1s=POSS husband PAST-die
     'So my husband (has) died in 1994.' (K051201nar01)

(8) *Suda buthul suuka **nyaari** siini **su**-dhaathang=nang.*
     thus   correct like   today   here PAST-come=DAT
     'So I appreciate a lot that you came/have come here today.' (G051222nar01)

6.2.1.2   past tense *anà-*

This prefix indicates past tense ⊙→15.5.2.1, p. 616. It is very difficult to find semantic
differences between *anà-* and *su-* (see below). *Anà-* is the morpheme with the most
extensive allomorphy in SLM. The full form of *anà-* is seldom heard. Normally,
phonetic reduction to a monosyllabic affix takes place. In Kandy, the predominant
realization seems to be *nà-* or *nyà-*, whereas in the higher regions of the upcountry,
eN- is heard more often, where N is a nasal homorganic to the following stop, or [n] in
the case that the next phoneme is a vowel. In the south finally the canonical realization
of this morpheme seems to be *e-*, regardless of environment (Ansaldo 2009b). The

total array of attested forms is *anà-, nà-, nyà-, em-, en-, eng-, n-, m-, ny-, e-.*[5]

When used in main clauses, *anà-* indicates the simple past and is interchangeable with *su-* in these contexts (cf. Smith & Paauw 2006:167). The following examples show the most common allomorphs.

(9) *Seelon=nang lai   hathu kavanan **anà-dhaathang**.*
     Ceylon=DAT other INDEF group   PAST-come
     'Another group came also to Ceylon.' (K060108nar02)

(10) *Kitham **em-pii**   86=ka=ke.*
     1PL   PAST-go 86=LOC=SIMIL
     'We arrived in about 1986.' (N060113nar04)

(11) *Baapa derang=pe kubbong=ka hatthu pohong **nya-poothong**.*
     father 3PL=POSS   garden=LOC INDEF tree   PAST-cut
     'My father cut a tree in their garden.' (K051205nar05)

Like *su-*, *anà-* can refer to distant past as in the examples above, or to an event in the past which is relevant to the speech situation, like the inquiry whether the addressee has understood what the speaker said in (12).

(12) a. *Gaathal su-kuurang kalu suda hatthu=ke   thraa.*
        itching PAST-few   if   thus INDEF=SIMIL NEG
        'When the itching diminishes, none is left.'
     b. *Ikang butthul. **Anà-mirthi**?*
        then   correct PAST-understand
        'Then it's OK. You got that?' (K060103cvs02)

6.2.1.3   Difference between *su -* and *anà-*

The differences between *su-* and *anà-* are very subtle (Smith & Paauw 2006:167), and in the great majority of contexts, either morpheme can be used. Most notably, there is not a difference between preterit and perfect, as in English, or between perfective and imperfective as in the Romance languages. In the sections above, examples (7), (8), (9) and (12) show that both *su-* and *anà-* can be used in contexts which are relevant to the speech situation, and in contexts which are not. The following four examples show that *su-* and *anà-* are interchangeable in many contexts, given that the same speaker in the same recording uses now *anà-*, now *su-* to convey exactly the same content. In

---

[5]Smith & Paauw (2006:164) give *ay(ng)* as allomorphs, but it is not entirely clear which sound is indicated by *ay*.

(13), the completing of a lecture of the Qur'an is first found with *su-*, then repeated with *anà-*.

(13) a. *Qur'an thiiga skali* **su-thamaam-king**. *[...]*
          Qur'an three times PAST-complete-CAUS
          'I completed the Qur'an thrice.'

       b. *Hathu thaaun didaalam thiiga skali quran* **nya-thamaam-king**.
          one   year   within   three time Qur'an PAST-complete-CAUS
          'Within one year, I completed the Qur'an thrice.' (K051213nar02)

Another instance of very similar contexts being first found with *su-*, and then repeated with *anà-* is found in the following two examples.

(14) *Se=ppe umma-baapa se=ppe maama=pe ruuma=nang* **su-kiiring**.
       1s=POSS mother-father 1s=POSS uncle=POSS house=DAT   PAST-send
       'One year my parents sent (me) to my uncle's house.' (K051213nar02)

(15) *See* **anà-kiiring** *se=ppe maama hatthu=pe ruuma=nang*.
       1s PAST-send   1s=POSS uncle   INDEF=POSS house=DAT
       'I was sent to an uncle of mine's.' (K051213nar02)

In declarative clauses with past reference, there seems to be no difference between the semantics of *su-* and *anà-* then. Where we do find differences is with polar constituent questions in the past, and in the pluperfect construction. Polar constituent questions as in (16) are only possible with *anà-*.

(16) **Daging baabi=si** **su-/anà- bìlli???*
       pork=INTERR   *su-anà-* buy
       'Did you buy PORK???' (K081105eli02)

Besides polar constituent questions, *su-* is also ruled out from content constituent questions. The following three examples show, that *su-* is not possible in constituent questions (17), but it is possible in declarative illocutions (18), and in predicate questions (19), marked by =*si* on the predicate, *giigith* 'bite' in this case. This contrasts with the use of =*si* on an argument as in (16), where it attaches to the argument *baabi* 'pork', entailing the impossibility of *su-*.

(17) **Mana** *binaathang lorang=yang* **anà/*su-***giigith.*
       which animal   2PL=ACC   PAST-bite
       'Which animal bit you?' (K081105eli02)

(18) *Itthu binaathan lorangyang **anà-/su**-giigith.*
    DIST animal   2PL=ACC   PAST-/PAST-bite
    'That animal bit you.' (K081105eli02)

(19) *Itthu binaathan lorangyang **anà-/su**-giigith=**si**?*
    DIST animal   2PL=ACC   PAST-/PAST-bite=INTERR
    'Did that animal bite you?' (K081105eli02)

While *su-* is ruled out from the contexts discussed above, in the pluperfect, it is the only form that can be used.

(20) *Baapa inni=nang duppang luvar   nigiri=nang asà-pii ***anà-/su**-aada.*
    father PROX=DAT before   outside country=DAT CP-go anà-su-exist
    'My father had already gone abroad earlier.' (K081201eml01)

This could suggest that *su-* is used to mark realized contexts. In a pluperfect construction, the first event is necessarily completed and realized before the subsequent event takes place. This contrasts with a question, where the reality of the event is not established yet. One could speculate that *su-* has a specification for [+realized], which makes it incompatible with (polar or content) constituent questions. *Anà-*, on the other hand, could have a specification for clauses with argument focus, since it often cooccurs with the emphatic marker *=jo*. Furthermore, constituent questions also have the questioned element in focus position. This distinction according to information structure would also make sense from a contact language perspective, because Sinhala has an 'emphatic' verb form used in focal contexts (Gair 1985a). Things seem to be more complicated in SLM than in Sinhala, though. While in Sinhala, the use of the emphatic form is obligatory in focal contexts, this is not the case with *anà-* in SLM. True, most argument focus constructions with past reference have *anà-* (21) (22), but there are some examples where we find *su-* as well (23). In that example, where we are dealing with a clear argument focus: the fooled person is not Andare, but rather the king.

(21) *TV=ka=**jo**   **anà**-kuthumung.*
    TV=LOC=EMPH PAST-see
    'It was on TV that we saw it.' (B060115nar02)

(22) *Mlaayu pada=**jo** inni pada=ka punnu pukurjan **anà**-girja.*
    Malay PL=EMPH PROX PL=LOC much work   anà-make
    'It was the Malays who did a lot of work there.' (K051222nar05)

(23)   *Suda kanabisan=ka raaja Andare=yang mà-enco-king asà-pii, **raaja=jo***
       thus last=LOC   king Andare=ACC INF-fool-CAUS CP-go king=EMPH
       ***su-**jaadi   enco.*
       PAST-become fool
       'So at last, the king had set out to fool Andare, but it was the king (himself)
       who turned out to be the fool.' (K070000wrt02)

6.2.1.4   conjunctive participle *asà-*

This is the 'conjunctive participle' prefix. It can also be realized as *as-* or *s-*, in which
case it is easily confounded with a homonymous allomorph of the past tense *su-*.[6]

   Conjunctive participles are a common category on the Indian subcontinent (Bloch
1934, Emeneau 1956, Masica 1976), and they are also found in Sinhala and Tamil.[7]
They are used when a speaker talks about several, normally successive events ⊡→13.3,
p. 529. If there are *n* events, the first *n*-1 will be marked as conjunctive, while only the
last gets regular TAM-marking (cf. Slomanson 2008a). An approximate rendering in
English is *having done X*. Example (24) shows the use of the conjunctive participle on
the first two of a total of three events.

(24)   *Oorang pada **s-pìrrang**,   derang=nang **asà-banthu**, siini=jo*
       man   PL   CP-wage.war 3PL=DAT   CP-help   here=EMPH
       su-ciiňǧgal.
       PAST-settle
       'The men, having waged war, having helped them, settled down right here.'
       (K051222nar03)

   The English *having done X* relegates the event denotated by X to a less central
position in the clause, an adjunct. This is not the case in SLM, where the conjunctive

---

[6]Saldin (2001:53,57) also lists *ai-*. Slomanson (2006, 2008a) has *abIs* as the main allomorph of this
form, which then contrasts with the vector verb *abis* and the full verb *abis*. In other Malay varieties, *habis*
indicates completion, so that it is semantically plausible to reanalyze this morpheme as conjunctive partici-
ple, as Slomanson seems to suggest. The semantic argument is convincing, yet there is one phonological
detail, the second *à* in *asà-*, which is difficult to explain if *habis* is taken to be the proto-form. Smith &
Paauw (2006:167) argue that *asà-* is a contraction of a double prefixal past marking < *ayng-su-*, where
ayng seems to be the same as *anà-* in the Upcountry. Given the semantics of *anà-/ayng-* and *su-*, which are
both 'PAST', this etymology is unlikely. It is unclear how double past marking could grammaticalize into
conjunctive participle. Furthermore, the etymology of *anà-/ayng-* is still a mystery, so that this proposed
origin does not clarify much. Combining the two proposed etymologies, a possible origin for *asà-* could by
**habis su-*, which contracts to *asà-*. A combination of perfective with past to yield a conjunctive participle
is semantically plausible, and the phonological development is also acceptable.

[7]Alternative names for similar morphemes in other Indian languages are 'gerund' or 'absolutive', as
well as some other qualifications of 'participle', like 'adverbial', 'verbal', 'indeclinable' etc. See Masica
(1976:110) for a discussion.

participle is used without implying any supremacy of one event over the other. In this regard, it is closer to the English conjunction *and*.[8] Example (25) shows that the conjunctive participle clause can actually be more central to the meaning that the 'main clause'.

(25) *Nyaakith oorang pada asà-pii, thaangan arà÷cuuci.*
sick   man   PL   CP-go   hand   NON.PAST÷wash
'The patients come and/to wash their hands.' (K060116nar03)

Asà- is used most often for past events, but it can also be used for sequencing events in the other tenses, like the general present in the following example.

(26) *Samma oorang school=nang **asà**-pii **arà**÷blaajar   cinggala.*
all   man   school=DAT   CP-go   NON.PAST÷learn Sinhala
'Everybody goes to school and learns Sinhala.' (B060115cvs01)

Example (27) shows the use of *asà-* in a future context.

(27) *Go   **asà**-niiňğgal, alla   go=nya[9]   **asà**-dhaathang, kuburan*
1S.FAMILIAR CP-die   Allah 1S.FAMILIAR=DAT CP-come   grave
***asà**-gaali, go=nya   kubur-king katha.*
CP-dig   1S.FAMILIAR=ACC bury-CAUS   QUOT
'[I said:]"When I will have died and Allah will have come for me and the grave will have been dug, bury me." ' (B060115nar05)

The events in (27) have not taken place yet since the speaker is still alive. It refers to the future, to the time of his death. Yet *asà-* is used to structure the sequence of events. Another example is (28) where neither the return nor the mischief of the monkeys have taken place yet.

(28) *Moonyeth pada=le **asà**-dhaathang creeveth **athi**÷kaasi.*
monkey   PL=ADDIT CP-come   trouble   IRR÷give
'The monkeys would certainly go and cause (some other) trouble.' (K070000wrt01)

Finally, there are some rare instances in which *asà-* does not carry a meaning of anteriority or subsequence. Example (29) shows that it can also be used to conjoin two events which are taking place at the same time.

---

[8]Note the similarity between the terms *conjunction* and *conjunctive participle*.
[9]The use of *=nya* here cannot be explained as of now.

(29)   *[Banthu-an **asà**-mintha **arà**÷naangis svaara] hatthu derang=nang*
       help-NMLZR CP-beg   SIMULT÷cry   sound   INDEF   3PL=DAT
       su-dìnngar.
       PAST-hear
       'They heard a sound of crying and begging for help.' (K070000wrt04)

In this example, the relative clause which modifies *svaara* 'sound' contains two
verbs, *mintha* 'beg' and *naangis* 'weep'. It is now not the case that the dwarf (who the
story is about) first begs and then cries. Rather, the two things happen simultaneously.
Note that the two events are conceptualized as different, but simultaneous. Conceptu-
alization as one event would mean that there is only one event, begging, which is done
in a weeping fashion. This can also be expressed in SLM, but the modification of a
verbal predicate by another verb asks for reduplication as in (30) or (31).

(30)   *[Banthu-an **naangis˜naangis** anà-mintha svaara] hatthu derang=nang*
       help-NMLZR weep˜RED   PAST-beg   sound   INDEF   3PL=DAT
       su-dìnngar.
       PAST-hear
       'They heard a sound of weepingly crying for help.' (K081103eli04)

(31)   *[Banthu-an **mintha˜mintha** anà-naangis svaara] hatthu derang=nang*
       help-NMLZR weep˜RED   PAST-beg   sound   INDEF   3PL=DAT
       su-dìnngar.
       PAST-hear
       'They heard a sound of weepingly crying for help.' (K081103eli04)

The conjunctive participle *asà-* is often found combined with the postposition
=*apa* 'after', which reinforces the meaning of subsequence.

(32)   *Pohong komplok duuva=yang **asà**-baa=**apa**, mliige=pe   duuva subla=ka*
       tree   bush   two=ACC   CP-bring=after palace=POSS two   side=LOC
       su-thaanàm.
       PAST-plant
       'The bushes were brought and planted on both sides of the palace.'
       (K070000wrt04)

The use of =*apa* is optional, as the following example shows, where we find *asà*-
twice but =*apa* only once.

(33) a. ***Asà**-pii=**apa***
       CP-go=after
       'He went there and'

  b. *sithu=ka=jo   kuburan samma **asà**-gaali ∅*
      there=LOC=EMPH grave   all   CP-dig
      'they dug his grave right there and'

  c. *karang itthu avuliya derang=pe zihaarath aada.*
      now   DIST saint   3=POSS   shrine   exist
      'now there is his shrine over there.' (B060115nar05)

6.2.1.5 infinitive *mà-*

This is a verbal prefix used in purposive clauses, in nominalizations and with certain modal particles (cf. Slomanson 2006:139f,144f). Sinhala and Tamil have analogous constructions, which are used in similar contexts and which are termed infinitives. Hence, it seems convenient to retain that name for this morpheme. It should be noted that this term is chosen to facilitate comparison with the adstrates and does not comply with the meaning of that term in Latin grammar, i.e. lack of person marking on a verb, person never being marked in SLM.[10]

   Mà- is a frequent morpheme, which is nearly always realized in the same way, but sometimes it is sometimes shortened to *m-*.

   The three functions of *mà-* (purposive, complement of modal, adclausal nominalization) are nicely exemplified in the following fragment.

(34) a. *Itthu* cave=*nang kithang=le pii aada **mà**_{purp}-liyath=nang.*
       DIST cave=DAT   1PL=ADDIT go exist INF-look=DAT
       'We have also gone to that cave to have a look.'

  b. *Bannyak jaau **mà**_{modcomp}-pii thàrboole, itthu=ka.*
      much   far INF-go   cannot   DIST=LOC
      'You cannot go far there.'

  c. *Daalang=ka **mà**_{nmlzr}-jaalang mlaarath.*
      inside=LOC INF-go   difficult
      'Walking inside is difficult.' (K051206nar02)

   The first sentence shows the use of *mà-* in a purposive construction. The purpose of the speaker's going to the cave was examination. The second sentence shows *mà-* used on the complement of the modal *thàrboole*. The third sentence finally shows

---

[10]Arden (1934:196) states for the comparable Tamil form: "The term 'Infinitive' is the best Western grammatical term for it", indicating that the decision to call this form 'infinitive' is not clear-cut, but that no better term could be found. Also cf. Beythan (1943:106).

a nominalization with *mà-*, of which the property 'difficult' is predicated. Note that the grouping of purposive, complement of modals and nominalization is actually not very different from English, where we can use the preposition *to* in all these contexts (*We went to the cave **to** have a look; it is impossible **to** go there; it is difficult **to** walk inside*).

The purposive use above differs from the other two uses in that the purposive clause also features the dative marker *=nang*. This is very often found, but by no means obligatory, as the following example shows.

(35)   *Blaakang Andare [Kandi=ka asduuduk Dikwella arà÷pii   jaalang]=ka*
        after   Andare Kandy=LOC from   Dikwella NON.PAST÷go road=LOC
        *aayer **mà**-miinong=∅ Udamalala kampong=ka su-birthi.*
        water INF-drink   Udamalala village=LOC PAST-stop
        'Then, Andare stopped on the street which leads from Kandy to Dikwella at the hamlet Udamalala to drink water.' (K070000wrt03)

Finally, the infinitive combined with the interrogative clitic *=si* is used to request permission.

(36)   *See **mà**-maakang=**si?***
        1s   INF-eat=INTERR
        'Shall I eat?' (K081106eli01)

In Sinhala and Tamil, the imperative form is used in this construction, but in Sinhala (and not in Tamil), infinitive and imperative are homophonous, so that the use of the infinitive here is a clear case of Sinhala influence.

Saldin (2001:57f), Smith & Paauw (2006:173) and Ansaldo (2008:29) propose that *mà* is cognate with the Std. Malay transitivizer *məN-*. The development from a transitivizing morpheme to an infinitive is not yet attested in the literature on grammaticalization paths. Slomanson (2006) proposes the modal *mɔ* from Ambonese Malay as a cognate of *mà*.

6.2.1.6   negative non-finite *jamà-*

This is the negation of *asà-* (Slomanson 2006, 2008a). As such it is used in subordinate clauses, for stating that a certain action was not completed before another action took place ⊡→13.3, p. 529. *Jamà-* can also be used as a negative imperative ⊙→17.1.5, p. 699.[11] Slomanson states that *jang* (the allomorph he found) cannot cooccur with the affirmative infinitive *mà-*. This is true also in the Upcountry for *jamà-*, but it might be the case that the second syllable of *jamà-* is actually a reflex of a former infinitive being fused, while the first is the historical form *jangang* (*jangang mà-V > jamà-V*).

---

[11]I would like to thank Peter Slomanson for explaining the behaviour of *jamà-* to me.

The negative imperative is the most common use of *jamà-* in the corpus. *Jamà-* can combine with verbs as in (37)(38) or with adjectives as in (39). The use with *maalu* 'shy' is especially common.

(37) *Jamà-kaaluth, biilang iithu sama oorang.*
      NEG.IMP-shout say   there all   man
      'Don't shout, but sing it properly.' (K061019sng01,K081105eli02)

(38) *Hatthu=le **jamà-gijja**, baapa ruuma=ka duuduk.*
      INDEF=ADDIT NEG.NONFIN- make father   house=LOC stay
      'Father, do not do anything and stay at home!' (B060115nar04)

(39) *Biilang, maalu, **jamà-maalu**.*
      say   shy   NEG.NONFIN-shy
      'Speak! You are shy, don't be shy.' (B060115prs07)

When used to negate the conjunctive participle, *jamà-* is often accompanied by the dative marker =*nang*.

(40) a. *Liivath aayer **jamà**-jaadi=**nang**,*
          much   water NEG.NONFIN-become=DAT
          'Without having too much water'
      b. *itthu aayer=yang hathu blaangan=nang luppas.*
          DIST water=ACC   INDEF amount=DAT   leave
          'leave that water for a while.' (K060103rec01,K081103eli04)

The occurrence of =*nang* could be explained by the fact that =*nang* is an adverbializer, among other things. Given that not doing something is difficult to situate in a chain of events because nothing happens, an interpretation as manner is preferred here, conveyed in the translation by *without*. It is not the case that first you do not do something and then you take a subsequent step. This would be the strict temporal interpretation. Rather, you follow the command expressed in the second clause *without* first performing the action expressed in the first clause. We are thus not dealing with a chain of events, but rather with a modification of a primary element by an indication of manner.

A related analysis can be given for (41).

(41)   *Itthu abbis maakang kalu kithang=nang bole=duuduk hatthu=le*
       DIST finish eat   if   1PL=DAT   can-stay   INDEF=ADDIT
       jamà-maakang=nang *two duuva two o'clock=ke   sangke bole=duuduk.*
       NEG.NONFIN-eat=DAT   two two   two o'clock=SIMIL until   can-stay
       'If we eat it up, we can stay up until 2 o' clock without eating anything.'
       (K061026rcp04)

Here, it is not the case that there is a sequence of events 'First I eat nothing and then I stay until 2 o'clock'. Rather, the staying until two o'clock is further characterized by the fact that no food is consumed. We are thus dealing with an indication of manner, and not with an indication of temporal subsequence.

A final example of the use of *jamà-V=nang* as meaning 'without' is the story of a dwarf liberated from being stuck in a thorn tree by two girls. The dwarf does not thank them, but scolds them. In other words, he scolds them without thanking them. The syntax of this sentence is somewhat complex, so that brackets are used to indicate the constituents. Four constituents are particularly important: 1) The agent *incayang* 'he' at first position, 2) the predicate *maaki* 'scold' at final position 3) the content of the scolding stretching over lines b. and c. and 4) the manner of the scolding in line a.

(42)   a.   *Incayang [[derang=nang thriima **jamà**-kaasi=nang],*
            3S.POLITE 3PL=DAT   thanks NEG.NONFIN-give=DAT
            'Without thanking them and'

       b.   *[anà-banthu vakthu=ka incayang=pe jeeňǧgoth=yang asà-thaarek=apa,*
            PAST-help   time=LOC 3S=POSS   beard=ACC   CP-pull=after

            'while they were helping him they had pulled his beard and'

       c.   *incayang=nang su-sakith-kang katha] anà-maaki].*
            3S=DAT   PAST-pain-CAUS QUOT PAST-scold
            'caused him pain, he scolded.' (K070000wrt04)

Note that (38) could be given a similar interpretation if =*nang* was present in that utterance. As it is now, the utterance consists of two commands 'Don't do this! Do that!' If *jamà* were present, we would be dealing with one command 'Do this without doing that!'

This morpheme seems to be more common in the speech of Slomanson's informants in Colombo. In the Upcountry, it is hard to find, and people often do not understand the point of the elicitation sessions. Slomanson also lists two additional uses of *jamà-*, negation of purposive clauses and negative nominalizations.

Negative purposive clauses would be formed by replacing the affirmative purposive marker *mà* with *jamà-*, but my informants were very reserved with regard to this

construction (43). While they would not reject it outright, they found it a bit weird and disrespectful. Instead, they preferred a construction with the negative imperative *thussa* in reported speech (44).

(43) *??se incayang=nang duvith anà-kaasi jamà-oomong katha.*
      1s   3s.POLITE=DAT   money PAST-give NEG.INF-talk   QUOT
      'I gave him money so that he would not talk.' (K081103eli04)

(44) *Se incayang=nang duvith anà-kaasi thussa mà-oomong katha.*
      1s 3s.POLITE=DAT   money PAST-give NEG.IMP INF-talk   QUOT
      'I gave him money so that he would not talk.' (K081103eli04)

   Slomanson also lists *jamà-* for nominalizations of negated predicates. This is not possible in the Upcountry (45). I have not been able to find out what would be used for this function instead, though.

(45) **luvar nigiri=nang kaapang=ke jamà-pii thàrbaae.*
      outside country=DAT when=SIMIL NEGINF-go bad
      '(It is not good to never go abroad).' (K081103eli04)

6.2.1.7   *kapang-* 'when'

This is a prefix indicating temporal coincidence (cf. Slomanson 2008b) ⊙→15.5.3, p. 625. It is often realized as *kaN-* with the nasal assimilating in place to the following consonant. The form *kam-* shows some similarities with Std. Malay *kapan* 'when', and *kalau* 'if'. Normally, *kaN-* is found in SLM, but *kapang-* or *kala-* can also be found in temporal context, without any obvious difference in meaning between the three forms. Finally, the further reduced form *ka-* is also possible. The following four examples show the use of *kapang- (46), kam- (47), kal- (48)* and *ka-* (49). *Kapang-* as a prefix is distinguished from interrogative pronoun *kaapang* 'when?' ⊡→5.7.4, p. 234 by its short vowel. The postposition *=kapang* ⊡→6.4.4.7, p. 349 is distinguished from *kapang-* by the position relative to the host.

(46) *Incayang **kapang**-dhaathang, cinggala incayang=nang thàrà-thaau.*
      3s.POLITE when-come   Sinhala 3s.POLITE=DAT   NEG-know
      'When he came, he did not know Sinhala.' (K060108nar02)

(47) *Mana=na mà-thiidor **kam**-pii, se maakan=nang kam-pii.*
      which=DAT INF-sleep   when-go 1s eat=DAT   when-eat
      'When I go to bed, when I eat [I think of my wife].' (K061120nar01)

(48)   *Paagi=nang* **kala**-*baavung sangke=soore busy.*
       morning=DAT when-rise   until=evening busy
       'When you get up in the morning until the evening, you are busy [in Colombo].' (B060115cvs01)

(49)   *Laskalli*   **ka**-*dhaathang mari Badulla=nang.*
       other.time when-come   come Badulla=DAT
       'Come to Badulla on your next visit!' (B060115cvs17)

While *kala-* bears some resemblance to the conditional particle *kalu*, in (48), a conditional interpretation is not possible: not getting up in the morning is not an option considered in the proposition. The meaning is rather 'as soon as you get up', highlighting the degree of stress one experiences in Colombo.

The time of the speech act has no bearing on the use of *kapang*, which can be used with present, past or future reference. Nevertheless, *kapang-* is in the same position as the other TAM-prefixes and blocks their use (cf. Slomanson 2008b). *kam-pii* can thus mean 'when he came', 'when he comes' or 'when he will come', as the case may be. The following two examples show the use of *kapang-* in a past situation (50) and in a future situation (51).

(50)   *Thaangang mà-saapu hatthu* paper *kapang-mintha baapa=yang*
       hand   INF-sweep INDEF paper when-ask   father=ACC
       su-kuthumung.
       PAST-see
       'When he asked for a paper to clean his hands, he saw daddy.'
       (K051205nar05)

(51)   *Aavi luvar=nang kapang-dhaathang itthu bambu=yang giini*
       steam outside=DAT when-come   DIST bamboo=ACC like.this
       angkath=apa pullang arà÷thoolak.
       lift=after   slow   NON.PAST÷push
       'When the steam comes out, lift the bamboo like this and push it slowly.'
       (K061026rcp04)

The line between *kapang-* (main meaning temporal) and *kalu* (main meaning conditional ⊡→5.14.2, p. 260) is blurred (cf. Slomanson 2008b), as the following two examples show. Example (52) is a clear instance of the temporal relation *whenever*. The fulfilment of the trips to the mosque is not hypothetical, the speaker regularly goes to the mosque and gives money to the needy.

(52) Mosque=*nang **kapang**-pii samma oorang=nang go   athi÷kaasi.*
    mosque=DAT when-go all   man=DAT   1S.FAMILIAR IRR÷give
    'When I go to the mosque, I give to everybody' (B060115nar04)

In (53), on the other hand, it is not sure whether the addressee will actually go to the bridge to have a look. While at first, *kapang-* is used, the afterthought rather has *kalu* as a more obvious conditional.

(53) *Mawanella **kapang**-liyath anthi-kuthumung, Mawanella* bridge=*ka*
    Mawanella when-see   IRR-see   Mawanella bridge=LOC
    *asàduuduk liyath **kalu**.*
    from   see if
    'When/If you look from Mawanella, you can see it, if you look from Mawanella bridge.' (K051206nar02)

6.2.1.8 superlative *anà-*

Homonymous to the past tense prefix, there is the superlative prefix *anà-*. It attaches to adjectives to mark superlative as in (54) ⊙→15.10, p. 664.

(54) *Seelon=ka **anà**-bìssar pohong.*
    Ceylon=LOC SUPERL-big tree
    'The biggest tree in Sri Lanka.' (K081104eli06)

Adjectives marked with *anà-* also often carry the emphatic marker =*jo*.

(55) *Ini dunia=ka **anà**-kaaya=**jo**   oorang Bill Gates*
    PROX world=LOC SUPERL-rich=EMPH man   Bill Gates
    'Bill Gates is the richest man of the world.' (K081103eli02)

6.2.2 Quasi-prefixes

This class comprises preverbal TAM morphemes. Quasi-prefixes are very similar to true verbal prefixes: They cannot stand on their own, cannot be moved around in the clause, have no free form, and are selective with regard to their host. What distinguishes them from true affixes is that they can be separated from their host by the emphatic morpheme *jo*, otherwise an enclitic (cf. Smith & Paauw 2006:164).[12] This

---

[12]Smith & Paauw (2006) also have other material intervening between the quasi-prefixes and the verb, like the adverb *konyam* 'few', but this is considered odd by my informants.

morpheme	gloss
su-	past
anà-	past
asà-	conjunctive participle
mà-	infinitive
jamà-	NEG.NONFIN
kapang-	when
arà÷	NON-PAST
anthi÷	IRR
masthi÷	must
thàrà÷	NEG.PAST
thama÷	NEG.IRR
marà÷	ADHORT

Table 6.2: Affixes and Pseudo-affixes

separability means that we have to distinguish an intermediate category between affixes and clitics in SLM, which I term *quasi-prefixes*. In this section, I will use the division sign (÷) to distinguish quasi-prefixes from both affixes, marked by a hyphen (-) and clitics, marked by an equal sign (=). In other parts of this book, the unfamiliar division sign is dispensed with and quasi-prefixes are marked just like affixes, i.e. with a hyphen. This is done to assure readability. It was felt that the loss of information incurred by not designating the phonological status of a bound element was not significant enough to burden the examples with this unfamiliar notation.

Six TAM morphemes have been found to be quasi-prefixes, namely the non-past marker *arà÷*, the irrealis marker *anthi÷*, the marker of obligation *masthi÷*, the negative past marker *thàrà÷*, the negative irrealis marker *thama÷*, and finally the adhortative *marà÷*. The other verbal prefixes shown in Table 6.2 do not show the possibility to combine with *jo* in sentences like (56) (position of the morpheme indicated by the underscore ___).

(56) *Incayang ___=jo   dhaathang*
       3s.POLITE ___=EMPH come
       'He definitely does/did/will/has (not) come.'

All quasi-prefixes have in common that they do not change the cases assigned by the verb they attach to (unlike modal particles, which assign dative). Furthermore, the quasi-prefixes are used with a bare verb, and not with the infinitive, which the modal particles can require.

The normal case for a quasi-prefix is to be used just like a prefix, as schematized in in (57).

(57) NP QUASIPREF÷V

But the emphatic clitic *jo* can intervene with quasi-prefixes, as shown in (58), but not with affixes (59)

(58) ┌─────────────────────────────┐
     │ NP QUASIPREF÷*jo*÷V   │
     └─────────────────────────────┘

(59) ┌─────────────────────────────┐
     │ * NP TRUEPREF-*jo*-V   │
     └─────────────────────────────┘

However, speakers seem to disprefer the above solution and often opt for the construction exemplified in (60).

(60) ┌──────────────────────────────────────┐
     │ N QUASIPREFIX÷jo QUASIPREFIX÷V   │
     └──────────────────────────────────────┘

I have been unable to find any hyphen-or-other notation that would capture the intriguing facts of (60). Basically, an enclitic =*jo* combines with a quasi-prefix, without a stem intervening, but the quasiprefix is copied to the normal, directly preverbal position.

6.2.2.1 non-past *arà÷*

Arà , allomorph *ar* [13], is an ubiquitous morpheme for which it is difficult to pin down a central meaning. It normally attaches to a verb directly, as shown in (61), but *jo* can intervene as shown in (62). If *jo* is used, the full form *arà* is obligatory, while in other contexts, the reduced form *ar* can be found.

(61) *Derang kaarang **arà÷blaajar** cinggala=le* English=*le.*
     3PL   now   NON.PAST-learn Sinhala=ADDIT English=ADDIT
     'Now they learn Sinhala and English.' (K051222nar06)

(62) *Se ini   pukurannang arà÷jo   dhaathang.*
     1s PROX work=DAT   NON.PAST÷EMPH come
     'To this work, I am definitely coming.' (K081105eli02)

Arà÷ is most often found referring to present events (Smith et al. (2004), Slomanson (2006:141), Smith & Paauw (2006:164), Ansaldo (2009b)) ⊙→15.5.2.2, p. 620. Second comes reference to future events (Smith & Paauw 2006:164) ⊙→15.5.2.3, p. 622, while rarely, *arà÷* is used to events taking place at the same time as the frame of reference ⊙→15.5.3.3, p. 630. Because of this, I assume that *arà÷* is actually polysemous, the first meaning being 'non-past', the second one, 'simultaneous'. We will discuss them in turn.[1415]

---

[13]Adelaar (1991:31), (Slomanson 2006:143) and Slomanson (2008a) also lists *ada* as a possible form. In my data, *a(a)da* is never found preverbally, but only after the verb, where it indicates a different tense, perfect ⊡→5.1.2.1, p. 165, ⊡→5.1.6.1, p. 192

[14]Adelaar (1991:31) lists this morpheme as 'progressive' in line with the semantics of cognates in other Malay varieties (Adelaar 2005), but the following discussion will show that the semantics in Sri Lanka differ from these other varieties.

[15]The present tense and simultaneous uses could possibly be unified under the gloss 'simultaneous', with

Non-past   The non-past reading of *arà÷* obtains in contexts of present reference as in (63), and in contexts of future time reference, as in (64). Furthermore, generic time reference is also expressed by this morpheme (65)(66). Reference to habitual activities can also be established by *arà÷* (67) but the irrealis marker *anthi÷* can also be used there (see below).

(63)   *Libbi   mlaayu samma Kluum̆bu=ka   **arà÷duuduk**.*
       remain Malay all   Colombo=LOC NON.PAST-stay
       'The remaining Malays live all in Colombo.' (K060108nar02)

At the time of speaking, it was true that more Malays live in Colombo than elsewhere, so we are dealing with a present time reference.

(64)   *Laskalli   **arà÷maakang**.*
       other.time NON.PAST-eat
       'Next time you will eat.' (B060115cvs16)

The word *laskalli* 'next.time' indicates that reference is not to the time of speaking, but rather to some point of time in the future. Nevertheless, *arà÷* is used in this context as well.

(65)   *[Itthu se **arà÷kirijja**]   mosthor=jo.*
       DIST   1s NON.PAST-make manner=EMPH
       'That's the way I do it.' (B060115rcp01)

When the speaker uttered this fragment of a recipe, she was not performing the cooking. The relative clause thus does not refer to the time of speaking, but is underspecified for tense. It has a generic time reference. This generic time reference is also expressed by *arà÷*.

(66)   Uncle*, mana=ka   **arà÷duuduk***?
       uncle   where=LOC NON.PAST-live
       'Where do you live?' (B060115cvs16)

In the question in (66), the information requested was not where I was staying at the time of speaking (that was just in front of the speaker), but where my residence was. This broader reference frame of time, which encompasses the moment of speaking as well as a sizeable period preceding and following it, is also referred to by *arà÷*.

---

the present tense reading emerging from implicature: simultaneous to the default point in time, which is the moment of speaking, necessarily implies present (Comrie 1985:63f). However, the future reading of *arà÷* cannot be accommodated under the simultaneous reading, so that we have to assume polysemy in any case. Two divisions are possible: non-past:simultaneous, or simultaneous:future. In this description, I choose the former.

(67)   *Itthu kalu, [...] Dubai=ka asduuduk laama kar pada kitham **arà÷baapi***
       DIST if   [...] Dubai-LOC from   old   car PL   1PL   NON.PAST÷take
       Iraq ports=*dang*
       Iraq ports=DAT
       'In that case, we bring old cars from Dubai to the ports in Iraq .'
       (K051206nar19)

Finally, the last example refers to a habitual activity (cf. Ansaldo 2009b). This sentence was uttered in the speaker's house in Sri Lanka, not on a ship in the Gulf. It refers to the speaker's work, which is in the cargo business in the Persian Gulf. At the time of speaking, he was thus not performing such a trip. Still, he habitually performs the trips depicted in the sentence, he did so before the time of speaking, and would do so afterwards. This habitual aspect is also expressed by *arà÷*.

Simultaneous   Simultaneous tense is the second reading of *ara÷*. It obtains in non-present contexts where an event in an embedded clause is simultaneous to an event in the main clause.[16] An example of this is (68), where a man looks at monkeys playing.[17]

(68)   *Blaakang=jo incayang anà-kuthumung [moonyeth pada thoppi asà-ambel*
       after=EMPH 3s.POLITE PAST-see   monkey   PL   hat   CP-take
       *pohong atthas=ka **arà÷maayeng**].*
       tree   top=LOC   SIMULT÷play
       'Then only he saw that the monkeys had taken his hats and were playing on the top of the trees.' (K070000wrt01)

The reference time is past tense, since the action already happened. This is why the verb *kuthumung* 'to see' is marked for past tense with the prefix *anà-*. The complement of this verb is an embedded clause. This clause inherits the reference time from the matrix clause, as it were. In the embedded clause, then, the only thing which is indicated is that the event of the monkeys playing was simultaneous to the man looking at them (rather than anterior or posterior). This simultaneity of the event expressed in the embedded clause with the event expressed in the matrix clause is expressed by *arà÷*. Note that at the same time, *asà-* in *asà-ambel* 'having taken' denotes anteriority with respect to *maayeng* 'play' and *kuthumung* 'see'. The order of events is first taking, then simultaneous playing and seeing.

---

[16]Technically, it also obtains in present tense contexts, where the action is simultaneous, but this is simply treated as present tense here, and not as relative tense simultaneous to present tense.

[17]Slomanson (2006:147) notes that in the Colombo variety, a case marker (dative =*nang* or another marker =*nya* of unclear semantics) is required for this type of subordination. This is not the case in the Upcountry.

Arà÷ forms part of the most basic grammar of SLM and is present in all varieties. It stems etymologically from the existential **ada*, which is attested as a preverbal marker of progressive aspect (Adelaar 2005:14). Note that there is another reflex of **ada*, namely *aada*, which indicates perfect tense on the verb.[18]

6.2.2.2   irrealis *anthi÷*

Anthi÷ is used for several kinds of irrealis reference. Besides the normal form *anthi* (69), it has the allomorphs *a(t)thi÷* and *thi÷* (70)(71) (cf. Adelaar 1991:31).[19] These latter forms cannot be separated from their host by *jo*, which is possible for *anthi*.

(69)   *Suda lorang=yang ruuma=nang **anthi÷**aaji.baapi*
       thus 2PL=ACC   house=DAT   IRR÷take.away.ANIM
       'So, I will take you back home.' (K070000wrt04)

(70)   *Ithukapang lorang=pe leher=yang kithang **athi÷**poothong.*
       then   2PL=POSS neck=ACC 1PL   IRR÷cut
       'Then we will cut your neck.' (K051213nar06)

(71)   *Kithan=le **thi÷**banthu.*
       1PL=ADDIT IRR÷help
       'We will help, too.' (B060115nar02)

Anthi÷ is used to refer to events which are to take place in the future (72)(73) ⊙→15.5.2.3, p. 622. Furthermore, it can refer to events the speaker is not sure of whether they will happen. This is the case for several kinds of epistemic modality (76)-(79) ⊙→15.7.2.4, p. 658, but also for the apodosis (80)(81) of conditional statements ⊙→15.8, p. 660. Furthermore, *anthi÷* can be used to refer to habitual actions as in (83) ⊙→15.6.2, p. 647.[20]

(72)   *Mpath* trips *kitham **atthi÷**pii.*
       four   trips 1PL   IRR÷go
       'We will make four trips.' (K051206nar19)

The speaker expresses his intention to perform a certain action in the future, and he is confident that these four trips will actually take place. There is no epistemic speculation or doubt involved, but still, the utterance is marked with *anthi÷*. A similar

---

[18]*aada* can also be used for other functions, see ⊡→5.1.2.1, p. 165.

[19]Smith & Paauw (2006:168) also give *ang-*, which is not found in the corpus. Ansaldo (2009b) gives *ai-*, which is not found in the corpus either.

[20]This clustering of functions is actually common cross-linguistically, as shown by Givón (1994).

situation obtains in (73), where the speaker is sure that she will provide information, albeit little, within moments after uttering that sentence.

(73) *Se=dang kalu bannyak mà-biilang thàrà-thaau; sdiikith see **athi**÷biilang.*
     1s=DAT if   much INF-say   NEG-know   few   1s IRR÷say
     'As for me, I cannot tell you a lot, I will tell you a little.' (K051205nar02)

Just as in example (72), the future action ('talking') is not subject to conditions or uncertainty, still, *anthi÷* is used. A borderline between future and irrealis meaning is (74), where the speakers try to convince the addressees to join their cause by offering them gifts. The transmission of the gifts is not sure to take place; it could be withheld in the event that the addressees do not join.

(74) *Kithang lorang=nang baaye mliiga **athi**÷kaasi, mà-kaaving panthas*
     1PL   2PL=DAT   good palace IRR÷give INF-marry beautiful
     *pompang pada **athi**÷kaasi, duvith **athi**÷kaasi.*
     girl   PL IRR÷give   money IRR÷give
     'We will give you beautiful palaces and women to marry and money.'
     (K051213nar06)

Another example on the borderline between future and irrealis is (75), where the fact of leaving is sure, but the time of departure is not.

(75) *Hathu duuva thiiga aari=dering see athi÷pii Middle East=nang.*
     INDEF two   three day=ABL   1s IRR÷go Middle East=DAT
     'In two or three days I will go to the Middle East.' (K061026prs01)

We now turn to the unequivocal irrealis examples.

(76) *[Duppang=ka inni samma **anthi**÷oomong=so], thàrà-thaau.*
     before=LOC PROX all   IRR÷speak=whether, NEG-know
     'Whatever they might have said before, (I) don't know.' (G051222nar02)

In this example, the speaker refers to something he does not know. Since the things he is ignorant of are not real for him, irrealis marking is used. Another example of *anthi÷* being used to indicate irrealis is (77), where a treatment for itching is explained. In the hypothetical case of itching, and in the hypothetical case that this treatment is applied, the itching would go away.

(77) a. *Minnyak   klaapa ini   raam̌buth=dering masà-goosok.*
coconut.oil coconut PROX hair=ABL   must-rub
'You must rub coconut oil (over the itching) with (human) hair.'

b. *Ithu=kapang ithu=ka   anà-aada itthu, raam̌buth ithu inni samma inni*
DIST=when,   DIST=LOC PAST-exist DIST hair   DIST PROX all   PROX
raam̌buth=nang anthi÷dhaathang.
hair=DAT   IRR÷come
'Then there is this there and it all comes on the hair.'

c. *Ithu=kapang gaathal **anthi**÷kuurang.*
DIST=when   itching IRR÷few
'Then the itching will become less.' (K060103cvs02)

Example (78) shows the use of *anthi÷* in a modal context. The speaker is not sure whether her son has already returned, but she speculates that this is the case. The speculative nature of the statement is marked by *anthi÷*.

(78) *Bìssar aanak asà-dhaathang **anthi**÷aada ruuma=nang.*
big   child CP-come   IRR÷exist house=DAT
'My eldest son should have come home by now.' (B060115cvs08)

This epistemic use is also possible for past events about the actual occurrence of which the speaker is not sure. Example (79) shows the use of *anthi÷* to mark the uncertainty or speculation of the speaker as to the veracity of the coming of the exiles in former times.

(79) *Hathu vakthu=nang, kithang=nang duppang dhaathang **athi**÷aada.*
INDEF time=DAT   1PL=DAT   before come   IRR÷exist
'Once, the [exiles] would have come before us' (K060108nar02)

Example (80) shows the use of *anthi÷* in the apodosis. The act of going will only take place if a certain condition, namely getting a chance, is met.

(80) Chance *kala daapath, **anthi**÷pii.*
chance if   get,   IRR÷go
'If (I) get the chance, (I)'ll go.' (B060115nar03)

The condition does not have to involve *kalu* 'if', but can be implicit, as in the following example.

(81) *Cinggala blaajar katha biilang thingka, ithu=kapang=jo gaaji*
Sinhala learn QUOT say middle DIST=when=EMPH salary
athi÷livath-king.
IRR÷much-CAUS
'When I say learning Sinhala, that means that your salary increases then.'
(K051222nar06)

Similar things hold for (82), which shows the use of *anthi* as a consequence of eating a certain heavy dish.

(82) *Itthu maakang=nang blaakang konnyong nanthok=ke **athi÷dhaathang.***
DIST eat=DAT after little sleepy=SIMIL IRR÷come
'After eating this, (you) get a little sleepy.' (K061026rcp04)

This use is actually on the borderline between the irrealis and the habitual reading of *anthi÷*. You get sleepy *if* you eat the dish, but also this is a thing which happens *habitually* after eating the dish. We now turn to more central examples of habitual reading conveyed by *anthi÷*. *Anthi÷* can also be used in habitual contexts (where it competes with *arà÷* then, see above).

In example (83) the speaker does not refer to people dancing at the time of speaking or at any precise time in the future. He refers to the general habit of mixed dances in Sri Lanka, as opposed to more rigid customs in other parts of the world where boys and girls may not dance together. This habitual way of performing the dances is expressed by *anthi÷*.

(83) *Siini klaaki prompang hamma **anthi÷thaaňdak**.*
here boy girl COMIT IRR÷dance
'Here, boys and girls will dance together.' (K060116nar10)

The habitual use of *anthi* is also possible with past reference, as examples (84)(85) show. The adverbs *kàthaama* and *pàrthaama*, both meaning 'before', clearly indicate reference to the past, but *anthi÷* is still used (cf. Givón 1994).

(84) *Punnu mlaayu pada **kàthaama** English=jona **anthi÷oomong.***
many Malay PL earlier English=PHAT IRR÷speak
'Many Malays would speak English in former times, wouldn't they.'
(K051222nar06)

(85) *Judge pada ini   ruuma=nang **athi**÷raathang **pàrthaama**.*
      judge PL   PROX house=DAT   IRR÷come   before
      'Judges used to come to this house before.' (B060115nar04)

The following example shows an ambiguous use of *anthi=*, which could be irrealis, but is more probably habitual because the giving is not subject to a hypothetical condition, but rather a habit which accompanies the regular visits to the mosque.

(86) Mosque=*nang kapang-pii samma oorang=nang go   **athi**÷kaasi.*
      mosque=DAT   when-go   all   man=DAT   1S.FAMILIAR IRR÷give
      'When I go to the mosque, I give everybody' (B060115nar04)

Finally, *anthi*÷ can also be used in expression to convey surprise, as in (87)

(87) *Itthu **saapa anthi**÷aada!.*
      DIST who   IRR÷exist
      'Who could that be!' (K070000wrt04)

Anthi is very common. The conflation of future and habitual is of Tamil influence. The etymological source is **nanti* 'soon' (Adelaar 1991, Bakker 2006).

Some speakers feel that the irrealis marker *anthi*÷ is not appropriate for use with first person and prefer the progressive *arà*÷ also for future reference, or a periphrastic construction with *arapi* 'going to'.[21]

6.2.2.3   debitive *masthi*÷

This morpheme is used to indicate obligation ⊙→15.7.2.2, p. 653and occasionally desire (Smith et al. 2004) ⊙→15.7.2.3, p. 656. Its allomorphs are *masà-* and *mas-*. The three allomorphs are shown in the following two examples.

(88) *Lai   aapa **masà**÷oomong, lai   aapa **masthi**÷oomong.*
      more what must÷speak,   more what must÷speak
      'What else should I say, what else should I say?' (K051206nar14)

---

[21]The semantics of this have a parallel in Sinhala, where the future marker =*nang* can only be used with 1st person. For third persons, the irrealis marker -*yi* has to be used, which in turn is never found on 1st person (Gair 2003:784).

(89)   *Incayang=pe* wife=*le*   wife=*pe* *baapa=le*   *masigith=pe bìssar*
     3s.POLITE=POSS wife=ADDIT wife=POSS father=ADDIT mosque=POSS chief
     mas÷panggel.
     must÷call
     'His wife and her father had to call the mosque's head.' (K051220nar01)

Just as with *anthi÷*, only the full form can be separated from the stem by *jo*. *Masthi÷* is the only quasiprefix for which the use of intervening *jo* is attested in naturalistic data (90).

(90)   *Derang su-biilang thraa,* Mrs. *Jaaya **masthi÷jo** raathang*
     3PL   PAST-say no   Mrs. Jaaya must÷EMPH come
     'They said "No! Mrs Jaaya MUST come! " ' (B060115prs01)

Normally, *masthi* leaves the case assigned by the verb unchanged (91).

(91)   *Duuva kaayu kithang*_{NOM} ***masà÷pii.***
     two   mile   1PL   must÷go
     'We have to walk two miles.' (K051213nar03)

In rare instances, *masthi* is found with the dative, as in (92)

(92)   *Se=**dang** karang ruuma=nang masà÷pii.*
     1s=DAT now   house=DAT   must÷go
     'I have to go home now.' (B060115cvs08)

Also, it normally combines with a bare verb stem. Smith et al. (2004) note that *masthi* combines with the infinitive in their data. This is not the case in the Upcountry, except for one instance:

(93)   *As-kumpul, blaakang laskalli as-*beat-*king, nni daalang=ka kithang aayer*
     CP-add   after   again   CP-beat-CAUS PROX inside=LOC 1PL   water
     masà mà-libbi-king.
     must INF-remain-CAUS
     'Having added (the ingredients), having then beaten them again, we must reserve the water in the middle.' (B060115rcp02)

This example is from Badulla, and the social background of the speaker is not known. It might be possible that she has some ties to the South, which is where the data in Smith et al. come from.

The primary function of *masthi* is to indicate obligation, as shown in (94)

(94) *Itthu* difference=*pe*   *atthas muuka masà÷thuuthup.*
      DIST difference=POSS about face   must÷close
      'Because of that difference, they have to cover their faces.' (K061026prs01)

In this case, we are dealing with social obligation, i.e. with the customs of the region which force the women to cover their faces. There is no logical need to do so, but rather a social need. This is similar in the following example, where the customary rules of conducting votes mandate that all voters be informed beforehand.

(95) a. *Badulla* Kandy *Matale, samma* association=*nang* **masà÷kasi-thaau.**
          Badulla Kandy Matale all   association=DAT   must÷give-know
          'Badulla, Kandy, Matale, all associations must be informed.'
      b. *Blaakang=jo* voting **masà-ambel.**
          after=EMPH   voting must-take
          'Only then must you take the vote.' (K060116nar06)

While in the above example, we were dealing with social necessity, the following example shows the use of *masthi* for event-oriented volitive modality ⊙→15.7.1, p. 650, a treatment for itchings (96). If the itchings are to go away, coconut oil must be rubbed, but there is no social obligation forcing the sufferer to apply coconut oil. Rather, it is generally desirable to treat itchings in the indicated manner.

(96) *Minnyak   klaapa ini   raam̆buth=dering masà÷goosok.*
      coconut.oil coconut PROX hair=ABL   must÷rub
      'You must rub coconut oil (over the itching) with (human) hair.'
      (K060103cvs02)

This morpheme is used by everybody to indicate obligation, but younger speakers also use it for desire as in (97)

(97) *Karang, se masà-maakang*
      now   1s must-eat
      'Now, I want to eat.' (overheard)

This can be explained by the adstrates Sinhala and Tamil, which do not distinguish obligation and desire (Lehmann 1989:84).[22]

This quasiprefix is very common in all registers. The etymon is *městi* 'must' (Smith et al. 2004, Smith & Paauw 2006).

---

[22]Karunatillake (2004:12) argues that nominative marking is used for obligation in Sinhala, and dative marking for desire. My Sinhala informants did not share that feeling, they found that a change in case did not change the semantics in that direction.

6.2.2.4 past negative *thàrà÷*

Thàrà is a prefix that indicates negative polarity in the past tense for verbs and negative polarity in any tense for certain adjectives. The final /-à/ is occasionally dropped (cf. Adelaar (1991:26), Slomanson (2006:141)).

Example (98) shows the use of *thàrà-* on a verb with past reference.

(98) *Derang thàrà÷dhaathang.*
  3PL NEG÷come
  'They did not come.' (K060116nar13)

Examples (99) and (100) show the temporally underspecified use of *thàrà-* on adjectives. The first example expresses a general judgment, which does not refer to the past in particular, while the second example refers to a gentleman who did not live anymore at the time of speaking, and therefore to the past. Without context, these examples could have any time reference ('is/was/will not (be) good/healthy').

(99) *Itthu thàrà-baae.*
  dist NEG-good
  'This is not good.' (K060116nar15)

(100) Mister *Yussuf thàrà-sìggar=le.*
  Mister Yussuf NEG-well=ADDIT
  'Mister Yussuf was also unwell.' (K0616nar15)

The verb *aada* is normally negated by replacing it with *thraa* ⊡→5.13.1, p. 254, but it is also possible to use *thraada* (101), which is a contraction of *thàrà-aada*. This latter full form is also sometimes found, as in (102), but even more rarely.

(101) *Giithu **thraada** kalu.*
  that.way NEG.exist if
  'If this is not available, ...' (K060103rec01)

(102) Confrontation *pada **thàrà÷aada**.*
  confrontation PL   NEG÷exist
  'There are no confrontations.' (K0151206nar20)

This morpheme is sometimes analyzed as an allomorph of *thraa*. Slomanson (2006) for instance discusses negation in SLM and does not distinguish between *thraa* and *thàrà*, while Adelaar (1991) seems to do so. In this description, a distinction is made between *thraa* on the one hand and *thàrà-* on the other hand. *Thraa* is a free word which is used for giving negative answers, for predications of non-existence, the

negated perfect and some other uses ⊡→5.13.1, p. 254. In clauses, it always follows
the predicate. It is monosyllabic. *thàrà-* on the other hand is a prefix which cannot
stand on its own. It is used to negate verbs and adjectives and always occurs before
the predicate. It is normally disyllabic. All in all, there are a number of phonological
and semantic differences distinguishing *thraa* from *thàrà-*. Because of the semantic
differences, treating them as allomorphs like for instance the pair *boole/bàr=* does not
seem advisable.

When *thàrà-* is used to negate predications in the past, it can negate either predi-
cations without relevance to the present, like (103), where the act of not sending some
papers occurred many years ago, or they can negate predications which have a rele-
vance for the speech situation, like the failure to come of the speaker's daughter, which
has the consequence that she is not here at the time of speaking in (104).

(103)   *See su-diya kithang=pe* president *[. . .]* subscription=*yang **thàrà÷kiiring***
       1s PAST-see 1s=POSS   president [. . .] subscription=ACC NEG.PAST÷send
       'I saw that our president had not sent the subscription.' (K060116nar10)

(104)   *Inni se=ppe* son-in-law. Daughter-in-law *nyaari **thàrà÷dhaathang***
       PROX 1s=POSS son-in-law daughter-in-law today NEG.PAST÷come
       'This is my son-in-law.   My daughter-in-law hasn't come today.'
       (B060115prs12)

The indistinction of preterit and perfect in the negative mirrors the same indis-
tinction in the use of the preterit *su-/anà-* or the perfect *V aada* in the affirmative
⊡→6.2.1.1, p. 271, ⊡→6.2.1.2, p. 272, ⊡→5.1.6.1, p. 192.

6.2.2.5 non-past negative *thama÷*

This morpheme is the verbal negator for non-past. As such, it is used to negate both
the non-past reading of *arà÷* and *anthi÷*. Allomorphs are *thuma÷* and *thamau÷*. All
allomorphs can combine with *jo*.

The following examples show the three allomorphs. They also show that the mor-
pheme can only be used preverbally, and with a bare verb.

(105)   *See lorang=nang **thama÷**sakith-kang.*
       1s 2PL=DAT   NEG.NONPAST-sick-CAUS
       'I will not hurt you.' (K070000wrt04)

(106)   *Inni bedahan jaadi=jo   **thuma÷**aada.*
       PROX difference come=EMPH NEG.NONPAST-exist
       'These differences would not at all have arisen.' (K060116nar03)

(107)   *Go   kaapang=le saala **thamau÷gijja.***
       1S.FAMILIAR when=ADDIT wrong NEG.NONPAST-do
       'I never do wrong.' (B060115nar04)

Thamau÷ and especially *thama÷* have to be distinguished from *thaama* (long vowel) meaning 'earlier' which is used in similar contexts, but then conveying affirmative past reference instead of negative non-past ⊡→10.4.1, p. 472.

Like the other quasi-prefixes, *thama÷* does not affect the case that the verb assigns. Example (108) shows that the verb *pii* 'go' continues to govern the nominative when used with *thama*, while the verb *luupa* 'forget' continues to govern the dative (109).

(108)   *Ithukapang derang nya-biilang: thraa, kithang*_{NOM} *giithu   thama÷pii.*
       then   3PL   PAST-say   no,   1PL   like.that NEG.NONPAST÷go
       'Then they said: "No, we will not go like that".' (K051213nar06)

(109)   *Suda itthu kithang=**nang**_{DAT} kaapang=pon thama÷luupa.*
       thus   DIST 1PL=DAT   when=any   NEG.NONPAST÷forget
       'We will never forget this.' (B060115nar02)

The two examples above show the use of *thama÷* in contexts with future reference. *Thama÷* can also be used to negate events with present or generic reference, as (110). For this example and related situations, it is difficult to distinguish whether the not-happening of the event is now or habitual, since something which does not happen habitually is normally also not happening at the time of speaking.

(110)   *Kitham=pe aanak pada thama÷oomong.*
       1PL=POSS child PL   NEG.NONPAST÷speak
       'Our children do not speak.' (G051222nar01)

Since *thama÷* is also the negation of *anthi÷*, it can be used in past contexts as well, if *anthi* would be possible in the affirmative. This is most notably true for habitual contexts. In (111), the speaker refers to the habitual way of things in former times. This habitual reading is expressed by *anthi÷* in the affirmative and by *thama÷* in the negative.as is the case in (111). The past reference is clear through the use of the temporal adverb *dovulu* 'before'.

(111)   a. *Non-Muslims pada=subbath kithang muuka konnyong arà÷cunjikang*
             non-Muslims PL=because   1PL   face   little   NON.PAST÷show
             siini.
             here
             'Because of the non-Muslims, we show our faces here.'

         b. *Thapi karang karang skarang skarang Sri Lanka=pe=le   Islam*
             but   now   now   now   now   Sri Lanka=POSS=ADDIT Islam
             pada muuka arà÷thuuthup karang.
             PL   face   NON.PAST÷close now
             'But now, Muslims in Sri Lanka cover their faces, now.'

         c. **Dovulu** *abbisdhaathang muuka* **thama÷thuuthup**
             before COPULA   face   NEG.NONPAST÷close
             'Before it was that they would not cover their faces.' (K061026prs01)

Another instance of the use of *thama÷* for habitual events in the past is (112),
where two army captains of a past century are said to have sticked to their word. This
habitual adherence to their words is not conveyed by the normal past negation *thàrà÷*,
but by *thama÷* since it is habitual.

(112)   *Derang pada kathahan* **thama÷thuukar.**
            3PL   PL   word   NEG.NONPAST÷change
            'They never changed their words.' (K051213nar06)

This habitual reading can be used to indicate the duration of an event, as in (113),
where the use of *thama÷* indicates that the flood had come to stay

(113)   *Itthu* **thama÷thuurung** *kiyang, aayer.*
            DIST NEG.IRR÷descend EVID   water
            'It would not recede, the water, it seems.' (K051206nar15)

Finally, *thama÷* can also be used in the apodosis of conditionals (114) (115).

(114)   *Aayer=ka* **kal**-*thaaro* **thama÷myaalak.**
            water=loc, when-put NEG.NONPAST÷catch.fire
            'When   they   put   it   into   water,   it   would   not   catch   fire.'
            (K051220nar01,K081103eli04)

(115) *Deram=pe* party **kal-***dhaathan* power=*nang, Seelon=ka,* English
   3PL=POSS Party if-come   Party=DAT,   Ceylon=LOC English
   thama÷*aada.*
   NEG.NONPAST÷be
   'If their party comes to power, there will be no English in Sri Lanka (any
   more).' (N060113nar02,K081103eli04)

Like *anthi÷*, *thama÷* can be used to express epistemic modality, also combined
with the emphatic clitic =*jo* ⊡→6.4.7.1, p. 378.

(116) *Derang Badulla=na sampe thama÷jo   aada.*
   3PL   Badulla=DAT reach NEG.IRR÷EMPH exist
   'They will probably not have reached Badulla yet.'

6.2.2.6 adhortative *marà÷*

Marà÷ is an adhortative quasi-prefix, used to encourage the addressee to perform an
action together with the speaker, similar to English *let's* ⊙→17.1.5, p. 699. It resembles
the imperative particle *mari*, and a common etymology is very likely. If this is the case,
it would also be cognate with the Std. Malay imperative *mari*. (117) gives a canonical
example.

(117) *Kitham **marà÷***maayeng.*
   1PL   ADHORT÷play
   'Let's play.' (K081104eli06)

When used on more than two actions, the first one is marked by the conjunctive
participle *asà-* (118), but the illocution holds for both.

(118) *Kithang asà-nyaanyi **marà÷***thaaňdak.*
   1PL   CP-sing   ADHORT÷dance
   'Let's sing and dance.' (K081104eli06)

Marà÷ can combine with =*jo*, as shown in (119), making it a quasi-prefix.

(119) *Kithang **marà÷***jo   pii.*
   1PL   ADHORT÷EMPH goo
   'Let's really go now.' (K081104eli06)

When combined with the interrogative clitic, *marà÷ ... =si* is used as a tentative
adhortative.

(120)   *Kithang marà÷maakang=si?*
        1PL   ADHORT÷eat=INTERR
        'Shall we eat?' (K081106eli01)

The following two examples are instances of *marà÷* used in naturalistic discourse. (121) shows the used of *marà÷* combined with *-la*.

(121)   *Karang criitha hatthu tsunaami=pe atthas **marà÷biilang-la**.*
        now   story   INDEF tsunami=POSS about ADHORT÷say-IMP
        'Now let's tell [Sebastian] a story about the tsunami.' (B060115nar02)

(122) shows the use of *marà÷* in an embedded sentence, where it translates as *should*.

(122)   *Itthusubbath see iingath [*moderate Islamization=*ka **marà÷pii katha**].*
        therefore   1s think   moderate Islamization=LOC ADHORT÷go QUOT
        'Therefore I think we should go towards a moderate Islamization.'
        (K051206nar18)

6.2.3   The semantics of affixes and quasi-prefixes

Many affixes and quasi-prefixes have more than one use. These uses are often semantically related. Figure 6.1 shows a semantic map for the meanings covered by these morphemes in the affirmative and the negative. A number of things are illustrated on this map. First, there is more differentiation in the affirmative than in the negative. Six semantic differences are expressed in the affirmative, and only three in the negative. While the affirmative affords a distinction between irrealis (*anthi÷*) and non-past (*arà÷*), this distinction is not found in the negative, where *thama÷* is used for both these functions (next to future marking). Furthermore, the concepts of anterior, infinitive and imperative are distinguished in the affirmative, but conflated in the negative (*jamà-*). For simultaneous tense finally, it is unclear how it is expressed in the negative, which is why it is not put within one of the domains of the morphemes.

6.2.4   Inflectional suffixes

Both inflectional suffixes in SLM are imperative markers. *-la* is a polite imperative, while *-de* is an impolite imperative.
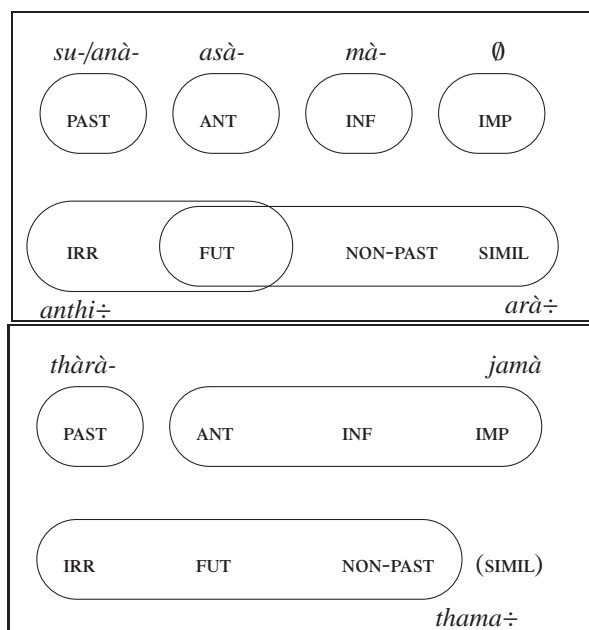
Figure 6.1: Semantic map of affirmative (left) and negative (right) verbal (quasi-)prefixes

6.2.4.1 imperative *-la*

This suffix is an imperative marker which attaches to a verbal host. It is used for com-
mand clauses ⊙→17.1.5, p. 699. It has no influence on vowel lengthening processes.
This morpheme can be used on its own on a verb (123), or be combined with the
imperative particle *mari*, which reinforces its meaning (124).

(123)   *Allah **diyath-la** inni pompang pada dhaathang aada.*
        Allah watch-IMP PROX female   PL   come   exist
        'Almighty, see, that woman has come!' (K061019nar02)

(124)   a. *Saayang se=ppe thuan **mari**   laari-**la**.*
            love   1s=POSS sir   come.IMP run-IMP
            'Come my beloved gentleman, come here.'

        b. *See=samma kumpul **mari**   thaaňdak-**la**.*
            1s=COMIT   gather come.IMP dance-IMP
            'Come and dance with me.' (N061124sng01)

This morpheme is reasonably frequent. It is cognate with Std. Malay *lah,* used for
imperatives and emphasis.

6.2.4.2 impolite imperative *-de*

This morpheme is an impolite imperative suffix. This morpheme was only discov-
ered at the evening of the very last day of elicitation in an informal setting without
recording device. It is not found in the corpus, which might be due to its very impolite
connotations. The following example is the one which was given to illustrate its use.

(125)   *Pii! ... pii!! ... pii-de!!!*
        go ... go   ... go-IMP.IMPOLITE
        'Go! Go now! Bugger off!!!' (not on recordings)

6.2.5 Derivational affixes

Sri Lanka Malay has two category-changing affixes: the nominalizer *-an* and the
causativizer *-king*. In the verbal domain, there is a further involitive prefix *kànà-*,[23]

---

[23]There is a possibility of the existence of a prefix used to derive accidental verbs, *tha-*. This was
overheard several times in the string *as-tha* , but at first misinterpreted as an allomorph of the conjunctive
participle *asà-*, instead of the probably correct parsing *asà+tha*. This prefix has been found with the verb
bunthur 'knock' and other highly transitive verbs. The morpheme might be related to the Standard Malay
prefix *ter-* used to derive "stative, accidental or habilitative verbs in [Standard Malay]" (Paauw (2004),
also see Adelaar (2005), and Adelaar (1985:181) on Jakartanese *tə-*). As far as I can remember, in all the

while in the nominal domain, we find a special suffix for kin of third person. There is notably no affix deriving adjectives, but the clitic =*ke* ⊡→6.4.6.3, p. 369 could be used for this function.

6.2.5.1 nominalizer *-an*

This suffix can be used to derive nouns from verbs like *maakang* 'eat' (*makanan* 'food') or adjectives like *maanis* 'sweet' (*manisan* 'sweets') (cf. Adelaar 1991:28) ⊡→7.3.1, p. 406. The semantic result is not predictable. It can be an item (*manisan, makanan*) or an action like *banthuan* 'help'. When attached to a stem ending in front vowel, *y* may be inserted (*pake-y-an* 'dress'). After a stem ending in *a*, *h* must be inserted (*suka-h-an* 'liking'). The result of the derivation is necessarily trisyllabic because all verbs are at least disyllabic.[24]

Words derived with *-an* never lengthen the penultimate vowel. This is not surprising, since there is normally enough material in the three syllables to construct a bimoraic foot even under extrametricality of the final syllable ($ma_\mu.ka_\mu.$ <nan>) (see Section 4.5.2, p. 126). A special case are the stems with a schwa in the initial syllables like /səgar/[$sı_\mu g_\mu.$<gar>]'healthy'. These normally raise the schwa to generate one mora, and geminate the consonant, to generate another one. When derived with *-an*, the gemination tends to go away, while the schwa remains raised, thus [$sı_\mu.ga_\mu.$<ran>]'health'. Note that the /g/ is resyllabified in the derived word.

This affix is directly inherited from the historical affix with the same form and function. There are also the related circumfixes *pàr-...-an* and *ka-...-an*, which are also nominalizers. These are no longer productive in Sri Lanka Malay and are treated as allomorphs of *-an* here.

The use of *-an* can be seen in several instances in (126) where the verbs *aajar* 'teach' and *thaksir* 'think' are derived by this suffix. The last form, *plajaran* 'lesson', bears resemblance to the expected form **blajaran*, derived from *blaajar* 'learn', but the initial consonant is voiceless. This is thus an irregular nominalization. The reason for this irregularity is the historic derivation *pe-...-an*, which has reflexes in Modern Indonesian *pelajaran*. This derivation by *pe-...-an* also has different semantics, since we are not dealing with acquiring knowledge, but with imparting knowledge.

---

contexts where *astha-* occurred, a reading as 'accidental' would have been semantically plausible. From the above description it should be clear that this morpheme, and its relations to *kìnna* ⊡→5.1.5.8, p. 187 and *kànà-,* ⊡→6.2.5.3, p. 312 need further research.

[24]With the exception of *pii* 'go', but this verb cannot combine with *-an*.

(126)   a. *Derang=pe **ajar-an**   bannyak baaye.*
        3PL=POSS   teach-NMLZR much   good
        'Their teaching was very good.'

        b. *Derang=pe **thaksir-an**=le bannyak baaye.*
        3PL=POSS   think-NMLZR   much   good
        'Their thoughts were very good.'

        c. *Derang=pe **plajaran** bannyak baaye.*
        3PL=POSS   lesson   much   good
        'Their lessons were very good.' (K051213nar03)

Nominalized verbs may continue to govern arguments, as example (127) shows.

(127)   *Lorang=nang **see=yang** ingath**-an**=si.*
        2PL=DAT   1S=ACC   remember-NMLZR=INTERR
        'Do you remember me?' (K070000wrt04)

In this example, the verb *iingath* 'think' governs the accusative, and this remains even after nominalization. This contrasts with English, where arguments of nominalized verbs obligatorily take the genitive (*Do you have remembrance/memories **of** me*).

An interesting case of the interplay of word classes and the nominalizer is the word *jaalang*. *Jaalang* can either be used as a verb with the meaning 'to walk', as in (128). It can also be used as a noun with the meaning 'street', which is obviously related semantically, but not enough to warrant an analysis as conversion because the meaning of 'street' is much more restricted than the meaning of 'walk'. Finally, the verb *jaalang* 'walk' can be derived by *-an* to yield the word *jalangan* 'trip' (130).

(128)   a. *Katugastota=nang train=ka   asà-dhaathang.*
        Katugastota=DAT   train=LOC CP-come
        'I arrived in Katugastota by train.'

        b. *Katugastota asàduuduk St.Anthony's=nang **arà-jaalang***
        Katugastota from   St.Anthony's=DAT   NON.PAST-walk
        '(Then) I used to walked from Katugastota (train station) to St. Anthony's (school).' (K051201nar02)

(129) *kàthaama kithang kiccil muusing=ka, inni* **Peeradheniya Jaalang**=*ka*
       before   1PL   small time=LOC   PROX Peradeniya   street=LOC
       samma an-aada   mlaayu.
       all   PAST-exist Malay
       'Before, when we were children, it was all Malays here on Peradeniya Rd.'
       (K051222nar04)

(130) **jalang-an**   *hatthu arà-pii   vakthu*
       walk-NMLZR INDEF   NON.PAST-go time
       'when we go on a trip' (K051213nar06)

In rare cases can *-an* be found on nouns, like *raja-han* 'king'+'NMLZR'='government'. This is another use of *-an*, also found in Standard Malay, and indicates 'collectivity' or 'similarity' when attached to nouns, according to Adelaar (1985:193). This meaning seems to be at hand here as well, where a government can be seen as a collection of kings, or similar to a king.

There is one instance of a nominalization after inflection, namely *thradahan* 'deprivation', which is composed of the negative quasi-prefix *thàrà÷*, the existential *aada,* and the nominalizer. The non-negated form *adahan* 'possession' also exists. One could argue that the negation takes place after derivation, however, *th(à)rà-* is not a morpheme which can attach to nouns, so that *th(à)rà-* must have been joined with *a(a)da* before the derivation.

6.2.5.2   causativizer *-king*

This causative morpheme can derive verbs from adjectives and verbs ⊡→7.3.2, p. 406. It has an alternative realization *-kang*, which seems to depend on phonological environment, but also on speakers. Adelaar (1991:29) proposes that this morpheme might either be a reflex of the transitivizing suffix *-kən* found in non-standard forms of Malay and Javanese, or derived from *bikiŋ* 'to do, make'. It might also be the case that *-kang* is related to *-kən* and *-king,* to *biking*.

In most cases, both *-king* and *-kang* are possible in SLM (cf. Saldin 2001:60). The cases where some speakers feel only one is possible are highly unsystematic and do not lend themselves to an easy categorization.

One and the same speaker can use then one allomorph, then the other with the same base in comparable contexts. The same writer uses *encoking* 'fool(V)' consistently in one text (K070000wrt02) and *encokang* 'fool(V)' in another one (K070000wrt05).

(131) *[Andare=yang mà-**enco-king**=nang] raaja su-biilang [itthu paasir katha].*
       Andare=ACC INF-fool-CAUS=DAT   king PAST-say DIST sand QUOT
       'To fool Andare, the king said that this was sand.' (K070000wrt02)

(132) *Kanabisan=ka=jo duva oorang=le anà-thaau ambel [Andare duva*
       last=LOC=EMPH   two   man=ADDIT PAST-know take   Andare two
       *oorang=yang=le asà-**enco-kang** aada] katha.*
       man=ACC=ADDIT CP-fool-CAUS   exist EMPH
       'At the very end, both women understood that Andare had fooled both of
       them.' (K070000wrt05)

Table 6.3 gives the instances of verbs derived by *-king* and *-kang* in the corpus,
excluding loanwords.

A special case is the word for 'to show' (*cunjikang, thunjiking*), which has a
causative morpheme, but whose base is synchronically not transparent. It is not clear
whether a swap of *king* for *kang* in this pair would be possible.

As for its phonological status, *-king* is in a middle position between an affix and
a clitic. Just like affixes, it is selective with regard of its host and its position. *-king* cannot be preposed or otherwise moved around in the clause, it cannot occur
on its own and it can only attach to adjectives and verbs. For some speakers, *-king* has the same metrical properties as *-an*, while for others *-king* is not parsed into the
same phonological word as its stem, which has as a consequence that the penultimate
syllable of the stem is lengthened. See ⌑→4.5, p. 120 for a discussion. On a diachronic
note, it appears that the original clitic status of *-king* is eroding and *-king* becomes
grammaticalized as an affix.

The changing status of *-king* can also be retraced in the literature. My informants
accept both short and long vowels, and produce both. On the other hand, Bichsel's
informants did not have *-king* within the word domain of the stem (i.e. they had vowel
lengthening), while Tapovanaye's informants had *-king* within the word domain of the
stem, preventing lengthening. This suggests that within the speaker population, there
are people who assign one status to *-king*, another group assigns the other status, and
a third one accepts and produces both.

As for its hosts, *-king* can be used on adjectives as in (133)(134), with intransitive
verbs as in (135)(136) and with transitive verbs as in (137)(138).

(133) *Itthuka   asà-thaaro, itthu=yang arà-**panas**_{adj}-**king**.*
       DIST=LOC CP-put   DIST=ACC NON.PAST-hot-CAUS
       'Having put (it) there, you heat it.' (B060115rcp02)

-king	-kang
abbis-king 'finish+CAUS=fulfill (a wish)'	*bava-kang* 'bring+CAUS=have s.o. bring'
bale(k)-king 'turn (intr)+CAUS=(re)turn (tr)'	*bavung-kang* 'rise+CAUS=lift'
biilang-king 'say+CAUS=make say'	*binthi-kang* 'stop (intr)+CAUS=stop (tr)'
blajar-king 'learn+CAUS=teach'	*buunan-kang* 'kill+CAUS=have killed'
buunung-king 'kill+CAUS=have s.o. killed'	*dudu-kang* 'sit+CAUS=put down'
enco-king 'fooled+CAUS=fool (V)'	*enco-kang* 'fooled+CAUS=fool(tr)'
jadi-king 'become+CAUS=grow, run (a business)'	*gaañdas-kang* 'connected+CAUS=connect'
kasi-king 'give+CAUS=make give'	*gijja-kang* 'make+CAUS=have s.o. make'
kaving-king 'marry+CAUS=marry off'	*hina-kang* 'embarrassed+CAUS=embarrass'
kubuur-king 'buried+CAUS=bury'	*iingath-kang* 'think+CAUS=recall'
mandi-king 'bathe+CAUS=wash'	*livath-kang* 'exceed+CAUS=increase'
mathi-king 'dead+CAUS=kill'	*jalang-kang* 'walk+CAIS=run (a business, country)'
salba-king 'escape+CAUS=liberate'	*mathi-kang* 'dead+CAUS=kill'
thau-king 'know+CAUS=inform'	*korban-kang* 'sacrifice(n)+CAUS=sacrifice'
thumpa-king 'spill (intr)+CAUS=spill (tr)'	*kubuur-kang* 'buried+CAUS=bury'
mlidi-king 'boil (intr)+CAUS=boil (tr)'	*luppas-kang* 'leave+CAUS=leave sth. behind'
panas-king 'hot+CAUS=heat'	*picca-kang* 'broken+CAUS=break'
thurung-king 'descend+CAUS=lower'	*punnu-kang* 'full+CAUS=fill'
	rubbus-kang 'boil (intr)+CAUS=boil (tr)'
	salba-kang 'escape+CAUS=liberate'
	sakith-kang 'sick+CAUS=hurt (tr)'
	vaasil-kang 'blessed+CAUS=bless'

Table 6.3: The use of *-king* and *-kang* on different bases in the corpus.

(134)   *Mà-mathi-king, mà-mathi_{adj}-king=nang, siithu=jo   anà-baapi.*
        INF-dead-CAUS   INF-dead-CAUS=DAT   there=EMPH PAST-bring
        'It was there that (they) brought (him) to make (him) dead.' (K051206nar02)

(135)   *Inni=ka   inni daalang=ka kithang aayer masà-mlidi_{intr}-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)

(136)   *Baaye meera caaya kapang-jaadi, thurung_{intr}-king.*
        good   red   colour when-become, descend-CAUS
        'When [the food] has turned to a nice rose colour, remove (it) [from the fire].'
        (K060103rec02)

(137)   *De   laaye hathu nigiri=nang anà-baapi, buunung_{tr}-king=nang.*
        3s.IMPOLITE other INDEF country=DAT PAST-bring kill-CAUS=DAT
        'They brought him to another country to have him executed.' (K051206nar02)

(138)   a. *Suda maven=subbath see anà-resign.*
            so   son=because 1s PAST-resign
            'So I quit because of my son.'

        b. *Maven masà-blajar_{tr}-king=jona.*
            son   must-learn-CAUS=EMPH=PHAT
            'Well,   somebody   has   to   teach   him,   isn't   it.'
            (B060115prs21,K081105eli02)

Besides the normal use of *-king* on verbs or adjectives, there are two less common bases in the corpus, a noun (139) and a particle (140). In (139), the noun *kafan* 'shroud' is derived with *-king* to yield *kafan-king* 'enshroud'.

(139)   *Spaaman=yang asà-kafan-king spaaman=yang sithu=ka   nya-kubuur-king.*
        3s.POLITE=ACC   CP-shroud-CAUS 3s.POLITE=ACC there=LOC PAST-buried-CAUS

        'The body was wrapped in cloth and then the body was finally buried.'
        (B060115nar05)

In (140), the negative particle *thraa* is used as a base of *-king* to yield *thraa-king* 'get rid of'.

	-king		-kang
capture-king	serve-king	admit-kang	organize-kang
bomb-king	dub-king	attack-kang	mix-kang
conduct-king	smash-king	boil-kang	introduce-kang
perform-king	break-king	celebrate-kang	print-kang
play-king	introduce-king	follow-kang	execute-kang
select-king	bat-king	force-kang	represent-kang
thanks-king	bowl-king	sacrifice-kang	mix-kang
set-king	spend-king	wrap-kang	curse-kang
issue-king			set-kang

Table 6.4: The use of *-king* and *-kang* on different loanwords in the corpus.

(140) *Itthu=nang, kithang arà-thraa-king, kithang=pe Seelon=pe*
       DIST=DAT 1PL NON.PAST-NEG-CAUS 1PL=POSS Ceylon=POSS
       mosthor=nang.
       manner=DAT
       'Therefore, according to the Sri Lankan way, we are getting rid of that.'
       (B060115rcp01,K081103eli02)

-king/-kang is also the standard way to integrate English loan verbs. Note that there is no causation involved, so *serve-king* means 'to serve' and not 'make serve'. The following example shows extensive use of this pattern in the explanation of the rules of sepaktakraw, which is a rather recent sport in Sri Lanka, and the terminology used (serve, break, set, smash) is borrowed from English volleyball terminology.

(141) a. *Itthu=yang* **serve-*king*** *atthyang subla=dering arà-**break-*king***.
          DIST=ACC serve-CAUS other side=ABL NON.PAST-break-CAUS
          'Serve the ball, on the other side they break the ball.'

       b. **Break-*king****=apa, kithang arà-**set-*king***.
          break-CAUS=after 1PL NON.PAST-set-CAUS
          'After breaking, we set the ball.'

       c. **Set-*king****=apa arà-**smash-*king***.
          set-CAUS=after NON.PAST-smash-CAUS
          'After setting, we smash the ball.' (N060113nar05)

Table 6.4 shows the loan verbs found in the corpus which bear this morpheme.

6.2.5.3   involitive *kànà-*

Kànà- is a morpheme used to derive involitive verbs from volitive verbs ⊡→7.3.3, p. 407. The verbs thus derived govern the dative. This morpheme is not very frequent, and it was difficult to get solid information about its use. The following two examples show the schematic use.

(142)   *Se naasi arà-maakang.*
        1s rice   NON.PAST-eat
        'I eat rice.' (K081104eli03)

(143)   *Se=dang naasi arà-**kànà**-maakang.*
        1s=DAT   rice   NON.PAST-INVOL-eat
        'I eat rice without wanting it/I compulsively eat rice/I was forced to eat rice by something beyond my control.' (K081104eli03)

Some semantically primary verbs, like *bangkas* 'sneeze' can also take *kànà-*. In this example, the verb does not seem to be derived from a volitive verb.

(144)   *Sedang arà-**kànà**-bangkas*
        1s.DAT   NON.PAST-INVOL-sneeze
        'I sneeze.' (K081104eli03)

It appears that *kànà-* is used to emulate the Sinhala involitive conjugation class in *-enavaa*, which also governs the dative (Gair 1971). The semantics of the verbs of this class in Sinhala are treated in detail in Gair (1971) as well, which might be an interesting point of departure for the semantic analysis of *kànà-*.

Kànà- can also be used with intransitive verbs (145). In this case, the dative marking can be left out.

(145)   *Incayang su-kànà-pii.*
        3s.POLITE PAST-INVOL-go
        'He left inadvertently/He left without wanting it/He left because of something beyond his control.' (K081104eli03)

It is possible that this prefix is related to the vector verb *kìnna*, which has very similar semantics. The difference is that *kànà-* changes the argument structure of the verb, which governs the dative after derivation with *kànà-*. *Kìnna* on the other hand, leaves the argument structure of the verb unaffected ⊡→5.1.5.8, p. 187. This suggests that these two morphemes are not allomorphs. *kànà-* changes the morphosyntactic properties of a verb, while *kìnna* only emphasizes a certain semantic aspect. The fact that these two morphemes are not allomorphs can also be seen by their possibility to

co-occur, as in (146).

(146) *Itthu haari=ka=jo aanak pompang duuva=nang hathu duuri*
DIST day=LOC=EMPH child female two=DAT INDEF thorn
*pohong=nang jeeňǧgoth=yang anà-**kànà**-daapath **kìnna** hathu Aajuth*
tree=DAT beard=ACC PAST-INVOL-get strike INDEF dwarf
hatthu=yang su-kuthumung.
INDEF=ACC PAST-see
'On the same day, the two girls saw a dwarf whose beard had got stuck in a
thorn tree.' (K070000wrt04)

In (146), *kànà-* marks the involitive nature of getting stuck, while *kìnna* adds a
shade of adversative surprise. It is nevertheless likely that both these morphemes are
cognate to the Std. Malay *kəna,* which can be used for adversative passives (Chung
2005).

6.2.5.4 kin term *-yang*

This is a derivational affix which can only attach to kin terms. The kin relation may not
involve the speaker or the addressee. It must not be confounded with the homophonous
direct object marker, the cliticized postposition =*yang* ⊡→6.4.4.1, p. 328. The former
is a derivational affix which indicates the kin relation, while the latter is an enclitic
postposition indicating the semantic role of patient. The affix can occur in places
where the postposition could not occur, e.g. with intransitives (147). Furthermore,
the suffix can be used on referents marked for other semantic roles, like in example
(148), where the father is marked for dative with the postposition =*nang*. It would not
be possible to stack the postposition =*yang* and the postposition =*nang*. Finally, the
affix *-yang* and the postposition =*yang* can stack, which clearly shows that they are
different morphemes (149).

(147) *[Bapa-yang]ₐg arà-dhaathang.*
father-KIN.3 NON.PAST-come
'His/her father is coming.' (K081103eli04)

(148) *Kake-yang bapa-**yang=nang** su-maaki.*
grandad-KIN.3 father-KIN.3=DAT PAST-scold
'His grandad scolded his father.' (K081103eli04)

(149) *Kake-yang bapa-**yang=yang** su-buunung.*
grandad father-KIN.3=DAT PAST-kill
'His grandad killed his father.' (K081103eli04)

Naturalistic examples of -*yang* being combined with a postposition are (150) and (151), where the genitive =*pe* follows the string *yang*, thereby excluding the possibility to analyze it as a postposition indicating patient.

(150)   *inni umma-**yang=pe**   dhaatha*
        PROX mother-KIN.3=POSS elder.sister
        'this mom's elder sister' (K051220nar01)

(151)   *derang=pe mama-**yang=pe** ruuma*
        3PL=POSS   uncle-KIN.3=POSS house
        'their uncle's house' (B060115nar05)

The use of -*yang* with intransitives in naturalistic speech is given in (152), where the only arguments of the two occurrences of *niiňǧgal* 'die' carry the string *yang*, and both are terms of kin (*biini* 'wife' and *umma* 'mother'). The verb *niiňǧgal* 'die' never assigns accusative case to its argument, so that the instance of *yang* we are dealing with here must be the affix -*yang* and not the postposition =*yang*. A further indication of this is the short vowel in *bini*, which can sometimes occur with affixes, but not with postpositions. For this see the analogous discussion of vowel length with -*king* above (p. 308.)

(152)   *Ithukang bini-**yang** niiňǧgal=le umma-**yang**   niiňǧgal=le aanak pada*
        Then   wife-KIN.3 die=ADDIT   mother-KIN.3 die=ADDIT   child PL
        caari   dhaathang.
        search come
        'So then the wife died and the mother died, still the children will come in search.' (K051220nar01,K081105eli02)

Sometimes, it is difficult to decide whether a given instance of *yang* is the kin affix or the accusative postposition. The reason for this is that both markers are optional. When a kin term is used as a patient, we cannot know whether the string *yang* indicates the kin relation of the person or its status as an undergoer without more context. An example of that is (153) where the father is introduced. Introduction is not a violent activity, making the use of the accusative marker unlikely, but still possible ⊙→15.1.2.2, p. 563. On the other hand, the speaker does refer to the father of a third person, so that the kin affix would be a possibility. Given the long vowel in *baapa*, the accusative reading is more likely, though.

(153)   *Blaakang see Zeenath=pe   baapa-**yang/=yang** nya-introduce-kang*
       After   1s   Zeenath=POSS father-kin/=ACC   PAST-introduce-CAUS
       Sebastian=nang.
       Sebastian=DAT
       'After that I introduced Zeenath's daddy to Sebastian.' (B060115cvs01)

6.2.6   Numeral affixes

In the numeral domain finally, there are three affixes, for '-ty', '-teen' and '-th'.

6.2.6.1   *-blas* '-teen'

This suffix is used to derive the numbers between 11 and 19. It is parsed within the same phonological word as the stem, so that no vowel lengthening takes place in *dhla(*a)pan* and *li(*i)ma* in the following two examples.

(154)   *Suda se=ppe thuuva anak klaaki asàdhaathang **dhlapan-blas** thaaun.*
       so   1s=POSS old   child boy   COPULA   eight-teen   year
       'So, my eldest son is eighteen.' (K060108nar02)

(155)   *Hathu **lima-blas** thaaun=nang duppang.*
       INDEF five-teen year=DAT   before
       'About fifteen years before.' (K061026prs01)

6.2.6.2   *-pulu* '-ty'

This suffix is used to derive the multiples of ten between 20 and 90. It is also parsed in the same phonological word domain as the stem, so that no lengthening takes place. Note the two occurrences of *du(u)va* in the following example, where one occurs with *-pulu* and is not lengthhened, while the other one occurs on its own and is lengthened. Also note the short *i* in *limapulu*.

(156)   ***Duva-pulu** ìnnam riibu   umpath raathus lima-pulu **duuva** votes*
       two-ty   six   thousand four   hundred five-ty   two   votes
       incayang=nang anà-daapath.
       3s.POLITE=DAT   PAST-get
       'He got 26,452 votes.' (N061124sng01)

6.2.6.3   ordinalizer *ka-*

Ka- is a prefix used to derive ordinals from cardinal numbers (Adelaar 1991:28). The vowel can be /a/ (157) or schwa (158).

(157)   *Se asdhaathangpa kitham=pe femili=ka   **ka**-duuva aanak.*
      1s copula   1pl=poss   family=loc ord-two   child
      'I am our family's second child.' (K060108nar01)

(158)   ***Kà**-thiiga member inni kitham=pe legislative council=ka.*
      ord-three member prox 1pl=poss   legislative council=loc
      'He was the third member in this legislative council here.' (N061031nar01)

    Ka-, like all prefixes, is ignored when calculating the moraic weight. This means that penultimate vowel lengthening takes place as normal. This is shown in *(ka-)thuuju* 'seven(-th)' in (159).

(159)   *Se karang thuuju, ka-th**uu**ju thaaun=ka.*
      1s now   seven, ord-seven year=loc
      'I am now in 7th grade.' (K060108nar01)

    The numerals *ùmpath* 'four' and *ìnnam* 'six' are realized by some speakers without the initial schwa, i.e. these numbers start with a syllabic nasal([m̩.pat̪, n̩.nam]). This nasal is *not* resyllabified when *ka-* is prefixed, so that the resulting word has three syllables ([ka.m̩.pat̪, ka.n̩.nam]).

(160)   *Se Seelon dìkkath karang **ka-(ù)mpath** generation.*
      1s Ceylon at   now   ord-four   generation
      'I am in Ceylon for the fourth generation.' (K060108nar02)

6.3   Simple clitics

Simple clitics are cliticized variants of words that can otherwise occur as a free form. In SLM, the cliticized forms of the modal particles, i.e. *bolle=, bol=, bàr(à)=* for *boole* 'can' and *mau=* for *(ka)mau(van)* fall under this category. Furthermore, the proclitic form *kal=* of the conditional particle *kalu* can also be analyzed as a simple clitic. Depending on the analysis of the quasi-prefixes, the number of simple clitics could increase. If the quasi-prefixes are regarded as free words because they can take the enclitic emphatic marker *=jo*, then the reduced forms *ar, at(t)hi, thi, mas(s)a, mas* and *thàr* would also have to be analyzed as simple clitics.

6.4 Bound words

SLM bound words are defined as words that cannot occur on their own, nor do they have a free form. They obligatorily need a host to which they can attach. What distinguishes them from affixes is that they are not selective with regard to their host. In SLM, all bound words attach to the right of the host, with the exception of the indefiniteness clitic *(=)atthu(=)* which can attach at both sides. No bound word interferes with the normal vowel lengthening pattern of the stem.

Bound words which attach to the NPs are

- the indefiniteness clitic ⊡→6.4.1, p. 319,

- the plural clitic ⊡→6.4.2, p. 323,

- postpositions ⊡→6.4.4, p. 327

- prepositions ⊡→6.4.5, p. 358.

Bound words which can attach to NPs, but also to predicates or clauses are

- the Coordinating Clitics ⊡→6.4.6, p. 359,

- the informations structure clitics ⊡→6.4.7, p. 378,

- and the evidential clitic ⊡→6.4.8.1, p. 386.

The last clitic, the quotative *katha*, attaches to the last word of an utterance ⊡→6.4.9.1, p. 389.
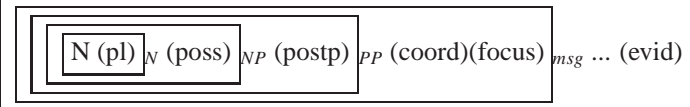
Bound words hail from a wide variety of semantic domains. Because of this, they do not compete for the same slots in the utterance and can be stacked quite well. The following example shows three clitics stacked on *Malaysia*.

(161) *Seelon=pe=lle   mlaayu=nang* Malaysia=***pe=lle=nang***   *konnyon*
       Ceylon=POSS=ADDIT Malay=DAT   Malaysia=POSS=ADDIT=DAT few
       different.
       different
       'Sri Lankan Malay and Malaysian Malay are a little bit different.'
       (B060115cvs09)

The following constructed example shows the theoretical maximum of seven clitics. This is highly unnatural.

(162) *?[incayang anà-pii [aanak =pada =pe   Ø(pàrhaal) =subbath] =le*
       3s.POLITE   PAST-go child   =PL   =POSS   =because =ADDIT
                                    deriv   term   prop.cont infstr

       =jo   =kiyang   =katha]   see su-biilang.
       =EMPH =EVID   =QUOT   1s PAST-say
       infstr   interpersonal utterance
       'I said that it was reportedly also because of the children's [problems] that he
       left.' (constructed) (K081103eli04)
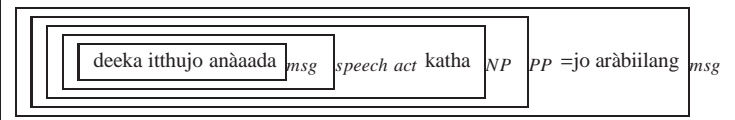
Note that the first three clitics (within the brackets) attach to N(P) and carry propo-
sitional content The following two (=*le* and =*jo*), attach also to the NP, but indicate no
propositional content but discourse function (additive focus and general focus). The
next clitic =*kiyang* indicates evidential modality, a speaker attitude, and has thus bear-
ing on the interpersonal leel, not indicating either propositional content or discourse
function, but the speaker's relation to the proposition he is using in his speech act. This
can be seen from the fact that *anàpii* 'left' can be placed between =*jo* and =*kiyang*.
Finally, *katha* is used to embed the whole reported utterance in a matrix sentence. We
can illustrate this as in (163), where we see at what levels the different clitics attach.

(163) | [[N (pl)]N (poss)]NP (postp)]PP (coord)(focus)]msg ... (evid)]spch act (quot)]

This schema should allow us to make neat predictions about the possible orders of
clitics we can find. A clitic on an outer layer should never precede a clitic on an inner
layer. This prediction is borne out. The only exception to this is the focus clitic =*jo*,
which can also be found attached to an utterance, i.e. at the rightmost position. An
example is given in (164). Note that there are two occurrences of =*jo* in this example,
we are only interested in the last one.

(164) *[Dee=ka   itthu=jo   anà-aada **katha**]ill=**jo** arà-biilang*
       3s.IMPOLITE=LOC DIST=EMPH PAST-exist QUOT=EMPH non.past-say .
       'It is said that it was that what he had with him.' (K051206nar02)

However, the utterance *deeka itthujo anàaada* 'it is that what he had with him' is
embedded as an argument of *biilang* 'say' in the superordinate clause by *katha*, so that
=*jo* in this case attaches to the argument as we would expect, and only indirectly to
the utterance. This is shown in (165).

(165) | [[[deeka itthujo anàaada]msg]speech act katha]NP]PP =jo aràbiilang]msg]spch act |

6.4.1 Indefiniteness clitic

The indefiniteness marker *hatthu* can be realized as *hat(t)hu* or *at(t)hu,* and can be either proclitic or enclitic. Occasionally, speakers use both the pro- and the enclitic, which yields double morphosyntactic encoding of semantic content. Care must be taken to distinguish the free numeral *sat(t)hu* ⊡→5.10, p. 242 from the clitic indefinite marker, especially because the numeral has two allomorphs *hatthu* and *atthu*, which are homophonous to the two forms of the clitic forms. However, the form with initial *s-* is not possible for the clitic, so that numerals can be identified in that manner.

Example (166) shows the different realizations of the clitic. In this section, the cliticized nature of *hatthu* will be indicated by an equal sign (=). This is not done in other sections for reasons of legibility.

(166) *Se **atthu**=aade,   se **hatthu**=aade.*
 1s INDEF=younger.sibling 1s INDEF=younger.sibling
 'I am a younger sibling, I am a younger sibling.' (K061120nar01)

The following example shows the use of a numeral *satthu,* identifiable by the *s-*. This *s-* could not occur in the example above, but the two forms used in (166) could very well also occur in (167).

(167) *Mlaayu=dring **satthu** oorang=jo se.*
 Malay=ABL   one   man=EMPH 1s
 'I am one of those Malays.' (K060108nar02)

The following examples to show the proclitic use (168), the enclitic use (169), and the doubled use (170) when referring to the concept of story.

(168) ***Hathu**=oorang=pe muuluth=dering **hathu**=criitha kal-dhaathang.*
 INDEF=man=POSS   mouth=ABL   INDEF=story   when-come
 'When a story comes out of a man's mouth.' (B060115prs15)

(169) *Giini   criitha=**hathhu**=le aada.*
 like.this story=INDEF=ADDIT exist
 'There is also a story like that.' (K051206nar07)

(170) *Itthu **atthu**=story=**atthu**.*
 DIST INDEF=story=indef
 'This is a story.' (B060115nar05)

Since the concept the clitic attaches to is the same in all three instances, lexical specification for pro- or enclisis is not a viable analysis. It is probable that the pref-

erence for one or the other is idiolectal, but speakers happily accept any of the two versions when asked.

The following utterances are three more instances of double marking on a loan word.

(171)   *Kitham=pe **atthu=**three-tonner**=atthu** aada, duppang=ka.*
        1PL=POSS   INDEF=three-tonner=INDEF   exist, front=LOC
        'There was a three-tonner of ours at the front.' (K051206nar16)

(172)   *Se **hatthu=butthul** moderate Muslim=atthu.*
        1s COPULA   INDEF=very moderate   Muslim=indef
        'I am a very moderate Muslim.' (K051206nar18)

(173)   *Itthu abbisdhaathang **hatthu=**traditional food**=hatthu**.*
        DIST   COPULA   INDEF=traditional   food=indef
        'That is one traditional food.' (K061026rcp01)

While double marking is most often found on loan words, there are also instances where it is found on native words, like *maccan* 'tiger' in (174).

(174)   *Sithu=ka **hatthu**=maccan=**hatthu** duuduk aada.*
        there=LOC INDEF=tiger=INDEF   stay   exist
        'A tiger stayed there.' (B060115nar05)

The following example shows that, while the use of double marking is common on loan words, it is not obligatory. The loanword *job* receives double marking, while the loan word *application* does not.

(175)   *Kithang=nang **hatthu=**job**=hatthu** mà-ambel=nang kithang=nang*
        1PL=DAT   INDEF=job=INDEF   INF-take=DAT   1PL=DAT
        ***hatthu=**application *mà-sign kamauvan vakthu=nang=jo kithang*
        INDEF=application INF-sign want   time=DAT=EMPH 1PL
        arà-pii   inni politicians pada dìkkath=nang.
        NON.PAST-go PROX politicians PL   vicinity=DAT
        'When we want to take a job, when we want to sign an application, we approach these politicians' (K051206nar12)

The fact that this double marking occurs mainly on loanwords can be explained by Sinhala influence, where English loanwords are integrated by means of the numeral *eka* 'one', like *kar eka* 'the car' or *bas eka* 'the bus' (Karunatillake 2004). Note the definite article in the English translation. If indefiniteness is to be marked on these

words, the affix *-k* (etymologically related to *eka*) is used (*kar eka-k*). Both *eka* and *-k* have *hatthu* as their SLM equivalent. Therefore this morpheme must occur twice in the SLM sentence if it was to copy the Sinhala pattern. In order to avoid repetition, one is placed before the noun, the other one after it. This is no problem, since both slots are available, as examples (168) and (169) show.[25] Finally, the fact that loan words are integrated can be seen by the fact that definite referents also can be marked with *hatthu* when they are loanwords. In example (176), the wedding is made definite by the demonstrative *inni*, still, *hatthu* is present.

(176) ***Inni*** mock wedding=***hatthu*** *mas-gijja.*
     PROX mock wedding=INDEF   must-make
     'I have to do this mock wedding.' (K060116nar10)

Similar arguments can be made about the following examples, where the deictic *ini* and *hatthu* cooccur on the same NP. *Atthu* in this case does not mark indefiniteness, but rather signals the loan word.

(177) *Suda deram* ***inni*** *political promise=****hatthu*** *derang=eng-kaasi 1958=ka.*
     so   3PL   PROX political promise=INDEF 3PL=PAST-give   1958=loc
     'So they made this political promise in 1958.' (N060113nar02)

(178) *Indonesia=dering Sri Lanka=nang kithang=pe ini   mlaayu pada*
     Indonesia=ABL   Sri Lanka=DAT 1PL=POSS   PROX Malay PL
     asà-dhaathang ***ini***   *Malay regiment=****atthu***.
     CP-come   PROX Malay regiment=INDEF
     'Our Malays came from Indonesia to Sri Lanka in this Malay regiment.'
     (G051222nar03)

(179) *Kàthaama* *police* ***oorang nya-maathi*** ***inni*** *terrorist=****hatthu***=*dering.*
     before   police man   PAST-dead PROX terrorist=INDEF=ABL
     'So the first police man died by (the hands of) this terrorist.' (K051206nar02)

The main use of this morpheme is to indicate indefiniteness, which is normally done to signal the introduction of new referents ⊙→16.1.1, p. 678. Furthermore, it is also used to mark categorial reference ⊙→15.1.4.3, p. 589. The latter use can be exemplified by example (180), repeated from above. The speaker has no specific referent for *oorang* 'man' in mind, any arbitrary referent will do. This is categorial reference. The same is true for the story.

---

[25]Note that in example (171), indefiniteness is marked between the host and the possessive modifier, which violates some claimed universals (Greenberg 1963, Hawkins 1994, Rijkhoff 2002). In (172), however, the order is reversed, which means that both orders are possible. This will be discussed in more detail in the section on NP-structure ▯→8.1.16, p. 440.

(180) ***Hathu**=oorang=pe muuluth=dering **hathu**=criitha kal-dhaathang.*
       INDEF=man=POSS   mouth=ABL   INDEF=story   when-come
       'When a story comes out of a man's mouth.' (B060115prs15)

This can be distinguished from the introduction of a specific referent, which is
expected to be unknown to the speaker as in (181), as well repeated from above.

(181) *Giini   criitha=**hatthu**=le aada.*
       like.this story=INDEF=ADDIT exist
       'There is also a story like that.' (K051206nar07)

A further use is to make nouns fit for use as predicates ⊡→10.4, p. 472. An example
of this is (182). Without *hatthu,* the sentence would not be grammatical in the intended
reading.

(182) *Itthu **hatthu**=kavanan laayeng.*
       DIST INDEF=group   different
       'That is a different group.' (K060108nar02)

Additionally, the proclitic is used to indicate the vagueness of numbers, like Span-
ish *unos veinte hombres* 'about twenty men'.[26] Only the proclitic is possible then. In
the following example, a vague quantity is talked about. This is marked by the numer-
als for the upper and the lower boundary, 10 and 15, respectively, but additionally by
the use of *hatthu* before the first numeral.

(183) ***Hatthu**=spuulu lima-blas thaaun=nang jaalang blaakang.*
       INDEF=ten   five-teen year=DAT   walk   after
       'After 10 or 15 years had passed.' (K060116nar01)

This use of *hatthu* in this function often cooccurs with the similative clitic =*ke*, as
in (184)

(184) ***Hatthu**=doblas thiga-blas thaaun=**ke**.*
       INDEF=twelve   three-teen year=SIMIL
       'About twelve, thirteen years.' (K051206nar17)

Hatthu can further indicate vagueness in reference to persons. Persons are of
course specific in reference, but the hearer is not expected to know that particular
person talked about in (185).

---

[26]Tamil has a similar use of the numeral one for indicating vagueness (Schiffman 1999:135).

(185) *Ithu=kapang [Aasura Mauluth katha]=**hatthu** aada.*
       DIST=when   Aasura Mauluth QUOT=INDEF   exist
       'There was a certain Aasura Mauluth then.' (B060115cvs01)

Hatthu is also used in the reciprocal construction, where it occurs twice. This is discussed in more detail in ⊡→8.6, p. 449.

The form of this clitic is clearly derived from Malayic **satu*, while the function is copied from the Sinhala affix *-ak~-ek*, which is used in precisely the same environments. Tamil, on the other hand, does not have a grammaticalized marker of indefiniteness.

Smith (2003:14) indicates that definiteness is not marked in SLM, but this is not true at least for the Upcountry data as shown above.

Speakers with high exposure to Sinhala seem to favour the enclitic use, while older speakers seem to use the proclitic more often. This should due to the fact that Sinhala marks definiteness postnominally, as in *pistoolaya-k* 'a pistol', and not prenominally.

6.4.2   Plural clitic

Pada is a morpheme indicating plurality or collectivity ⊙→15.6.1.1, p. 639. It is either realized with a voiced stop or a flap. This seems to depend on speech tempo. Ansaldo (2005b:14) gives *pada* as a suffix, but the fact that it can combine with hosts from a variety of word classes, like nouns, adjectives, quantifiers, numerals and pronouns and even headless relative clauses suggests that it is a clitic. The clitic nature of *pada* is furthermore underscored by (186), where we find metalinguistic material intervening between *pada* and its host.

(186) *Derang samma jaau **uudik** – village area – **pada**=nang su-pii.*
       3PL   all   far   village   PL=DAT   PAST-go
       'They all went to remote villages.' (K051222nar04)

The most common use of *pada* is on nouns, as in (187)

(187) *Itthu vatthu=ka itthu **nigiri**$_N$ **pada**=ka arà-duuduk.*
       DIST time=LOC DIST land   PL=LOC   NON.PAST-stay
       'At that time, (they) lived in those countries.' (N060113nar01)

Postnominal modification of the noun does not prevent the use of *pada*, as the following example shows.

(188)   *Se=pe   oorang_N thuuva_{ADJ} pada anà-biilang kitham pada*
        1s=POSS man   old   PL   PAST-say 1PL   PL
        Malaysia=*dering anà-dhaathang katha.*
        Malaysia=ABL   PAST-come   QUOT
        'My ancestors told (me) that we had come from Malaysia.' (K060108nar02)

Pada can also combine with plural pronouns (189) and quantifiers (190). Since these are all inherently already plural, *pada* only serves to emphasize that fact.

(189)   *Itthu=nam blaakang=jo, kitham pada anà-bìssar.*
        DIST   after=EMPH 1PL   PL   PAST-big
        'After that, we grew up.' (K060108nar02)

(190)   *Spaaru pada bannya baae=nang anthi-duuduk.*
        some   PL   very   good=DAT IRR-stay
        'Some are well off.' (K061122nar01)

This emphasizing function can also be used in enumerations as in (191).

(191)   *Mr Dole=pe   Mr Samath=pe   Mr Yusu pada=le interview=nya thraa.*
        Mr Dole=POSS Mr Samath=POSS Mr Yusu PL=ADDIT interview=DAT NEG
        'Mr Dole, Mr Samath and Mr Yusu were not selected for the interview.'
        (K060116nar05)

This reinforcing use of *pada* parallels the use of *-gal* in Tamil, as pointed out by Smith (2003). Sinhala, on the other hand, has obligatory plural marking which is normally not done by a suffix (cf. Nitz & Nordhoff forthcoming).

Pada can also be used on NPs marked with the possessive postposition =*pe* as shown in (192).

(192)   *Itthu=pe pada=jo bannyak mlaayu pada karang siini aada.*
        DIST=POSS PL=FOC much   Malay PL   now   here exist
        'It's their folks we get a lot of today here.' (K051205nar04)

Finally, *pada* can also be used to modify the entities denoted by headless relative clauses as in (193)(194).

(193)   *[[Seelon=nang anà-dhaathang] ∅ pada] mlaayu pada.*
        Ceylon=DAT   PAST-come   PL   Malay PL
        'Those who had come to Ceylon were the Malays.' (N060113nar01)

(194)   *[[[Neene   pada] anà-biilang] ∅ pada]=jo itthu.*
        grandmother PL   PAST-say   PL=EMPH DIST
        'What the grandmothers said was this' (K051206nar03)

(193) consists of two NPs, a referential NP *[Seelon=nang anà-dhaathang] pada* and a nominal predicate *mlaayu pada*. The referential NP in turn consists of a headless relative clause *Seelon=nang anà-dhaathang ∅*, which maximizes reference to all entities which had come to Ceylon, and *pada*. *Pada* in this case highlights the quantity of people who had come to Ceylon. This is semantically speaking redundant, since the headless relative clause already implies plurality, as does the predicate *mlaayu pada*, so that the first *pada* would not be needed for disambiguating purposes. It thus serves more pragmatic function to emphasize the great numbers. The argumentation for (194) is analogous.

Two occurrences of *pada* after clauses can be found in (195), also in equational sentences.

(195)   a.   *[[Se=dang nya-boole] pada] se nya-ambel*
             1s=DAT   PAST-can   PL   1s PAST-take
             'I took what I could' (K051213nar01)

        b.   *[[Se=dang thàrà-boole] pada] se thàrà-ambel.*
             1s=DAT   PAST-cannot   PL   1s NEG.PAST-take
             'What I could not take, I did not take.' (K051213nar01)

The fact that *pada* can follow verbs, as in this example, but also nouns or adjectives, as seen above, is the motivation to treat it as a clitic, rather than an affix. Since it cannot occur on its own, it is also not possible to see it as a free word.

The use of *pada* is optional, but frequent. As for its origin, Adelaar (1991:26) states: "Jakartanese has *pada* preceding the predicate and indicating plurality of subject. The syntactically different SLM *-paḍa* must be borrowed from Jakartanese, which in turn probably borrowed it from Javanese."

6.4.3   The semantics of *pada* and *hatthu*

Pada might actually be better analyzed as expressing *collective nominal aspect,* rather than nominal number. Collective aspects signals 'that the set consists of multiple individual entities which together form a collective' (Rijkhoff 2002:102). This interpretation of highlighting the collective interpretation of a set is supported by the following example, where a group of three gentlemen should give interviews. The three gentlemen are named and coordinated, but then the plural marker is added.

(196)   *Mr Dole=pe   Mr Samath=pe   Mr Yusu **pada**=le* interview=*nya thraa.*
        Mr Dole=POSS Mr Samath=POSS Mr Yusu PL=ADDIT interview=DAT NEG
        'Mr Dole, Mr Samath and Mr Yusu were not selected for the interview.'
        (K060116nar05)

It is clear that there is only one specimen of each of the named persons, and the group only exists once, so *pada* cannot indicate cardinality greater than 1 in the strict sense. Rather, it emphasizes that we are dealing with a collectivity, the group is not seen as monolithic, but as composed of several members ('multiple individual entities'), and the cardinality of the members is greater than one. This interpretation as collective actually fits well with the optionality of *pada* according to Rijkhoff's presentation of *set nouns*.

A logical extension of the analysis of *pada* as a 'collective aspect marker' would be to analyze *hatthu* (called 'indefiniteness marker' here) as a 'singulative aspect marker', indicating that the set is conceptualized as a whole, and not as 'consisting of multiple individual entities'. SLM *ruuma* 'house' would then be transnumeral, *ruuma pada* 'house COLLECTIVE' would mean 'the concept "house" interpreted as consisting of multiple entities', and *hattu ruuma* 'SINGULATIVE house' would mean 'the concept "house" interpreted as consisting of a singleton entity'. This 'singularizing' function of *hatthu* finds support in example (197), where the concept of 'parents', which is ontologically necessarily of cardinality greater than one, is modified with *hatthu*, indicating that it should be conceptualized holistically, and that the internal constituency of the concept does not matter.

(197)   *Kithang samma **hatthu umma+baapa**=pe   aanak pada, kithang samma*
        1PL   all   INDEF mother+father=POSS child PL   1PL   all
        sudaara pada.
        brother PL
        'We are all the same parents' children, we are all brothers' (B060115cvs01)

A reanalysis of the indefinite article as singulative has been proposed for Turkish by Schroeder (1999), and can be used to explain the uncommon order of the 'indefinite article' *bir* in this language (Rijkhoff 2002:319). In Turkish, *bir* can intervene between an adjective and a noun as in *meşur bir şair* 'famous a poet'. If *bir* is an indefinite article, this order would violate a universal that the order ADJ INDEF N does not exist (Greenberg 1963, Hawkins 1994). If it is analyzed as a nominal aspect marker, on the other hand, the universal would not apply. The interesting thing is now that SLM shows the same structure as Turkish ⊙→8.1.16, p. 440. An example would be *bàrnaama hatthu oorang* 'famous a man'. One could speculate that what is good enough for Turkish could also do for SLM.

6.4.4 Postpositions

Another important class of bound words are postpositions. These indicate the semantic roles of the NPs they attach to. The use of postposition is ubiquitous, but of course there are some semantic roles like GOAL which are used more often than, say, COMITATIVE. Furthermore, some postpositions can be used for more than one role, which increases their frequency. This is especially true of =*nang*, which combines a plethora of semantic roles. All postpositions are used by everybody.

The change of the historic prepositions to postposition is frequently noted in the literature. Adelaar (1991:30f) states their cliticized nature. Smith et al. (2004) treat them as "case suffixes and postpositions", but admit that the boundary between the two is fuzzy in postposing languages. It is clear that the members of this class are used to indicate case, yet their phonological status is rather one of clitics than one of suffixes. This can be seen from the fact that they can attach to different hosts and that a pause between the host and the case marker is possible, something which is expected for clitics, but not for suffixes.

A typical example of the use of the postpositions is found in (198), where we find three postpositions, to indicate the semantic roles of PATIENT/THEME, SOURCE and GOAL. The semantic role of AGENT is not marked, which is typographically indicated by ∅ here.

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

These postpositions are clitics because they can attach to any host.[27] Nominal hosts are given above, a verbal host is given in (199).

(199) *Suda butthul suuka **asà-dhaathang=nang**.*
       thus very like CP-come=DAT
       'So am pleased very much that you have come.' (G051222nar01)

Postpostions can be stacked. This most commonly occurs if the first of them is the possessive, but other possibilities, like the combination of locative and ablative in (200) is also possible.

---

[27]Already Bakker (2006) notes that "the postnominal markers seem to be connected more loosely to the nouns [. . .] than in Tamil."

(200) *Itthu=nang blaakang inni oorang lìkkas˜lìkkas thoppi pada=yang*
      dist   after   PROX man   fast˜RED   hat   PL=ACC
      *asà-kumpul ambel sithu**=ka=dering** su-pii.*
      CP-collect   take   there=LOC=ABL   PAST-go
      'After that, the man quickly picked up his hats and left from that place.'
      (K070000wrt01)

In the following, the postpositions are loosely ordered according to increasing semantic content. The postpositions treated first mainly serve to indicate quite general roles functions and do not carry a lot of semantic information, while the semantic content of the postpositions further down the list is less bleached.[28]

6.4.4.1 accusative *=yang*

The postposition *=yang* is most often used to indicate patient (Ansaldo 2008:24) ⊙→15.1.2.2, p. 563. It must not be confounded with the derivational affix of kin - *yang* ⊡→6.2.5.4, p. 313. Smith et al. (2004) report conflation of accusative and dative in one marker, but this statement needs to be revised; all varieties of SLM have a distinction between the two cases (Ansaldo 2005b, 2008, 2009b, Slomanson 2006). *=yang* is a bound word rather than an affix because it is not selective with regard to ist host. (201) shows the use on a noun while (202) shows the use on a verb.

(201) *Itthu=nang blaakang inni oorang lìkkas˜lìkkas **thoppi**$_N$ **pada=yang***
      dist   after   PROX man   fast˜RED   hat   PL=ACC
      asà-kumpul ambel sithu=ka=dering su-pii.
      CP-collect   take   there=LOC=ABL   PAST-go
      'After that, the man quickly picked up his hats and left from that place.'
      (K070000wrt01)

(202) *Ini   oorang thoppi arà-**kumpul**$_V$**=yang** asà-kuthumung=apa, ...*
      PROX man   hat   SIMIL-collect   CP-see=after
      'After they had seen the man collect the hats, ...' (K070000wrt01)

Furthermore, *=yang* can combine with NPs headed by pronouns (203).

---

[28]Slomanson (2008b) reports the use of a further temporal postposition, *=ambe(l)* in the Southern dialect, which would stem from *sambil* 'while'. This postposition is then homophonous to the verb *ambel* 'take', which has complicated the analysis. This postposition is not used in the Upcountry.

(203) **See*PRON*=yang** *Tony katha arà-panggel.*
      1s=ACC   Tony QUOT NON.PAST-call
      'I am called "Tony".' (K060108nar01)

Additional evidence for the status of =*yang* as a bound word comes from the fact that metalinguistic commentary can intervene between the host and =*yang* as shown in (204), which should not be possible if =*yang* were an affix.

(204) *Davong karri –daavong* means leaf– =*yang campur.*
      leaf   curry   =ACC mix
      'Then you mix the curry leaves.' (K060103rec01)

The main use of =*yang* is patient (Ansaldo 2008:24) as in (205), but theme can also be indicated by =*yang* as in the three examples above (201)(202)(204).

(205) *Itthusubbath deram pada jaalang arà-kijja*   *butthul ruuma pada=yang*
      therefore   3PL   PL   road   NON.PAST-make correct house PL=ACC
      arà-picca-kang.
      NON.PAST-broken-CAUS
      'Therefore, they build the street, they demolish many houses.' (K051222nar04)

There is a lot of variation in the use of =*yang* between speakers, but also within the same idiolect (cf. Slomanson 2006:148). The use of =*yang* is inconsistent (Ansaldo 2005b, 2008), as the following example, taken from a letter, shows.

(206) *Lai*   *se* computer=*nang baaru* optical mouse *atthu*=∅=**le** Encarta2006
      other 1s computer=DAT new   optical mouse INDEF=ADDIT Encarta2006
      software=**yang**=**le** *su-bìlli.*
      software=ACC=ADDIT PAST-buy
      'Then I also bought a new optical mouse and the Encarta 2006 software for the computer.' (Letter 26.06.2007)

This sentence informs us about the purchase of two items, a computer mouse and a software program. These are coordinated by the *X=le Y=le* construction. Of the coordinated items, one, the software, is marked by =*yang*, and the other one, the mouse, is not. Obviously, the semantic role (THEME) of these two participants is exactly the same, given that they have taken part in the very same event, the purchase. One could argue that the entity hosting the clitic =*yang* is the set formed by the mouse and the software. Then, it would not be surprising to find only one indication of the semantic role. This hypothesis is disproved by =*yang* occurring between *software*

and the coordinator =*le*. If the set was marked with =*yang*, then =*yang* should occur
outside of the *X=le Y=le* construction, which it does not. We have to conclude that the
mouse and the software are individuated items participating in the act of buying, but
one is marked with =*yang* and the other one is not. The only conclusion is that the use
of =*yang* is optional.

The probability of =*yang* surfacing is affected by the following criteria:

- affectedness favours =*yang*

- topicality and definiteness favour =*yang*

- singular reference favors =*yang*

- animacy favors =*yang*

Still, none of these criteria are on their own sufficient to predict the presence or
absence of =*yang*. Examples for this will be given in more detail below.

The difference in affectedness corresponds to the difference between patient and
theme. Patients are significantly affected by the action, while themes are not affected
in an important way. A conjecture would be that =*yang* is used on patients but not
on themes. While there is certainly some truth to this analysis, the following four
examples show that there is no neat correspondence between patient and presence of
=*yang* on the one hand, and theme and absence of =*yang* on the other hand. Example
(207) shows a patient with =*yang*, (208) shows a patient without =*yang*, (209) shows
a theme with =*yang* and (210) shows a theme without =*yang*.

(207) *Incayang see$_{pat}$=**yang** hathu buruan mà-jaadi   su-bale-king.*
       3s.POLITE 1s=ACC   INDEF bear   INF-become PAST-turn-CAUS
       'He turned me into a bear.' (K070000wrt04)

(208) *Snow-white=le   Rose-red=le   pinthu$_{pat}$=∅ su-bukka.*
       Snow-white=ADDIT Rose-red=ADDIT door   PAST-open
       'Snow-white and Rose-red opened the door.' (K070000wrt04)

(209) *Oorang su-baavung, thoppi$_{theme}$=pada=**yang** anà-caari.*
       man   PAST-rise,   hat=PL=ACC   PAST-find
       'The man got up and looked for (his) hats.' (K070000wrt01)

(210) *Hathu haari hathu oorang thoppi$_{theme}$=∅ mà-juval=nang kampong=dering*
       INDEF day   INDEF man   hat   INF-sell=DAT village=ABL
       kampong=nang su-jaalang pii.
       village=DAT   PAST-walk go
       'One day, a man walked from village to village to sell hats.' (K070000wrt01)

The second criterion is that topicality favors =*yang*, while non-topical undergoers should not be marked with =*yang*. Ansaldo (2008, 2009b) argues that the presence of =*yang* implies definiteness, while its absence implies indefiniteness. Impressionistically, there is such a tendency, but it is by no means absolute, as the following four examples show, where topical referents are found with (211) and without =*yang* (212). The topicality is indicated by the deictics *itthu* and *ini*. Non-topical referents, identifiable by the indefiniteness marker *atthu,* can also occur with (213) and without =*yang* (214).

(211) **_Itthu_** *aayer=**yang** baaye=nang arà-boil-kang.*
DIST water=ACC good=DAT NON.PAST-boil-CAUS
'You boil that water well.' (K061026rcp04)

(212) *Incayang [**ini** [Seelong=ka anà-aada lakuan] baathu=∅] asà-caari.*
3S.POLITE PROX Ceylon=LOC PAST-exist valuable stone cp-search
'He looked for these valuable stones from Ceylon.' (K060103nar01)

(213) *Derang **hathu** papaaya=**yang** asà-poothong*
3PL INDEF papaw=ACC cp-cut
'They cut a papaw' (K051220nar01)

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

The third criterion is that singular referents are more likely to be marked with =*yang* than plural referents. Again, this is the case impressionistically, but counterexamples can be found. (213) and (214) show the presence and absence of =*yang* on singular referents. (215) and (216) show the presence and absence of =*yang* on nouns with plural reference.

(215) *Itthusubbath deram pada jaalang arà-kijja butthul ruuma **pada=yang***
therefore 3PL PL road NON.PAST-make correct house PL=ACC
arà-picca-kang.
NON.PAST-broken-CAUS
'Therefore, they build the street, they demolish many houses.' (K051222nar04)

(216) British government Malaysia Indonesia *ini   nigiri   **pada=∅** samma*
       British government Malaysia Indonesia PROX country PL   all
       anà-peegang ambel.
       PAST-catch   take
       'The British government captured all these countries.' (K051213nar06)

The fourth criterion, animacy, states that animate referents are more likely to be marked by *=yang*. Again, this is probably true as a probabilistic observation, but not an absolute rule. (215) and (216) show that *=yang* can be present or absent on inanimate referents, while (217) and (218) show the same for animate referents.

(217) *Kanabisan=ka=jo duva oorang=le anà-thaau ambel [Andare duva*
       last=LOC=EMPH   two man=ADDIT PAST-know take   Andare two
       ***oorang**=yang=le asà-enco-kang aada] katha.*
       man=ACC=ADDIT CP-fool-CAUS   exist EMPH
       'At the very end, both women understood that Andare had fooled both of them.' (K070000wrt05)

(218) *Kumaareng=le   thuuju=so   dhlaapan=so **oorang**=∅ asà-buunung.*
       yesterday=ADDIT seven=UNDET eight=UNDET man   CP-kill
       'Again yesterday, seven or eight people were killed.' (K051206nar11)

The above discussion should have shown that the occurrence of *=yang* does not follow from hard and fast rules but rather is a result of the complex interplay of several factors. This can be captured in rule (219).

(219) =yang *is the more likely to occur the more of the following features have a positive value in the context of the utterance: topicality of the referent, definiteness of the referent, singular reference, high animacy, affectedness.*

Besides the most common use as a kind of affected object marked, *=yang* also has some other less central uses, namely undergoer marking on intransitive verbs, sentential nominalization, and some uses with monovalent predicates which are difficult to capture.

There are occurrences of *=yang* in a monovalent predication, where it signals the undergoer.[29] One example is given in (220). Crucially, the verb *thìnggalam* 'sink' is intransitive, the transitive verb 'to drown' is *cullop*.

---

[29]Compare the use of 'accusative subjects' in Sinhala described in (Gair 1976a, 1991b).

(220) Titanic *kappal=yang su-thìnggalam.*
Titanic ship=ACC   PAST-sink[intr.]
'The ship "Titanic" sank.' (K081104eli05)

Sentential nominalization is formed by adding =*yang* to the end of the clause to be nominalized.

(221) *[Se arà-maakang]$_{cls}$=**yang** lorang=nang atthu creeveth=si?*
1s NON.PAST-eat=ACC   2s=DAT   INDEF problem=INTERR
'Do you have a problem with my eating?' (K081103eli04)

Note that in this example, the act of eating is not affected, not topical, not patient, and overall a very bad candidate for marking with =*yang*. In this case it is not a semantic need which motivates the use of =*yang*, but the syntactic need to show that the clause is used in referential function here.

Another somewhat surprising use of =*yang* is in questions where it combines with the WH-word to query for things, *aapa*, to form *aapeyang* ⊡→5.7.1, p. 231. This is used to query for themes and patients, and this comes as no surprise, but it can also be used to query for existence, like *aapeyang aada?* 'What is there?'

=*yang* can also be used on the only arguments of some adjectival predicates, where its occurrence is difficult to explain, as in (222).

(222) *Hatthu komplok bannyak=jo puuthi caaya, **hathyeng=yang** meera=jo*
INDEF bush   much=EMPH white colour, other=ACC   red=EMPH
meera caaya.
red   colour
'One bush was very white, the other one was of the reddest red.'
(K070000wrt04)

We see that there a very similar predicates applied to two bushes, one is white, the other one is red. Yet the first one is not marked by =*yang,* but the second one is. The reason for this could have to something to do with contrastive topics, but this will need further research.

The second example of =*yang* being used on an atypical referent is (223), where the location of a cave is given, and the cave is marked with =*yang*.

(223) Three miles *cara jaau=ka aada [dee   anà-sbuuni duuduk cave]=**yang**.*
three miles way far=LOC exist 3s.IMPOLITE PAST-hide sit   cave=ACC
'Three miles away from here is the cave where he stayed hidden.'
(K051206nar02)

It is not clear why =*yang* is used here, but it might have to do with the fact that *cave* is a loanword, and loanwords behave differently in morphosyntax in other domains as well, namely indefiniteness marking ⊡→6.4.1, p. 319 and the quotative ⊡→6.4.9.1, p. 389.

The relevance of this marker was first described by Ansaldo (2005b). Previously, Smith (2003) had seen =*yang* and =*nang* (see below) as allomorphs. Ansaldo describes =*yang* as a definite object marker in the Kirinda dialect.

Slomanson (2006:149) traces the origin of =*yang* to question formation patterns in Jakarta Malay, which require =*yang* for queries for patient. From there, =*yang* would have spread as a generalized patient marker.

6.4.4.2   dative =*nang*

This is the most frequent postposition with a variety of meanings, of which the dative stands central (Ansaldo 2005b, 2008, 2009b, Slomanson 2006, 2008a). =*nang* is often reduced to one of *na, nà*. In combination with *se* '1s.polite', *go* '1s.FAMILIAR', *lu* '2s.FAMILIAR', *de* '3s.IMPOLITE' it is always realized as *dang*. The latter realization is also used for other lexemes by some speakers (cf. Saldin 2001:66). The following examples show the standard form (224) and the form used for monosyllabic pronouns (225).

(224)   *Kithang lorang=nang baaye mliiga athi-kaasi.*
        1PL   2PL=DAT   good   palace IRR-give
        'We will give you beautiful palaces.' (K051213nar06)

(225)   *Lorang se=dang mà-hiidop thumpath kala-kaasi.*
        2PL   1s=DAT INF-stay   place   if-give
        'If you give me a place to stay.' (K070000wrt04)

In order to distinguish the environments where the allomorphs =*nang* and =*dang* are used, different accounts are possible. Let us start with the descriptive facts: =*nang* is much more frequent than =*dang*, but some speakers only have =*dang*. For this latter group, the generalization is obvious and trivial: =*dang* is always used, =*nang* is never used. The remaining speakers (the vast majority) use =*dang* with monosyllabic singular pronouns (*se, go, lu, de*), and use =*nang* elsewhere. Ansaldo (2005b, 2008, 2009b) suggests that a possible generalization could be found along the lines of the Animacy Hierarchy (Silverstein 1976). Pronouns are higher on the hierarchy than other nouns, and hence the split between hosts taking =*dang* (i.e. pronouns) and hosts taking =*nang* (all others) would reflect a semantic difference. However, there are two problems with this account. The first one is that the plural pronouns *kithang, lorang, derang* do not take =*dang* but =*nang*. It is not obvious why the plural forms should pattern with the lower-animacy class, rather than with the singular pronouns. This can

be captured by the question 'why should the first person plural referent be considered less animate than the second person singular referent?'[30] The second problem is that the impolite third person pronoun *de* takes *=dang* in the Upcountry (Ansaldo has *dia* and *nang*), while the polite third person pronoun *incayang* takes *=nang*. If this is to be explained along the lines of the Animacy Hierarchy, one would have to argue that referring to a person in a polite way makes her less animate (causing *=nang*) than referring to her in an impolite way (entailing *=dang*).

Another possible generalization would be a phonological one: 'All monosyllabic words take *=dang*, the other ones take *=nang*.' This is appealing at first sight, since there are not many monosyllabic words, and most of them are the pronouns. However, when the verb *pii* 'go' combines with the dative marker, the allomorph *=nang* is used, as in (226)

(226) *Daalang=ka **pii=nang** blaakang, dhraapa   puukul=le thama-kuthumung.*
       inside=LOC go=DAT after   how.many hit=ADDIT NEG.NONPAST-see
       'After going inside, how much he hit them, you could not see.'
       (K051206nar02, K081103eli04)

The phonological account thus makes wrong predictions, too. We need both phonological information to explain why *incayang* takes *=nang*, and we need morphological information to explain why *pii* does not take *=dang*. Which leaves us with the generalization 'monosyllabic pronouns take *=dang*, all other words take *=nang*'. This 'generalization' is actually quite weak. Two criteria are needed to describe a set of four forms. It remains an open question whether the speakers of SLM actually use productive rules of this sort for the dative forms of the pronouns, or whether they simply store *sedang* '1s.DAT' etc. as four ready-made forms in the lexicon.

Just like the other postpositions, *=nang* is a clitic, which can attach to different hosts. The following examples show a noun (227) and a verb (228) hosting *=nang*.

(227) *Hathu haari hathu oorang thoppi mà-juval=nang kampong=dering*
       INDEF day INDEF man   hat   INF-sell=DAT   village=ABL
       *kampong_N=**nang** su-jaalang pii.*
       village=DAT   PAST-walk go
       'One day, a man walked from village to village to sell hats.' (K070000wrt01)

---

[30]This account could be saved by incorporating the notion of number into the animacy hierarchy even if it lacks immediate appeal. A complicating factor is the following: While number is normally not a factor in reflexes of the hierarchy, when it is, it appears that plural is *higher* on the scale (Silverstein 1976:126), rather than lower. However, we would need plural to be lower for the SLM case.

(228)   *Suda buthul suuka nyaari siini su-dhaathang_V=nang* .
      thus   correct like   today   here PAST-come=DAT
      'So I very much liked that you came here today.' (G051222nar01)

Metalinguistic commentary can intervene between =*nang* and its host, giving further evidence of its status as a bound word rather than an affix.

(229)   ***Lummas*** – **soft** – ***=nang*** *blaakang minnyak   klaapa=ka   inni=yang*
      soft   –   – DAT   after   coconut.oil coconut=LOC PROX=ACC
      gooreng.
      fry
      'After it has become tender, fry this in coconut oil.' (K060103rec02)

=*nang* has a multitude of functions in SLM, being used for

- goal of motion ⊙→15.1.2.7, p. 571,

- recipient ⊙→15.1.2.4, p. 568,

- beneficiary ⊙→15.1.2.10, p. 573,

- experiencer ⊙→15.1.2.5, p. 569,

- purpose ⊙→15.1.2.12, p. 576,

- some patients ⊙→15.1.2.2, p. 563,

- manner ⊡→10.1, p. 461,

- and argument of modals ⊡→10.3, p. 471.

This wide use of the dative is typical of South Asia (Masica 1976, Sridhar 1976a,b, 1979, Verma & Mohanan 1990, Abbi 1994, Bhaskararao & Subbarao 2004a,b). Historically, =*nang* was used to indicate goal of motion. This is still the case, as the following example shows.

(230)   *Kitham=pe Badulle   sudaari sudaara pada, kitham em-pii   ruuma*
      1PL=POSS   Badulla.LOC sister   brother PL   1PL   PAST-go house
      *saakith**=nang**.*
      sick=DAT
      'All of us Malays from Badulla went to the hospital.' (B060115nar02)

When used to indicate goal of motion, =*nang* cannot attach directly to human referents. Instead, a relator noun like *dikkath* 'nearby' must be used.

(231) *Se=ppe **dìkkath**=nang mari.*
     1s=POSS side=DAT   come.IMP
     Come close to me. (K081103eli04)

=*nang* is also used to indicate recipient. In this case, it can directly attach to human referents, as (232) shows.

(232) ***Oorang=nang** thumpath masà-kaasi.*
     man=DAT   place   must-give
     '(They) must give land to the people.' (K051222nar05)

The semantic role of beneficiary is very close to recipient, and is also expressed by =*nang*.

(233) *Derang derang=pe **umma=nang** butthul saayang=kee=jo samma ruuma*
     3PL   3PL=POSS mother=DAT correct love=SIMIL=EMPH all   house
     pukurjan=nang=le anà-banthu.
     work=DAT=ADDIT   PAST-help
     'They also helped their mother with all the housework.' (K070000wrt04)

Furthermore, =*nang* is used to indicate experiencer, as in (234) (cf. Ansaldo 2005b:19).

(234) ***Derang=nang** byaasa svaara=hatthu su-dìnngar.*
     3PL=DAT   habit noise=INDEF   PAST-hear
     'They heard a familiar sound.' (K070000wrt04)

Semantically quite close to experiencer is the use of =*nang* on the arguments of modal particles (Ansaldo 2005b:19), which is shown in the following two examples.

(235) ***Se=dang** karang jaau mà-pii thàràboole.*
     1s=DAT   now   far   INF-go cannot
     'I cannot go far.' (K061120nar01)

(236) ***Kithang=nang** baaye=nang mulbar bole=baaca.*
     1PL=DAT   good=DAT   Tamil   can=read
     'We can read Tamil well.' (K051222nar06)

Next comes the use of =*nang* to indicate purpose, on either nominal arguments (237)(238) or clausal arguments (239), which then typically have the verb in the infinitive. Complements of modal verbs also sometimes take dative marking (240).

(237) Second world war time=*ka   Kluum̌bu=nang* Japanese *arà-*bomb-*king*
second world war time=LOC Colombo=DAT   Japanese NON.PAST-bomb-CAUS
***thakuth-an=nang**.*
fear-NMLZR=DAT
'During the second world war, the Japanese bomb Colombo to cause fear.'
(N060113nar03)

(238) *Se=ppe aade   ... **inni=nang** su-pii*
1s=POSS younger.sibling ... PROX=DAT   PAST-go
'My younger sister went in order to do that.' (B060115nar01)

(239) *Itthu cave=nang kithang=le pii aada **mà-liyath=nang**.*
DIST cave=DAT   1PL=ADDIT go exist INF-look=DAT
'We have also gone to that cave to have a look.'

(240) *Derang pada=nang atthu=le   mà-**kijja=nang** thàràboole=subbath ....*
3PL   PL=DAT   one=ADDIT INF-do=DAT   cannot=because
'Because they couldn't do anything.' (N060113nar01)

Furthermore, =*nang* is used to indicate the semantic roles of manner (241) and
exchange value (242).

(241) *Karang see siithu pukurjan arà-jalang-kang   **hathu** engineer*
now   1w there work   NON.PAST-walk-CAUS INDEF engineer
***mosthor=nang**.*
way=DAT
'Now I run the work over there, like an engineer.' (K061026prs01)

(242) *Laayeng nigiri=pe   soojor   pada=nang **baae lakuvan=nang***
different country=POSS European PL=nang   good value=DAT
anà-juuval.
PAST-sell
'(He) sold (the stones) to the Europeans from abroad for a good price.'
(K060103nar01)

In some instances, =*nang* can also be used on arguments which are very patient
like, for instance *puukul* 'hit' (243).

(243) **Dutch=*nang*** *mà-puukul=jo cinggala raaja pada pii aada.*
Dutch=DAT INF-hit=EMPH Sinhala king PL go exist
'(The Malays) came to the Sinhalese kings to hit (=fight) the Dutch.'
(K051206nar04)

On the other hand, it can be argued that the conceptualization of 'hit' in SLM involves, next to the agent, an entity which receives something, namely blows. In that case, there is no need to include patient in the array of semantic roles that =*nang* can fulfill. It is the semantic representation of the event which is different, not the morphosyntactic encoding of a certain semantic role.

The last function of =*nang* is to indicate possession when combined with an existential as in (244).

(244) ***Se=dang*** *liima anak klaaki pada **aada**.*
1s=DAT five child male PL exist
'I have five sons.' (K060108nar02)

It should have become clear from the discussion above that =*nang* is pervasive in the grammar of SLM. The following excerpt is a nice example of this, where 5 words out of 16 are marked by =*nang*.

(245)   a. *Ithu   oorang=**nang**$_{exp}$ baaye=**nang**$_{adv}$ nanthok pii=**nang**$_{postp}$*
      DIST=DAT man=DAT   good=DAT   sleep   go=DAT
      blaakang
      after
      'After the man had well fallen asleep.'

  b. *pohong=dering baava=**nang**$_{goal}$ asà-thuurung.*
      tree=ABL   down=DAT   CP-descend
      '(the monkeys) climbed down from the tree and'

  c. *[oorang anà-baava] samma thoppi=pada asà-ambel*
      man   PAST-bring all   hat=PL   CP-take
      'took all the hats the man had brought and'

  d. *mà-maayeng=**nang**$_{compl}$ su-mulain.*
      INF-play=DAT   PAST-start
      'started to play.' (K070000wrt01)

Another example where te use of =*nang* is pervasive is (246), where =*nang* occurs on three words out of five,

(246)   *Cinggala=**nang** buthul=**nang** thàrà-thaau inni mà-kirja=**nang**.*
         Sinhala=DAT   correct=DAT   NEG-know   PROX INF-make=DAT
         'Sinhalese don't really know how to prepare this (dish).' (K061026rcp01)

This morpheme seems to stem from a Javanese allative preposition =*nang* (Slomanson 2006:151). From the original meaning of GOAL, the use would have spread to RECIPIENT and then to the other instances described above. Ansaldo (2005a) suggests that =*nang* stems from reanalysis of *nya*-structures in (South East Asian) Malay, indicating possessive. *Nya* would then get a paragogic ŋ, (see Section 4.6.2.5 for the development of paragogic velar nasals), and change ɲ to n. While this development cannot be ruled out, it appears that tracing contemporary =*nang* to the homonymous Javanese allative adposition is a more straightforward explanation, especially since the immigrants used to spend a lot of time in Batavia on Java, if they were not even Javanese themselves.

6.4.4.3   locative =*ka*

This postposition is a locative marker ⊡→13.2.3.7, p. 525 (Smith et al. 2004, Ansaldo 2008, 2009b). It has only one form, which attaches nearly always to a noun (247), a relator noun (248) or a pronoun (249). =*ka* is used to indicate the location of referents in space and time,[31] as can be seen in (247), where it attaches to the temporal noun *vatthu* 'time' and the noun *nigiri* 'country' referring to a spatial location.

(247)   *Itthu **vatthu=ka** itthu **nigiri**   **pada=ka** arà-duuduk.*
         DIST time=LOC   DIST country PL=LOC   NON.PAST-stay
         'Then, they lived in those countries.' (N060113nar01)

---

[31]Ansaldo (2005b:23) states that =*ka* can be used to indicate possessor and gives the following example (original orthography, repeated as well in Ansaldo (2008:30))

Nembak orang-ka ada snapan bae.
hunt   man-LOC have rifle   good

'The hunter has a good rifle'
However, this example appears to have a different information structure than what is indicated in the translation.This sentences suggests a prenominal relative clause modifying *snapan* 'rifle' and an adjectival predicate *bae* 'good'. In the relative clause, a subtype of the existential predication with *aada* ⊡→10.2.2, p. 470 is used, the temporary possession type with =*ka*. Note that *aada* means 'exist' and not necessarily 'have' ⊡→5.1.2.1, p. 165. The translation which reflects that information structure would be 'The rifle which the hunter has with him is good' or 'The rifle which is with the hunter is good.' In this reading, =*ka* does not indicate the semantic role of possessor, but the semantic role of location, which aligns nicely with the general use of =*ka*.

(248) *Blaakang=jo incayang anà-kuthumung moonyeth pada thoppi asà-ambel*
  After=ᴇᴍᴘʜ 3s.ᴘᴏʟɪᴛᴇ ᴘᴀsᴛ-see monkey ᴘʟ hat ᴄᴘ-take
  *pohong **atthas=ka** arà-maayeng.*
  tree top=ʟᴏᴄ sɪᴍᴜʟᴛ-play
  'After that he saw that the monkeys had taken the hats and were playing with
  them on top of the tree.' (K070000wrt01a)

(249) ***Incayang=ka** ... bìssar beecek caaya hathu bag su-aada.*
  3s.ᴘᴏʟɪᴛᴇ=ʟᴏᴄ ... big mud colour ɪɴᴅᴇғ bag ᴘᴀsᴛ-exist
  'He had a big brown bag with him.' (K070000wrt04)

Still, a case can be made that it is a bound word rather than an affix because metalinguistic material can intervene between the host and =*ka* as shown in (250).

(250) *Minnyak **klaapa** – coconut oil – =**ka** gooreng.*
  coconut.oil coconut –   – =ʟᴏᴄ fry
  'Fry it in coconut oil.' (K060103rec02)

This analysis is supported by the fact that there is one instance in the corpus where =*ka* attaches to an adjective.

(251) *[Dee arà-sbuuni duuduk cave] asaraathang sini=ka asàduuduk*
  3s.ɪᴍᴘᴏʟɪᴛᴇ sɪᴍᴜʟᴛ-hide sit cave] ᴄᴏᴘᴜʟᴀ here=ʟᴏᴄ from
  *hathu three miles cara **jaau=ka**.*
  ɪɴᴅᴇғ three miles way far=ʟᴏᴄ
  'The cave where he remained hidden is about three miles from here.'
  (K051206nar02)

Furthermore, =*ka* can also attach to deictics, like *sini* in (252).

(252) *See athi-thiidor **sini=ka**=jo.*
  1s ɪʀʀ-sleep here=ʟᴏᴄ=ᴇᴍᴘʜ
  'I will sleep here.' (K051205nar05)

Next to the essive meaning, =*ka* can occasionally been found in lative meanings, where it attaches to the goal of motion (253) (Smith et al. 2004) ⊙→15.1.2.7, p. 571.

(253)   *Derang samma oorang [hatthu hatthu thumpath pada]=**ka** asà-pii pukurjan*
          3PL   all   man   INDEF   INDEF   place   PL=LOC   CP-go   work
          su-gijja
          PAST-make
          'All those people go to one place or another and work.' (B060115cvs06)

Besides the temporal and the spatial use mentioned above, =*ka* is used to indicate
means of transport (254) (Ansaldo 2005b:23) and the person from whom information
is requested[32] (255).

(254)   **Bus=*ka*** *kapang-pii cumma anà-pii.*
          bus=LOC   when-go   idle   NON.PAST-go
          'When I went with the bus, I did not do anything.' (K061125nar01)

(255)   *Incalla   [lai   thaau **sudaara sudaari pada**]=**ka** bole=caanya ambel*
          Hopefully other   know   brother sister   PL=LOC   can-ask   take
          [nya-gijja lai   saapa=kee aada=si   katha].
          PAST-make other   who=SIMIL exist=INTERR QUOT
          'Hopefully, you can enquire from another person you know whether there is
          someone else who did something.' (N061031nar01)

This morpheme is very frequent and probably stems from the proto-form **dəkat*.
Another reflex of this form is the relator noun *dìkkath* 'proximity', ⊡→5.2.2.7, p. 206,
but this is not transparent to the speakers, because *dìkkath* and =*ka* can be combined,
as examples (256) shows

(256)   Fifth mile post ***dìkkath=ka***.
          fifth mile post proximity=LOC
          'Close to the Fifth Mile Post.' (K051206nar16)

6.4.4.4   ablative =*dering*

This postposition is an ablative marker. It has the forms *deri* and *d(e)ring* (257)(258).
Smith et al. (2004) give *dari*, which is not found in the corpus, Ansaldo (2005b, 2008,
2009b) gives *ring*, which is also not found in the corpus.

---

[32]Smith et al. (2004) found the relator noun *dìkkath* for the latter function, which is etymologically
related.

(257) *[Itthu nigiri=**deri** asà-dhaathang anà-thiiňǧgal oorang pada]=jo kithang.*
DIST country=ABL CP-come PAST-stay man PL=EMPH 1PL
'The people who have come from those countries and stayed (here) are we.'
(K051222nar03)

(258) *Itthu kaaki=**dering**=jo arà-kirja.*
DIST leg=ABL=EMPH NON.PAST-make
'We make it from that leg.' (N060113nar05)

=*dering* is normally found attached to nouns. Elicitation shows that it is possible to attach =*dering* to verbs (259).

(259) *Derang pada su-/anà-dhaathang=dering kitham pada su-suuka.*
3PL PL PAST-/PAST-come=ABL 1PL PL PAST-like
'We were happy as soon as they came.' (K081103eli04)

The possibility to insert metalinguistic commentary between =*dering* and the host (260) also shows that =*dering* is not an affix but rather a bound word.

(260) *Kaaya oorang* – monied people – =*dering*=*jo arà-cuuri.*
rich man =ABL=EMPH NON.PAST-steal
'It was from rich people that he stole.' (K051206nar02)

This morpheme is used to indicate source as shown in (261)(262), and instrument ⊙→15.1.2.9, p. 573, as show in (263). Both uses are already noted in Ansaldo (2005b:23) and Ansaldo (2008:30).

(261) *Spaaru **Indonesia=dering** dhaathang aada.*
some Indonesia=ABL come exist
'Some came from Indonesia.' (K060108nar02)

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

(263)   ***Thaangang=dering*** *bukang*   ***kaaki=dering*** *masà-maayeng.*
        hand=ABL   NEG.NONV leg=ABL   must-play
        'You must play not with the hands but with the feet.' (N060113nar05)

Both source and instrument can receive a liberal interpretation. In example (264),
the source is an employer, which is not a spatial entity. In (265), *daavon* 'leaves' are
given as instrument, but there is no one who would have manipulated that instrument.

(264)   *See asà*-retire *aada* **police=*dering***.
        1s CP-retire   exist police=ABL
        'I have retired from the police.'

(265)   *[**Daavon=deri** thuuthup aada gaaja]*   *hatthu asà-dhaathang.*
        leaf=INSTR   close   exist elephant INDEF CP-come
        'An elephant, which had been hidden by leaves, appeared.' (B060115nar05.)

By extension of the semantic role of instrument, codes of communication are also
coded by *=dering*.

(266)   *Arà-biilang*   ***mlaayu=deri****.*
        NON.PAST-speak Malay=ABL
        'They speak in Malay.' (K060103nar01)

(267)   *Allah! se=dang thàrà-thaau; derang* **Arabic=*dering*** *arà-caanya.*
        Allah 1s=DAT NEG-know   3PL   Arabic=ABL   NON.PAST-ask
        'Oh my god! I didn't know; they were asking in Arabic!' (K061019nar02)

An interesting use of the instrumental, which is also found in Sinhala (Gair &
Paolillo 1997:31) is the marking of institutional agents with the instrumental, rather
than zero, which is the normal marking for agents ⊙→15.1.2.1, p. 562. Examples (268)
and (269) show two instances of this institutional instrumental.

(268)   **British Government=*dering*** Malaysia Indonesia*, inni nigiri   pada*
        British Government=ABL   Malaysia Indonesia PROX country PL
        samma peegang.
        all   catch
        'The British Government captured Malaysia and Indonesia, those countries.'
        (K051213nar06)

In example (268) the entity performing the capturing is not a person but an institu-

tion, which is why it is marked with the instrumental =*dering*. Like a government, the police department is an institution, and the agency of the police department in (269) is also coded by =*dering*.

(269)   *See=yang* **police=*dering*** *nya-preksa.*
        1s=ACC   police=ABL   PAST-enquire
        'I was questioned by the police.' (K051213nar01)

In some cases, *dering* is used on individual agents as well, in a manner comparable to the *by-agent* in English, although there is no passive construction involved. This use is shown for the intransitive verb *maathi* 'die' in (270) and (271).

(270)   *kàthaama* police *oorang nya-maathi inni* **terrorist *hatthu=dering*.**
        first   police man   PAST-dead PROX terrorist INDEF=ABL
        'So the first police man died by (the hands of) this terrorist.' (K051206nar02)

(271)   *[***Terrorist *hatthu=dering*** *anà-maathi kàthaama oorang]=jo incayang.*
        terrorist   INDEF=ABL   PAST-dead first   man=EMPH 3s.POLITE
        'The first man to die by (the hands of) a terrorist was him.' (K051206nar02)

This morpheme stems from **dari*, a preposition (Bakker 2006:141). Ansaldo (2005b) notes that this morpheme shows an ablative/instrumental syncretism, as also found in Sinhala.

6.4.4.5 possessive =*pe*

This postposition is a possessive marker (Adelaar 1991, Smith et al. 2004, Ansaldo 2005b, 2008, 2009b). When it follows a monosyllabic pronoun (i.e. *se* '1s.polite', *go* '1s.familiar', *lu* '2s.familiar', *de* '3s.familiar') it is realized with a geminate consonant =*ppe*. The following two examples show the two allomorphs.

(272)   *Kithang=**pe*** *baapa=**pe*** *naama Mahamud.*
        1PL=POSS   father=POSS name Mahamud
        'Our father's name is Mahamud.' (B060115nar03)

(273)   *Se=**ppe*** *naama Mohomed Imran Salim.*
        1s=POSS name   Mohomed Imran Salim
        'My name is Mohomed Imran Salim.' (K060108nar01)

=*pe* normally attaches to a noun or a pronoun as shown in above and in (274), but it can also attach to verbs as in (275), adverbs (276)(277),[33] and to interrogative

---

[33]Adelaar (2005) contains examples of a very liberal use of the 'linker' morpheme on adverbs in other

pronouns, as in (278).

(274) **Lu=ppe**   *muuluth=ka=le*   *paasir,* **se=ppe** *muuluth=ka=le*
       2s.FAMILIAR=POSS mouth=LOC=ADDIT sand   1s=POSS mouth=LOC=ADDIT
       paasir
       sand
       'There is sand in your mouth and there is sand in my mouth.' (K070000wrt02)

(275) **Anà-liyath=pe** *mosthor thàrà-baae.*
       PAST-come=POSS manner NEG-good
       'The way he looked was not good.' (K081103eli04)

(276) **Karam=pe** *mosthor=nang, mpapulu aari=ka=jo*   *sunnath=le*
       now=POSS manner=DAT   forty   day=LOC=EMPH circumcision=ADDIT
       arà-kijja.
       NON.PAST-make
       'For today's way of doing (it), it is on the fortieth day that they also do the
       circumcision.' (K061122nar01)

(277) **Dovulu=pe** *oorang pada.*
       before=POSS man   PL
       'People in former times.' (K061026rcp04)

(278) *Lorang pada asàdhaathang [***saapa=pe***=ke   baa   thaangang=ka=jo*
       2PL   PL   COPULA   who=POSS=SIMIL under hand=LOC=EMPH
       pukurjan mà-gijja athi-jaadi].
       work   INF-make IRR-become
       'You will always have to work under someone's command' (K051206nar07)

Like the other bound words, *=pe* can be separated from the host. In example (279)
this is done by the ritual formula to be spoken after naming the Islamic prophet.

(279) Prophet **Mohomed sallallaahu.alai.wasallam=pe**
       [Prophet Mohomed May.Allah.bless.him.and.grant.him.peace]=POSS
       sunna   pada samma kithang arà-follow-*kang*
       religious.path PL   all   1PL   NON.PAST-follow-CAUS
       'We all follow the paths of the Prophet Mohomed (s.a.w.).' (K061026prs01)

Malay varieties.

=*pe* is exclusively used for attributive possession ⊙→15.11, p. 668, but this can be interpreted loosely, as example (276) shows, where *today's way of doing things* is not an instance of prototypical possession, since temporal periods like *today* cannot own anything. Other instances of non-canonical possessors are *thiiga di buulang* 'third of the month' in (280), *Jaapna* 'Jaffna' in (281) and *lunch* in (282).

(280) *Kithang arà-biilang kithang=nang suurath* third **thiiga di buulang=pe**
      1PL   NON.PAST-say 1PL=DAT   letter   third three of month=POSS
      suurath doblas di buulang nya-daapath.
      letter   twelve of month   PAST-get
      'We will say that we got the letter from the third of the month only on the twelfth.' (K060116nar10)

(281) *Itthu muusing bannyak* teacher *pada* **Jaapna=pe**.
      DIST time   many   teacher PL   Jaffna=POSS
      'Back then, many teachers were from Jaffna.' (K051213nar03)

(282) **Lunch=pe** arrangement *kithang mà-kirja kithang sama oorang*
      lunch=POSS   1PL   INF-make 1PL   all   man
      nya-caanya.
      PAST-ask
      'We all inquired whether we should make arrangements for lunch.' (K060116nar07,K081105eli02)

The following example shows a naturalistic sentence with five occurrences of =*pe*.

(283) *Itthu=kaapang [se**=ppe** baapa] [se**=ppe** kaake]   [[se**=ppe***
      DIST=when   1s=POSS father 1s=POSS grandfather 1s=POSS
      *kaakee**=pe**]   baapa] kithang samma oorang [Seelon**=pe**   oorang]*
      grandfather=POSS father, 1PL   all   man   Ceylon=POSS man
      pada.
      PL
      'Then my father and my grandfather and my grandfather's father, all of us people became Ceylon people.' (K060108nar02)

This morpheme stems from **puɳa* (Adelaar 1991, 2005, Saldin 2001, Ansaldo 2005b). *Puunya* is also found synchronically as a verb 'to own' with the lexicalized relativized form *anà-puunya* 'one who owns=owner'. (284) gives an example of this. Note the occurrence of both =*pe* and *puunya*.

(284)   British *oorang=hatthu, incayang [thumpath=**pe** anà-**puunya** ∅].*
        British man=INDEF   3s.POLITE place=POSS   PAST-own
        'There was a British person, he was the owner of the place.'
        (K051220nar01,K081103eli04)

6.4.4.6   =*subbath* 'because'

This postposition is used to indicate cause and reason ⊙→15.1.2.13, p. 578 (Slomanson
2008b). It can attach to a noun (285)(286), a pronoun (287) or a clause (288)-(290). It
is also found frequently attached to the distal deictic *itthu*, where a lexicalized meaning
of *therefore* emerges (291)(292).

(285)   a.   *Suda **maven=subbath** see anà-resign.*
             so   son=because   1s PAST-resign
             'So I quit because of my son.'

        b.   *Maven masà-blajar-king=jona.*
             son   must-learn-CAUS=PHAT
             'I have to educate my son, don't I?' (B060115prs21)

(286)   ***Non-Muslims pada=subbath** kithang muuka konnyong arà-cunji-kang*
        non-Muslims PL=because   1PL   face   little   NON.PAST-show-CAUS
        siini.
        here
        'Because of the non-Muslims, we show our faces here.' (K061026prs01)

(287)   *Lorang **see=subbath** ithu Aajuth=yang su-salba-king.*
        2PL   1s=because   DIST dwarf=ACC   PAST-safe-CAUS
        'You saved that dwarf because of me.' (K070000wrt04)

(288)   *[Derang pada=nang atthu=le   mà-kijja=nang thàràboole]=**subbath** ....*
        3PL   PL=DAT   one=ADDIT INF-make=DAT cannot=because
        'Because they weren't able to do anything.' (N060113nar01)

(289) *[Ini oorang giini   kapang-jaalang pii caape]=**subbath** jaalang hathu*
     PROX man   this.way when-walk   go tired=because   road   INDEF
     piiňǧgir=ka anà-aada   hathu pohong baava=ka su-seeňder.
     border=LOC PAST-exist.inanim INDEF tree   down=LOC PAST-rest
     'Because he was tired from walking then, this man sat down under a tree which
     stood at a side of the street.' (K070000wrt01)

(290) *[Umma buthul miskiin]=**subbath** aanak su-laari kluuling.*
     mother correct poor=because   child PAST-run roam
     'Because the mother was very poor, the child ran away.' (K061019sng01)

(291) *Suda **itthu=subbath**=jo, se laile* Marine Engineering *asà-kijja ambel*
     SO   DIST=because=EMPH 1s again Marine Engineer   CP-make take
     arà-pii.
     NON.PAST-go
     'So, it was because of that that I took up again the Marine Engineering work
     and went (away).' (K051206nar20)

(292) *Suda **itthu=subbath** derang konnyong* westernize.
     thus DIST=because 3PL   few   westernize
     'So, therefore, they get a little westernized.' (K061026prs01)

Subbath is not as frequent as the postpositions discussed before, which might be related to the fact that one speaks about location and possession more often than about cause.

This morpheme might stem from *səbab* 'reason; because' (Adelaar 1985:35), with raising of schwa and gemination of the following consonant as usual. The substitution of the final labial by a dental would still be in need of explanation.

6.4.4.7 =*kapang* 'when'

This temporal postposition meaning *when* indicates a vague relation in time between figure and ground ⊙→15.5.3, p. 625. It must be distinguished from the related interrogative pronoun *kaapang* 'when?' ⊡→5.7.4, p. 234. The postposition is often reduced to =*kang*. This is never done with the interrogative pronoun. Furthermore, there is the prefix *kapang-* ⊡→6.2.1.7, p. 283, which must not be confounded with the postposition either. The allomorphs =*kapang* and =*kang* are given in the following two examples.

(293) ***Itthu=kapang*** *se=ppe baapa, se=ppe kaake,   se=ppe*
       DIST=when   1s=POSS father 1s=POSS grandfather 1s=POSS
       kaake=pe   baapa, kithang samma oorang Seelon=pe   oorang pada.
       grandfather=POSS father, 1PL   all   man   Ceylon=POSS man   PL
       'Then my father and my grandfather and my grandfather's father, all of us
       people (became) Ceylon people.' (K060108nar02)

(294) ***Ithukang*** *ithu bambu giithu=jo   luvar=nang arà-dhaathang.*
       then   DIST bamboo like.that=EMPH outside=DAT NON.PAST-come
       'Then that bamboo comes out like that.' (K061026rcp04)

=*kapang* can attach to a deictic (see above), a noun or an adjective (see examples
below).

(295) *See aanak_N=**kapang**, se=dang thàràsìggar.*
       1s child=when,   1s=DAT sick
       'I was sick when I was a child.' (K081103eli04)

(296) *Kiccil_ADJ=**kapang**, kithang sudaara pada samma* cricket *arà-maayeng.*
       small=when   1PL   siblings PL   all   cricket NON.PAST-play
       'When we were small, us children used to play cricket.' (K051201nar02)

This postposition also combines often with the distal deictic *itthu*, yielding a mean-
ing of 'then', which is frequently employed to structure discourse (see (293)(294)).
Because of this high frequency, more erosion takes place and a form like *i(k)kang* can
also be heard. This form has to be distinguished from the homophonous *ikkang* 'fish'.
The following example shows both *ikang* 'then' and *ikkang* 'fish'.

(297) ***[Ikang*** *Seelong=nang dhaathang aada mlaayu] oorang ikkang.*
       then   Ceylon=DAT come   exist Malay man   fish
       'The Malays who had come to Ceylon then were fishermen.' (K060108nar02)

6.4.4.8 =*sangke* 'until'

This postposition is used for terminal boundaries in time, like English *until* ⊙→15.5.3.5,
p. 633. There are no further allomorphs. =*sangke* normally attaches to an adverb
(298), a noun (299)(300), or a numeral (301)(302).

(298) ***Nyaari=sangke*** *se inni ruuma=ka=jo arà-duuduk.*
today=until 1s PROX house=LOC=EMPH NON.PAST-live
'It is in this house that I have been living up to today.' (K060108nar01)

(299) *Thuju-pul-liima **thaaun=sangke** incayang anà-iidop.*
seven-ty-five year=until 3s PAST-stay
'He stayed until he was 75.' (K060108nar02)

(300) *Suda derang=nang [**hathyang muusing**]=sangke mà-duuduk su-jaadi.*
thus 3PL=DAT other time=until INF-stay PAST-become
'So they had to stay until the next time .' (K051213nar01)

(301) **1948=*sangke***, *Independence kan-daapath, Seelon=ka baae thumpath*
1948=until independence when-come Ceylon=LOC good place
pada=ka baae=nang anà-duuduk.
PL=LOC good=DAT PAST-live
'Until 1948, when (we) got the independence, (they) lived well in the nice places in Ceylon.' (K081105eli02)

(302) Two **o'clock=*ke*=*sangke*** *bole=duuduk.*
two o'clock=SIMIL=until can=stay
'You can stay until about two o'clock.' (K061026rcp04)

Example (302) shows that =*sangke* is not an affix, because additional material, in this case the clitic =*ke*, can intervene between the host and *sangke*.

=*sangke* can also attach to clauses, as the following example shows.

(303) *Buruan [diinging abbis]=**sangke** siithu su-siinga.*
bear cold=finish until there PAST-stay
'Bear stayed there until the cold was over.' (K070000wrt04)

A very intriguing fact is that *sangke* is sometimes used *before* the ground noun, which is against the general left-branching structure of SLM. Peter Slomanson (p.c.) suggests that this might be an iconic effect. Since the ground noun is the terminal boundary, it should also be the final element in the PP. The following two examples show the prenominal use of *sangke*.

(304) *See=yang lorang=susamma diinging muusing **sangke**=habbis anà-simpang*
        1s=ACC   2PL=COMIT   cold   season until=finish   PAST-keep
        ambel.
        take
        'You have kept me with you until the cold season was over.' (K070000wrt04)

(305) *Itthu muusing asàduuduk **sangke**=nyaari see pukuran arà-gijja.*
        DIST time   from   until=today   1s work   NON.PAST-make
        'From that time until now I have been working.' (K060108nar01)

6.4.4.9   comitative (=sà)saama

This postposition is used for comitative (Smith et al. 2004) ⊙→15.1.2.11, p. 576. The
first syllable is optional. When it is used, it can be either *sà, sa* or *su*. The *saama* part
can either be *saama* or *samma*. The following examples show different combinations
of the two components. In all these examples, the postposition indicates comitative.

(306) *Soojer   pada **incayang=sàsaama** Seelon=nang asà-dhaathang, ...*
        Europeans PL   3s=COMIT   Ceylon=DAT CP-come
        'The Europeans came together with him to Sri Lanka and ...' (K060103nar01)

(307) *Ithu=nang blaakang kithang=nang santham=**sasaama**   baae=nang*
        DIST=DAT after   1PL=DAT   coconut.milk=COMIT good=DAT
        asà-mix-*kang=apa.*
        CP-mix-CAUS=after
        'After having mixed it with the coconut milk ...' (K061026rcp04)

(308) *See=yang lorang=**susamma** diinging muusing sangke-habbis anà-simpang*
        1s=ACC   2PL=COMIT   cold   season until-finish PAST-keep
        ambel.
        take
        'You have kept me together with you until the cold season was over.'
        (K070000wrt04)

(309) *Mlaayu oorang pada bannyak pukurjan anà-kirja soojor pada=**saama**,*
Malay man PL much work PAST-make European PL=COMIT
thaau=si soojor, English *oorang pada=**sasaama**.*
know=INTERR European English man PL=COMIT
'The Malays did a lot of work together with the Soojors, you know "Soojor",
together with the Europeans.' (K061026prs01)

(310) *See=**samma** kumpul=apa nyaanyi.*
1s=COMIT gather=after sing
'Gather and sing along with me.' (K061019sng01)

(311) *See=**samma** kumpul mari thaaňdak-la.*
1s=COMIT gather IMP dance-IMP
'Come and dance with me.' (N061124sng01)

The postposition *(sà)saama* must not be confounded with the quantifier *samma*
'all, every'[34] ⊡→5.9, p. 240.

6.4.4.10   =*apa* 'after'

This postposition is used to indicate subsequence of events, similar to English *after*
⊙→15.5.3.2, p. 627. It attaches to the right edge of the clause which precedes in time.
Very often, the conjunctive participle *asà-* underscores the 'subsequence' meaning of
=*apa* in the same clause (312). However, other tenses like in (313)(314), or no tense
marking at all like in (315)(316) are also possible.

(312)   a. *Siithu **asà**-blaajar=**apa**,*
there CP-learn=after
'After having learned there,'

  b. *thaaun nnamblas=ka se skuul **asà**-luppas=**apa**,*
year sixteen=LOC 1s school CP-leave=after
'after having left the school at 16,'

  c. *pukuran asà-caari anà-pii.*
work CP-find PAST-go
'I looked for work and went (away).' (K060108nar01)

---

[34]Interestingly, while *sama has undergone a functional split into *saama* and *samma* in SLM, the exact
reverse is true for Riau Indonesian, where *sama could tremendously enlarge its functional domain (Gil
2004).

(313)   *Suda kithang karang inni Gampola=nang anà-dhaatham=apa hatthu*
        thus 1PL   now   PROX Gampola=DAT PAST-come=after   INDEF
        thuuju thaaun=ka su-jaadi.
        seven year=LOC PAST-become
        'So, now, it has become about seven years after we moved to Gampola.'
        (G051222nar01)

(314)   a. *Seelong* independent state ***anà*-*jaadi=nang=apa,***
            Ceylon independent state PAST-become=DAT=after
            'After Sri Lanka had become independent,'

        b. *Kithang=nang* independence ***anà*-*daapath=nang=apa,***
            1PL=DAT   independence PAST-get=DAT=after
            'after we had obtained the independence,'

        c. *laile derang anà-duuduk* under the Commonwealth.
            still 3PL   PAST-stay   under the commonwealth
            'they still were under the commonwealth.' (K051222nar06)

(315)   *Laiskali* Netherland *pii=**apa**, laiskali* September *arà-raathang.*
        again   Netherland go=after again   September NON.PAST-come
        'After he will have gone back to the Netherlands, the will come back again in
        September.' (B060115prs15)

(316)   *Oorang pada thiikam=**apa** oorang pada=nang theeṁbak=**apa** se=dang*
        man   PL   stab=after   man   PL=DAT   shoot=after   1s=DAT
        bannyak creeveth pada su-aada.
        much   trouble PL   PAST-exist
        'People were stabbed, people were shot, I had a lot of problems.'
        (K051213nar01)

Combined with *dhaathang* 'come', the resulting form *(asà)dhaathampa* has gram-
maticalized into a copula ⊡→5.5, p. 218.
    When combined with the existential *duuduk*, =*apa* can be used to indicate source
as in (317).

(317)   1970=*ka* ***duuduk=apa*** 1987=*sangke* Sports Ministry=*ka ae-duuduk.*
        1970=LOC from   1987=until   Sports Ministry=LOC PAST-stay
        'From the 1970s (onward), I stayed in the Sports Ministry until 1987.'
        (N060113nar04)

The fact that the combinations with *dhaathang* and *duuduk* yield these meanings,

and the fact that combinations of the conjunctive participle *asà-* with these verbs yield the very same meanings ⊡→6.2.1.4, p. 276 suggests that *asà-* and *=apa* are semantically very close.[35] Since the main use of the conjunctive participle *asà-* is to indicate subsequence of events, which is also a prime function of *=apa*, this makes even more sense. It might be possible to see *asà-* as a proclitic form of *=apa*, much in the same way as *bàrà=* is a proclitic form of postverbal *boole*.[36]

One problem with this analysis is that *asà-* and *=apa* can cooccur, as in (312). Such cooccurrence is not found for the other free-word/clitic pairs like *bàrà/boole*. Furthermore, *=apa* can combine with other TAM-markers, as for example *anà-* in (313)(314). Another reason for not treating *=apa* as verbal inflection is that the string *itthu=nang=apa* 'after that' is frequently heard, although it does not occur in the corpus. It cannot be excluded that in the future a unified analysis of *asà-* and *=apa* can be found, for the time being, they are treated as separate morphemes in this grammar.

The following example might shed light on the interplay between *asà-* and *=apa*. Both morphemes are present in the last line and it is unclear as of now why then *=apa*, then *asà-* is chosen.

(318)   a.   *Andare aanak=nang su-biilang:*
             Andare child=DAT   PAST-say
             'Andare said to his child:'

        b.   *Aanak.*
             child
             ' "Son, " '

        c.   *['lu=ppe umma   su-maathi' katha bithàràk=apa] [asà-naangis]*
             2s=POSS mother PAST-die   QUOT scream=after CP-weep
             mari.
             come.IMP
             ' "come and cry and weep 'My mother has died!' " '

6.4.4.11   *=sikin* 'because'

This postposition indicates reason ⊙→15.1.2.13, p. 578. It has been found attached on nouns as in (319) and deictics (320).

(319)   *Kithang=pe Riverstan* trip=*yang* **uujang=siking** *su*-cancel-*king.*
        1PL=POSS   Riverstan trip=ACC   rain=because PAST-cancel-CAUS
        'We had to cancel our trip to Riverstan because of rain.' (K071029eml01)

---

[35] Smith et al. (2004) actually have *=apa* as the only form for the conjunctive participle. Also cf. Bakker (2006:143).

[36] The adstrate Sinhala actually has a conjunctive participle *-la* and an 'prior temporal form' *=aama* (Gair 1976b, 2003), which are very similar in function and seem to be parallel to SLM *asà-* and *=apa*.

(320)   *Inni aari pada=ka kithang=nang* test. **Itthu=siking**=*jo   see thàrà-kiiring*
        PROX day PL=LOC 1PL=DAT   test DIST=because=EMPH 1s NEG.PAST-send

        'We are having tests these days. That is why I have not sent (anything).'
        (K071203eml01)

    This postposition is not used very often. The etymology is unclear.

6.4.4.12   =*lanthran* 'because'

Like *sikin*, this postposition indicating reason ⊙→15.1.2.13, p. 578. It is not very frequent. It has been found attached to a deictic (321), a noun (322), and a verb (323).

(321)   a. *Beeso   luusa   lubaarang arà-dhaathang.*
           tomorrow later.in.the.future festival   NON.PAST-come
           'The day after tomorrow is the festival.'

        b. **Itthu**=*lanthran kithang=pe ruuma see arà-cuuci.*
           DIST=because 1PL=POSS   house 1s NON.PAST-clean
           'That's why I am cleaning the house.' (K061019prs01)

(322)   *Se=ppe* argument=**lanthran** *ten days=jo   aada* Hill Country Malay *inni*
        1s=POSS argument=because   ten days=EMPH exist Hill Country Malay PROX
        Malaka chief minister visit=*yang kithang=nang kaasi.*
        Malaka chief minister visit=ACC 1PL=DAT   give
        'It was due to my argument that we received the Malaka chief minister's visit
        within ten days.' (K060116nar06,K081103eli04)

(323)   *Derang hathu suurath nya-kiiring [see ini   Kandi Mlaayu*
        3PL   INDEF letter   PAST-send 1s   PROX Kandy Malay
        Association=*dering nya-kiisar]*=**lanthran**.
        association=ABL   PAST-go.aside=because
        'They had written a letter because I had left the Kandy Malay Association.'
        (K061122nar03)

6.4.4.13   =*thingka* 'when'

This postposition indicates the meaning of *when* ⊙→15.5.3, p. 625. This is a temporal postposition. It can attach to a noun as in (324), or to a verb as in (325). The verb does

not carry prefixes in that case.

(324) *Inni habbar=**thingka**.*
PROX news=middle
'When we heard about these news.' (B060115nar02)

(325) *Paanas muusing ∅-**dhaathang=thingka** see siini=dering arà-pii.*
hot season -come=middle 1s here=able NON.PAST-go
'When the hot season will have come, I will leave from here.' (K070000wrt04)

Example (326) shows an ambiguous use, where the nature of the host cannot be determined because *kaaving* can both mean 'marry' and 'wedding'.

(326) *See **kaaving=thingka** husband su-biilang pukujan asà-luppas mari*
1s marry=when husband PAST-say work CP-leave come
Navalapitiya=nang.
Nawalapitiya=DAT
'When I got married, my husband told me to leave work and come to Nawalapitiya.' (K051201nar01)

In the above examples, =*thingka* has a meaning of *as soon as*. This is not the only meaning, as (327) shows, where =*thingka* has a durative meaning, which should not be glossed by *as soon as* but rather by 'while'.

(327) *Ithu ship=ka kithang duuduk=thingka kithang arà-baapi general*
DIST ship=LOC 1PL stay=while 1PL NON.PAST-bring general
cargo.
cargo
'When we stay on that ship, we carry general cargo.' (K051206nar19)

The two temporal meanings are not easy to distinguish, as in the following example, where both interpretations (*as soon as, while*) are possible.

(328) *Mà-bavung giithu=jo thaarek=thingka mà-baaung thàràboole.*
INF-raise like.that=EMPH pull=when INF-raise cannot
'When he tried to pull himself up like that, he could not get up.' (K051205nar05)

=*thingka* is very similar in meaning to *(=)kapang(-)* in its use as a postposition or a prefix, which can be seen from the following examples, where first *kapang-* and then =*thingka* are used to refer to the same temporal relation.

(329) a. *Suda giini   **kapang**-duuduk,*
           thus   like.this when-stay
           'So when he was staying there like that,'

       b. *spaaman duuduk=**thingka**,*
           3s.POLITE stay=when
           'when he was staying there,'

       c. *spaaman su-iilang.*
           3s.POLITE PAST-disappear
           'he disappeared.' (B060115nar05)

The following example shows again both *kapang-* and *=thingka* in the same utterance, but the interpretation is less clear-cut in this case.

(330) a. *Nigiri=pe   oorang pada dhaathang biilang.*
           village=POSS man   PL   come   say
           'The villagers came and said:'

       b. *Allah incayang=yang siaanu asà-buunung thaaro=apa katha.*
           Allah 3s.POLITE   3s.prox CP-kill   put=after QUOT
           ' "Allah, he has been killed" '

       c. ***Biilang=thingka**,*
           say=when
           'On saying this,'

       d. *oorang pada **kapang-laari** dhaathang ini   daara sgiithu=le*
           man   PL   when-run   come   PROX blood that.much=ADDIT
           suusu su-jaadi.
           milk   PAST-become
           'when people came running, the blood had turned into milk.'
           (K051220nar01)

6.4.5   Prepositions

While close to the totality of all adpositions in SLM are postnominal, there are three prenominal adpositions. The first is *sangke*, which can also be used as a postposition and is discussed in that section ⊡→6.4.4.8, p. 350. The second one, *dari* is only found in registers influenced by Standard Malay, where it replaces the postposition *dering*. In those registers, which are used for instance in inaugural speeches, *dari Colombo* is used instead of *Kluum̆bu=dering* for 'from Colombo'. This use is almost exclusively found with toponyms, when the participants are listed. The third preposition is genuinely prenominal: *di*. It is used for indicating a subpart of a time period, like 'first of

the month' or 'third month of the year' (331)(332).

(331) *Ka-thiiga haari di buulang/miiňǧgu/thaaun.*
ORD-three day   of month/week/year
'The third day of the month/week/year.' (K081111eli01)

(332) *Ka-thuuju* hour=*ka   di haari.*
INF-seven   hour=LOC of day
'At the seventh hour of this day.' (K081111eli01)

Di can only be used with temporal nouns in the strictest sense. Nouns which have a temporal extension, but have additional semantics to them, like *pukuran* 'job' cannot be used with *di* (333). Instead, a periphrastic construction has to be used (334).

(333) **pàrthaama haari di se=ppe pukuran.*
first   day   of 1s=POSS work
'(The first day of my job).' (K081111eli01)

(334) *Se=ppe pukuran=dika pàrthaama haari buthul creeveth.*
1s=POSS work=vicinity first   day   correct trouble
'The first day at my job was troublesome.' (K081111eli01)

Another use of *di* is still in need of fuller analysis: *Di* is used in some contexts with *atthas* 'about', as shown in (335).

(335) *Saapa=pe=ke **di** atthas hatthu omong-an   arà-jaadi*
who=POSS=SIMIL di about INDEF   speak-NMLZR NON.PAST-become
'There is a rumour going round about somebody.' (K081111eli01)

6.4.6 Coordinating Clitics

There are a number of bound words in SLM which are used for various types of coordination. These are termed *Coordinating Clitics* here. Besides the coordinating use when attached to both of two coordinands, all have additional non-coordinating uses, e.g. in negation or questions. No Coordinating Clitic can ever occur as a free word, but they all can attach to a variety of hosts.

We can distinguish five Coordinating Clitics. These are summarized in Table (6.5) and will be discussed in turn.

- =*si* 'INTERR' ⊡→6.4.6.1, p. 360

- =*le* 'ADDIT' ⊡→6.4.6.2, p. 363

	gloss 1	gloss 2	label	Sinhala	Tamil
=si	INTERR	or	INTERR	=da	=aa
=le	too	and	ADD	=(u)t, =yi	=um
=ke	similar	or	SIMIL	vagee,	=aavadu
=so	whether	or	UNDET	=vat	=oo
=pon	not even one	neither ... nor	any	=(u)t	=um

Table 6.5: Coordinating Clitics. Information for the adstrates is from Karunatillake (2004) for Sinhala and Lehmann (1989) for Tamil.

- =*ke* 'SIMIL' ⊡→6.4.6.3, p. 369

- =*so* 'whether' ⊡→6.4.6.4, p. 374

- =*pon* 'any' ⊡→6.4.6.5, p. 376

These clitics can be stacked on top of each other, as shown in the following example. The three clitics =*le* 'ADDIT', =*ke* 'SIMIL' and =*si* 'INTERR' stack. =*le* occurs twice and is used to coordinate the two referents. =*ke* is used to express the similarity between the coordinated set and another referent, which is not mentioned. =*si* finally is used to give the illocutionary force of 'question' to the expression.

(336)   *Baapa=le   aanak=le=ke=si?*
        father=ADDIT child=ADDIT=SIMIL=INTERR
        'Are they like father and son?' (not on recordings)

Different subsets of these clitics participate in a number of constructions. Every clitic participates in some of these constructions, but not necessarily in all

- X=CLT Y=CLT coordination: all ⊡→14.3.2, p. 548

- formation of a pronoun with WH(=POSTP)=CLT: =*le, =ke, =so, =pon* ⊡→8.4, p. 447

- universal quantification with WH N=CLT: =*le* ⊡→8.1.12, p. 424

- quantification with WH(~WH) PRED=CLT: =*le, =so* ⊡→8.1.12, p. 424 as well

6.4.6.1   interrogative =*si*

This clitic denotes a questioned element (De Silva Jayasuriya 2002:50). On the South coast this clitic is realized as =*sin* (Slomanson 2008b). When occurring only once, a yes-no-question is formed, when occurring on more than one item, an alternative question is formed. This alternative reading can also be used in affirmatives, giving an reading of undeterminedness.

In yes-no-questions, =*si* can be used to query the sentence as in (337), or a constituent, as in (338) ⊙→17.1.2, p. 697.

(337) *Se=ppe uumur **masà-biilan=si**?*
      1=POSS age   must-tell=INTERR
      'Do I have to tell my age?' (B060115prs01)

(338) *Saapa? **See=si**?*
      who   1s=INTERR
      'Who? Me?' (B06015prs18)

=*si* can be used with any tense reference of the verb. The following three examples show the use of =*si* with the present tense, the perfect tense and the past tense.

(339) *Sebastian pùddas **arà-maakang=si**?*
      Sebastian spicy   NON.PAST-eat=INTERR
      'Do you eat spicy food, Sebastian?' (B060115cvs02)

(340) *Saathe maakang **aada=si**?*
      sate   eat   exist=INTERR
      'Have you eaten sate?' (B060115cvs02)

(341) *Sebastian **su**-kaaving=**si**?*
      Sebastian PAST-marry=INTERR
      'Are you married, Sebastian?' (B060115cvs03)

When attached to two or more items, an alternative question is formed. The alternatives can be specified as in (342), or be a simple restatement of polarity as in (343).

(342) *Piisang=si   maaňǧga=si maau?*
      plantain=INTERR mango=INTERR want
      'Is it plantain or mango that you want?' (K081105eli02)

(343) *Kithang **arà-baapi=si**   **thraa=si**?*
      1PL   NON.PAST-bring=INTERR not=INTERR
      'Shall we bring (it) or not?' (K051206nar16)

This alternative reading is also available for affirmatives, when the speaker wants to express uncertainty, rather than free choice or arbitrariness. In (344), the speaker emphasizes the fact that the distinction of who brought food and who brought clothes

is not central.

(344) *Ketham pada **makanan pada=si pakeyan pada=si** su-baavang.*
1PL   PL   food   PL=disj clothing PL=disj PAST-bring
'We all brought food or clothing.' (B060115nar02)

=*si* can also be used in subordinates, which can be interrogatives (345)(346), ig-noratives (348) or other predicates of knowledge (347). In these uses, =*si* competes with =*so* (see below) (348) (349).

(345) *[Aashik=nang hathu soldier mà-jaadi   suuka=**si**   katha] arà-caanya.*
Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT   NON.PAST-ask
'He asks if you want to become a soldier, Ashik.' (B060115prs10)

(346) *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
Hopefully other know brother sister   PL=LOC   can-ask   take
*[nya-gijja lai   saapa=kee aada=**si**   katha].*
PAST-make other who=SIMIL exist=INTERR QUOT
'Hopefully, you can enquire from another person you know whether there is someone else who did something.' (N061031nar01)

(347)   a. *Malaysia samma oorang=pe naama pada Maas.*
           Malaysia all   man=POSS name PL   Maas
           'People from Malaysia are all called "Maas".'
    b. *Suda itthu=dering=jo kithang=nang ini Indonesia=pe oorang=**si***
        thus DIST=ABL=EMPH 1PL=DAT   PROX Indonesia=POSS man=DISJ
        *giithu   kalthraa Malaysian oorang=**si** katha bàrà=thaau ambe.*
        that.way if-NEG   Malaysian man=DISJ QUOT can=know take
        'So with that we can come to know whether someone is Indonesian or otherwise Malaysian' (K060108nar02)

(348) *Se thàrà-thaau baapa anà-dhaathang=si katha.*
1s NEG-know   father PAST-come=INTERR QUOT
'I do not know if father has come.' (K081103eli04)

(349) *Se thàrà-thaau baapa anà-dhaathang=so.*
1s NEG-know   father PAST-come=UNDET
'I do not know whether father has come.' (K081103eli04)

A common occurrence of =*si* is following *thaau* 'know', giving a meaning of

English *you know?* which is used to check whether the addressee is keeping track of the story.

(350) *Bannyak pukurjan anà-kirja soojor   pada-saama, **thaau=si**   soojor?*
      much   work   PAST-do European PL=COMIT   know=INTERR European
      'He worked together with "soojors". You know "soojor"?' (K061026prs01)

When combined with the infinitive, the negative imperative, or the adhortative, =*si* is used to check for orders, like English *Shall I do X?*.

(351) *See mà-dhaathang=si?*
      1s   INF-come=INTERR
      'Shall I come?' (K081105eli02)

(352) *See jamà-dhaathang=si?*
      1s   NEG.INF-come=INTERR
      'Shall I not.come?' (K081105eli02)

(353) *Kithang marà-maakang=si?*
      1PL   ADHORT-eat=INTERR
      'Shall we eat?' (K081106eli01)

De Silva Jayasuriya (2002) sees this morpheme as a particle, although she writes it -*si* and does not justify this theoretical and notational choice. She interprets it as a question tag used to "mark the information that the [speaker] expects the hearer to agree with". However, the examples she shows also allow an interpretation as an open question. In all the data I have gathered, *..., bukang* ⊡→5.13.2, p. 258 or =*jona* ⊡→6.4.7.2, p. 385 are the usual way to form question tags for affirmatives, not =*si* ⊙→17.1.3, p. 699.

Bakker (2006) gives the etymological source of this morpheme as **siapa* 'who' without further explanation. This morpheme is very frequent. The functions of Sinhala =*da* and Tamil =*aa* are very similar.

6.4.6.2   additive =*le*

This is an additive clitic (cf. Smith & Paauw 2006:168). Ansaldo (2005b, 2008, 2009b) analyzes this marker as comitative case, but the morpheme for comitative case is rather *(sà)saama* ⊡→6.4.4.9, p. 352. =*le* is in a class with the other Coordinating Clitics, as can be seen from the list on page 360. Furthermore, =*le* is not used to indicate semantic role. In the following example, the semantic role of the father is clearly patient, not comitative. Still, =*le* is used, so that the presence of =*le* must be conditioned by something else than semantic role.

(354)   Heart attack *asà-peegang, baapa=**le**   su-niiňǧgal.*
        heart attack CP-catch   father=ADDIT PAST-die
        'After having got a heart attack, my father died as well.' (K051205nar05)

=*le* is used as a device on the level of information structure in (354). It indicates
that other referents of which the predicate *suniiňǧgal* 'died' is true had been mentioned
before. Like the other Coordinating Clitics, it attaches thus at the layer of information
structure, whereas postpositions attach at the propositional layer (cf. (162) on page
318). Another argument to treat =*le* as a Coordinating Clitic is that informants con-
sistently equate it with Tamil =*um* and Sinhala =*(u)t* or =*yi*, which are traditionally
analyzed as coordinators (Lehmann (1989:151) for Tamil, Karunatillake (2004:26) for
Sinhala).[37]

The fact that =*le* is found on the level of information structure, and not on the level
of the proposition, also explains why the dative marker =*nang* and =*le* can be found
stacked. This is not an instance of double case marking, but rather a combination of a
case marker and an element resembling a conjunction. If =*le* marked a semantic role,
combination with =*nang* would confer two semantic roles to a referent, which is not
possible.[38] Example (355) illustrates this.

(355)   *Oorang mlaayu siithu=dering dhaathang=apa cinggala raaja=**nang**=**le***
        man   Malay there=ABL   come=after   Sinhala king=DAT=ADDIT
        anà-banthu
        PAST-help
        'The Malays came from there and helped the Sinhalese king, too.'
        (K051206nar04)

As far as the king is concerned, he clearly has the role of beneficiary in (355),
indicated by =*nang*. If =*le* marked comitative, the king would end up with two se-
mantic roles, which is not what the semantics suggest. Furthermore, the king clearly
does not have the same role in the event as the Malays. Comitative suggests that a
certain action is performed by one participant in company of the other one. This is not
what we find here, where it is not the case that the Malays, accompanied by the king,
engage in an act of helping. True, the Malays do engage in an act of helping, but the
king does not accompany them in their helping, rather, he is on the receiving end of
the action of helping. The Malays and the king have thus different macroroles, which
does not agree with comitative semantics.

To sum up, =*le* patterns with the other Coordinating Clitics, occurs in sentences
where comitative is semantically ruled out (354)(355), and does not compete for the

---

[37]The distribution of =*le* is actually exactly identical with a pattern that is common on the Indian sub-
continent (Skrt. *ápi* , Marathi *-hī̄* , Maithili *-ō̃*, Hindi *-bhī̄*) and goes back to Vedic times (Emeneau 1974).
Emeneau gives 'additive' as overall gloss for the related meanings of the pattern, not 'comitative'.

[38]Unless reflexives or reciprocals are used, but this is irrelevant for the dative and the comitative.

slot where semantic roles are normally assigned (355). Furthermore, native speakers equate it with adstrate morphemes which are not analyzed as comitative, but as co-ordinators. There is thus ample evidence to call into question the analysis of =*le* as comitative.

When =*le* is used on one item, it translates as 'too' (356)(357), when used on more than one item, it translates as 'and' (358) ▱→14.3.2, p. 548, as shown in the following examples.

(356) *Se=ppe nyoonya=le su-mniiňğgal.*
 1s=POSS lady=ADDIT PAST-die
 'My wife died as well.' (B060115prs14)

(357) *Itthu suuňbu aayer=ka=le anthi-myaalak.*
 DIST wick water=LOC=ADDIT IRR-burn
 'That wick would burn in water, too.' (K051220nar01)

(358) *See siini Kluuňbu=ka=le Kandi=ka=le bannyak paasar pada=ka*
 1s here Colombo=LOC=ADDIT Kandy=LOC=ADDIT many shop PL=LOC
 transcriber software=*yang su-caari.*
 transcriber software=ACC PAST-search
 'I have looked for transcriber software all over here, in Colombo and Kandy.'
 (20070930eml01)

=*le* can also be used to coordinate negative statements, in which case it translates as *either*. The first occurrence of =*le* in (359) shows this. A second use of =*le* in negative statements is to emphasize the totality of the negation like English *any*. This use is also given in (359), as the second occurrence of =*le*.

(359) *Incayang=le kithang=pe mlaayu pada hatthu oorang=nang=le thumpath*
 3s.POLITE=ADDIT 1PL=POSS Malay PL INDEF man=DAT=ADDIT place
 thàrà-kaasi.
 NEG.PAST-give
 'He did not not give any position to any of our Malays either.' (N061031nar01)

Next to the associative and the coordinating use, =*le* can also be used to indicate exhaustivity, predominantly for pairs of two, where it translates as *both of them*.

(360) *Kanabisan=ka=jo **duva oorang=le*** *anà-thaau ambel [Andare **duva***
       last=LOC=EMPH   two   man=ADDIT PAST-know take   Andare two
       oorang=yang=le *asà-enco-kang aada] katha.*
       man=ACC=ADDIT CP-fool-CAUS   exist EMPH
       'At the very end, both women understood that Andare had fooled both of
       them.' (K070000wrt05)

This use does not have to rely on the occurrence of the word *duva* 'two' or a head
noun. The following example shows that the head noun can be missing in the first
part, and that the numeral can be missing, in the second part.

(361) *Ini   duuva=le   laayeng laayeng; [kithang=pe* Hill Country Malay
       PROX two=ADDIT other   other   1PL=POSS   hill country Malay
       *Club]=**le**   [Kandy Malay Association]=**le**.*
       club=ADDIT Kandy Malay association=ADDIT
       'These two are different, our Hill Country Malay Club and the Kandy Malay
       Association.' (K060116nar07)

The lexical word for 'every', *samma* is also often combined with =*le* in the ex-
haustive meaning.

(362) *Suda incayang=pe   aanak pada **samma=le***   musicians *pada=jo.*
       thus 3s.POLITE=POSS child PL   every=ADDIT musicians PL=EMPH
       'So all his children are musicians.' (G051222nar01)

=*le* is seldom found on the predicate and normally attaches to an NP, even if se-
mantically it has scope over the clause. In the following example, it is clear that,
semantically, it is the clauses that are coordinated by the two occurrences of =*le*.
However, morphologically, =*le* does not attach to the right edge of the clause, the
predicate *paasir* 'sand', but rather to an argument NP, in this case *muuluthka* 'in the
mouth'.

(363) *Lu=ppe   muuluth=ka=**le**   paasir, se=ppe muuluth=ka=**le***
       2S.FAMILIAR=POSS mouth=LOC=ADDIT sand   1s=POSS mouth=LOC=ADDIT
       paasir.
       sand
       'There is sand in your mouth and there is sand in my mouth.' (K070000wrt02)

This constraint against attaching =*le* to the predicate entails that it is very seldom
found on verbs.
    Besides NPs, it is also possible to attach =*le* to modifiers, as the following example

shows.

(364) *Skarang [Sri Lanka=pe=**le***   *Islam]* *pada muuka arà-thuuthup.*
     now   Sri Lanka=POSS=ADDIT Muslim PL   face   NON.PAST-close
     'Now also Sri Lanka's [female] Muslims cover their face .' (K061026prs01)

=*le* can also combine with interrogatives to form the meaning of every (365). This is discussed in more detail in ⊡→8.1.12, p. 424.

(365) *Derang=yang arà-iingath*   ***maana** vakthu=**le**, **maana** thaaun=**le**.*
     3=ACC   NON.PAST=think which time=ADDIT, which year=ADDIT
     '(The commemorators) think of him all the time, every year.' (K051206nar02)

When used in negative contexts, =*le* emphasizes the totality of the negation. In example (366), a verbal predicate is negated. The emphasis of the negation by =*le* yields the meaning that not one single man would step forward. Example (367) shows the affirmative use of *kaapangle* 'everywhen=always'. The negated usethereof ('never') is shown in (368).

(366) *Hatthu oorang=**le*** forward ***thama**-pii.*
     INDEF   man=ADDIT forward NEG.NONPAST-go
     'No one comes forward' (K051213nar01)

(367) Girls High School *Kandi=ka* *se=dang **kaapang=le** udahan*   *hatthu*
     girls high school Kandy=LOC 1s=DAT   when=ADDIT invitation INDEF
     arà-kiiring.
     NON.PAST-send
     'I am always invited to the Girls High School in Kandy.' (K061127nar03)

(368) *Go*   ***kaapang=le** saala **thamau**-gijja.*
     1s.FAMILIAR when=ADDIT wrong NEG.NONPAST-make
     'I never do wrong.' (B060115nar04)

This reinforcing meaning of =*le* can also be seen in (369), where it underscores the total absence of need to eat until 2 o'clock.

(369) *Itthu abbis maakang kalu kithang=nang bole=duuduk **hatthu=le***
       DIST finish eat   if   1PL=DAT   can-stay   INDEF=ADDIT
       ***jamà**-maakang=nang two duuva two o'clock=ke   sangke bole=duuduk.*
       NEG.NONFIN-eat=DAT   two two   two o'clock=SIMIL until can-stay
       'If we eat it up, we can stay up until 2 o' clock without eating anything.'
       (K061026rcp04)

=*le* is also used in concession (Smith et al. 2004),[39] very often combined with
a clitic as *itthule* ⊡→7.6.4, p. 412 ⊙→16.4, p. 692. The use of the additive clitic for
concession is also found in Sinhala (Jayawardena-Moser 2004:48).

(370) ***Itthu=le***   *kitham samma oorang asà-pii inni sudaari=pe saanak=ke.*
       DIST=ADDIT 1PL   all   man   CP-go PROX sister=POSS relative=SIMIL
       'But we have all become like this sister's relatives.' (B060115nar02)

The deictic can carry a postposition when used for concession, as shown in (371).

(371) a. *Duva-pulu ìnnam riibu   ùmpath raathus lima-pulu duuva* votes
           two-ty   six   thousand four   hundred five-ty   two   votes
           incayang=nang anà-daapath.
           3S.POLITE=DAT PAST-get
           'He got 26,452 votes.'
       b. ***Itthu=nang=le*** *incayang=nang=le   inni thumpath thàrà-daapath.*
           DIST=DAT=ADDIT 3S.POLITE=DAT=ADDIT PROX place   NEG.PAST-get
           'In spite of that, he did not get the seat either.' (N061031nar01)

The case of concession is the only one where =*le* can be found attached to verbs,
which might be due to the fact that the verbs are not in a predicate position there.
Example (372) shows the concessive use of =*le* on a clause, where it attaches to *kaala*
'lose'.

---

[39]Smith et al. (2004) note that =*le* cooccurs with the conditional *kalu* when used for concession, this was
not found in the corpus.

(372) a. *Sudaara TB Jayah inni* state council election *pada=nang duuduk*
Brother TB Jayah ᴘʀᴏx state council election ᴘʟ=ᴅᴀᴛ   exist.ᴀɴɪᴍ
aada.
exist
'Brother TB Jayah was in that state council election.'

     b. *Thiga-pulu ìnnam=ka incayang itthu=dering su-kaala.*
three-ty   six=ʟᴏᴄ   3s.ᴘᴏʟɪᴛᴇ ᴅɪsᴛ=ᴀʙʟ   ᴘᴀsᴛ-lose
'He lost in the '36 elections.'

     c. **Kaala=le** *thàrà=na=apa incayang=nang* appointed member=*pe*
lose=ᴀᴅᴅɪᴛ ɴᴇɢ=ᴅᴀᴛ=after 3s.ᴘᴏʟɪᴛᴇ=ᴅᴀᴛ   appointed member=ᴘᴏss
hathu thumpathan=yang government=*ka anà-kaasi.*
ɪɴᴅᴇꜰ post=ᴀᴄᴄ   government=ʟᴏᴄ ᴘᴀsᴛ-give
'In spite of having lost, he was given a post as appointed member in the government.' (N061031nar01)

6.4.6.3 similative =*ke(e)*

This clitic denotes similarity ⊙→15.10.2, p. 664. It can occur with a long vowel or with a short vowel. This distinguishes it from the other Coordinating Clitics, which always have a short vowel. It has been found to attach to a noun like in (373)-(375) or to a verb (376)(377).

(373) *Se=dang **baapa=ke*** soldier *mà-jaadi suuka.*
1s=ᴅᴀᴛ   father=sɪᴍɪʟ soldier ɪɴꜰ-become like
'I want to become a soldier like daddy.' (B060115prs10)

(374) *Kumbang rooja komplok duuva=pe naama=**kee**=jo.*
flower   rose bush   two=ᴘᴏss name=sɪᴍɪʟ=ᴇᴍᴘʜ
'Just like the names of the two rose bushes.' (K070000wrt04)

(375) *Hathu raaja=**ke** asà-paake=apa kampong=nang mà-pii maau.*
ɪɴᴅᴇꜰ king=sɪᴍɪʟ ᴄᴘ-dress=after village=ᴅᴀᴛ   ɪɴꜰ-go want
'I want to go to the village dressed up as a king.' (K070000wrt03)

(376) *Itthu pada [sraathus binthan pada **arà-kiilap**]=ke su-kiilap.*
ᴅɪsᴛ ᴘʟ 100   star   ᴘʟ sɪᴍᴜʟᴛ-shin=sɪᴍɪʟ ᴘᴀsᴛ-shine
'They shone like a hundred stars.' (K070000wrt04)

(377)   *Thiiga oorang=le [. . .]* volleyball *arà-**play-king=ke**.*
  three   man=ADDIT [. . .] volleyball NON.PAST-play-CAUS=SIMIL
  'Three men play [Sepaktakraw], like volleyball.' (N060113nar05)

It is also possible to use *=kee* on a proper noun.

(378)   *Karang **Sebastian=kee**=jo   spaaru oorang bannyak suuka=nang*
  now   Sebastian=SIMIL=EMPH some   man   much   like=DAT
  arà-blaajar.
  NON.PAST-learn
  'Now some people like to learn, just like you' (B060115cvs01)

The elements of the similarity have to be inferred by the addressee. In the following example, the dying of Andare is said to be similar to the dying of a king, but the king's dying is not expressed. Rather, the addressee has to infer that the state-of-affairs of dying can not be similar to a person, hence the dying must be similar to the mentioned person's dying instead.

(379)   *Andare=le   anà-maathi **hathu raaja=ke**.*
  Andare=ADDIT PAST-dead   INDEF king=ADDIT
  'And Andare died like a king.' (K070000wrt03)

Like the other clitics, *=ke* can combine with an interrogative pronoun, in which case it yields the meaning of 'some=WH'. In example (380) *aapa* 'what' + *=ke* yields the meaning 'something'; *saapa* 'who' + *=ke* in examples (381)(382) yields the meaning of 'someone'. Note that all three examples have irrealis semantics, with realis examples, *(s)aapa=so* would be more likely (see below).

(380)   *Lai   saapa mlaayu kuthumung=le, **aapa=ke**   connection hatthu aada.*
  other who   Malay see=ADDIT   what=SIMIL connection INDEF exist
  'If you see any other Malay, there will always be some kind of connection.'
  (K051206nar07)

(381)   *Lorang pada asà-raathang **saapape=ke** thaangan=ka=jo pukurjan*
  2PL   PL   CP-come   whose=SIMIL hand=LOC=EMPH work
  mà-gijja athi-jaadi.
  INF-make IRR-come
  'You will always have to work under somebody.' (K051206nar07)

(382) *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
       Hopefully other know brother sister   PL=LOC   can-ask   take
       *[nya-gijja lai **saapa=kee** aada=si   katha].*
       PAST-make other who=SIMIL exist=INTERR QUOT
       'Hopefully, you can enquire from another person you know whether there is
       someone else who did something.' (N061031nar01)

When combined with *mana* 'where', =*ke* yields the meaning 'somewhere' in affirmative contexts, combined with *kaapang* 'when' it yields 'someday'. Note that *maanake* in (383) refers to a place, but is not translated as 'somewhere' in English because the place name is specified. This has an influence on the expression of indefiniteness in English, but not in SLM.

(383) ***Maana=ke** hathu government=pe   hathu thumpath=ka asà-pii pukurjan*
       where=SIMIL INDEF government=POSS INDEF place=LOC   CP-go   work
       bole=girja.
       can-do
       'They can go to some government place and work there.' (K051222nar05)

(384) *See lorang=nang arà-simpa   **kaapang=ke** see lorang=nang ithu*
       1s 2PL=DAT   NON.PAST-promise when=SIMIL 1s 2PL=DAT   DIST
       uuthang arà-baayar   katha.
       debt   NON.PAST-pay QUOT
       'I promise you that I will pay back that debt some day.' (K070000wrt04)

Like =*le*, =*ke* can be used in negative contexts when combined with indefinite expressions or an interrogative pronoun. The use of *hatthuke* 'none' is shown in examples (385)(386).

(385) *Gaathal su-kuurang kalu, suda **hatthu=ke thraa**.*
       itching PAST-few   if   thus INDEF=SIMIL NEG
       'When the itching has diminished, none will be left.' (K060103cvs02)

(386) *Snow-white=nang=le Rose-red=nang=le   ini **hatthu=ke***
       Snow.white=DAT=ADDIT Rose.Red=DAT=ADDIT PROX INDEF=SIMIL
       ***thàrà**-mirthi.*
       NEG.PAST-understand
       'Snow White and Rose Red did not understand a thing.' (K070000wrt04)

=*ke* combined with interrogative pronouns gives the meaning of *no-WH* in negated

contexts. This is shown for *maanake* 'nowhere', *saapayangke* 'no one's' and *kaa-pangke* 'never'.

(387) *Itthu blaakang se=dang karang **maana=ke**   pii **thàràboole**.*
      DIST after   1s=DAT   now   where=SIMIL go cannot
      'I cannot go anywhere now.' (K061120nar01)

(388) *[Bannyak aari=dering **saapa=yang=ke thàrà**-enco-kang   katha]*
      many   day=ABL   who=ACC=SIMIL NEG.PAST-fool-CAUS QUOT
      anà-iingath Andare.
      PAST-think   Andare
      'Andare thought that he had not fooled anybody for a long time.'
      (K070000wrt05)

(389) *Siini duuduk kalu **kaapang=ke** lorang=nang lorang **thàràboole** kaaya.*
      here stay   if   when=SIMIL 2PL=DAT   2PL   cannot   rich
      'If you stay here you can never become rich' (K051206nar07)

    =ke can also be used to indicate professions, as in the following two examples, where it would be translated as English *as*.

(390) **Visiting agent=*kee*** *dhraapa=so*   estates *pada=nang se=ppe baapa*
      visiting agent=SIMIL how.many=UNDET estates PL=DAT   1s=POSS father
      anà-bagijja.
      PAST-work
      'My father worked on so many estates as a visiting agent.' (K051201nar01)

(391) *Kithang=pe inni kaake   mooyang pada samma oorang **soldiers=kee***
      1PL=POSS   PROX grandfather greatgrand PL   all   man   soldiers=SIMIL
      su-dhaathang.
      PAST-come
      'These our forefathers all came as soldiers.' (B060115nar01)

    The profession can be interpreted with some leeway, so that functions are also possible in this context. In (392), the speaker is talking about nominating someone, i.e. giving him the function of nominee.

(392) *Sudara TK Azoor* Sri Lanka Muslim Congress=*dering* **hathu**
Brother TK Azoor Sri Lanka Muslim Congress=ABL   INDEF
nominate=*kee* *thaaro aada.*
nominate=SIMIL put   exist
'Brother TK Azoor was a designated as a nominee for the Sri Lanka Muslim Congress.' (N061031nar01)

=*ke* is also used in the equative comparative constructions ⊙→15.10, p. 664. In this case, the emphatic clitic =*jo* is normally present.

(393) *Se=ppe aanak loram=pe aanak=ke=jo*   *bìssar.*
1s=POSS child   2PL=POSS child=SIMIL=EMPH big
'My child is as big as your child.' (K081103eli04)

If *keejo* is present on the first element of the comparison, the additive clitic =*le* must be present on the second one (394)(395).

(394) *Kithang=**kee=jo** deram pada=**le** thaau.*
1PL=SIMIL=EMPH 3PL   PL=ADDIT know
'They know (it) as well as we do.' (K081105eli02)

(395) *Kithang=**kee=jo** derang=**le** kaaya.*
1PL=SIMIL=EMPH 3PL=ADDIT rich
'They are as rich as we.' (K081105eli02)

=*ke* is also used for expressing vagueness of amounts as in the following three examples.

(396) *Hatthu dooblas thigablas **thaaun=ke**.*
one   twelve thirteen year=simil
'About 12 or 13 years.' (K051206nar17)

(397) *Hathu buulang=nang hathu **15,000=kee*** *go=dang arà-kiiring.*
one   month=DAT   INDEF 15,000=SIMIL 1s=DAT NON.PAST-send
'(They) send about 1500 (Rupees) a month.' (B060115nar04)

(398)   *Punnu=le Seelong=ka arà-duuduk,*   *hathu **ìnnam-pulu liima***
        full=ADDIT Ceylon=LOC NON.PAST-exist.ANIM INDEF six-ty   five
        riibu=ke   *mlaayu pada.*
        thousand=SIMIL Malay PL
        'A lot are in Sri Lanka, about sixty-five thousand Malays.' (K060108nar02)

In example (396), the speaker is not sure about the exact number of years. This is indicated by giving two numbers, *dooblas* 'twelve' and *thigablas* 'thirteen', but also additionally by =*ke*. In example (397) and (398), the speaker indicates that the indicated figure is only a rough estimate.

6.4.6.4   indeterminate =*so*

This clitic is glossed as 'undetermined' here. When it attaches to two items, it indicates that any of them would fulfill the truth value of the predication. This is the case in (399), where people of different professions are said to know the speaker, and, in fact, the profession does not matter because everybody knows the speaker. In (400), the same speaker asserts that he knows everything about tailoring, and that it does not matter whether it is sewing or cutting since he has complete knowledge.

(399)   Doctors *pada=**so*** police ASP=***so***   judge=***so,***   *samma oorang thaau*
        doctors PL=UNDET police ASP=UNDET judge=UNDET all   man   know
        see=yang
        1s=ACC
        'Whether they be doctors, police, A(ssistent) S(uperintendents of) Police, or judges, all people know me.' (B060115nar04)

(400)   *Suda go buthul baaye=nang bole=thaau, mà-jaaith=**so***   *mà-poothong=**so**.*
        thus 1s correct good=DAT   can=know   INF-sew=UNDET INF-cut=UNDET
        'So I know (those things) very well, be it sewing, be it cutting.' (B060115nar04)

In the examples above, =*so* attached to an NP. =*so* can also attach to a subordinate clause if the clause expresses a question or ignorance. In this respect, =*so* resembles English *whether*. Examples (401) and (402) show the use of =*so* on a subordinate clause conveying ignorance about the fact stated in the subordinate

(401)   *Duppang=ka inni samma anthi-oomong=**so** thàrà-thaau.*
        before=LOC   PROX all   IRR=say=UNDET   neg-know
        'What they might have said before, I don't know.' (G051222nar02)

(402) *Karang jaadi karang kitham=pe aanak pada=nang inni duppang=ka*
now   become now   1PL=POSS child PL=DAT   PROX front=LOC
inni samma anthi-oomong=so thàrà-thaau.
PROX all   IRR-speak=UNDET NEG-know
'Now it happens that our children, after us, they do not know how to say all
this.' (G051222nar02)

=*so* can also attach to interrogative pronouns. In that case, the combination yields
an indefinite pronoun with an arbitrary aspect to its meaning, like *saapaso* 'someone'
in (403).

(403) **Saapa=so** Malay exam *arà-girja.*
who=UNDET Malay exam NON.PAST-make
'Someone is taking a Malay exam.' (K060103cvs01)

This use is also possible with other interrogative pronouns, as the following exam-
ples attest.

(404) *Uuthang=ka asà-pii=apa see picakang, ithu* **aapa=so,**   *ithu daavong*
jungle=LOC CP-go=after 1s break   DIST what=UNDET DIST leaf
pada.
PL
'I went to the woods and I was breaking these what's-their-name, these leaves.'
(K061125nar01)

(405) *Thapi* **aapacara=so** *itthu samma asà-iilang   su-aada.*
But   how=UNDET   DIST all   CP-disappeared PAST-exist
'But somehow it had all disappeared.' (K20070920eml01)

The combination of WH=*so* must be distinguished from the use of =*so* on WH-
words which just happen to occur in a subordinate marked by =*so*. In these cases,
the conjunction of the pronoun with =*so* is accidental. This is shown in the following
example, where the subordinate 'how many children grandfather had' contains both
the interrogative *dhraapa* 'how.many' and the clitic =*so*, but the meaning is not 'how
ever many', rather it is an indirect question to which the answer is not known.

(406)   *[Kaake=nang* **dhraapa=so**   *bannyak aanak pada] se=dang kalu*
        grandfather=ᴅᴀᴛ how.many=ᴜɴᴅᴇᴛ much   child ᴘʟ   1s=ᴅᴀᴛ if
        blaangang thàrà-thaau.
        amount   ɴᴇɢ-know
        'As for me, I do not know the exact amount of how many children my grand-
        father had.' (K051205nar05,K081105eli02)

The use of *=so* in subordinates of question or ignorance on the one hand and
its use as indefinite pronoun with an arbitrary component to its meaning are related
semantically and could be captured under the rubric 'ignorative'. In all three cases,
we are ignorant of the precise reference, be it because we do not care (pronoun),
we are asking (question) or we assert our ignorance (subordinate). Sometimes, these
meanings conflate, as in (407), where the occurrence of *=so* on *dhraapa* 'where' can
either be a subordinate rhetorical question or an indication of indefinite quantity.

(407)   *See=le   pii aada,* **dhraapa=so**,   *duuva thiiga skali.*
        1s=ᴀᴅᴅɪᴛ go exist how.many=ᴜɴᴅᴇᴛ two   three time
        'Me too I went there about two or three times.' (B060115nar05)

As a special case, *=so* can be used in a construction involving a reduplicated in-
terrogative pronoun, yielding the meaning of 'Wh-ever'. In that case, *=so* attaches to
the predicate, as seen below.

(408)   *Inni* **saapa~saapa**=*ka inni mlaayu pakeyan pada aada*=**so**, *lorang pada*
        prox who~red=ʟᴏᴄ   ᴘʀᴏx Malay dress   ᴘʟ   exist=disj 2ᴘʟ   ᴘʟ
        ini   mlaayu pakeyan=samma ini   kaving=nang mà-dhaathang bannyak
        ᴘʀᴏx Malay dress=with   ᴘʀᴏx wedding=ᴅᴀᴛ ɪɴꜰ-come   much
        uthaama.
        honour
        'Whoever owns such Malay dresses, your coming together with this Malay
        dress to the wedding will be greatly appreciated.' (K060116nar04)

6.4.6.5   *=pon* 'any'

This is a negative clitic with about the same meaning as the English *not any* ⊙→15.12,
p. 671. When occurring on one item it translates as something like 'not even one'(409)-
(411).

(409) *Kithang **hatthu**=oorang=**pon** thàrà-iingath.*
     1PL   INDEF=man=any   NEG.PAST-think
     'We cannot think of any person.' (B060115nar02)

(410) *See pukuran **hatthu=pon** thama=gijja,   ruuma=ka arà-duuduk.*
     1s   work   INDEF=any   NEG.NONPAST-make house=LOC NON.PAST-stay
     'I don't do any work, I stay at home.' (B060115prs03)

(411) ***Hatthu** aanak=**pon** thàrà-baae katha thraa.*
     one   child=any NEG-good   QUOT NEG
     'There is not one child that you would want to call naughty.' (B060115cvs13)

When occurring on more than one item, it translates as 'neither . . . nor' (412).

(412) *Se maaňǧga=pon piisang=pon thàrà-suuka.*
     1s mango=any   banana=any NEG-like
     'I like neither mangoes nor bananas.' (K081103eli04)

=*pon* can only be used in negative contexts in the cases mentioned above. An exception to this is the WH=*pon* V-construction ⊡→8.4, p. 447, which is used to indicate the arbitrariness of a referent. This construction can be used both in negative (413) and in positive contexts (414).

(413) *Suda itthu kithang=nang **kaapang=pon** thama=luupa.*
     thus DIST 1PL=DAT   when=any   NEG.NONPAST=forget
     'So, we will never ever forget this.' (B060115nar02)

(414) *Kithang **craapa=pon** kithang=pe kappal asà-ambel, kithang su-baalek*
     1PL   how=any   1PL=POSS ship CP-take   1PL   PAST-turn
     kithang=pe harbour=nang.
     1PL=POSS   harbour=DAT
     'Anyhow, we managed to take our ship and return to he harbour.'
     (K051206nar20)

Note that *craapapon* above conveys a stronger meaning of incredulity than *craapaso*, which would also have been possible.

Next to nouns and WH-words, =*pon* can also attach to the indefiniteness marker *hatthu*, where the resulting string *hatthupon* means 'not anything, nothing'.

(415)   Bus=*ka* **hatthu=pon** *mà-kirja thàràboole.*
        bus=LOC INDEF=any   INF-make cannot
        'You can't do anything on the bus.' (K061125nar01)

This morpheme is not very frequent. Is has to be distinguished from its lexical
homophone *pon* 'bride'. The etymology is unclear.

6.4.7   Information structure clitics

There are two information structure clitics. The first one, =*jo* is very frequent and used
for many different but related functions. The second one, =*jona* ⊡→6.4.7.2, p. 385, is
used to request confirmation for the element it attaches to. It is a lot less frequent than
=*jo*.

6.4.7.1   emphatic =*jo*

This is the ubiquitous emphatic clitic (cf. Bakker 2006:143). It is used on various
emphasized constituents, such as constituents in argument focus, superlatives and el-
atives, contrastive topics etc. The phonetic form is invariant.

=*jo* can attach to noun phrases, postpositional phrases, clauses and utterances, as
the following examples show:

(416)   *Itthu kumpulan=dang [**derang**]*NP=*jo bannyak arà-banthu.*
        DIST association=DAT 3PL=EMPH   much   NON.PAST-help
        'They helped the association a lot.' (B060115cvs01)

(417)   *[**Inni nigiri=ka**]*PP=*jo   kithang=pe aanak.buuva pada=yang asà-simpang*
        PROX country=LOC=EMPH 1PL=POSS children   PL=ACC   CP-keep
        'We wanted to keep our children here in this country.' (K051222nar04)

(418)   *Itthu abbis dhaathang [**custard powder=dering**]*PP=*jo arà-kirja.*
        DIST compl come   custard powder=ABL=EMPH   NON.PAST-make
        'This one is such that it is made with custard powder.' (K061026rcp02)

(419)   *Suda [**derang pada dhaathang abbis**]*CLS=*jo, derang pada itthu=jo*
        thus 3PL   PL   come   finish=EMPH 3PL   PL   DIST=EMPH
        arà-biilang.
        NON.PAST-say
        'So when they had finally arrived, this is what they said.' (K051206nar05)

(420) *[Itthu katha]*_{UTT}*=jo Mahindha arà-biilang.*
      DIST   QUOT=EMPH   Mahinda NON.PAST-say
      'That's what Mahinda [Rajapksa, President of Sri Lanka] is saying.'
      (K051206nar11)

The first major function of *=jo* is indicating argument focus (narrow focus). In example (421), the father has already spoken about the schools his eldest children attend, one of them being Swarnamali College. The presupposition is then that all his children study in some school, and the new information as for the last child is simply which school that is. The focus domain in (421) only covers the argument *Swarnamali=ka* and not the predicate *aràblaajar* 'learn'. Note that we are dealing with argument focus and not contrastive focus in (421) since the other children also attend Swarnamali College.

(421) *Kanabisan aanak incayang=le **Swarnamali=ka=jo**   arà-blaajar.*
      last   child, 3s=ADDIT   Swarnamali=LOC=EMPH NON.PAST-learn
      'As for my last child, she also attends Swarnamali College' (G051222nar01)

Another example of argument focus is found in (422), taken from a story about a thief. The presupposition is that a thief steals, so the new information is only the victim of the theft. This smaller focus domain is indicated by *=jo*.

(422) *Kaaya oorang — monied people — =dering=jo arà-cuuri.*
      rich   man   monied people   =ABL=EMPH NON.PAST-steal
      'He steals from the rich.' (K051206nar02)

The presupposition is exactly stated in (423) below, where the king tried to fool Andare. But contrary to the presupposition, it was not Andare who got fooled but the king himself. We are thus dealing with argument focus, again indicated by *=jo*.

(423) *Suda kanabisan=ka raaja Andare=yang mà-enco-king asà-pii, **raaja=jo***
      thus last=LOC   king Andare=ACC INF-fool-CAUS CP-go   king=EMPH
      su-jaadi   enco.
      PAST-become fool
      'So finally the king had tried to make a fool out of Andare, but it was the king who turned out to be the fool.' (K070000wrt02)

A double use of *=jo* for argument focus in two adjacent clauses is found in (424).

(424) ***Guunung=ka=jo***   *kithang arà-duuduk;* ***guunung=nang=jo*** *kithang*
       mountain=LOC=EMPH 1PL   NON.PAST-stay mountain=DAT=EMPH 1PL
       arà-pii.
       NON.PAST-go
       'It is in the hills that we live, it is to the hills that we go.' (B060115prs01)

Above, *=jo* was used only on new information. It is also possible to use *=jo* on given information, as in (425), where the content provided so far is stated as the reason for the son's being born there. The information is not new, which is obvious from it being referred to by a deictic anaphora. But the relation between the birth and aforementioned events is new, it is a causal relationship, expressed by *subbath* 'because'. This new information is not the predicate (which would be *anà-braanak* 'give birth' in this case), but rather an argument. As above, the argument in focus is marked by *=jo*.

(425) ***Itthu=subbath=jo***   *incayang=yang siithu anà-braanak.*
       DIST=because=EMPH 3S.POLITE=ACC   there PAST-give.birth
       'It is because of that that he was born there' (K060108nar02)

=jo can also be used to indicate contrastive focus, as in (426), where the big associations in the first sentence are contrasted with the small ones in the second.

(426)   a. ***Bìssar*** *atthu kumpulan   thraa.*
            big   one   association NEG
            'A big association doesn't exist.'

        b. ***Kiccil*** *kumpulan   pada=**jo**.*
            small association PL=EMPH
            '(But) what there is, (are) small associations.' (N060113nar01)

Another example of *=jo* in contrastive focus is (427), where English and Malay contrast in being present and absent, respectively.

(427) *Itthu=ka **mlaayu thraa**, bannyak=nang* **English=jo**   *aada.*
       DIST=LOC Malay   NEG   much=DAT   English=EMPH exist
       'There is no Malay. What there is, is a lot is English.' (B060115prs15)

Another use of *=jo* is to indicate specificational information structure. (428) is a case in point. The preceding stretch of discourse established that Malays in the past used to be brought to protect the king. The members of that set are unknown, but (428) informs us that the speaker is a person of whom the predicate brought.to.protect(X) is true. The ascriptional information structure would be *as for me, it is true that I am one*

of the people who ... whereas the specificational information structure is *As for the set of people who came to protect, I am a member of that set.*

(428) *[[Itthu mà-jaaga=nang anà-baa   mlaayu=dering] [anà-aaji baa*
      DIST   INF-protect=DAT PAST-bring Malay=ABL   PAST-bring bring
      mlaayu=dering] satthu oorang]=jo see.
      Malay=ABL   one   man=EMPH 1s
      'One of the Malays brought to protect him is me' (K060108nar02)

Very similar to the specificational context are identificational contexts. While in the former, the missing argument of a predicate is specified, in the latter, the identity of the referents of two arguments is asserted. An example would be *The murderer was the butler*, where the pre-established terms of *the murderer* and *the butler* turn out to refer to the same person. This differs from the specificational contexts above in that there are already *two* terms, whose references are said to be identical, while in specificational contexts, we are dealing with only one argument, and one predicate.

In (429) the identity of the the grandfather's younger brother, who has been established before, is asserted to be identical to the identity being referred to as *baapa* 'my father', given in the context of the discourse as well.

(429) *Suda itthu kaake=pe   aade=pe   aanak=jo baapa.*
      thus DIST grandfather=POSS younger.sibling=POSS child=FOC father
      'So that grandfather's younger sister's child is my father.' (K051205nar05)

The same is true of (430), where a certain discourse referent is established beforehand, and that discourse referent is asserted to be identical to the speaker and his fellows. The pre-established, referential, nature of *itthu oorang pada* 'those people' is visible by the distal deictic *itthu*, which functions as anaphora here. The referential nature of *kithang* 'we' is given since speech act participants cannot be predicates. Again, =*jo* is attached to the first of the equated referents.

(430) *Itthu oorang pada=jo kithang.*
      DIST man   PL=EMPH 1PL
      'We are those people' (K051222nar03)

Note the difference between (428) and (430). In the former, the presupposition is that there is a set of referents with certain properties. The assertion is that the speaker forms part of that set. In (430), the presupposition is that there are some people, who are *identical* with the speaker and his fellows. In the former case, we are dealing with class-membership while in the latter case, we are dealing with identity.

As a final focalizing function, =*jo* can also be used to emphasize the TAM-meaning of a quasi-prefix, as in (431)

(431) *Derang su-biilang thraa Mrs. Jaaya **masthi=jo** raathang*
       3PL   PAST-say   no   Mrs. Jaaya must=EMPH come
       'They said "No! Mrs Jaaya MUST come! " ' (B060115prs01)

So much for the use of *=jo* in focal contexts. But *=jo* can also be used on topics, namely when they are contrastive, as in (432)

(432) *[Pon=pe   ruuma=ka=**jo**   thaama duuva thiiga aari athi-duuduk.*
       bride=POSS house=LOC=EMPH earlier two   three day IRR-stay
       'At the bride's house, then, they would stay two, three days back then.'
       (K061122nar01)

This contrasts with preceding information, which was about the time of stay in the bridegroom's house. In (432), it is not the case that the presupposition was that people stayed somewhere for two or three days, but the precise location was unknown. This would be the case for a specificational context. Rather, the presupposition was that some people stayed at the bridegroom's house. The sentence at hand then establishes a new, contrastive, topic (*ponpe ruumaka* 'at the bride's house'), which is marked by *=jo*, and asserts that one used to stay there two or three days.

These are the information structure uses of *=jo*. Besides this, *=jo* also has some other uses which are not directly related to information structure. *=jo* can be used to emphasize the fact that a certain referent is a particularly good member of the set denoted by the predicate. This is the case for the elative (433) and the superlative (434).

(433) *Hatthu komplok **bannyak=jo** puuthi caaya.*
       one   bush   very=EMPH white colour
       'One bush was very, very white.' (K070000wrt04)

(434) *Seelon=ka   **bìssar=jo** pohong.*
       Ceylon=LOC big=EMPH tree
       'The biggest tree in Sri Lanka.' (K081103eli04)

In the following example, it is hard to tell whether the occurrence of *=jo* is due to the superlative (*kàthaama* 'first') or to the equational context. It could be the case that *=jo* is used to indicate superlative, similar to English *very first*, and analogous to the structure found in (434). Another analysis would be that in discourse, a referent 'first victim of terrorism' was established, which is then equated with the currently active referent *incayang* 'he', analogous to example (429) where the grandfather's younger sibling's son is equated with the father.

(435) Terrorist *hatthu=dering anà-maathi kàthaama oorang=**jo** incayang.*
      terrorist INDEF=ABL   PAST-dead first   man=EMPH 3S.POLITE
      'The first man to be killed by a terrorist was him.' (K051206nar02)

=jo can also be used on quantifiers to stress their scope. This can be done to both increase (436) and diminish (437) the number of entities.

(436) Cricket *arà-maayeng **samma=jo** kithang.*
      cricket NON.PAST-play every=EMPH 1PL
      'EVERYbody played cricket.' (K051201nar02)

(437) *Sri Lankan oorang pada **sdiikith=jo** arà-duuduk.*
      Sri Lankan man   PL   few=EMPH NON.PAST-stay
      'There are only/very few Sri Lankans.' (K061026prs01)

Note that this last example can also have the interpretation of 'only' (see below).

Furthermore, *=jo* can be used to indicate precision in quantity, like in the following example

(438) *Thiiga oorang, thiiga oorang=le, **thiiga oorang pada=jo** itthu ini*
      three   man,   three man=ADDIT, three man   PL=EMPH DIST PROX
      volleyball *arà-play-king=kee.*
      volleyball NON.PAST-play-CAUS=simil
      'Three men, three persons, three people play it, like volleyball.'
      (N060113nar05)

Other instances of the use of *=jo* for precision are give in the following two examples:

(439) 58=*ka=**jo***   *anà-mulain.*
      58=LOC=EMPH PAST-start
      'It started in '58.' (B060115prs17)

(440) ***Lima raathus** vote=**jo**   itthu time=ka daapath aada.*
      five   hundred vote=EMPH PROX time=LOC get   exist
      'He got 500 votes at that time.' (N061124sng01)

Example (440) can actually also be interpreted as not referring to 'precisely' 500, but to 'only' 500. This is another use of *=jo*, of which another example is given in (441). In this example, it is not clear whether there were four or five days left, hence

the 'precision' interpretation is not possible. =*jo* in this context rather indicates that
the amount was less than expected.

(441)   ***Ùmpath liima*** *aari=**jo**   aada,* exam=*nang blaajar.*
        four   five   day=EMPH exist exam=DAT   learn
        'There were only four or five days left to learn for the exam.' (K051220nar01)

Other instances of this 'only'-reading are (442)-(445).

(442)   *Kapang-baavung **spuulu** aari=**jo**   su-aada   exam=nang.*
        when-rise   ten   day=EMPH PAST-exist exam=DAT
        'When she got better only ten days remained until the exam' (K051220nar01)

(443)   *Lai   **sdiikith** aari=**jo**   go=dang bolle=duuduk.*
        more little   day=EMPH 1s=DAT   can=stay
        'I only have some more days to stay (on Earth).' (B060115nar04)

(444)   ***Konnyong*** *mlaayu=**jo**   Seelong=ka thiiňǧgal aada.*
        few   Malay=EMPH Ceylon=LOC settle   exist
        'Few Malays have settled down in Sri Lanka.' (K051222nar06)

(445)   a. ***Hatthu*** *lancar=**jo**   English.*
            one   subject=EMPH English
            'Only one subject was taught in English.'

        b. *Layeng samma mulbar.*
            other   all   Tamil
            'All others were in Tamil.'
            [. . .] [In 6th grade]

        c. *Siithu samma* English=*ka,* ***hatthu=jo*** *mulbar.*
            there   all   English=LOC one=EMPH Tamil
            'There, everything was in English, only one subject was in Tamil.'
            (K051213nar02)

This can then further be reinforced by using it on the lexical item *sajja* 'only', as
in (446) (447).

(446)   *Lubaarang pada=nang **sajja=jo**   kithang arà-gijja.*
        festival   PL=DAT   only=EMPH 1PL   NON.PAST-make
        'We prepare it only for festivals.' (K061026rcp01)

(447) *Itthu baaye mosthor=nang bole=kirja oorang mlaayu=nang **sajja=jo**.*
DIST good manner=DAT can-make man Malay=DAT only=EMPH
'The Malays are the only ones who can do it in such a good way.'
(K061026rcp01)

The 'only-reading is also available in the temporal domain when =*jo* focusses a temporal relator noun like *blaakang* 'after' in (448).

(448) *Siini Seelong=nang dhaathang=nang **blaakang=jo** incayang cinggala*
here Ceylon=DAT come=DAT after=EMPH 3s.POLITE Sinhala
asà-blaajar=apa ...
CP-learn=after
'Only after having come here he learned Sinhala and ...' (K060108nar02)

The 'only'-reading of *hatthujo* as in (445) must not be confounded with another reading, namely *hatthujo* 'same', as given in (449). In this example, an 'only'-reading is not possible. We are rather dealing with a 'precision'-reading: they are exactly one, i.e. they are the same.

(449) *Igaama arà-muuji mosthor samma **hatthu=jo**.*
religion NON.PAST-pray manner all one=EMPH
'The religion and the way of praying are all the same.' (K061026prs01)

6.4.7.2 phatic =*jona*

This enclitic expresses that the speaker thinks that the information is common ground between the speaker and the addressee, similar to English *y'know*. It is used as a mild request for confirmation ⊙→17.1.3, p. 699 and does not require a reaction from the addressee. It thus serves to get confirmation about the status of the conversation channel and is glossed as 'phatic' here. An example is (450), where the speaker thinks that the addressee is probably aware of the Malays' linguistic preferences of that period.

(450) *Punnu mlaayu pada kàthaama English=**jona** anthi-oomong.*
many Malay PL earlier English=PHAT IRR=speak
'Many Malays would speak English, y'know, in former times.'
(K051222nar06)

Another example is (451), where the speaker asserts that he goes to the mosque for prayers. This is not surprising given his religion. The fact that the speaker regards this as common ground is again marked by =*jona*.

(451)   a. Mosque=*nang kapang-pii samma oorang=nang go   athi-kaasi.*
            mosque=DAT when-go   all   man=DAT   1S.FAMILIAR IRR-give
            'When I go to the mosque, I give everybody'

       b. *Mà*-pray *arà-pii=**jona**   ithu vakthu go   athi-kaasi.*
            INF-pray NON.PAST-go=PHAT DIST time 1S.FAMILIAR IRR-give
            'I go to pray, y'know, and then I give them' (B060115nar04)

As seen in the two examples above, =*jona* can attach to nouns and verbs, hence
it is a clitic. The following two examples provide two more instances of the use of
=*jona*.

(452)   *Muuka asà-thuuthup=**jona** arà-dhaathang.*
        face   CP-close=PHAT   NON.PAST-come
        'They come with their faces covered, y'know.' (K061019nar02)

(453)   *Ikang see nya-caanya 'lorang=nang bole=oomong=**jona** English=dering.'*
        then 1s PAST-ask   2PL=DAT   can=speak=PHAT   English=ABL
        'Then I asked them: "You can speak in English, can't you".' (K061019nar02)

The following stretch of discourse shows first the use of =*jona* to request confir-
mation and then the use of =*jo* to assert certainty.

(454)   a. *Saapa anà-biilang anà-dhaathang ini   Malay mà-girja katha.*
            who   PAST-say   PAST-come   PROX Malay INF-make QUOT
            'Who had said that Sebastian had come to do this Malay thing?'

       b. *Izi=**jona** anà-biilang.*
            Izi=PHAT PAST-say
            'It was Izi who said it, no?'

       c. [pause]

       d. *Izi=**jo**   anà-biilang.*
            Izi=EMPH PAST-say
            'Izi was the one who said it.' (K060103cvs01)

6.4.8   Evidential clitic

6.4.8.1   evidential *kiyang*

Kiyang is the evidential marker (Bakker 2000a, 2006). It is used to indicate second-
hand information (Aikhenvald 2004:31) ⊙→15.7.2.5, p. 659. In my corpus, I have
kiyang and *keyang*, Saldin (2001:69) and Smith & Paauw (2006:175) report *kanyang*.

It is very difficult to prompt informants to produce this morpheme. Constructing a situation where the use of an evidential marker would be felicitous can be done, but evidentiality marking seems to be optional and informants then do not encode this information in their response. Elicitation through English was also not fruitful. Sri Lankan English has a grammaticalized evidentiality marker *it seems*, but when asked to translate sentences containing it, informants used the lexical solution ... *katha arà-biilang* '..., it is said' and not the grammatical marker. Sinhala as a priming language finally elicited (in the literal sense) *kiyang*, but the chance for interference or word-for-word *ad hoc* translations is extremely high here, so that the following comments must be taken with a grain of salt.

Like in Sinhala, the evidential clitic attaches to the argument of which the speaker only has second-hand information. The Sinhala sentences are given for comparison.

(455)   *Haturaa*   *balahatkaarayen* *lamun*   *baňdavaaganavaa*
       Satthuru   *paksa*   *aanak.pada* *arà-*   *kumpulkang*
       enemy   PL force   children   NON.PAST recruit
       =lu.
       =kiyang
       =EVID
       'The enemy is recruiting children, it seems.'

(456)   **Haturaa** *=lu*   *balahatkaarayen* *lamun*   *baňdavaaga*
       Satthuru *=kiyang* *paksa*   *aanak.pada* *arà-*   *kumpulkang*
       enemy   =EVID   force   children   NON.PAST- recruit
       -nnee.

       EMPH
       'The enemy, it seems, is recruiting children.'

(457)   *Haturaa* *balahatkaarayen* **lamun**   *=lu*   *baňdavaaga*
       Satthuru *paksa*   **aanak.pada** *=kiyang* *arà-*   *kumpulkang*
       enemy   force   children   =EVID   NON.PAST- recruit
       -nnee.

       =EMPH
       'The enemy is recruiting CHILDREN, it seems.'

Kiyang is used to mark information as hearsay. For instance, in (458), the main Sri Lankan airport was closed during nighttime after an LTTE attack. Some time later, I received a letter which included (458).

(458)   *Seelong* Airport=*yang duva-pulu-umpath vakthu*=*le   asà-bukka*
        Ceylon Airport=ACC   two-ty-four   hour=ADDIT CP-open
        *arà-simpang **kiyang**.*
        NON.PAST-stay EVID
        'The Ceylon Airport is open 24h, it seems.' (Letter 26.06.2007)

The writer reports some knowledge he has gathered from the newspapers. He was not at Katunayaka Airport himself to check whether the airport is actually open during night time now. This lack of personal verification is the reason why *kiyang* is used. Another example is the explanation of the linguist's work to another Malay (459), where the speaker reports what he thinks he has understood of my work. Since he is not completely sure, he uses *kiyang*.

(459)   *Arà-blaajar   mlaayu ini   buk   mà-thuulis **kiyang**.*
        NON.PAST-learn Malay PROX book INF-write   EVID
        '(Sebastian) is learning Malay to write this book of his, it seems.'
        (K051222nar07)

In a similar vein, the report of my planned travel activities also receives *kiyang*.

(460)   *Karang incayang arà-pii   **kiyang**, laiskalli Netherland pii=apa, laiskali*
        now   3S.POLITE NON.PAST-go EVID   again   Netherland go=after again
        September *arà-raathang   katha anà-biilang.*
        September NON.PAST-come QUOT PAST-say
        '(He) told (me) that he would apparently leave and go back to the Netherlands and come back in September.' (B060115prs15)

An example without the researcher as the topic of the text is (461), where the hiding place of a talisman is discussed. Since this place is of course secret, knowledge of it is not directly trustworthy, so usage of *kiyang* seems appropriate

(461)   *Incayang=pe   kaaki=ka anà-aada **kiyang** hatthu talisman.*
        3S.POLITE=POSS leg=LOC PAST-exist EVID   INDEF talisman
        'In his leg there was a talisman it seems.' (K051206nar02)

Another rumour which is indicated by *kiyang* is that some Malaysians had come after WWII to force the Sri Lankan Malays to return to South East Asia. The speaker is not too sure about this, hence the use of *kiyang*.

(462)   a. *derang pada panggel=nang blaakang [thàrà-dhaathang oorang*
              3PL   PL   call=DAT   after   NEG.PAST-come   man
              *pada]=nang nya-force-kang **kiyang**.*
              PL=DAT   PAST-force-CAUS EVID
              'After they had called (them), (they) apparently forced the people who had
              not come (to join).'

       b. *Blaakang thàrà-pii   **kiyang**.*
              after   NEG.PAST-go EVID
              'But still (they) did not go it seems' (K051206nar07)

In the above examples the past tense marker *anà-/nya-* is used, which could be indicative of *kiyang* triggering a special verb form, like Sinhala *=luu*, which triggers the 'emphatic' form ((456)) (Gair 2003:797). In SLM, this is not the case; there is an example of *kiyang* combining with *su-* rather than *anà-* (463).

(463) *Itthu blaakang inni maccan su-baavung **kiyang**.*
       DIST after   PROX tiger   PAST-rise   EVID
      'After that the tiger apparently got up.' (B060115nar05)

Finally, a legend about the water of the sea rising and not receding also gets *kiyang*-marking.

(464) *Itthu thama-thuurung **kiyang**, aayer.*
       DIST NEG.IRR-descend EVID   water
      'It would not recede, it seems, the water.' (K051206nar15)

Note that in all the naturalistic examples, *kiyang* follows the predicate, whereas the elicited examples in (455)-(457) can have *kiyang* at other positions as well. The naturalistic examples are closer to Tamil then, where the evidential marker *=aam* must attach to the predicate (Smith & Paauw 2006:165). This is not obligatory in Sinhala, which was used to elicit examples (455)-(457).

6.4.9 Utterance final clitic

6.4.9.1 quotative *katha*

Katha is a quotative particle (Bakker 2000a), which is used to mark direct speech (Smith et al. 2004) ⊡→13.8, p. 536. Its primary use is to report sentences (465), but shorter strings like names (466) interjections (467)(468) or onomatopoetics (469)(470) can also be marked by *katha*.[40]

---

[40]While I have no information about the possibility to combine the quotative marker with onomatopoetics in Sinhala or Tamil, at least the Dravidian language Kota does use the quotative for onomatopoetics, e.g.

(465)   *Se=pe   oorang thuuva pada anà-biilang [kitham pada* Malaysia=*dering*
        1s=POSS man   old   PL   PAST-say   1PL   PL   Malaysia=ABL
        anà-dhaathang] katha.
        PAST-come   QUOT
        'My ancestors told me that we had come from Malaysia.' (K060108nar02)

(466)   *See=yang **Tony katha** arà-panggel.*
        1s=ACC   Tony QUOT   NON.PAST-call
        'I am called "Tony".' (K060108nar01)

(467)   **Yes *katha** m-biilang.*
        yes QUOT   PAST-say
        'He said "yes".' (K060116nar11)

(468)   *Aanak pompang duuva=nang [**slaamath**] **katha** su-biilang.*
        child girl   two=DAT   goodbye   QUOT PAST-say
        'He said "Goodbye" to the two girls.' (K070000wrt04)

(469)   ***Dam dam dam katha** su-aada.*
        dam dam dam QUOT   PAST-exist
        '(The rain) went like "dam dam dam".' (overheard)

(470)   Question *pada samma buthul baaye=nang ingath-an;   **tak tak katha***
        question PL   all   correct good=DAT   think-NMLZR tak tak QUOT
        su-thuulis ambel.
        PAST-write take
        'She could remember all the questions very well; and then she wrote them
        down tak tak' (K051222nar03,K081105eli02)

The reported string need not be literal, as the following example shows, where the
actual content of Mahinda's utterance is replaced by a deictic.

(471)   *[Itthu katha]=jo Mahindha arà-biilang.*
        DIST   QUOT=EMPH Mahinda   NON.PAST-say
        'That's what Mahinda [Rajapaksa, President of Sri Lanka] is saying.'
        (K051206nar11)

---

dop 'noise of falling' (Emeneau 1969).

While it is tempting to see *katha* as a subordinator, this is not the case, as the following example shows, where *katha* is attached to a main clause.

(472)   *Itthu muusing kithang=nang anà-biilang giithu=jo:*   **blood thirsty**
        DIST time   1PL=DAT   PAST-say   like.that=EMPH blood thirsty
        people *katha=jo!*.
        people QUOT=EMPH
        'At that time they called us the following: "blood-thirsty people!"   '
        (K051206nar04)

Another argument against the analysis of *katha* as a subordinator is that it can have scope over more than one clause. In (473), there is a mismatch between the number of clauses (3) and the number of *katha*s (2). The first *katha* has scope over both the first clauses, while the second instance of *katha* has scope over the last clause.

(473)   a.   *[Sama oorang masà-thaksir kithang=pe nigiri=ka   kanaapa kithang*
             all   man   must-think   1PL=POSS   country=LOC why   1PL
             ini   bedahan arà-simpang.
             PROX difference NON.PAST-keep
             'All men must think why we maintain these differences in our country.'
        b.   *Cinggala laayeng mulbar laayeng Moor laayeng mlaayu laayeng*
             Sinhala   different Tamil   different Moor different Malay   different
             sraani   laayeng]. ***katha.***
             Burgher different   QUOT
             'Sinhalese, Tamils, Moors, Malays and Burghers are all different.'
        c.   *[Itthu caara igaama pada=ka* catholic *laayeng* Protestant *laayeng*
             DIST way   religion PL=LOC   catholic different protestant different
             Hindu igaama laayeng Buddha igaama laayeng Islaam pada
             Hindu religion different Buddhist religion different Islam   PL
             laayeng] ***katha.***
             different QUOT
             'That way, Catholicism, Protestantism, Hinduism, Buddhism and Islam
             are different.' (K061127nar03)

Katha can not only be used for *verba dicendi*, but also for psych verbs like *iingath* 'think', as in (474)-(476).

(474) *Skarang biini arà-**iingath***   *[puthri thuuli **katha**]; Puthri arà-**iingath***
       now   wife NON.PAST-think queen deaf QUOT   queen NON.PAST-think
       *[biini thuuli **katha**].*
       wife deaf QUOT
       'Now the wife thought the queen was deaf, and the queen thought the wife
       was deaf.' (K070000wrt05)

(475) *[Bannyak haari=dering saapa=yang=ke thàrà-enco-kang] **katha***
       much   day=ABL   who=PAT=SIMIL NEG.PAST-fool-CAUS QUOT
       *anà-**iingath** Andare.*
       PAST-think   Andare
       'Andare thought of all the days that he had not taken anyone for a fool.'
       (K070000wrt05)

(476) *Aanak pada* class *pada=ka=le   masà-**iingath** ambel [bedahan thàràbaae*
       child PL   class PL=LOC=ADDIT must-think   take   difference bad
       *simpang] **katha**.*
       keep   QUOT
       'The children in the classes must also start thinking that maintaining these
       differences is bad.' (K061127nar03)

Even verbs which have nothing to do with cognition can be used with *katha* if
their complement can be interpreted as an utterance. An example is (477), where the
question of the essence of the saint is the object of a query, expressed by *caari* 'look
for'. This query could be formulated as a question, and hence marking by *katha* is
possible.

(477) *Incayang arà-**caari***   *[inni avuliya aapa mosthor **katha**].*
       3S.POLITE NON.PAST-search PROX saint   what manner QUOT
       'He is looking for how that saint was.' (B060115cvs04)

Katha has been shown above to be used in declaratives, but it can also be used if
the reported clause is a content interrogative as in (477) (478)(479).

(478) *Andare raaja=ka su-**caanya** [inni mà-kìrring simpang aada **aapa=yang***
       Andare king=ʟᴏᴄ ᴘᴀꜱᴛ-ask   ᴘʀᴏx ɪɴꜰ-dry   keep   exist what=ᴀᴄᴄ
       ***katha**.*
       ǫᴜᴏᴛ
       'Andare inquired from the King that what was there left [on the mat] to dry.'
       (K070000wrt02)

(479) *Se=ppe oorang pada [see **saapa**] **katha thàrà-thaau** subbath see=yang*
       1s=ᴘᴏss man   ᴘʟ   1s who   ǫᴜᴏᴛ ɴᴇɢ-know   because 1s=ᴀᴄᴄ
       su-uubar.
       ᴘᴀꜱᴛ-chase
       'Because my folks did not know who I was, they chased me.' (K070000wrt04)

For polar interrogatives in reported clauses, =*si* ⊡→6.4.6.1, p. 360 or =*so* ⊡→6.4.6.4, p. 374 are alternatives, but use of both =*si* and *katha* is also possible, as shown in (480).

(480) *[Aashik=nang hathu soldier mà-jaadi   suuka=si   **katha**] arà-**caanya**.*
       Aashik=ᴅᴀᴛ   ɪɴᴅᴇꜰ soldier ɪɴꜰ-become like=ɪɴᴛᴇʀʀ ǫᴜᴏᴛ   ɴᴏɴ.ᴘᴀꜱᴛ-ask
       'He asks if you want to become a soldier, Ashik.' (B060115prs10)

It is also possible to use *katha* in reported clauses expressing ignorance. Again, marking subordinates with =*so* instead of *katha* is a possibility.

(481) *Se=dang mà-biilang **thàrà-thaau** [maana thaaun **katha**].*
       1s=dat inf-say   ɴᴇɢ-know   which year ǫᴜᴏᴛ
       'I can't tell you which year that was.' (K051205nar04)

Besides verbs, nouns can also trigger the use of *katha*. This is the case for the noun *habbar* 'news' in (482). Note that the *katha*-clause is separated from the 'head noun' *habbar* by the first occurrence of *kithangnang* 'for us'.

(482) *[Lorang pada pukurjan arà-gijja   **katha**] kithang=nang **habbar***
       2ᴘʟ   ᴘʟ   work   ɴᴏɴ.ᴘᴀꜱᴛ-make ǫᴜᴏᴛ 1ᴘʟ=ᴅᴀᴛ   news
       kithang=nang bannyak suuka.
       1ᴘʟ=ᴅᴀᴛ   much   like
       'We like a lot to get the news that you do that work.' (N061031nar01)

The presence of a lexeme denoting a speech act in the matrix clause is not obligatory. *Katha* can also be used in non-verbal predications, often when indicating the

name of something. It can then be glossed as quotation marks,[41] as in the following
example.

(483) *Baava=ka* **Kaasim katha** *hatthu* **family**.
       bottom=LOC Kaasim ""   INDEF   family
       'Below there is a family called "Kaasim".' (K051206nar06)

This usage of *katha* can be seen as introducing metalinguistic information. This is
also at hand in some contexts where loanwords are used, which are often marked by
katha.

(484) *Drang* **defunct** *katha* *arà-biilang*.
       3PL   defunct QUOT   NON.PAST-say
       'They were said to be defunct.' (K060116nar03)

(485) **Jumbo Sale** *katha* *hatthu su-aada*.
       Jumbo Sale QUOT   INDEF   PAST-exist
       'There was a so-called "Jumbo Sale".' (B060115cvs01)

Both of the aforementioned functions combine in the following example.

(486) *Itthu* **Masonic Hall** *katha* *aada*.
       DIST Masonic Hall QUOT   exist
       'There is the so-called "Masonic Hall".' (K051220nar01)

This metalinguistic overtone can be used to distance oneself from an utterance as
in (487) and (488).

(487) *Karang* '**mlaayu siigith**' **katha** *thraa, samma Islam oorang pada=jo atthu*.
       now   Malay   mosque QUOT NEG   all   Moor people PL=EMPH one
       'There is no mosque that you could possibly call "Malay mosque", it is all to
       the Moors that a mosque belongs.' (B060115cvs04)

(488) *Hatthu aanak=pon* **thàrà-baae katha** *thraa*.
       one   child=any NEG-good   QUOT NEG
       'There is not one child that you would want to call naughty.' (B060115cvs13)

Besides the distancing function, *katha* can also be used to give definitions of the
metalinguistic implications proper of a term. This is the case in (489), where the
speaker explicates what *kaùmpath generation* 'fourth generation' actually means.

---

[41]I would like to thank Eike Nitz for this suggestion.

(489) *Ka-ùmpath* generation *katha arà-biilang se=dang duppang lai thiiga*
ORD-four   generation QUOT NON.PAST-say 1s=DAT before other three
generation *pada ini   nigiri=ka   anà-duuduk.*
generation PL   PROX country=LOC PAST-exist.ANIM
' "Fourth generation" is to say that before me another three generations existed
in this country' (K060108nar02)

A similar use is found in (490), where the meaning of *Haj* is explicated.

(490) *Itthu kithang=pe igaama=pe   mosthor=nang, ithu **Haj katha** arà-biilang*
DIST 1PL=POSS   religion=POSS manner=DAT   DIST Hajj QUOT NON.PAST-say
Mecca arà-pii.
Mecca NON.PAST-go
'According to our religion, "Hajj" means to go to Mecca.' (B060115cvs01)

Both the defining and the reporting function of *katha* are found in the following
string.

(491) ***Aavi katha*** *arà-biilang, thaau=jona **aapa katha** aavi?*
steam QUOT   NON.PAST-say know=PHAT what QUOT   steam
' "Aavi" is, you know what that is, "aavi", don't you?' (K061026rcp04)

It is possible to have an afterthought occurring after *katha*, even if it semantically
pertains to the quoted string. In (492), the destination of the cars, Iraq, is given after
katha, even if it clearly belongs to the predicate *masàbaa* 'must bring', which pre-
cedes *katha*. The expected position within the subordinate clause is marked by an
underscore.

(492) *Skarang derang biilang aada [baaru* car *sajja=jo* ___ *masà-baa katha],*
now   3PL   say   exist new   car only=EMPH   must-bring QUOT
Iraq=nang.
Iraq=DAT
'Now they have said that we must bring only new cars to Iraq.'
(K051206nar19)

The following example shows a variety of uses of *katha*. In the first clause, *katha*
marks the content of the reported imperative *biilang* 'say!' In the second and third
clause, the utterance marked by *katha* is used to premodify the noun *balasan* 'answer'
in the manner of a relative clause ⊡→13.2, p. 516.

(493)   a. *[Luu=nya   jadi-kang   rabbu saapa lu=ppe nabi   pada*
            2s.FAMILIAR=ACC become-CAUS prophet who   2s=POSS prophet PL
            *saapa **katha**] biilang.*
            who   QUOT   say
            'Say who the prophet is who made you, who are your prophets.'

        b. *[[Lu=ppe rabbu   saapa **katha**] buthul balas-an]   asà-biilang*
            2s=POSS   prophet who   QUOT   correct answer-NMLZR CP-say
            'Having given the correct answer as to who your prophets are,'

        c. *[[lu=ppe nabi   saapa **katha**] buthul balas-an]   asà-biilang ...*
            2s=POSS   prophet who   QUOT   correct answer-NMLZR CP-say
            'having given the correct as to who your prophet is, ...'

Katha stems from **kata* 'word', and is modeled on Sinhala *kiyalaa* and Tamil *endru*, both the conjunctive participle of the verb 'to say', which is used as a quotative. It is curious that SLM has opted for *katha* instead of using the conjunctive participle *asbiilang*, which does exist and would perfectly fit the pattern of the other Lankan languages. *Katha* is very frequent and used by anybody.

6.5   Nominal and verbal morphology

This chapter has presented the lexical and functional morphemes of Sri Lanka Malay. The order in which the items were discussed was of phonological nature. First came the affixes, then the quasi-prefixes, then the clitics. There is of course a tradition in linguistics to treat morphology in terms of nominal and verbal morphology. This often tends to confuse nominal and NP morphology on the one hand, and verbal morphology, VP morphology and clause morphology on the other hand. Furthermore, simple clitics with their free form and the corresponding cliticized form are difficult to incorporate into such an approach. This is why the presentation of morphemes here followed a phonological order, but, for convenience, Table 6.6 gives the morphemes discussed in this section and indicates to which traditional domain they belong.

	Affix		Quasi-Affix	Simple Clitic	Bound Word
N	-an -yang		–	–	–
V	su- anà- asà- mà- jamà- kapang-	kànà- -king -la -de	arà÷ anthi÷ masthi÷ thàrà÷ thama÷ marà÷	bàrà= mau=	–
ADJ	anà-				
NUM	-blas -pulu ka-		–	–	see NP
NP	–		–	–	(=)atthu(=) =pada adpositions =si =le =so =ke =pon =jo =jona =kiyang
VP	–		–	–	=le =si =so =ke =jo =kiyang
Clause	–		–	–	=si =so =jo =jona =kiyang
Utterance	–		–	–	=katha

Table 6.6: Morphemes in nominal and verbal morphology

Chapter 7

Word formation

This section describes how one or more morphemes discussed in the previous chapters can be combined to a well-formed morphological word.[1] Several types can be distinguished: Compounding joins two lexical words ⊡→7.1, p. 399, while inflection ⊡→7.2, p. 405 and derivation ⊡→7.3, p. 406 join lexical and functional morphemes. Conversion changes the category of a lexeme without morphological expression ⊡→7.4, p. 407, while reduplication causes a difference in meaning by using the same morpheme twice ⊡→7.5, p. 408. A well-developed word formation process in SLM is the combination of a deictic with an enclitic, which is discussed in a special section ⊡→7.6, p. 411. The reason for the inclusion of a process involving a clitic within word formation is that many of these pairs are highly lexicalized.

7.1 Compounding

Compounding is the creation of a new word by combining two or more stems. Compounding is a morphological operation and must be distinguished from syntactic operations. This distinction is not always easy to make in SLM. For the purpose of this discussion, we will treat compounds as morphological if they form exactly one phonological word, whereas syntactic constructions form more than one word. Under this criterion, it appears that modification of nouns generally takes place in syntax; only few modifications of a nominal head are found in morphology. Those which are found are normally lexicalized and show the historical order of head-modifier; *ad hoc* formations take place in syntax and show the order of modifier-head. An illustration of this is the word for 'garlic', which is a morphological compound, *bavampuuthi*. It can be analyzed as *baavang* 'onion' + *puuthi* 'white'. The same order is found in the Std. Malay cognate *bawang putih*. It seems reasonable to assume that this pairing of form and meaning was already established when the first immigrants came to Sri

---

[1]For phonological words, see ⊡→4.4.3, p. 115.

Lanka. Given the lexicalization of the concept, it underwent phonological integration into one word, which can be seen today by the presence of only one long vowel (uu); the first vowel in /baʋaŋ/ does not undergo lengthening. If /baʋaŋ/ was parsed into a phonological word of its own, we would expect the first vowel to lengthen (see Section 4.5.2.1 p. 127). This contrasts with an ad hoc formation like *gaaram aayer* 'water with salt', which consists of the words *gaaram* 'salt' and *aayer* 'water'. This string is parsed into two separate words, as can be seen from the two long vowels. Furthermore, it has the order modifier-head. This string does not seem lexicalized; the concept 'water with salt' is a lot less salient or important, which distinguishes it from 'garlic'. It is useful to have a lexical entry for garlic; it is less necessary to have a lexical entry for 'water with salt'. This string was also not brought with the first immigrants. The Std. Malay cognate is 'air garam', with inverted word order as compared to the Sri Lankan construction.

For expository reasons, head-modifier constructions are all treated in the section on word formation in this grammar, while all nominal modifier-head constructions are treated in the section on NP formation. The reason for this division of exposition is that it relies on the principle of headedness, which is easy to apply. Phonological parsings, which would be a sounder criterion, are more difficult to establish, are more prone to misanalyses, and are subject to greater variance. Therefore, it was decided against phonological parsings as the criterion to tell morphological from syntactic combinations, and the short-hand of headedness was used instead.

The endocentric compounds alluded to above will be elaborated in more detail in Section 7.1.1 for N+N compounds and in Section 7.1.2 for N+ADJ compounds. A marginal word formation process is the combination of a noun with a verb, discussed in Section 7.1.3. Next to these endocentric compounds, SLM can also form exocentric compounds, which are discussed in 7.1.4. Furthermore, verbal compounds can be formed, which are discussed in Section 7.1.5.

7.1.1   Compounds involving two nouns

Morphological combinations of two nouns (compounds) are always left-headed as in (2). In this section, compounding will be indicated with a plus sign (+).

(1) $\boxed{\text{N}_{head}+\text{N}_{mod}}_N$

(2)   *Seelong=nang duppang duppang anadhaathang mlaayu asadhaathang*
       Ceylon=DAT   before   before   PAST-come   Malay   COPULA
       orang⌒+ikkang.
       man⌒+fish
       'The Malays who had come a long time ago were fishermen' (K060108nar02)

The combination in (3) is right-headed, and analyzed as syntactic. This is represented without the plus sign.

(3) *Gaaram⌒ aayer asà-thuvang=apa.*
    salt   water cp-pour=after
    'Having poured salt water (on the wound).' (K061125nar01)

7.1.2   Compounds involving a noun and an adjective

Next to *bavam+puuthi* 'onion'+'white'='garlic', the words *aer + meera* 'water' +
'red' = 'tea' and, more rarely, *aer + iitham* 'water' + 'black' = 'coffee'[2] are lexicalized
left-headed compounds of a noun and an adjective. An example without a colour word
is (4), where we find the lexicalized compound *oorang+thuuva* 'man'+'old'='fore-
fathers, elders'.[3]

(4) *Se=ppe oorang$_N$+thuuva$_{ADJ}$ pada anà-biilang [kitham pada*
    1s=poss man+old   PL   PAST-say   1PL   PL
    Malaysia=*dering anà-dhaathang katha].*
    Malaysia=ABL   PAST-come   QUOT
    'My elders told me that we had come from Malaysia.'
    (K060108nar02,K081105eli02)

7.1.3   Compounds involving a noun and a verb

These compounds are not found in SLM, with two exceptions: *kappal+thìrbang*
'ship'+'fly'='airplane' and *kappal+bìrnang* 'ship'+'swim'='submarine'. It is very
likely that these are actually unanalyzed loans from Std. Malay. Both the left-headed-
ness and the compounding of a noun and a verb make this compound an improbable
innovation on Lankan soil. Furthermore, it is highly unlikely that the somewhat pecu-
liar semantics of this compound ('ship' for 'plane') originated independently in two
different places. It can also be excluded that this word came to Sri Lanka with the
immigrants because at that point in time neither airplane nor submarine had been in-
vented yet.

7.1.4   Exocentric compounds

A special kind of compound is the exocentric N+N-compound.[4] An example for
this is *ummabaapa* 'mother'+'father'='parents'. The resulting compound is neither a

---

[2]It is unclear why the glide present in *aayer* 'water' is not retained in the compounds.

[3]This compound shows the mismatch between headedness, lexicalization and phonological parsing.
While it is left-headed and lexicalized, it is still parsed into two phonological words, which is clear from
the long vowel in *oorang*.

[4]This type is common in South Asia and also found in Sinhala (Karunatillake 2004:131) and Tamil
(Arden 1934:96).

modification of the first part, nor of the second part, but a new term meaning the set of the two components. The fact that we are dealing with a morphological operation here, and not simply with a lexicalized zero-marked coordination ⊡→14.3.1.1, p. 545, can be seen from (5).

(5) *Kithang samma **hatthu umma+baapa**=pe   aanak pada, kithang samma*
     1PL   all   INDEF mother+father=POSS child PL   1PL   all
     sudaara pada.
     brother PL
     'We are all the same parents' children, we are all brothers' (B060115cvs01)

In this example, *ummabaapa* is combined with the indefiniteness marker *hatthu*, which indicates that we are dealing with only one entity here, not with several ones. This meaning is difficult to render in English, because the singular does not combine well with parenthood due to the ontological necessity of two people participating in procreation. An approximate translation could be *We are all the same ancestry's children.* Example (5) shows that *ummabaapa* can be treated as one referent, which rules out its being a zero-marked coordination. For purposes of comparison, the following example shows a similar sentence, but with other kinship terms which cannot enter the exocentric compound pattern.

(6) **kithang samma hatthu kaaka+dhaatha=pe   aade*
     1PL   all   INDEF elder.brother+elder.sister=POSS younger.sibling
     (pada).
     PL
     'We are all an elder sibling's younger sibling.' (K081104eli06)

Instead of the impossible exocentric compound in (6), zero coordination of two PPs must be employed, where every NP is individually marked as possessive.

(7) *Kithang samma kaaka   pada=pe dhaatha   pada=pe aade*
     1PL   all   elder.brother PL=POSS elder.sister PL=POSS younger.sibling
     pada.
     PL
     'We are all the younger siblings of elder brothers and elder sisters.' (K081104eli06)

In addition to *ummabaapa, sudaara+sudaari* 'brothers and sisters' given in (8) is another common exocentric compound.

(8) *Se=dang **sudaara+sudaari** pada thuuju oorang aada.*
     1s=DAT brother+sister   PL   seven man   exist
     'I have seven siblings.' (B060115prs14)

7.1.5   Compounds involving two verbs

SLM has a small number of V+V compounds, which must be distinguished from serial verbs involving two full verbs ⊡→10.1.3, p. 466 or a verb and a vector verb ⊡→10.1.2, p. 466. Verbal compounds consist of one phonological word, whereas the serial verbs consist of two phonological words. This can be seen by the fact that a verbal compound like (9) can only have one long vowel, whereas serial verbs can have two (10). Section ⊡→5.1.4, p. 171 contains a more detailed discussion of multi-verb constructions.

(9) *Ini   kavanan=pe aanak pompa=nya asà-**kasi+kaaving*** (compound).
     PROX group=POSS child female=ACC CP-give+marry
     'They gave this group's girls into marriage.' (K051222nar03)

(10) *Itthu muusing Islam igaama nya-**aajar kaasi** Jaapna Hindu teacher.*
     DIST time   Islam religion PAST-teach give Jaffna Hindu teacher
     'Those who taught us Islam back then were Hindu teachers from Jaffna.'
     (K051213nar03)

   kasi-kaaving 'given in marriage' in (9) has only one long vowel (in ***kaa**ving*) and forms one phonological word. *Aajar kaasi* 'educate' in (10), on the other hand, has a long vowel in both verbs, which indicates that we are dealing with two phonological words. In this description, we treat the former case (*kasikaaving*) as an operation on the word level, whereas the latter case (*aajar kaasi*) is an operation on the phrase level and will be treated in more detail in ⊡→10.1.2, p. 466. Note that the difference is reflected in orthography by the intervening space for serial verbs, which is not present for verbal compounds.

   Furthermore, the two constructions differ in their internal structure, which can be seen from the position of *ka(a)si* 'give'. In the compound, it is initial, whereas in the vector verb construction, it is final.

   There are not very many verbal compounds in SLM, and all of those which could be found involve *ka(a)si* 'give'. The following examples illustrate these words.

(11)   *Badulla* Kandy *Matale samma* association=*nang **masà-kasi+thaau**.*
       Badulla Kandy Matale samma association=DAT   must-give+know
       'Badulla, Kandy, Matale, we must inform all other associations.'
       (K060116nar06)

In (11) the TAM-marker precedes the verbal compound *kasithaau* 'inform', indicating that we are dealing with a one-predicate sentence, rather than two.[5] The vowel in *kasi* is short, as usual.

The same predicate, *kasithaau* 'inform' is also used in (12), where it occurs between the agent *lorang pada* 'you' and *aada,* indicating perfect tense.

(12)   *Lorang=pe* Amsterdam university=*ka   athu   aada katha kithang=nang*
       2PL=POSS   Amsterdam University=LOC INDEF exist QUOT 1PL=DAT

       *lorang pada **kasi+thaau** aada.*
       2PL   PL   give-know   exist
       'You have informed us that there is one [archive] in your university in Amsterdam.' (N061031nar01)

There is no additional material, indicating again that we are dealing with only one predicate and not several. Again, the vowel is short in *kasi*, but not in *thaau* 'know', corroborating the one-word-analysis.

The verbal compounds found so far do not carry aspectual or attitudinal information, which distinguishes them from serial verb constructions involving vector verbs ▢→10.1.2, p. 466, where the purpose of the vector verb is precisely to convey that kind of information. Another difference is that the semantics of verbal compounds is transparent, as in the examples above, where 'giving knowledge' is equated with 'inform' and 'giving marriage', with 'marry off'. This transparency distinguishes verbal compounds again from serial verb constructions with vector verbs, where the lexical meaning of the vector verb can be bleached. As an example of a bleached meaning of a vector verb in a serial verb construction (not a verbal compound), consider (13).

(13)   *Incayang=pe   kàpaala=ka anà-aada thoppi=dering moonyeth pada=nang*
       3S.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
       *su-**buvang puukul**.*
       PAST-throw hit
       'He took the hat from his head and threw it violently at the monkeys.'
       (K070000wrt01)

In this example, the meaning of *puukul* 'hit' is bleached and only the violent aspect

---

[5]If we had two predicates, we would expect the conjunctive participle *asà-* ▢→6.2.1.4, p. 276▢→13.3, p. 529.

remains. The simultaneous contact between agent, instrument and patient, which is obligatory for the literal use of *puukul* 'hit', is not respected, since we are dealing with throwing, where contact between agent and instrument and instrument and recipient are successive and not simultaneous.

The only instances of verbal compounds found up to date all involve the verb *ka(a)si* 'give'. The small size of the set of possible verbs points to a closed class, so that it might be the case that the constructions with *kasi-V* are actually fossilized instances of valency-changing derivations (cf. similar uses in other varieties of Malay in Adelaar & Prentice 1996, Bakker 2006, Paauw 2008).

7.2 Inflection

After compounding, inflection is another productive process of word formation. There are only two word formation constructions in SLM involving inflection, a prefixal one which combines verbal (quasi-)prefixes with a verb (14) and a suffixal one with the imperatives *-la* or *-de* attached to a verb (15).

(14) $\left\{\begin{array}{l} \text{VPREF–} \\ \text{QUASIPREF÷} \end{array}\right\} \text{V}$

(15) $\text{V}\left\{\begin{array}{l} \text{–la} \\ \text{–de} \end{array}\right\}$

Example (16) shows the use of an inflection prefix ⊡→6.2.1, p. 271, (17) shows the use of an inflectional quasi-prefix ⊡→6.2.2, p. 285, while (18) shows the use of the imperative suffix ⊡→6.2.4, p. 302, an inflection as well.

(16) *Baapa=le   aanak=le   guula su_{pref}-maakang$_V$.*
father=ADDIT child=ADDIT sugar PAST-eat
'Father and child ate the sugar.' (K070000wrt02)

(17) *Kithang caabe $arà÷_{quasipref}$maakang$_V$.*
1PL   spicy   NON.PAST-eat
'We eat spicy.' (B060115rcp01)

(18) *Diya$_V$-la_{suf} uncle, se=ppe SSC* exam *mà-ambel=nang duva-pul aari*
see-IMP   uncle 1s=POSS SSC exam INF-take=DAT   two-ty   day
su-aada.
PAST-exist
'See, uncle, there are (only) twenty days left until I will have take the SSC exam.' (K051220nar01)

7.3   Derivation

Derivation is not very productive in SLM. There are three general patterns of derivational word formation: nominalization, causativization and involitive derivation. Three further patterns are only used for numeral expressions ▭→7.3.4, p. 407. There is furthermore a possibility to use *-yang* on a kin term, which is described in ▭→6.2.5.4, p. 313.

7.3.1   Nominalization

A noun can be formed by attaching the suffix *-an* to adjectives or verbs.

(19) $\left[\left\{\begin{array}{c} \text{V} \\ \text{ADJ} \end{array}\right\} \textit{-an}\right]_N$

An example of this is given in (20), where *oomong* 'speak' is not used predicatively, but as a term used in the predication headed by *biilang* 'say', which is why it must be nominalized by *-an*, a derivational process yielding *omongan* 'speech'.

(20)   *Se=ppe laayeng **omong-an**   samma see anà-biilang.*
       1s=POSS other   speak-NMLZR all   1s PAST-say
       'I had said everything in my other speech.' (K061127nar03)

An example with deadjectival noun is given below.

(21)   *[Manis-an   maakang]=nang go   suuka bannyak.*
       sweet-NMLZR eat=DAT   1s.FAMILIAR like   much
       'I like very much to eat sweets.' (B060115prs20)

Nominalization is common, and the suffix *-an* is inherited unchanged from former stages of the language (Adelaar 1991).

7.3.2   Causativization

The second general suffixal derivation pattern is causativization by *king/kang* ▭→6.2.5.2, p. 307. This pattern is possible with verbal or adjectival hosts.

(22) $\left[\left\{\begin{array}{c} \text{V} \\ \text{ADJ} \end{array}\right\}\left\{\begin{array}{c} \textit{-king} \\ \textit{-kang} \end{array}\right\}\right]_V$

(23) shows the use of *-king* on an adjectival host. (24) shows the use of *-king* on a verbal host.

(23) *Itthuka   asà-thaaro, itthu=yang arà-**panas**adj-**king**.*
     DIST=LOC CP-put   DIST=ACC   NON.PAST-hot-CAUS
     'Having put (it) there, you heat it.' (B060115rcp02)

(24) *Baaye meera caaya kapang-jaadi, **thurung**intr-**king**.*
     good   red   colour when-become, descend-CAUS
     'When [the food] has turned to a nice rose colour, remove (it) [from the fire].'
     (K060103rec02)

Recursive causativization ('cause someone to cause someone to do something') is generally considered weird. The informants did not really know what to do with it. Solutions like *thiithi-king-kang* 'cause someone to cause someone to feed someone' with two causatives were proposed by some but ultimately rejected after discussion.

7.3.3   Involitive derivation

The only derivational prefix is *kànà-*, used to derive an involitive verb.

(25) $\boxed{\text{kànà-V}}_V$

Kànà- is different from inflectional (quasi-)prefixes and does not compete for the same slot, as shown in (26), where we find both a quasi-prefix (*arà-*) and *kànà-*.

(26) *Se=dang naasi arà-kànà-maakang.*
     1s=DAT   rice   NON.PAST-INVOL-eat
     'I compulsively ate (all) the rice.' (K081104eli03)

7.3.4   Derivation of numerals

Numeral derivation comprises *-blas* to derive the numbers from 13 to 19, *-pulu* to derive multiples of 10 from 20 to 90, and *kà-* to derive ordinals.

(27) $\boxed{(\text{kà-}) \text{NUM} \begin{pmatrix} -\text{pulu} \\ -\text{blas} \end{pmatrix}}_{NUM}$

7.4   Conversion

When a lexeme can be used in syntactic positions prototypically associated with different word classes, the question arises whether this is to be treated as double class-membership or as conversion (zero-derivation). I use 'conversion' only when the process is fully generalized across a word class. 'Double class-membership', on the other

hand, is reserved for lexemes that happen to have the facility to be employed in other functions, while most of the members of the class they are in cannot. Under this definition, all adjectives ⊡→5.3, p. 210 in SLM can convert to nouns or to verbs, but no other word class can convert. Isolated elements of other word classes can have double class membership, though, for instance *jaalang* 'walk(V)' and *jaalang* 'street'.

Conversion on a lexical level has to be distinguished from referential use of entire clauses at phrase level. It is actually possible to find verbs in referential position, prototypically associated with nouns. While this is indeed an instance of conversion, it is not on the word level, but on the phrase level. In those cases, the conversion is not V+∅→N, but S+∅→NP, S consisting only of a verb. See ⊡→8.8, p. 450 for discussion.

7.4.1   Adjectives to nouns or verbs

All members of the adjective class can convert to verbs or nouns. The converted adjectives can then be used in all the constructions where other verbs/nouns could be used.

(28) $\boxed{\text{ADJ-}\emptyset}_V$

(29) $\boxed{\text{ADJ-}\emptyset}_N$

The following examples show the use of the adjective *bìssar* 'big' as an adjective (30), a noun (31), and a verb (32).

(30)   *Se=ppe **bìssar** lai   ruuma aada.*
       1s=POSS big   other house exist
       'There is another big house of mine.' (B060115cvs09)

(31)   *Incayang=pe* wife=le   wife=pe baapa=le   masigith=pe **bìssar**$_N$
       3s.POLITE=POSS wife=ADDIT wife=POSS father=ADDIT mosque=POSS chief
       mas-panggel.
       must=call
       'His wife and her father had to call the mosque's head.' (K051220nar01)

(32)   *Aanak pada asà-**bìssar** skuul=nang anà-pii.*
       child PL CP-big   school=DAT PAST-go
       'Having grown up, the children went to school' (K051222nar04)

7.5   Reduplication

There are three productive processes of reduplication in SLM to form new words:

1. Reduplication of quantifiers to get an intensive reading

2. Reduplication of adjectives to get a prototypical/focal reading

3. Reduplication of a verb to form the simultaneous conjunctive participle

Other word classes, like nouns, cannot be reduplicated, but see below for fossilized nominal reduplications.

7.5.1 Quantifier reduplication

Quantifier reduplication is reasonably common. The quantifiers *konnyom* 'few' and *bannyak* 'many' are often reduplicated to intensify the meaning.

(33) | *QUANT ~ QUANT* |

(34) ***Deram pada arà-duuduk   konnyom˜konnyom** kiccil˜kiccil kavanang=ka.*
     3PL   PL   NON.PAST-stay little˜RED   small˜RED   family=loc
     'They live in a family which is a very little bit small.' (N060113nar01)

7.5.2 Adjectival reduplication

Adjectival reduplication is not found often. When adjectives are reduplicated, they yield a focal meaning.

(35) | *ADJ ~ ADJ* |

kiccil 'small' and *pullam* 'slow' are found reduplicated, but *bìssar* 'big' and *kìrras* 'fast' are not. For an example, see (34) above.

7.5.3 Reduplicated roots

There are a number of words which are clearly the result of historical reduplication but whose elements do not occur any more on their own. Examples for this are *lavalaava* 'cobweb', *cumicuumi* 'squid', *thangathaanga* 'ladder', *bathabaatha* 'bricks'.

A somewhat different case is *barambaarang* 'furniture', where the non-reduplicated form *baarang* does exist, but has a less specific meaning, 'item, good'.

7.5.4 Verbal reduplication

Verbal reduplication is an operation to derive the simultaneous conjunctive participle.[6]
Saldin (2001:56) attributes this to influence from an analogous Sinhala construction
(cf. Garusinghe 1962:88).

(36) | $V \sim V$ |

An example is given in (37), where the flight of the rabbit happens in a jumping
manner. This verbal modification of the verb *pii* 'go' has to be in the simultaneous
conjunctive participle, which is formed by reduplicating the verb *lompath* 'jump'.

(37)   *Kancil **lompath˜lompath** arà-laari.*
       rabbit jump˜RED   NON.PAST-laari
       'The rabbit runs away jumping.' (K081104eli06)

A similar case is found in (38), where the laughing of the Andare happens in a
bursting manner, and *lompath* 'jump' is reduplicated to indicate this modification of
the predicate *thathaava* 'laugh' by the verb *lompath* 'jump'. Note that the English
translation suggests an aspectual value of suddenness which is not there in the SLM
example.

(38)   *Siini Andare=atthas **lompath˜lompath** arà-thàthaava.*
       here Andare=about jump˜RED   NON.PAST-laugh
       'Here, he burst into laughter/laughed in a bursting manner about Andare.'
       (K070000wrt05)

A final example of a predicate (in this case *maakang* 'eat') being modified by
another predicate (*biilang* 'say') is (39), where the reduplicated modifier also has a
quite elaborate complement.

---

[6]Reduplicative formation of the simultaneous participle seems to be a common feature of Indian lan-
guages, cf. Abbi (1994:10,35), or Singh (2005), and is also one of the many similarities with Japanese
(Shibasaki 2005). Bakker & Parkvall (2005) compiled a list of reduplication patterns found in creole lan-
guages. While quantifier reduplication and adjectival reduplication as discussed above are common there,
the pattern of verbal reduplication we observe here is not listed in the compilation of Bakker & Parkvall
(2005).

(39) *[Lu=ppe muuluth=ka=le paasir, se=ppe muuluth=ka=le*
  2s.FAMILIAR=POSS mouth=LOC=ADDIT sand 1s=POSS mouth=LOC=ADDIT
  *paasir katha] **biilang˜biilang** baaye=nang baapa=le aanak=le guula*
  sand QUOT say˜RED   good=DAT father=ADDIT child=ADDIT sugar
  su-maakang.
  PAST-eat
  'Saying "There is sand in your mouth and there is sand in my mouth" both
  father and son ate the sugar.' (K070000wrt02)

This construction has to be distinguished from constructions with *kapang-* 'when', *=kapang* 'when' and *vatthu* 'time, while'. These forms introduce a temporal adjunct, while the reduplicated construction is not an adjunct, but a modifier of the verbal predicate ⊡→10.1, p. 461. The meaning of (39) is thus not 'While saying ..., they ate' but rather 'They ate in a talking manner'.

7.6 Deictic+X

Due to the absence of conjunctions and the scarcity of adverbs, SLM has to resort to other means to fulfill the linking functions ⊡→14.1, p. 541 taken care of by these elements in other languages. Very often, this involves a combination of the distal deictic *itthu* with one other element, normally an enclitic. These combinations are often lexicalized and not necessarily analyzed by the speakers. This is the reason why they are treated with word formation processes, rather than on the phrase level, where they historically originated.

(40) $\boxed{itthu \left\{ \begin{array}{l} = \text{POSTP} \\ = \text{COORDINATINGCLT} \end{array} \right\}}_{ADV}$

Of these combinations, the most salient are *itthunam* 'therefore', *itthusubbath* 'because of that', *itthukapang* 'then', and *itthule* 'but'.

7.6.1 *itthunam*

The combination of *itthu* with the dative marker *=nang* is used in discourse to mark that the following event was a consequence of the preceding one. The use of the dative in this construction might have to do with the fact that *=nang* is generally used to indicate purpose ⊙→15.1.2.12, p. 576 and consequence, and that this use has generalized to this discourse marker.

(41) ***Itthunam***, *samma baae oorang pada inni nigiri=ka=jo   deram*
      therefore all   good man   PL   PROX country=LOC=EMPH 3PL
      sa-duuduk.
      PAST-stay
      'It is because of that that all good men settled down in this country.'
      (N060113nar01)

(42) ***Itthunam***, *mlaayu pada=pe* tradition=*ke   aada.*
      therefore Malay   PL=POSS tradition=SIMIL exist
      'That's why there are the Malays' traditions.' (N060113nar01)

7.6.2   *itthusubbath*

The combination of *itthu* with the causal postposition =*subbath* is used to indicate a
reason, similar to English *therefore*. The causal component is stronger in *itthusubbath*
than in *itthunam*.

(43) ***Itthusubbath***=*jo incayang=yang siithu anà-braanak.*
      therefore=EMPH   3S.POLITE=ACC   there PAST-be.born
      'Because of that, he was born here.'

7.6.3   *itthuka(apa)ng*

The combination of *itthu* and *kaapang* 'when' is frequently used to advance to a fol-
lowing event in discourse. In the recipe in (44), the list of ingredients is structured by
means of *itthukang*.

(44) a. *Duuva raathus lima-pulu* gram *buula maau.*
          two   hundred five-ty   gram flour want
          'You need 250g flour.'
      b. ***Itthukang*** *nni   santhang   maau.*
          then   PROX coconut.milk want
          'And then you need this coconut milk.' (B060115rcp02)

7.6.4   *itthule*

The combination of *itthu* with the additive clitic =*le* is used to cancel implications
⊙→16.4, p. 692, like English *but*, or *this being the case, ... nevertheless ....*[7] Example

---

[7]The same combination of distal deictic and additive clitic is also used in Sinhala (Jayawardena-Moser
2004:55).

(45) shows the use of this marker, where its occurrence in the second clause cancels the implicature that the ransom referred to in the first clause would cause the speaker to be returned in his original state. Instead, and contrary to expectations, he is turned into a bear.

(45) a. *Se=ppe baapa incayang=nang ummas su-kaasi.*
          1s=POSS father 3s.POLITE-DAT   gold   PAST-give
          'My father gave him the gold.'

      b. ***Itthule**, [see=yang mà-kiiring=nang duppang] incayang see=yang*
          But   1s=ACC   INF-send=DAT   before 3s.POLITE 1s=ACC
          hathu Buruan mà-jaadi   su-bale-king.
          INDEF bear   INF-become PAST-turn-cause
          'But before he send me back, he turned me into a bear.' (K070000wrt04)

This use must not be confounded with the use of =*le* as an additive clitic meaning *also, too, as well* on a deictic. In example (46), we are not dealing with the adversative use of *itthule* but rather with a transparent combination of *itthu*, referring to a dish mentioned earlier, and =*le*, which indicates that the mentioned dish is tasty, just like the other dishes mentioned before.

(46) ***Itthu=le***   *oorang mlaayu=pe baaye hathu* traditional food *hatthu.*
      DIST=ADDIT man   Malay=POSS good INDEF tradtional food INDEF
      'That is also one of the Malays' good traditional foods.' (K061026rcp04)

Chapter 8

The noun phrase

In SLM, a nominal phrase can be built around any one of the following: Nouns ⊡→8.1, p. 415, personal pronouns ⊡→8.3, p. 446, interrogative pronouns ⊡→8.4, p. 447, deictics ⊡→8.2, p. 445, and numerals ⊡→8.5, p. 448. A special case are reciprocal noun phrases ⊡→8.6, p. 449.

Besides NPs headed by words as given above, NPs can also be headed by higher order syntactic structures, namely postpositional phrases ⊡→8.7, p. 450 and clauses ⊡→8.8, p. 450. These possibilities will be discussed in turn, before I argue in the last section ⊡→8.9, p. 455 that the SLM NP is non-configurational and has an appositional structure.

8.1 Noun phrases based on a noun

The noun is a very common base for constructing a noun phrase to be used in a clause. The noun can be modified by a number of other items: adjectives, nouns, deictics, quantifiers, numerals, the indefinite article, possessors, relative clauses, and the plural marker. We will discuss NPs involving only one of these items in turn, before we turn to the order of the elements in more complex noun phrases in the post- ⊡→8.1.13, p. 427 and prenominal domain ⊡→8.1.14, p. 427.

Table 8.1 gives an overview over the position of these elements within the NP. The postnominal field offers less possibilities than the prenominal field. Fewer classes are represented there, and stacking of modifiers is not possible in the postnominal field. This differs from the prenominal field, where members of more classes can be found, and where modifiers can be stacked. Another important feature of the SLM NP is the free position of the indefinite article (symbolized by ↓), which can occur at several places and multiple times in the same NP. The full structure of the NP is given as an orientation in (1).

left		right	far right
INDEF		INDEF	PL
ADJ		(ADJ)	
N	Noun	(N)	
QUANT		QUANT	
NUM		NUM	
DEIC			
POSS			
RELC			

Table 8.1: Position of elements within the NP. The general occurrence of modifiers to the left of the head word is noted by Adelaar (1991, 2005), Smith et al. (2004). Elements in parentheses are generally morphological compounds on the level of the word, rather than syntactic postmodifications on the level of the NP, although this distinction is fuzzy.

$$
(1) \quad \downarrow \left\{ \begin{array}{c} \text{DEIC} \\ \text{POSS} \end{array} \right\} \text{RELC} \left\{ \begin{array}{c} \text{POSS} \\ \text{DEIC} \end{array} \right\} \downarrow \text{QUANT} \left\{ \begin{array}{c} \text{NUM} \\ \text{ADJ} \end{array} \right\}* \downarrow \text{N} \downarrow \textbf{Noun} \begin{array}{c} \downarrow \\ \text{(N)} \\ \text{(ADJ)} \\ \text{QUANT} \\ \text{NUM} \end{array} \text{PL} \Big|_{NP}
$$

In the following, we will discuss modifications of a noun by different elements ⊡→8.1.1, p. 416-⊡→8.1.12, p. 424, then the relative order of postnominal modifiers ⊡→8.1.13, p. 427, then the relative order of prenominal modifiers ⊡→8.1.14, p. 427, before we turn to the position of the indefinite article in ⊡→8.1.16, p. 440. The full structure of the NP is discussed in more theoretical detail in ⊡→8.1.17, p. 444.

8.1.1   NPs containing an adjective

NPs containing an adjective are very common. The following two examples show several NPs with prenominal adjectival modification.

(2) ***Laayeng***$_{ADJ}$⌢ ***nigiri***$_N$*=pe   soojor   pada=nang* ***baae***$_{ADJ}$⌢
     different   country=POSS European PL=DAT   good
     lakuvan$_N$*=nang anà-juuval.*
     price=DAT   PAST-sell
     '(He) sold (it) to the Europeans from the other countries for a good price.'
     (K060103nar01)

(3) **Baaru**_ADJ_⌢ **oorang**_N_ *pada masà-thaaro.*
new   man   PL   must-put
'(We) must put new people.' (K060116nar11)

The order of the adjective and the noun in the NP has been described by Adelaar (1991), De Silva Jayasuriya (2002), Smith (2003) and Slomanson (2006), who all list prenominal adjectives as the only possibility. This ADJ N order is the most common one, but the inverse order is also possible, but less often heard. The ordering of adjective and noun is not lexically specified nor dependent on the speaker, as (4) shows. In this example, raw beef is being referred to twice, first in the order ADJ N, then in the order N ADJ.[1]

(4) a. **Mìntha**_ADJ_ **daaging**_N_=*yang cuuci.*
raw   beef=ACC   wash
'Wash the raw beef.'

   b. *asà-cuuci laada=le   gaaram=le bathu giling-an=ka   giiling.*
CP-wash   pepper=ADDIT salt=ADDIT stone grind=NMLZR=LOC grind
'Having washed it, grind salt and pepper in a grinding stone.'

   c. *asà-giiling* **daaging**_N_ **mùntha**_ADJ_=*yang baathu=ka asà-thaaro, giccak.*
CP-grind   beef   raw=ACC   stone=LOC CP-put   smash
'Having ground and having put the raw beef on a stone, smash it.'
(K060103rec02)

Given the low frequency of N+ADJ strings and the possibility to exlain them on the level of morphology, all N+ADJ strings are treated as compounds ▣→7.1, p. 399 rather than phrase formation for ease of exposition.

8.1.2   NPs containing another noun

While many N+N sequences can be analyzed as compounds on the word level, this compounding analysis is not possible if other material intervenes, as can be the case with the indefinite article *atthu*. This article can separate the modifier from its head noun, as in (5) and (6).

(5) *Ini   pohong atthas=ka* **moonyeth hathu=kavanan** *su-aada.*
PROX tree   top=LOC monkey   INDEF=group   PAST-exist
'On top of this tree was a group of monkeys.' (K070000wrt01)

---

[1] Also note that this example shows the different realization of raised schwa (*ì* in the first clause, *ù* in the last one).

(6) *[Ini oorang [...] caape subbath] [[**jaalang hathu=piiňǧgir**]=ka*
     PROX man   [...] tired because road   INDEF=border=LOC
     anà-aada   hathu pohong] baava=ka su-seeňder.
     PAST-exist.inanim INDEF tree   down=LOC PAST-rest
     'Because he was tired, this man sat down under a tree which stood at the side
     of the street.' (K070000wrt01)

This separation of modifier and head indicates that we are not dealing with a mor-
phological process on the word level, but with a syntactic process on the level of the
NP. *Moonyeth* 'monkey' modifies *kavanan* 'group' in (5) in very much the same way
as the adjective *lai* 'more' modifies *kavanan* in (7).

(7) *Ikang Seelon=nang **lai**   **hathu=kavanan** anà-dhaathang.*
     then Ceylon=DAT other INDEF=group   PAST-come
     'Then, (yet) another group came to Sri Lanka.' (K060108nar02)

This parallel structure indicates that nouns can premodify other nouns on the syn-
tactic level. Because of this, nominal premodification of nouns is always analyzed
as syntactic in this description, while postmodification is treated as a morphological
operation ▢→7.1.1, p. 400.

An example for nominal premodification without the indefinite article is given in
(8).

(8) *Aapcara=ke incayang ini   **ciina oorang** Islam=nang asà-dhaathang*
     how=SIMIL   3s.POLITE PROX China man   Islam=DAT CP-come
     asà-kaaving=apa karang mà-siigith=nang arà-pii.
     CP-marry=after   now   mosque=DAT   NON.PAST-go
     'Somehow he, this Chinaman converted to Islam and married and now goes to
     the mosque.' (K051220nar01)

Another instance of NPs containing two nouns is comparison. In (9), *soldier* is
modified by *baapake* 'like daddy'. The intervening similative clitic =*ke* makes it
impossible to treat this as a morphological pattern on the level of the noun; rather, we
are dealing with a construction on the level of the noun phrase.

(9) *Se=dang [[baapa]$_N$=ke soldier$_N$]$_{NP}$ mà-jaadi   suuka.*
     1s=DAT father=SIMIL   soldier   INF-become like
     'I want to become a soldier like daddy.' (B060115prs10)

Postnominal modification with nouns (*orang ikkang* 'man'+ 'fish'= fisherman') is
also possible, but is treated as a morphological phenomenon in this grammar, not as a

syntactic phenomenon.

8.1.3 NPs containing a quantifier

This type of NP is are also common. The quantifier is normally preposed (10), but can also be floated. This is especially true for *samma* 'every' (11)-(14).

(10) *Inni sudaari=pe femili=ka* **bannyak⁀** *oorang tsunami=dang spuukul*
PROX sister=POSS family=LOC many people tsunami=DAT CP-hit
su-pii.
PAST-go
'In this sister's family, many people were swept away by the tsunami.'
(B060115nar02)

In (11)-(13), the quantifier occurs after the plural marker *pada*, which is the right-most element in any NP (see below, ⊡→8.1.10, p. 423). This is an indication that the quantifier occurs outside of the NP over which it has scope.

(11) *[Go=ppe aanak pada]* **samma** *baaye.*
1S.FAMILIAR=POSS child PL all good
'My children are all good.' (B060115cvs13)

(12) *[Kafan kaayeng pada]* **samma** *asà-ambel.*
shroud cloth PL all CP-take
'Having taken all the tissue for the shroud, ...' (B060115nar05)

(13) *[Kithang=pe oorang thuuva pada]* **bannyak** *dhaathang aada.*
1PL=POSS man old PL many come exist
'Our ancestors came in great quantities.' (K060108nar02)

In (14) the quantifier occurs even after the predicate, clearly outside the NP over which it has scope.

(14) *Kitham=pe association itthu vatthu* **duvith** *thraa* **bannyak**.
1PL=POSS association DIST time money NEG much
'At that time, our association did not have much money.' (K060116nar01)

8.1.4 NPs containing numerals

Numerals are commonly found in NPs. The numeral normally precedes the noun, as in (15)(16), but may also follow on rare occasions as in (17).

(15) **Thiiga** *oorang,* **thiiga** *oorang=le,* **thiiga** *oorang pada=jo itthu ini*
      three   man,   three man=ADDIT, three man   PL=EMPH DIST PROX
      volleyball *arà-play-king=kee.*
      volleyball NON.PAST-play-CAUS=SIMIL
      'Three men, three persons, three people play it, like volleyball.'
      (N060113nar05)

(16) **Duva-pulu ìnnam riibu**   **ùmpath raathus lima-pulu duuva** votes
      two-ty   six   thousand four   hundred five-ty   two   votes
      incayang=nang anà-daapath.
      3s.POLITE=DAT PAST-get
      'He got 26,452 votes.' (N061031nar01)

(17) *[Panthas rooja kumbang pohong komplok]* ⌢*duuva asà-jaadi su-aada.*
      beautiful rose flower tree   bush   two   CP-grow PAST-exist
      'Two beautiful rose bushes had grown.' (K070000wrt04)

8.1.5 NPs containing the indefinite article

If the indefinite article is present within the NP, it can either be preposed (18)(19),
postposed (20)(21), or both pre- and postposed (22)(23), as the following six examples
illustrate.

(18) **Hatthu** *avuliya aada kitham=pe ruuma dìkkath.*
      INDEF   saint   exist 1PL=POSS house vicinity
      'There is a saint close to our house.' (K060108nar02)

(19) **Hathu** *haari,* **hathu**=*oorang [thoppi mà-juval]=nang kampong=dering*
      INDEF   day   INDEF=man   hat   INF-sell=DAT   village=ABL
      kampong=nang su-jaalang pii.
      village=DAT   PAST-walk go
      'One day, a man walked from village to village to sell hats.' (K070000wrt01)

(20) *See avuliya* **atthu** *su-jaadi.*
      1s saint   INDEF PAST-become
      'I have become a saint.' (K051220nar01)

(21) *Mà-blaajar=nang see anà-kiiring se=ppe maama **hatthu**=pe ruuma=nang.*
INF-learn=DAT   1s PAST-send 1s=POSS uncle   INDEF=POSS house=DAT
'I was sent to an uncle of mine's to study.' (K051213nar02)

(22) *Sithu=ka **hathu** maccan **hathu** duuduk aada.*
there=LOC INDEF tiger   INDEF stay   exist
'A tiger stayed there.' (B060115nar05)

(23) ***Hathu** kaaving **hatthu**=nang kapang-pii*
INDEF   wedding INDEF=DAT   when-go
'When we go to a wedding' (G051222nar04)

Within one idiolect, the positioning of *(h)at(t)hu* can vary. In the following stretch of discourse, we find two instances of nouns with following *hatthu* and one instance of pre- and postnominal *hat(t)hu*.

(24) *Kaaving **hatthu**=nang arà-pii   vakthu, jalang-an   **hatthu** arà-pii*
wedding INDEF   NON.PAST-go time   walk-NMLZR INDEF NON.PAST-go
*vakthu, **hatthu** mayyeth **hatthu** arà-mnaaji.*
time   INDEF corpse INDEF NON.PAST-pray
'When we go to a wedding, when we go on a trip, when we recite for a dead person.' (K051213nar06)

Hat(t)hu is actually found quite often more than once in an NP. This is often once preposed and once postposed, but double occurrences in the prenominal field can also be found. This is the case in (25), where the indefiniteness marker for *makanan* 'food' occurs three times.

(25) *Itthu=le   [**hathu** [mlaayu oorang pada=pe]poss **hathu** [baae]ADJ **hathu***
DIST=ADDIT INDEF Malay man   PL=POSS   INDEF good   INDEF
makanan]NP=jo.
food=EMPH
'That is also a good dish of the Malays.' (K061026rcp02)

This flexibility of the position of the indefiniteness marker suggests that it does not have a fixed position in the NP but can occur between any two constituents in the NP. Whereas the possessor, the adjective, the relative clause etc. have more or less predictible positions, the indefiniteness marker defies such predictions. This will be developed in more detail below ⊡→8.1.16, p. 440.

8.1.6   NPs containing deictics

Deictics always precede the noun (Adelaar (1991:29), Adelaar (2005:214), Slomanson (2006:137)). The following two examples show this for the proximal deictic *in(n)i* and for the distal deictic *it(t)hu*.

(26) ***Inni** maccan su-baavung kiyang.*
      PROX tiger   PAST-rise   EVID
      'The tiger apparently got up.' (B060115nar05)

(27) ***Ithu** jaalang=ka mà-pii thàràboole.*
      DIST road=LOC   INF-go cannot
      'You could not take that road.' (B060115nar05)

8.1.7   NPs containing possessors

Possessors always precede the head noun (Adelaar 1991, De Silva Jayasuriya 2002, Slomanson 2006). Examples (28) and (29) shows this for simple possession, while (30) shows more complex recursive possessive relationships.

(28) *[Se=ppe$_{POSS}$ naama$_N$]$_{NP}$ Mohomed Imran Salim.*
      1s=POSS   name   Mohomed Imran Salim
      'My name is Mohomed Imran Salim.' (K060108nar01)

(29) *[Incayang=**pe** wife]=lè, [wife=**pe** baapa]=le   [masigith=**pe** bìssar]*
      3s.POLITE=POSS w.=ADDIT w=POSS   father=ADDIT mosque=POSS chief
      mas-panggel.
      must-call
      'His wife and her father had to call the mosque's head.' (K051220nar01)

(30) *[[Kithang=**pe** baapa]=**pe** naama] Mahamud.*
      1PL=POSS   father=POSS name   Mahamud
      'Our father's name is Mahamud.' (B060115nar03)

8.1.8   NPs containing locations

These are rare. Normally, the location is put in a relative clause as in (31).

(31) *Meeja=ka *(aada) maaňğga=yang kaasi.*
      table=LOC exist   mango=ACC   give
      'Give me the mango (which is) on the table' (K081118eli01)

8.1.9   NPs containing relative clauses

Relative clause always precede the head noun. Example (32) illustrates this. The topic of relative clauses will be treated in more detail in ⊡→13.2, p. 516.

(32) *[Seelon=nang dhaathang aada ∅] mlaayu oorang ikkang.*
      Ceylon-DAT   come   exist   Malay man   fish
      'The Malays who came to Sri Lanka were fishermen.' (K060108nar02)

8.1.10   NPs containing the plural marker

The plural marker is always at the rightmost position. Example (33) illustrates this.

(33) *[Spaaru oorang **pada**] su-pii.*
      some   man   PL   PAST-go
      'Some men left.' (B060115nar01)

It is impossible to have anything pertaining to the NP after *pada* with the exception of floated quantifiers ⊡→8.1.3, p. 419, of which one example is repeated here for convenience.

(34) *Go=ppe   aanak **pada samma** baaye.*
      1S.FAMILIAR=POSS child PL   all   good
      'My children are all good.' (B060115cvs13)

8.1.11   NPs containing indefinite expressions

There is no indefinite pronoun strictly speaking in SLM, but the combination of an interrogative pronoun with the clitic *=ke* can be used for that end. This construction precedes the noun.

(35) Important occasion *pada kala aada,* **aapa=ke**   **festival** *pada laayeng*
      important occasion PL   if   exist what=SIMIL festival PL   other
      wedding ceremony*, [...] ithu vakthu kithang arà-kirja*
      wedding ceremony [. . .] DIST time   1PL   NON.PAST-make
      'If there is an important occasion, some festival or otherwise a wedding cere-
      mony, we will prepare it.' (K061026rcp01,K081105eli02)

8.1.12 NPs containing interrogative pronouns used for universal quantification

An interrogative pronoun can occur in the left field of an NP if the additive clitic =*le* is
present at the right edge of the NP. If a postposition is used on the NP, then the additive
clitic attaches to the right of that postposition, and technically speaking is outside of
the NP itself. This combination of interrogative pronoun and additive clitic yields a
universally quantified NP [2]

(36)   WH N(=POSTP)=*le*

The use in an NP is given in (37), the use of a locative PP in (38).

(37) *Skarang* **maana** *aari=le*   *atthu atthu oorang=yang arà-buunung.*
      now   which day=ADDIT one   one   man=ACC   PAST-kill
      'Now people kill each other every day.' (K051206nar11)

(38) **Maana** *thaahun=ka=le ini   rooja pohong komplok duuva*
      which   year=LOC=ADDIT PROX rose tree   bush   two
      kumbang=dering arà-punnu.
      flower=ABL   NON.PAST-push
      'Every year, these two rose bushes grow flowers.' (K070000wrt04)

If the additive clitic attaches to a predicate rather than to the NP/PP, then the uni-
versal quantification only holds for the subset of entities which happen to have a pos-
itive truth value in the proposition formed by the predicate and its arguments. An
example for this is (39). In this example, the proposed connection does not hold for
any Malay. It only holds for the Malays which the addressee happens to meet.

---

[2]This construction never carries the meaning of *whichever*. For this meaning, the WH˜WH ... =*so*
construction is used ⊡→8.1.12, p. 424.

(39) *Lai* **saapa** *mlaayu kuthumung=**le** aapa=ke* connection *hatthu aada.*
other who Malay see=ADDIT what=SIMIL connection INDEF exist
'Any other Malay you see, there will always be some kind of connection.'
(K051206nar07,K081105eli02)

Another example of this structure is (40).

(40) *Daalang=ka pii=nang blaakang,* **dhraapa** *puukul=**le** thama-kuthumung.*
inside=LOC go=DAT after how.many hit=ADDIT NEG.NONPAST-see
'After going inside, how ever much he hit them, one would not see.'
(K051206nar02,K081105eli02)

This structure can be formalized as follows, where reference of N_i is restricted by the predicate V_i before application of the universal quantifier.

(41)
$$\left[\left[WH \ N_i (= POSTP) \left\{ \begin{array}{l} NP \\ PP \\ ADJCT \end{array} \right\} * \ V_i = le \right] PRED \right]$$

It is possible to use other interrogative pronouns in this construction, as given below.

(42) **Aapa** *bìlli=**le**, baae.*
what buy=ADDIT good
'Anything you buy is good.' (K081104eli05)

(43) **Mana**=ka=**le** *pii, panthas.*
where=LOC=ADDIT go beautiful
'Anywhere you go, it's beautiful.' (K081104eli05)

While the construction with the additive clitic *=le* has a maximizing function, it is also possible to use the 'undetermined' clitic *, =so*, in which case the meaning is more that of an arbitrary referent than of universal quantification. While *=le* is used with a bare verb, a finite verb must be used with the *=so*-construction.

(44) **Saapa** *arà-nyaanyi=**so**, bannyak upaama.*[3]
who NON.PAST-sing=UNDET much honour
'Whoever sings, it will be a big honour.' (K081104eli05)

---

[3]The word for 'honour' was given as both *upaama* and *uthaama*.

The interrogative pronoun can be reduplicated in the =*so*-construction, which is not possible in the =*le*-construction. This yields an exhaustive meaning, analogous to the normal reduplicated question words ⊡→5.7, p. 229.

(45)   ***Saapa˜saapa** arà-nyaanyi=**so,**   bannyak upaama.*
        who˜red   NON.PAST-sing=UNDET much   honour
        'Whoever all sing, it will be a big honour.' (K081104eli05)

Different TAM-prefixes can be used in this construction, as shown in (46) and (47).

(46)   *Aapa˜aapa **su**-bìlli=so,   itthu baae.*
        what˜red   PAST-buy=UNDET DIST   good
        'Whatever you have bought, it is good.' (K081104eli05)

(47)   *Aapa˜aapa **anthi**-bìlli=so, itthu baae.*
        what˜red   IRR-buy=UNDET DIST   good
        'Whatever you will buy, it will be good.' (K081104eli05)

The formalization of this is given in (48).

(48) $\left[WH(\sim \text{RED})\ N_i(= POSTP) \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * TAM - V_i = so \middle| PRED \right]$

The only naturalistic example of this construction is (49).

(49)   *Inni **saapa~saapa**=ka inni   mlaayu pakeyan pada aada**=so,** lorang pada*
        prox who~who=LOC   PROX Malay dress   PL   exist=DISJ 2PL   PL
        ini   mlaayu pakeyan=samma ini   kaving=nang mà-dhaathan bannyak
        PROX Malay dress=with   PROX wedding=DAT INF-come   much
        uthaama.[4]
        honour
        'Whoever owns such Malay dresses, your coming together with this Malay dress to the wedding will be greatly appreciated.' (K060116nar04)

The examples above have shown the use of =*le* on an item other than the question word. It is also possible to attach =*le* directly to the question word if the interrogative pronoun provides some quantifiable semantic content, as is the case for *saapa* 'who' in the following two example.

(50) $\boxed{\text{WH(=POSTP)=le}}$

---

[4]The word for 'honour' was given as both *upaama* and *uthaama*.

(51) *Cinggala bahasa* **saapa=nang=le** *bole=bicaara siini.*
      Sinhala   language who=DAT=ADDIT can-speak   here
      'Sinhala   could   be   spoken   by   anybody   around   here.'
      (K051206nar14,K081105eli02)

8.1.13   Relative order of postnominal modifiers

We have seen above that the indefinite article, adjectives, numerals, and the plural
marker can follow the noun. Relatively little can be said about the order of these
items. The plural marker is always at the rightmost positions, and can cooccur with
any of the other postnominal modifiers but the indefinite article. It is not possible to
have more than one of the other modifiers in postnominal position. We can schema-
tize the postnominal field as in (52). Postnominal adjectives and nouns are listed in
this schema for illustrative reasons as well, but are actually analyzed as being com-
pounds in this grammar ⊡→7.1, p. 399. The plural marker cannot cooccur with the
indefiniteness marker. This is not represented in the schema.

$$
(52) \quad \boxed{\text{Noun} \begin{Bmatrix} (N) \\ (ADJ) \\ INDEF \\ QUANT \\ NUM) \end{Bmatrix} \text{PL}}
$$

8.1.14   Relative order in the prenominal field

There are many more prenominal modifiers than postnominal modifiers. The possible
order between them are indicated in Table 8.2. The following sections give the logi-
cal possibilities and also give an example, if available. Sections without an example
indicate that this combination was not found.

RELC RELC N   It is very difficult to find contexts where a head noun could be
modified by two clauses of the Relative Clause type, which comprises both true rel-
ative clauses and fact clauses in SLM (see ⊡→13.2, p. 516 for a discussion). Even if
this occurs, normally, the two clauses are coordinated and put into *one* relative clause,
itself consisting of two lower clauses. Example (53) shows such a case. The two
verbs *mintha* 'beg' and *naangis* 'cry' both depend on the head noun *svaara* 'sound'.
But they are found in *one* clause.

	REL	DEIC	POSS	QUANT	NUM	ADJ	N
REL	?	+	+	+	+	+	+
DEIC	+	/	+	+	+	+	+
POSS	(+)	+	/	+	+	+	+
QUANT	–	–	+	/	/	+	+
NUM	–	–	–	/	(–)	+	+
ADJ	–	–	–	–	(+)	+	+
N	–	–	–	–	–	–	+

Table 8.2: Order of prenominal modifiers. The cells indicate whether the item at the left end of the row can precede the item at the top of the column. A + indicates that this is possible, a – indicates that it is impossible, a / indicates that the combination is ruled out semantically, e.g. combination of a numeral and a quantifier.

(53) *[[Banthu-an asà-mintha]$_{CLS}$ [arà-naangis]$_{CLS}$]$_{RELC}$ [svaara]$_N$ hatthu*
      help-NMLZR cp-beg   SIMULT-cry   sound   INDEF
      derang=nang su-dìnngar.
      3PL=DAT   PAST-hear
      'They heard a sound of crying and begging for help.' (K070000wrt04)

To sum up, there is no case of two clauses preceding a head noun as far as syntax is concerned. Semantically speaking, it is possible to have two propositions depending on a head noun, as in (53), but these are packaged into one clause in syntax before they attach to the head noun. The structure is then [[[...]$_{infiniteclause}$ ⋯]$_{finiterelativeclause}$]$_{mainclause}$

RELC DEIC N   This order is found in two examples in the corpus.

(54) *[Seelong=ka su-mniiňğgal]$_{RELC}$ [ithu]$_{DEIC}$ thiiga [oorang]$_N$=pe naama*
      Ceylon=LOC PAST-die   DIST   three man=POSS   name
      kithang thàrà-thaau.
      1PL   NEG-know
      'We do not know the names of the three people who died in Ceylon.'
      (K060108nar02)

(55) *Laiskali, inni   Aajuth [incayang=pe jiiva anà-salba-king]*_{RELC}
   again   DEM.PROX dwarf 3s.POLITE=POSS life PAST-escape-CAUS
   *[itthu]*_{DEIC} *[aanak]*_N *pompang duuva=]nang thriima thàrà-kaasi.*
   DEM.DIST   child   female   two=DAT   thanks NEG.PAST-give
   'Once again, the dwarf did not thank the two girls who had saved his life.'
   (K070000wrt04)

RELC POSS N   There is one instance of a possessive pronoun intervening between
the relative clause and the head noun, (56).

(56) *[Ini British government=samma pii=apa mà-oomong]*_{RELC}
   PROX British government=COMIT   go=after INF-talk
   *[kithang=pe]*_{POSS} *[statesmen]*_N *pada [...]*
   1PL=POSS   statesmen   PL   [...]
   'Our statesmen who had gone to negotiate with the British Government ...'
   (N061031nar01)

RELC QUANT N   There is one instance of this order.

(57) *[Anà-libbi]*_{RELC} *[konnyong]*_{QUANT} *mlaayu*_N*=jo Seelong=ka (asà-)thiiňǧgal*
   PAST-remain   few   Malay=EMPH Ceylon=LOC CP-settle
   aada.
   exist
   'The remaining few settled in Sri Lanka.' (K051222nar06,K081105eli01)

RELC NUM N   The following two examples give this order with a cardinal (58) and
an ordinal numeral (59).

(58) *[Seelong=ka su-mniiňǧgal]*_{RELC} *ithu [thiiga]*_{NUM} *[oorang]*_N*=pe naama*
   Ceylon=LOC PAST-die   DIST three   man=POSS   name
   kithang thàrà-thaau.
   1PL   NEG-know
   'We do not know the names of the three people who died in Ceylon.'
   (K060108nar02)

(59) *[Terrorist hatthu=dering anà-maathi]*_{RELC} *[kàthaama]*_{NUM} *oorang*_N*=jo*
      terrorist   INDEF=ABL   PAST-die   first   man=EMPH
      incayang.
      3s
      'The first man killed by a terrorist was him.' (K051206nar02)

RELC ADJ N   This order is common. There are several instances of it in the corpus.
A simple example is (60), where *bìssar* 'big' is found between the relative clause and
the head noun.

(60) *[Anà-kijja]*_{RELC} *[bìssar]*_{ADJ} *[thumpath]*_N *pada.*
      PAST-make   big   place   PL
      'The big lands that were made.' (N060113nar02)

Bìssar is also found as the intervening adjective in (61).

(61) *[Sithu=ka aada]*_{RELC} *[bìssar]*_{ADJ} *[oorang]*_N *pada=yang asà-attack-kang ...*
      there=LOC exist   big   man   PL=ACC   cp-attack-CAUS
      'He attacks the leaders who are there and ... ' (K051206nar02)

The head of a relative clause does not have to be a noun actually, an interrogative
pronoun is possible as well, as shown in (62), where additionally the adjective *lai*
'other' intervenes between the relative clause and the interrogative pronoun.

(62) *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
      Hopefully other know brother sister   PL=LOC   can-ask   take
      *[[nya-gijja]*_{RELC} *[lai]*_{ADJ} *[saapa=kee]*_{HEAD} *aada=si   katha].*
      PAST-make   other   who=SIMIL   exist=INTERR QUOT
      'Hopefully, you can enquire from another person you know whether there is
      someone else who did something.' (N061031nar01)

A somewhat different case is given in (63), where we are dealing with a purposive
clause, which precedes the adjective.

(63) *Kithang lorang=nang baaye mliiga athi-kaasi, [mà-kaaving]*_{RELC}
    1PL   2PL=DAT   good palace IRR-give INF-marry
    *[panthas]*_{ADJ} *[pompang]*_N *pada athi-kaasi, duvith athi-kaasi.*
    beautiful   female   PL   IRR-give   money IRR-give
    'We will give you nice palaces, we will give you beautiful girls to marry, we
    will give you money.' (K051213nar06)

RELC N N   A relative clause can precede a sequence of nouns.

(64) *[Ruuma duuva subala=ka su-aada]*_{RELC} *[rooja]*_N *[pohong komplok]*_N *duuva*
    house two   side=LOC PAST-exist rose   tree   bush   two
    'The rose bushes that stood at both sides of the house' (K070000wrt04)

DEIC RELC N   Both the proximal (65) and the distal deictic (66) have been found
to precede a relative clause.

(65) *Incayang [ini]*_{DEIC} *[Seelong=ka anà-aada]*_{RELC} *[lakuan]*_N *[baathu]*_N
    3S.POLITE PROX   Seelon=LOC PAST-exist   wealth   stone
    asà-caari.
    CP-find
    'He was looking for the gems which were present in Ceylon.' (K060103nar01)

(66) *[Ithu]*_{DEIC} *[spaaman anà-niiňǧgal]*_{RELC} *[thumpath]*_N*=nang=le [Passara*
    DIST   3S.POLITE PAST-die   place=DAT=ADDIT   Passara
    katha arà-biilang nigiri]=nang=le dìkkath.
    QUOT NON.PAST-say village=DAT=ADDIT vicinity
    'The place where he died and the village called Passara are close to each other.'
    (B060115nar05)

DEIC DEIC N   This order was not found, which is probably due to the opposite
meanings of 'proximal' and 'distal', which do not combine well.

DEIC POSS N   The deictics can be found preceding a possessor and a noun.

(67) *[Ini]$_{DEIC}$ [indonesia=pe]$_{POSS}$ [oorang]$_N$=si giithu   kalthraa* Malaysian
      PROX   Indonesia=POSS   man=DISJ   that.way if.not   Malaysian
      oorang=si.
      manDISJ
      'These Indonesians or otherwise Malaysians.' (K060108nar02)

(68) Eleventh=*ka su-aada   [itthu]$_{DEIC}$ [kithang=pe igaama=pe]$_{POSS}$*
      eleventh=LOC PAST-exist DIST   1PL=POSS   religion=POSS
      [mosthor]$_N$=nang.
      manner=DAT
      'It was on the 11th, according to the customs of our religion.' (B060115cvs01)

DEIC QUANT N   Deictics and quantifiers precede the noun, in that order.

(69) School=*nang arà-pii=subbath=jo   [inni]$_{DEIC}$ [samma]$_{QUANT}$*
      school=DAT   NON.PAST-go=because=EMPH PROX   all
      [seksa]$_N$.
      problem
      'It is because we go to school that there are all these problems.'
      (B060115cvs01)

DEIC NUM N   Deictics and numerals precede the noun, in that order.

(70) *[Ini]$_{DEIC}$ [duuva]$_{NUM}$ [army captain]$_N$ thàrà-maathi.*
      PROX   two   army captain   NEG.PAST-dead
      'These two army captains did not die.' (K051213nar06)

DEIC ADJ N   Deictics precede the adjective in the noun phrase.

(71) *[Ini]$_{DEIC}$ [laama]$_{adj}$ [car]$_N$ pada=jo kithang arà-baapi.*
      PROX   old   car   PL=EMPH 1PL   NON.PAST-bring
      'It is these old cars we take [to Iraq].' (K051206nar19)

DEIC N N   The following two examples show deictics preceding two nouns. If these
strings of two nouns are analyzed as syntactic combinations, we get the sequence
DEIC N N, i.e. a deictic preceding a noun modifying a noun.

(72) *Kitham=pe=le [inni]ᴅᴇɪᴄ [prompang]ɴ [kumpulan]ɴ bannyak samma=ka*
1ᴘʟ=ᴘᴏss=ᴀᴅᴅɪᴛ ᴘʀᴏx   woman   association much   every=ʟᴏᴄ
arà-banthu.
ɴᴏɴ.ᴘᴀsᴛ-help
'This women's association of ours has also helped a lot with everything.'
(B060115cvs01)

(73) *Aapcara=ke incayang [ini]ᴅᴇɪᴄ [ciina]ɴ [oorang]ɴ Islam=nang*
how=sɪᴍɪʟ   3s.ᴘᴏʟɪᴛᴇ ᴘʀᴏx   China   man   Islam=ᴅᴀᴛ
asà-dhaathang ...
ᴄᴘ-come   ...
'Somehow he, this Chinaman converted to Islam and ...' (K051220nar01)

POSS RELC N   Very short relative clauses can be preceded by a possessor, as in
(74), where the possessor *se=ppe* 'my' immediately precedes the relative clause *nya-
laaher* 'being born'. Note that this example has a lot of code-mixing with English in
it, so that its value could be contested.

(74) *[Se=ppe]ᴘᴏss [nya-laaher]ʀᴇʟᴄ [date]ɴ duuva duuva 1960.*
1s=ᴘᴏss   ᴘᴀsᴛ-be.born   date   two   two   1960
'My birthday is 2-2-1960.' (K061019prs01)

POSS DEIC N   Normally, the location of possessed items is retrievable for the
hearer, and the deictics *ini* and *itthu* do not combine with possessed items then. In
the following two examples, the deictics are used for emphasis.

(75) *Kandi=pe   raaja=nang [kitham=pe]ᴘᴏss [inni]ᴅᴇɪᴄ [banthu-an]ɴ*
Kandy=ᴘᴏss king=ᴅᴀᴛ 1ᴘʟ=ᴘᴏss   ᴘʀᴏx   help-ɴᴍʟᴢʀ
asà-kamauvan se-aada.
ᴄᴘ-want   ᴘᴀsᴛ-exist
'The Kandyan king had needed this help of ours.' (K060108nar02)

(76) *[Kithang=pe]_{POSS} [ini]_{DEIC} [pompang kumpulan]_N bannyak samma=nang*
     1PL=POSS   PROX   female   association   much   all=DAT
     arà-banthu.
     NON.PAST-help
     'This women's association of ours helps a lot with everything'
     (B060115cvs01,K081105eli02)

In examples (77), the possessor is used to recall that the Malays being talked about are actually forefathers of the speaker.

(77) *Indonesia=dering Sri Lanka=nang [kithang=pe]_{POSS} [ini]_{DEIC} [mlaayu]_N*
     Indonesia=ABL   Sri Lanka=DAT   1PL=POSS   PROX   Malay
     pada asà-dhaathang ini   Malay regiment *atthu.*
     PL   CP-come   PROX Malay regiment INDEF
     'Our Malays came from Indonesia to Sri Lanka in this Malay regiment.'
     (G051222nar03,K081105eli02)

POSS POSS N   There is strictly speaking no modification of a noun by two possessors, since an item normally has only one possessor. What is possible is recursive possession, but this is not a modification of one head noun by several possessors, but a modification of the head noun on level *n*-1 by the possessor on level *n* (78)(79).

(78) *[Se=ppe_{POSS} dhaatha]_N=pe]_{POSS} thiiga [aanak]_N=le Dubai=ka*
     1s=POSS   elder.sister=POSS   three   child=ADDIT Dubai=LOC
     arà-duuduk.
     NON.PAST=stay
     'My elder sister's three children also live in Dubai.' (B060115prs21)

(79) *[[Se=ppe]_{POSS} bìssar aade=pe]_{POSS}   [manthu]_N.*
     1s=POSS   big   younger.sibling=POSS child.in.law
     'My elder younger brother's son-in-law.' (K060116nar02)

POSS QUANT N   This order of prenominal modifiers is possible and is given in the following three examples.

(80) *Itthu blaakang [kithang=pe]ₚₒₛₛ [bannyak]_QUANT [sudaari-sudaara]_N pada,*
DIST after 1PL=POSS much sister-brother PL
kithang anà-pii ruma saakith=nang.
1PL PAST-go house sick=DAT
'After that, many of our brothers and sisters, we all went to the hospital.'
(B060115nar02)

(81) *[Ini kaaving=nang aada] haadath-saadath pada, [kitham=pe]ₚₒₛₛ*
PROX wedding=DAT exist traditions PL 1p=POSS
[bannya]_QUANT [ooram]_N pada arà-kijja.
many people PL NON.PAST-make
'The traditions that were at this wedding, many of our people follow them.'
(K061122nar01,K081105eli02)

(82) *Derang=nang [Kluum̌bu=pe]ₚₒₛₛ [samma]_QUANT [association]_N=le*
3PL=DAT Colombo=POSS all association=ADDIT
support.
support
'All Colombo associations supported them.' (K060116nar06,K081105eli02)

POSS NUM N This order is given in the following two examples.

(83) *[Mliige=pe]ₚₒₛₛ [duuva]_NUM [subla]_N=ka su-thaanàm.*
palace=POSS two side=LOC PAST-plant
'They planted them on both sides of the palace.' (K070000wrt04)

(84) *Se=ppe [dhaatha=pe]ₚₒₛₛ [thiiga]_NUM [aanak]_N=le Dubai=ka*
1s=POSS elder.sister=POSS three child=ADDIT Dubai=LOC
arà-duuduk.
NON.PAST=stay
'My elder sister's three children also live in Dubai.' (B060115prs21)

POSS ADJ N The possessor can precede the adjective preceding a noun.

(85) *[Se=ppe]ₚₒₛₛ [bìssar]_ADJ [aade]_N=pe manthu.*
1s=POSS big younger.sibling=POSS child.in.law
'My elder younger brother's son-in-law.' (K060116nar02)

(86) *Itthu=le*   *[oorang mlaayu=pe]*_{POSS} *[baaye]*_{ADJ} *hathu* *[traditional food]*_N
    DIST=ADDIT man   Malay=POSS   good   INDEF tradtional food
    hatthu.
    INDEF
    'That is also one of the Malays' good traditional food.' (K061026rcp04)

POSS N N   A possessor can precede a sequence of two nouns.

(87) *[Se=ppe]*_{POSS} *[aanak]*_N *[pompang]*_N=*nang duuva aanak klaaki.*
    1s=POSS   child   female=DAT   two child male
    'My daughter has two sons.' (K051201nar01)

QUANT RELC N   This order was not found.

QUANT DEIC N   This order was not found.

QUANT POSS N   The noun can be preceded by a quantifier and a possessor in that order.

(88) *[Samma]*_{QUANT} *[kithang=pe]*_{POSS} *[mlaayu]*_N, *hathu muusing su-aada,*
    all   1PL=POSS   Malay   INDEF time   PAST-exist
    samma cinggala=dering=jo athi-oomong.
    all   Sinhalese=ABL=EMPH IRR-talk
    'There once was a time when all Malays would speak in Sinhala ([n our homes, but we have taken measures against it].' (B060115cvs01)

QUANT QUANT N   This order was not found.

QUANT NUM N   This order was not found.

QUANT ADJ N   The quantifier precedes the adjective, as shown in the following example.

(89) *[Spaaru]*_{QUANT} *[bìssar]*_{ADJ} *[ruuma]*_N *pada aada.*
    some   big   house   PL exist
    'There are some big houses.' (K081105eli02)

QUANT N N A quantifier can precede a string of two nouns, as shown in the following example.

(90) *Derang derang=pe umma=nang butthul saayang=kee=jo [samma]*_{QUANT}
      3PL   3PL=POSS mother=DAT correct love=SIMIL=EMPH all
      [ruuma]_N [pukurjan]_N=nang=le anà-banthu.
      house   work=DAT=ADDIT   PAST-help
      'They also helped their mother with all the housework.' (K070000wrt04)

NUM RELC N This order was not found.

NUM DEIC N This order was not found.

NUM POSS N This order was not found.

NUM QUANT N This order was not found.

NUM NUM N A head noun can be modified by two adjacent numerals, which conveys uncertainty about the exact amount. In the following example, the speaker is not sure whether his stay in the Navy lasted twelve or thirteen years. The two numeral *doblas* 'twelve' and *thigablas* 'thirteen' jointly modify the noun.

(91) *Oman* Navy=*ka se-duuduk hatthu [doblas]*_{NUM} *[thiga-blas]*_{NUM}
      Oman Navy=LOC PAST-stay one   twelve   three-teen
      [thaaun]_{year}=ke.
      year=SIMIL
      'I stayed in the Oman Navy for about twelve or thirteen years, something like that.' (K051206nar17)

However, it could be argued that in this case, we are not dealing with two modifications, but rather with one complex modification. It is not the case that the years had a cardinality of twelve and that the years at the same time had a cardinality of thirteen. Rather, they had only one cardinality, which is vague and therefore expressed by the complex numeral modifier [doblas thigablas]_{NUM}.

NUM ADJ N The numeral precedes the adjective and the noun.

(92) *thiiga kiccil aanak pada aada.*
      three small child PL   exist
      'There are three small children.' (K081105eli02)

NUM N N   A string of two nouns can be preceded by a numeral.

(93) *Se=dang duuva*$_{NUM}$ *pompang*$_N$ *aade*$_N$=le   *hathu*$_{NUM}$ *klaaki*$_N$
      1s=DAT two   woman   younger.sibling=ADDIT one   man
      aade$_N$=le   *anà-duuduk.*
      younger.sibling=ADDIT PAST-exist
      'I had two younger sisters and one younger brother.' (B060115prs03)

ADJ RELC N   This order was not found.

ADJ DEIC N   This order was not found.

ADJ POSS N   This order was not found.

ADJ QUANT N   This order was not found.

ADJ NUM N   Very rarely, a numeral can be found between an adjective and its head noun.

(94) *[Mlaayu]*$_{ADJ}$ *[thiga-pulu tuuju]*$_{NUM}$ *[baasa]*$_N$ *aada.*
      Malay   three-ty   seven   language exist
      'There are 37 Malay languages [in the world].' (K060116nar02)

(95) *Malay thiiga aanak pada arà-duuduk.*
      Malay three child PL   NON.PAST-stay
      'There are three Malay children [in this house].' (G051222nar01)

   While above, we are dealing with a genuine adjective, the two examples below show this pattern with *hathyang* 'next' and *lai* 'additional', where it can be argued that the scope relations are slightly different, similar to English *Another three men*, as compared to *three other men*.

(96) *[Hathyang]*$_{ADJ}$ *[thiiga]*$_{NUM}$ *[oorang]*$_N$=*yang* Malaysia=*nang su-ambel*
      next   three   man=ACC   Malaysia=DAT PAST-take
      baapi.
      bring
      'They took three more men to Malaysia.' (K060108nar02)

(97) *Kitham=nang duppang [lai]_{ADJ} [duuva]_{NUM} [bàrgaada]_N asà-dhaathang*
     1PL=DAT   before   additional two   group   NON.PAST-come
     aada.
     exist
     'Before us, there are two more families.' (K060108nar02)

ADJ ADJ N Two adjectives can stack before a noun. Two instances of this (*puuthi paanjang* 'white long' and *kiccil jillek* 'small ugly') can be found in the following example.

(98) *Aanak pompang duuva=nang [[[hathu duuri pohong]=nang [[**puuthi**]_{ADJ}*
     child   female   two=DAT   INDEF   thorn tree=DAT   white
     *[**paanjang**]_{ADJ} [**jeeňǧgoth**]_N]=yang anà-kànà-daapath kìnna] hathu*
     long   beard=ACC   PAST-patfoc-get   strike INDEF
     *[**kiccil**]_{ADJ} [**jillek**]_{ADJ} [**Aajuth**]_N hatthu]=yang su-kuthumung.*
     small   ugly   dwarf   INDEF=ACC   PAST-see
     'The two girls saw a small ugly dwarf whose long white beard had got stuck in a thorn tree.'

ADJ N N An adjective can precede a string of two (99) or more (100) nouns.

(99) *Suda se=ppe [thuuva]_{ADJ} [anak]_N [klaaki]_N asàdhaathang dhlapan-blas*
     so   1s=POSS old   child   male   COPULA   eight-teen
     thaaun.
     year
     'So my eldest son is eighteen' (K060108nar02)

(100) *[Panthas]_{ADJ} [rooja]_N [kumbang]_N [pohong]_N [komplok]_N duuva asà-jaadi*
     beautiful   rose   flower   tree   bush   two CP-grow
     su-aada.
     PAST-exist
     'Two beautiful rose bushes had grown.' (K070000wrt04)

N RELC N This order was not found.

N DEIC N This order was not found.

N POSS N   This order was not found.

N QUANT N   This order was not found.

N NUM N   This order was not found.

N ADJ N   This order was not found.

N N N   Strings of three or more nouns are also possible:

(101)   *Panthas [rooja]*$_N$ *[kumbang]*$_N$ *[pohong]*$_N$ *[komplok]*$_N$ *duuva asà-jaadi*
        beautiful rose   flower   tree   bush   two   CP-grow
        su-aada.
        PAST-exist
        'Two beautiful rose bushes had grown.' (K070000wrt04)

8.1.15   Preliminary summary of prenominal modifications

We can schematize the findings about prenominal modifications as in (102):

(102) $\left\{ \begin{array}{c} \text{DEIC} \\ \text{POSS} \end{array} \right\}$ RELC $\left\{ \begin{array}{c} \text{POSS} \\ \text{DEIC} \end{array} \right\}$ QUANT $\left\{ \begin{array}{c} \text{NUM} \\ \text{ADJ} \end{array} \right\}$* N Noun

    Up to now, we have omitted discussion of the position of the indefiniteness marker
(h)a(t)thu. The position of this marker within the NP will be discussed in the next
section.

8.1.16   The position of the indefinite modifier

Above, we have treated the order of some prenominal modifiers. There is one modifier
which we have not treated, which is the indefiniteness marker *(h)a(t)thu*. This marker
can intervene between any two markers in (102), and additionally can also occur more
than once (cf. Slomanson 2006). Example (103) gives a sentence with multiple oc-
currence of *atthu* in one NP. The possibility to occur more than once points to an
appositional nature of the NP, which will be discussed in more detail in ⊡→8.9, p. 455.

(103)   *Itthu=le   **hathu** mlaayu oorang pada=pe **hathu** baae **hathu** makanan=jo.*
        DIST=ADDIT INDEF Malay man   PL=POSS INDEF good INDEF food=EMPH
        'That is also a good dish of the Malays.' (K061026rcp02)

***hatthu* preceding a relative clause**   This order was not found.

***hatthu* following a relative clause**   The following four examples show the use of *hatthu* after a subordinate clause modifying the noun. This is a relative clause in (104) to (106) and a purposive clause in (107).

(104)   *[anà-birthi-king]ᵣₑₗc [hatthu]ᵢₙₒₑₓ [paapang]ₙ=ka*
        PAST-stand-CAUS   INDEF   pole=LOC
        'on a plank put upright' (K081105eli02)

(105)   *[Kìrras pinthu=nang arà-thatti]ᵣₑₗc [hathu]ᵢₙₒₑₓ [svaara]ₙ su-dìnngar.*
        strong door=DAT   SIMULT-hammer] INDEF   noise   PAST-hear
        'They heard a noise of hard hammering at the door.' (K070000wrt04)

(106)   *[Seelon=le   kithang=pe mlaayu=nang=le hatthu bagiyan anà-aada]ᵣₑₗc*
        Ceylon=ADDIT 1PL=POSS Malay=DAT=ADDIT INDEF part   exist
        [hatthu]ᵢₙₒₑₓ [nigiri]ₙ su-jaadi
        INDEF   country PAST-become
        'Ceylon became a country where our Malays also have a part in.'
        (K051222nar04)

(107)   *[Thaangang mà-saapu]ᵣₑₗc [hatthu]ᵢₙₒₑₓ [paper]ₙ kapang-mintha,*
        hand   INF-sweep   one   paper   when-ask
        baapa=yang su-kuthumung.
        father=ACC PAST-see
        'When he asked a paper to clean his hands, he saw father.' (K051205nar05)

***hatthu* following or preceding a deictic**   Given the semantics of *hatthu* marking indefiniteness, and the deictics marking definite referents, these morphemes cannot cooccur.

***hatthu* preceding a possessor**   *Hatthu* can precede the possessor (108), but is more likely to follow it as will be discussed in the next section

(108)   *Itthu=le   [hathu]ᵢₙₒₑₓ [mlaayu oorang pada=pe]ₚₒₛₛ hathu baae hathu*
        DIST=ADDIT INDEF   Malay man   PL=POSS   INDEF good INDEF
        [makanan]ₙ=jo.
        food=EMPH
        'That is also a good dish of the Malays.' (K061026rcp02)

***hatthu* following a possessor**   To indicate that within a set of possessed items, an indefinite one is talked about, *hatthu* and the possessor can cooccur. In those cases, *atthu* normally follows the possessor. In example (109), the speaker has many sons, but only one of them, who is at the time of speaking not known to the hearer, is on the estates.

(109)   *Karang [se=ppe]ₚₒₛₛ [hathu]ᵢₙDₑF [aanak]ₙ estate=ka.*
        now   1s=POSS   INDEF   child   estate=LOC
        'Now one of my sons is on the estates.' (K051201nar01)

Similar things can be said about (110), where the identity of the Malay woman's child is unknown to the hearer.

(110)   *Lai=le   hatthu [sudaari=pe]ₚₒₛₛ [atthu]ᵢₙDₑF [aanak]ₙ, kiccil aanak*
        more=ADDIT INDEF   sister=POSS   INDEF   child   small child
        atthu puruth=ka se-mniiňǧgal.
        INDEF womb=LOC PAST-die
        'A child of yet another sister, a small child, died in the womb.'
        (B060115nar02)

The third example for this patterns has again to do with indicating that the speaker is not supposed to be able to identify which one of the children the speaker is talking about.

(111)   *[Itthu kaake=pe]ₚₒₛₛ   [hatthu]ᵢₙDₑF [aanak]ₙ=jo se=ppe umma.*
        DIST grandfather=POSS once   child=EMPH 1s=POSS mother
        'One   of   that   grandfather's   children   is   my   mother.'
        (K051205nar05,K081105eli02)

***hatthu* following or preceding a quantifier**   These orders were not found.

***hatthu* preceding a numeral**   *Hatthu* can precede a numeral, in which case it marks uncertainty about the exact amount. Still, this is an instance of *hatthu* modifying the numeral expression rather than the noun.

(112) *Oman* Navy=*ka se-duuduk [hatthu$_{INDEF}$ doblas$_{NUM}$ thiga-blas$_{NUM}$]$_{NUM}$*
Oman Navy=LOC PAST-stay INDEF twelve three-teen
[thaaun]$_N$=ke.
year=SIMIL
'I stayed in the Oman Navy for about 12 or thirteen years, something like that.'
(K051206nar17)

***hatthu* following a numeral** This order was not found.

***hatthu* preceding an adjective** *Hatthu* can precede an adjective modifying a noun.

(113) *Hathu muusing=ka ... [hathu]$_{INDEF}$ [kiccil]$_{ADJ}$ [ruuma]$_N$ su-aada*
INDEF time=LOC ... INDEF small house PAST-exist
'Once upon a time, there was a small house.' (K07000wrt04)

***hatthu* following an adjective** Just like *hatthu* can precede the adjective, it can also follow, as shown in (114).

(114) *Itthu [bannyak [laama]$_{ADJ}$ [hathu]$_{INDEF}$ [ruuma]$_N$].*
DIST very old INDEF house
'That one was a very old house.' (K070000wrt04)

***hatthu* preceding a modifying noun** It is the normal case for *atthu* to precede a modifying noun.

(115) *Sithu=ka, [hathu]$_{INDEF}$ bìssar [beecek caaya]$_N$ [Buruan]$_N$ su-duuduk.*
there=LOC INDEF big mud colour bear PAST-exist
'There was a big brown bear.' (K070000wrt04)

***hatthu* following a modifying noun** In some instances, *hatthu* can intervene between the two nouns of a two-noun sequence. The following two examples show this for a monkey group and a street bend.

(116) *Ini pohong atthas=ka [moonyeth]$_N$ [hathu]$_{INDEF}$=[kavanan]$_N$ su-aada.*
PROX tree TOP=LOC monkey INDEF=group PAST-exist
'On top of this tree was a group of monkeys.' (K070000wrt01)

(117) *[[Jaalang_N hathu_INDEF=piiňǧgir_N]=ka anà-aada   hathu=pohong]*
road   INDEF=border=LOC   PAST-exist.inanim INDEF=tree
baava=ka su-seeňder.
down=LOC PAST-rest
'Because he was tired, he sat down under a tree which stood at a side of the street.' (K070000wrt01)

8.1.17 The final structure of the noun phrase

By combining the order of postnominal modifiers, presented in (52), with the order of the prenominal modifiers presented in (102) and the possible occurrences of the indefiniteness marker *atthu*, we get the full templatic structure of the NP, represented in (118). The arrows represent positions where the indefinite marker *atthu* can occur. The semantic (in)compatibility of certain items is not reflected in (118).

(118)
$$\downarrow \begin{Bmatrix} \text{DEIC} \\ \text{POSS} \end{Bmatrix} \text{RELC} \begin{Bmatrix} \text{POSS} \\ \text{DEIC} \end{Bmatrix} \downarrow \text{QUANT} \begin{Bmatrix} \text{NUM} \\ \text{ADJ} \end{Bmatrix} * \downarrow \text{N} \downarrow \textbf{Noun} \quad \begin{matrix} \downarrow \\ \text{N(PL)} \\ \text{ADJ(PL)} \\ \text{QUANT(PL)} \\ \text{NUM(PL)} \end{matrix} \quad {}_{NP}$$

The only order which cannot be captured by this schema is QUANT POSS N ⊡→88, p. 436, but this order could be explained by quantifier floating ⊡→88, p. 436.

Analyzing this schema, we can observe a certain number of points: The relative clause belongs to the leftmost elements. This can be explained by the desire to have little material separating heads from modifiers (Hawkins 1994). Given that relative clauses are generally heavier than the other modifiers, an obvious solution to minimize separation is to put them at the margins of the NP (Rijkhoff 2002:298f), the left edge in the case of SLM. In schema (118), the relative clause is listed together with the deictics and the possessors because of isolated examples which permit the order DEIC RELC and POSS RELC. It might also be possible to discount these examples and have the relative clause as the leftmost element, which is the position it assumes in the near totality of cases. The position of the deictics and the possessors with regard to the other modifiers is unremarkable. What is more remarkable is the possibility to arrange ADJ and NUM in both possible orders. It seems that the order ADJ NUM violates semantic scope; the numeral should have scope over the adjective, but its syntactic position suggests that it has not. Another interesting aspect to note is that the postnominal field only allows one element, while the prenominal field allows an arbitrary number of elements. One could argue that postmodifications are archaic constructions which are somehow lexicalized, like *orang ikkang* 'man'+'fish'=fisherman' or *aer meera* 'water'+'red'='tea'. These lexicalized postmodifications persist, but generally do not involve more than one modifier. More elaborate modifications are ad hoc, and

prenominal.[5] The limited possibility of postmodification is then a consequence of the limited lexicalization of collocations with more than two elements.

A final puzzle is the position of the indefiniteness marker *hatthu*, the theoretical implications of this will be discussed in more detail in ⬚→8.9, p. 455.

8.2 Noun phrases based on a deictic

Next to nouns, NPs can also be formed based on a deictic, which replaces the lexical content. The following examples show the use of an NP based on the proximal deictic *ini* (119) and the distal deictic *itthu* (120). The postposition =*yang* following the deictics shows that there is no other content in the NP.

(119) *Thraa thraa **inni=yang** masà-picca-kang katha biilang.*
no   no   PROX=ACC must=broken-CAUS QUOT say
' "No, no", he said, "you must break this one." ' (K051220nar01)

(120) *Asà-cuuci, **itthu=yang** baaye=nang asà-rubbus, ...*
CP-wash   DIST=ACC good=DAT   CP-boll,   ...
'You wash (it), then you boil it well and ... .' (B060115rcp01)

NPs based on a deictic can only be modified by the plural marker *pada,* as in (121) (122). Other modifications are not possible.

(121) *Mlaayu pada=jo **inni pada**=ka punnu pukurjan anà-girja.*
Malay PL=EMPH PROX PL=LOC much work   PAST-do
'It was the Malays who did a lot of work in these (jobs) [i.e army, navy, police].' (K051222nar05)

(122) *Iiya, **itthu pada**=jo su-dhaathang.*
yes DIST PL=EMPH PAST-come
'Yeah, those [people] came.' (K051201nar02)

A special case is (123), where the NP consisting of a deictic hosts the possessive marker =*pe*, which in turn is made into a new NP by conversion.

---

[5]See Saldin (2001:53) for a diverging view.

(123) *[[[Itthu_DEIC]_NP=pe]_PP=∅]_NP pada=jo bannyak mlaayu pada karang siini*
       DIST=POSS   PL=FOC much   Malay PL   now   here
       aada.
       exist
       'It's their folks we get a lot of today here.' (K051205nar04)

8.3   Noun phrases based on a personal pronoun

NPs can also be formed based on a personal pronoun. Normally, personal pronouns
occur unmodified, as in (125).

(124) | PERSPRON |_NP

(125) ***Kithang**_NP=nang baaye=nang mulbar bole=baaca.*
       1PL=DAT   good=DAT Tamil can=read
       'We can read Tamil well.' (K051222nar06)

It is possible to use appositions of number for plural pronouns. These can be the
plural marker *pada* as in (127), a numeral as in (128) or a whole expression as in
(129)(130).

(126) $\left[\text{PERSPRON} \left\{ \begin{array}{c} \text{PL} \\ \text{NUM} \\ \text{EXPR} \end{array} \right\} \right]_{NP}$

(127) *Itthu=nam blaakang=jo, **kitham pada** anà-bìssar.*
       DIST   after=EMPH   1PL   PL   PAST-big
       'It was after that that we grew up.' (K060108nar02)

(128) *Mr Sebastian aada, se aada, **kitham duuva** arà-oomong.*
       Mr Sebastian exist 1s exist 1PL   two   NON.PAST-speak
       'You are here, I am here, the two of us are talking.' (K060116nar05)

(129) ***Derang duuva oorang**=pe naama pada kalu Snow-white hattheyang*
       3PL   two   man=POSS name PL if Snow.White other
       Rose-red.
       Rose.Red
       'The names of the two of them were, if you ask, Snow White, the other one
       Rose Red.' (K070000wrt04)

(130) *Itthu=kaapang se=ppe baapa se=ppe kaake   se=ppe*
      DIST=when   1s=POSS father 1s=POSS grandfather 1s=POSS
      kaake=pe   baapa, **kithang samma oorang** *Seelon=pe   oorang*
      grandfather=POSS father, 1PL   all   man   Ceylon=POSS man
      pada.
      PL
      'Then my father and my grandfather and my grandfather's father, all of us people became Ceylon people.' (K060108nar02)

The use of these appositions is optional, as the following example shows, where *pada* is now present, now absent.

(131) **Lorang**=*le,* **lorang pada**=*pe umma=le   see=yang baaye=nang*
      2PL=ADDIT 2PL   PL=POSS mother=ADDIT 1s=ACC good=DAT
      anà-kuaather.
      PAST-look.after
      'You and your mother took good care of me.' (K070000wrt04,K081105eli02)

8.4 Noun phrases based on interrogative pronouns

There are three ways to form NPs based on interrogative pronouns, bare, reduplicated and combined with clitics.

The simplest one is to use only the pronoun. This means that the referent is unknown and implies a question, as in (133).

(132) | WH | _{NP}

(133) *Itthu=nang blaakang* **aapa** *nya-gijja?*
      DIST=DAT   after   what PAST-make
      'What did (they) do then?' (K051206nar07)

An NP can be formed by a reduplicated interrogative pronoun as well. In this case, an exhaustive answer is required.

(134) *[Aapa˜aapa]_{NP} kitham Kandi=pe   cultural show atthu=le   thaaro?*
      what˜red   1PL   Kandy=POSS cultural show one=ADDIT put
      'What did we also put on the Kandy Cultural show?' (K060116nar11)

An interrogative pronoun can be combined with the clitics =*so, =ke* or =*pon* to yield the reading of an indefinite pronoun.

(135) $\text{WH}=\left\{\begin{array}{l}= so \\ = ke \\ = pon\end{array}\right\}_{NP}$

(136) ***Saapa=so***   Malay exam *arà-girja.*
who=UNDET Malay exam NON.PAST-make
'Someone is taking a Malay exam.' (K060103cvs01)

(137) *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
Hopefully other know brother sister   PL=LOC   can-ask   take
*[nya-gijja lai   **saapa=kee** aada=si   katha].*
PAST-make other who=SIMIL exist=INTERR QUOT
'Hopefully, you can enquire from another person you know whether there is
someone else who did something.' (N061031nar01)

(138) *Kithang **craapa=pon** kithang=pe kappal asà-ambel, kithang su-baalek*
1PL   how=any   1PL=POSS   ship   CP-take   1PL   PAST-turn
kithang=pe harbour=*nang.*
1PL=POSS   harbour=DAT
'Anyhow, we managed to take our ship and return to the harbour.'
(K051206nar20)

There are subtle differences between the meanings conveyed by the three clitics. In
example (139), the speakers are lost in the woods, but finally arrive at their destination.
If the similative clitic =*kee* is used, the meaning is one of concession, or canceling
of implicatures, similar to English *anyhow*. If the 'undetermined' clitic =*so* is used
instead of =*kee*, the meaning is that the walkers were not aware of their having reached
their destination.

(139) *Kithang **caraapacara**=kee/so [mà-pii anà-aada thumpath]=ka su-sampe.*
1PL   how=SIMIL/UNDET   INF-go PAST-exist place=LOC   PAST-reach
'We arrived at our destination unknowingly (=kee)/anyhow (=so).'
(K081105eli02)

8.5   Noun phrases based on a numeral/quantifier

Numerals and quantifiers can constitute a noun phrase on their own, even if they still
need some quantifiable content in discourse to relate to.

(140) $\boxed{\text{(DEIC) NUM (PL)}}_{NP}$

(141) | (DEIC) QUANT (PL) |_{NP}

Example (142) shows the use of a quantifier as the only element of an NP.

(142) *Mr Yusuf **samma** asà-ambel=apa,* Commercial Bank=*ka su-thaaro*
Mr Yusuf all   CP-take=after   Commercial Bank=LOC past-put
'Mr Yusuf took everything and deposited it at the Commercial Bank.'
(K060116nar09)

Example (143) shows the use of a numeral as the only element of an NP.

(143) *'Siking' katha arà-biilang* 'that-wise/So', *inni **duuva**=le buthul.*
siking   QUOT NON.PAST-say that-wise/So   PROX two=ADDIT correct
' "siking" means "that-wise/so"; both of them are correct.' (K071113eml01)

NPs based on quantifiers can also be modified by the plural marker, as shown in
(144).

(144) ***Spaaru pada** bannyak suuka arà-blaajar.*
some   PL   much   like   NON.PAST-learn
'(Only) some like to study a lot.' (B060115cvs01)

8.6 Reciprocal noun phrases

The reciprocal construction is formed by adding *hatthunang hatthu* to a noun with
plural reference[6]

(145) *Oorang pada **hatthu=nang hatthu** maara.*
man   PL   INDEF=DAT   INDEF anger
'People are angry with each other.' (K081106eli01)

The dative marker =*nang* is present in the construction, even if the verb used would
normally subcategorize for another case marker, like *buunung* 'kill' in (146), which
normally governs the accusative.

(146) *Oorang pada **hatthu=(nang/*yang) hatthu** arà-buunung ambel.*
man   PL   INDEF=DAT/ACC   INDEF NON.PAST-kill take
'The people kill each other.' (K081106eli01)

---

[6]See Beythan (1943:177) and Malten (1989:11) for an analogous Tamil construction.

8.7 Noun phrases based on a postpositional phrase

Postpositional phrases can convert into noun phrases, as shown in (147) and schematized in (148).

(147) ***Itthu=pe pada**=jo bannyak mlaayu pada karang siini aada.*
       DIST=POSS PL=FOC much Malay PL now here exist
       'It's their folks we get a lot of today here.' (K051205nar04)

(148) [[[*itthu*$_N$]$_{NP}$ = *pe*]$_{PP}$ ∅ *pada*]$_{NP}$

8.8 Noun phrases based on a clause

In SLM, clauses can be used as noun phrases as they are. No further morphological flagging of this use is necessary. While in English, NPs consisting of clauses are indicated by special means, such as the complementizer *that* in *I appreciated that you came*, this is not the case in SLM. Clauses can be used as they are as complements of verbs ⊡→8.8.1, p. 450, or as NPs (headless relative clauses) ⊡→8.8.2, p. 453.

8.8.1 Argument Clause

SLM clauses can be used as the head of a term without further measures (like nominalizations or complementizers) being taken.

(149) | CLAUSE |$_{NP}$

    Example (150) shows the use of a finite clause as a complement of the verb *suuka* 'like'.

(150) *Kitham=pe baapa su-biilang [[**lorang suurath=yang mlaayu=dering***
       1PL=POSS father PAST-say 2PL letter=ACC Malay=ABL
       ***anà-thuulis**]$_{CLAUSE}$=nang bannyak arà-suuka].*
       past=write=DAT much simult-like
       'Daddy said that he liked very much that you wrote the letter in Malay.' (Letter 26.06.2007)

    Postpositions can then attach to this new NP like to any other NP. This produces the curious situation that a verbal lexeme can have a verbal prefix on the left side and a case marker on the right side, like *anà-* 'past' and *=nang* 'DAT' both attaching to the verb *thuulis* 'write' in (150).[7]

---

[7]It would also be possible to analyze these embedded clauses as subordinates, and the postpositions as conjunctions. But this needlessly increases the number of categories without adding to our understanding.

An alternative to using finite clauses is to put the verb into the infinitive, which yields slightly different semantics, often purposive as in (151).

(151) *Hathu haari, hathu oorang [[**thoppi mà-juval**]₍CLS₎]₍NP₎=nang*
     INDEF day INDEF man hat   INF-sell=DAT
     kampong=dering kampong=nang su-jaalang pii.
     village=ABL   village=DAT PAST-walk go
     'One day, a man walked from village to village to sell hats.' (K070000wrt01)

We have seen in (150) that a clause whose verb is inflected with *anà-* can be used as an NP. (151) shows the same for a verb in the infinitive. The use of *arà-* in its use as simultaneous tense marker is also possible, as the following three examples show.

(152) *Blaakang=jo incayang anà-kuthumung [[moonyeth pada thoppi asà-ambel*
     after=EMPH 3s.POLITE PAST-see   monkey   PL hat CP-take
     pohong atthas=ka arà-maayeng]₍CLS₎]₍NP₎.
     tree   top=LOC SIMULT-play
     'Then only he saw that the monkeys had taken his hats and were playing on the top of the tree.' (K070000wrt01)

(153) *Derang su-kuthumung [[ithu buurung=pe kuuku=ka Aajuth asà-sìrrath*
     3PL   PAST-see   DIST bird=POSS claw=LOC dwarf CP-stuck
     kìnna arà-duuduk]₍CLS₎]₍NP₎.
     strike NON.PAST-stay
     'They saw that the dwarf sat stuck in the claws of the bird.' (K070000wrt04)

(154) a. *Thàrà-kalu [[ini oorang thoppi arà-kumpul]₍CLS₎]₍NP₎=**yang***
        NEG-if   PROX man hat   NON.PAST-collect=ACC
        asà-kuthumung=apa
        CP-see=after
        'Furthermore, when (they) had seen the man collect the hats'

     b. *moonyeth pada=le asà-dhaathang creeveth athi-kaasi katha.*
        monkey PL=ADDIT CP-come   trouble IRR=give QUOT
        'the monkeys would certainly go and cause (some other) trouble.' (K070000wrt01)

The use of clauses as NPs where the verb is inflected with *anà-* is shown in the following two examples. *Su-* is shown in (157).

On the contrary, it obscures the fact that lexical and clausal arguments are treated exactly alike when it comes to assigning semantic roles.

(155)   *[[School=nang anà-pii]CLS]NP=nang blaakang.*
        school=DAT   PAST-go=DAT   after
        'After having gone to school.' (K051222nar04)

(156)   *[[Seelong independent state anà-jaadi]CLS]NP=nang=apa [[kithang=nang*
        Ceylon   independent state PAST-become=DAT=after   1PL=DAT
        independence *anà-daapath]CLS]NP=nang=apa.*
        independence PAST-get=DAT=after
        'After Ceylon had become an independent state, after we had obtained inde-
        pendence.' (K051222nar06)

(157)   *[[Se=ppe uncle=ka   su-dhaathang]CLS]NP=nang blaakang*
        1s=POSS   uncle=LOC PAST-come=DAT   after
        'After I had come to my uncle's' (K051201nar02)

The use of *asà-* as inflection on the clause serving as NP is exemplified by (158).
But the use of an uninflected verb as in (159) (160) is also possible.

(158)   *Suda butthul suuka [[asà-dhaathang]CLS]NP=nang.*
        thus very   like   cp-come=DAT
        'So, I am very pleased that you have come.' (G051222nar01)

(159)   *[[Manis-an   maakang]CLS]NP=nang go   suuka bannyak.*
        sweet-NMLZR eat=DAT   1s.FAMILIAR like   much
        'I like very much to eat sweets.' (B060115prs20)

(160)   *Kithang=nang maau, [[kitham=pe mlaayu looang ∅-blaajar, lorang=pe*
        1PL=DAT   want 1PL=POSS   Malay 2PL   learn   2PL=POSS
        mlaayu kitham ∅-blaajar]CLS]NP.
        Malay 1PL   learn
        'We want that you learn our Malay and that we learn your Malay.'
        (K060116nar02)

To sum up, we see that clauses which function as an NP can be headed by verbs
in various tenses. There seem to be no restrictions on the character of the verb or the
TAM-prefix.

8.8.2 Headless relative clauses

The last possibility to form NPs is the headless relative clause. On the surface, it looks exactly like the argument clause above, but the semantics are different.

(161)   [[CLAUSE]∅] $_{NP}$

The headless relative clause has the same form as any other relative clause ⊡→13.2, p. 516. The difference between the two is that the head noun is not expressed. Headless relative clauses then do not have a *restrictive* function (among all head nouns, select those which comply with the relative clause), but a *maximizing* function (among *all* nouns, select those which comply with the relative clause). Since there is no head noun, it is impossible to claim that the headless relative clause is a modifier in the NP; it must be the head. Example (162) gives a normal relative clause with a head. Example (163) gives a headless relative clause, which fulfills the function of the NP.

(162)   *[Lorang anà-maasak ikkang] eenak.*
         2PL   PAST-cook fish   tasty
         'The fish you cooked is tasty.' (K081105eli02)

(163)   *[Lorang anà-maasak ∅] eenak.*
         2PL   PAST-cook   tasty
         'What you cooked is tasty.' (K081105eli02)

These NPs formed by headless relative clauses can take case markers, as in (164).

(164)   *[Lorang=ka aada=∅]=yang kaasi.*
         2PL=LOC   exist=∅=ACC   give
         'Give me whatever you have.' (K081105eli02)

NPs based on headless relative clauses can only be modified by the plural marker *pada* (165), but by nothing else.

(165)   *[Lorang=ka aada ∅] pada(=yang) kaasi.*
         2PL=LOC   exist   PL(=ACC)   give
         'Give me all you have.' (K081105eli02)

Examples from the corpus with headless relative clauses as NPs are given in (166) to (169). (166) shows an equational sentence, where the headless relative clause is without any doubt the first term, which is furthermore modified by *pada*.

(166) *[Seelon=nang anà-dhaathang ∅] pada mlaayu pada.*
      Ceylon=DAT   PAST-come   PL   Malay PL
      'Those who had come to Ceylon were the Malays.' (N060113nar01)

Example (167) shows a parallel construction with two headless relative clauses for verbal predications.

(167)   a. *[Se=dang nya-boole ∅] pada see nya-ambel.*
            1S=DAT   PAST-can   PL   1S PAST-take
            'I took what I could.'

        b. *[Se=dang thàràboole ∅] pada see thàrà-ambel.*
            1S=DAT   cannot   PL   1S NEG.PAST-take
            'What I couldn't take, I didn't take.' (K051213nar01)

Example (168) shows again a non-verbal predication, this time supported by the copula *asàdhaathang*.

(168)   *Itthu vakthu [kithang=nang nya-aada ∅] asàdhaathang ini   JVP katha*
        DIST time 1PL=DAT   PAST-exist COPULA   PROX JVP QUOT
        hathu problem hatthu=jo.
        INDEF problem INDEF=EMPH
        'What we had at that time was the so-called JVP problem.' (K051206nar10)

Example (169) shows the use of a headless relative clause marked for case.

(169)   *[Derang anà-kuthumung ∅] pada=**nang** asà-thaakuth ruuma=nang mà-laari*
        3PL   PAST-see   PL=DAT   CP-fear   house=DAT INF-run
        kapang-pii derang=nang byaasa svaara hatthu su-dìnngar.
        when-go   3PL=DAT   habit sound INDEF PAST-hear
        'They feared what they saw and when they went running back to their home, they heard a familiar voice.' (K070000wrt04)

The most audacious use of a headless relative clause is probably (170), where a clause containing the verb *biilang* 'say' is used to refer to the person of the name given as an argument for *biilang*.

(170)   *[Andare katha arà-biilang ∅] raaja mliiga=ka **hathu oorang koocak**.*
        Andare QUOT NON.PAST-say] king palace=LOC INDEF man   joke
        '(The man) called Andare was jester at the royal palace.' (K070000wrt05)

8.9 The SLM NP as fundamentally appositional

As discussed in Section 8.1.14, the order of elements in the SLM NP is quite free. Furthermore, just about any item can constitute a NP on its own, without the need for dummy elements (like English *a big one*, ⊡→8.1.1, p. 416-⊡→8.1.9, p. 423. Representing these facts in a hierarchical structure is difficult. Which element should be the head, if any element could be the head? How to account for the many possible permutations? Furthermore, how can one explain the multiple occurrences of *hatthu* in one NP ⊡→8.1.16, p. 440? All this suggests, that a hierarchical structure might not be the best analysis of the SLM NP. If we assume an appositional structure on the other hand, the three problems mentioned above can be resolved (cf. Rijkhoff 2002:275).

I will exemplify this with the following example.

(171) *[Itthu=le] [hathu mlaayu oorang pada=pe hathu baae hathu makanan]=jo.*
       DIST=ADDIT INDEF Malay man PL=POSS INDEF good INDEF food=EMPH
       'That is also a good dish of the Malays.' (K061026rcp02)

This example is an ascriptive sentence which attributes to the anaphoric referent *itthu* membership in the class of tasty Malay foods. This class of foods is expressed by the string *hathu mlaayu oorang pada=pe hathu baae hathu makanan* 'Malay people's good food', an NP. We will take this (rather complex) NP as a point of departure for our analysis. In a first step, we will disregard all occurrences of the indefinite article. We will return to it afterwards.

The following two trees show the difference between a hierarchical representation (172) and an appositional representation (173).[8]

(172) hierarchical structure
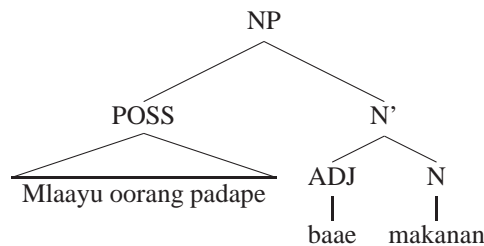
(173) appositional structure

---

[8]The trees follow Givón's (2001) argument for the usefulness of trees based in the style of Chomsky (1957, 1965) to represent constituency, hierarchy, category labels, and linear order, in an elegant way. Generative grammar has used, and modified, trees and tree structure in a number of ways since then, but these more recent types of trees would not serve the illustrative purpose they are intended for here any better than their more traditional counterparts, regardless of whether one finds the theoretical arguments for the more modern trees compelling or not.

```
                                NP
                      /          |          \
                   POSS         ADJ          N
               /        \        |           |
        Mlaayu oorang padape   baae       makanan
```
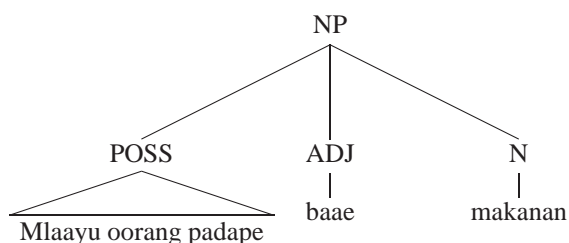
The hierarchical structure is more nested, and clear dependency and government relations hold between the nodes of the tree. In the appositional structure, this is not the case. All elements are of equal importance, and none governs another one. Since in SLM, just about anything can head an NP (see above), a difference in prominence of the elements in the NP does not seem to exist. A theoretical representation which does not imply such difference in prominence, like the appositional structure in (173), is superior to a representation which makes additional assumptions, which are unwarranted.

If one node is left out of the appositional structure, the structure remains intact, whereas in (172), leaving out certain nodes leaves the structure ill-formed. If the order of elements is changed in (172), the hierarchy will have to be modified to accommodate the new linear order. In (173) on the other hand, the structure for the new order will resemble the structure for the former order very much.

Turning to the position of the indefiniteness marker, the multiple occurrences of *hatthu* are difficult to justify within one hierarchical NP. Under the appositional account, multiple occurrences can be explained if we tweak the structure a bit, as show in (174).

(174)   tweaked appositional structure

```
                                      NP
                     /                |                    \
                   NP                NP                    NP
              /         \         /      \            /        \
          INDEF       POSS    INDEF     ADJ       INDEF         N
            |        /     \     |        |         |           |
         hatthu  Mlaayu oorang  hatthu  baae     hatthu      makanan
                    padape
```
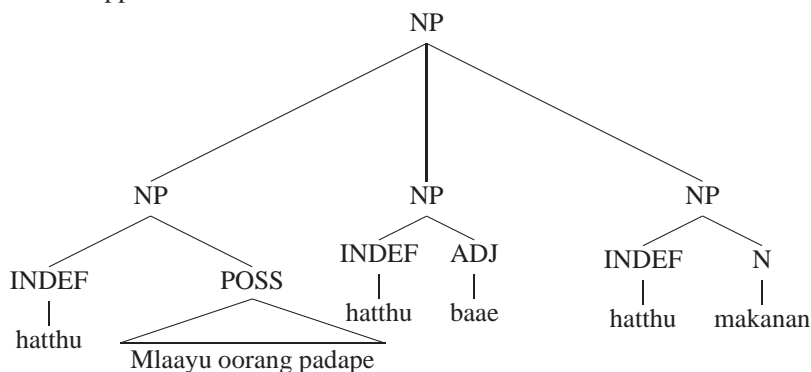
Instead of an apposition of POSS, ADJ and N, as in (173), we now have an apposition of three NPs, which are in turn headed by said elements. Since we are dealing with three complete NPs (four if we include the overarching NP), the three occurrences of *hatthu* are no problem since they modify three different entities.

The above discussion has shown that an appositional representation of the SLM NP can account for the permutation of elements, arbitrary heads, and multiple occurrence of *hathu*. These three aspects could not be explained satisfactorily by a hierarchical analysis. There is one thing, however, which posits a problem for the appositional analysis: the obligatory position of the noun in final or prefinal position (if a noun occurs in the NP). It is impossible to have more than one postmodifying element in an SLM NP (excluding *pada* for the moment), illustrated in (175). If the structure was fully appositional, this is not what we would expect. The noun should then be able to occur in any position within the NP. I have no good explanation at the moment for this semi-free structure we find in the SLM NP.

(175)
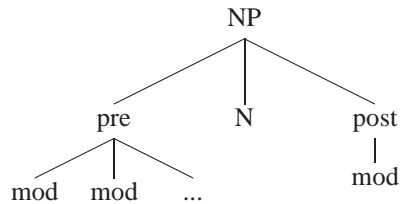
Chapter 9

The postpositional phrase

A postpositional phrase consists of an NP plus a postposition.

(176) | NP=POSTP |$_{PP}$

There are no restrictions on the character of the NP (Noun, pronoun, numeral, ...) or the character of the postposition as examples (177) to (180) show, as long as semantic interpretability is possible.

(177)   *[Seelong$_{NP}$=nang dhaathang]$_{CLAUSE}$=nang blaakang=jo incayang*
         Ceylon=DAT   come=DAT   after=EMPH   3s.POLITE
         [cinggala asà-blaajar]$_{CLAUSE}$=apa sini=pe   raaja$_{NP}$=nang mà-banthu
         Sinhala   CP-learn=after   here=POSS king=DAT   INF-help
         anà-mulain.
         PAST-start
         'It was after that he had come to Ceylon that he learned Sinhala and began to help the local king.' (K060108nar02)

(178)   *Incayan$_{PRON}$=nang baae$_{ADJ}$=nam mlaayu mà-oomong butthul suuka.*
         3s.POLITE=DAT   good=DAT   Malay INF-speak   very   like
         'He likes very much to speak Malay well.' (K051222nar01)

(179)   *Samma hatthu$_{NUM}$=na mas-aada.*
         all   one=DAT   must-exist
         'We must all go together.'

(180)   *Se=dang kalu suda bannyak thàrà-thaau **inni**_DEIC=pe atthas*
        1s=DAT if   thus much   NEG-know   PROX=POSS about
        mà-biilang=nang.
        INF-say=DAT
        'So, as for me, I cannot tell you much about this.' (K051205nar04)

A relator noun ⊡→5.2.2, p. 199 can be used instead of a pure postposition (cf.
Adelaar 2005). The use of the possessive marker =*pe* between the host and the relator
noun is optional.

(181) | NP(=pe) RELN=POSTP |_PP

(182)   *Andare [[hathu pohong]_NP=**pe baava**=ka]_PP kapang-duuduk.*
        Andare INDEF   tree=POSS   bottom=LOC when-sit
        'When Andare sat down below a tree.' (K070000wrt03)

(183)   *[[Ini pohong]_NP=∅ **atthas**=ka]_NP moonyeth hathu kavanan su-aada.*
        PROX tree   top=LOC   monkey INDEF group PAST-exist
        'On top of this tree was a group of monkeys.' (K070000wrt01)

The relator noun can also attach to clausal NPs (184).

(184)   *[[Mlaayu pada anà-dhaathang]_CLS]_NP=**pe atthas** se=dang hatthu=le*
        Malay   PL PAST-come=POSS   about 1s=DAT INDEF=ADDIT
        mà-biilang thàràboole.
        INF-say   cannot
        'I cannot tell you anything about the coming of the Malays.' (K081105eli02)

Chapter 10

Predicates

Having discussed NPs and PPs, we now turn to different types of predicates, which can take NPs or PPs as arguments. While English basically has only two different predicate constructions, the verbal one and the non-verbal one supported by the copula *to be*, six types can be distinguished in SLM, with some further subdivisions. These are:

1. The verbal predicate ⊡→10.1, p. 461,

2. The existential predicate ⊡→10.2, p. 468

3. The modal predicate ⊡→10.3, p. 471

4. The nominal predicate ⊡→10.4, p. 472

5. The circumstantial predicate ⊡→10.5, p. 474

6. The adjectival predicate ⊡→10.6, p. 477

The different predicate types can be identified by the way they are negated, as shown in Figure 10.1. The differentiation between main type and subtypes is discussed in the respective sections themselves.

10.1 Verbal predicates

Verbal predicates are very frequent in SLM. They consist of a verb (or a converted adjective), typically marked for TAM.[3] This structure is given in (1). An illustrative

---

[1]Identificational predicates always have generic time reference. It is impossible for my father's brother to be the same person as my uncle at point A and no longer to be the same person at point B.

[2]Circumstantial predicates which are true in the present but will no longer be true in the future are difficult to conceive, and for those, there is no information available.

[3]Although TAM-marking is optional (Slomanson (cf. 2006:143), Smith & Paauw (2006:169)).

	PAST	PERFECT	PRESENT	FUTURE
VERBAL	thàrà-V	V thraa	thama-V	thama-V
EXISTENTIAL	thraa(da)	thraa(da)	thraa	thama-aada
MODAL				
boole	thàrboole	thàrboole	thàrboole	thàrboole
maau	thussa	thussa	thussa	thussa
kamauvan	thàrkamauvan	thàrkamauvan	thàrkamauvan	thàrkamauvan
NOMINAL	bukang	bukang	bukang	bukang
				thama-jaadi
ident.[1]			bukang	
CIRC.[2]	bukang	bukang	bukang	?
LOCATION	thraa	thraa	thraa	thama-aada
ADJECTIVE				
ADJ1	ADJ thraa	ADJ thraa	ADJ thraa	TAM-ADJ
ADJ2	thàrà-ADJ	thàrà-ADJ	thàrà-ADJ	TAM-ADJ

Table 10.1: Negation of different types of predicative phrases

example is (2). More detailed discussion can be found below ⊡→10.1.1, p. 463.

(1) | TAM-V | *VPRED*

(2) *Ini gaaja   **su-pii**.*
     PROX elephant PAST-go
     'This elephant left.' (B060115nar05)

It is also possible to use two verbs in one verbal predicate. Two cases can be distinguished and will be discussed below: combination of a verb with a vector verb (4) ⊡→10.1.2, p. 466 and combinations of two full verbs (5) ⊡→10.1.3, p. 466.

(3) | TAM-V { V / VectorV } | *VPRED*

(4) *Itthukapang=jo derang **nya-thaau**fullverb **ambel**vectorverb derang pada*
     then=EMPH   3PL   PAST-know   take   3PL   PL
     politic=nang suuka katha.
     politic=DAT like QUOT
     'Only then will they come to know that they like politics' (K051206nar12)

(5) *Hathu haari, hathu oorang thoppi mà-juval=nang kampong=dering*
    INDEF day   INDEF man   hat   INF-sell=DAT   village=ABL
    *kampong=nang **su-jaalang**$_{fullverb}$ **pii**$_{fullverb}$.*
    village=DAT   PAST-walk   go
    'One day, a man walked from village to village to sell hats.' (K070000wrt01)

10.1.1 Standard verbal predicate

In all verbal predicates, the verb can be modified by an adjective (7), a postpositional phrase with =*nang* (8)(9) or =*ka* (10)(11), a comparative with =*ke* (12) or a secondary predicate with a reduplicated verb (13). This is formalized in (6).

(6)
$$
\left[
\begin{Bmatrix}
\text{ADJ} \\
\text{NP} = nang \\
\text{NP} = ka \\
\text{CLS} = ke \\
\text{V}_i \sim \text{V}_i
\end{Bmatrix}
\text{TAM-V}
\right]_{VPRED}
$$

(7) *Kithang **baaye** mlaayu arà-oomong   katha incayang biilang thraa.*
    1PL   good Malay NON.PAST-speak QUOT 3S.POLITE say   NEG
    'He has not said that we speak good Malay.' (B060115prs15)

(8) *Bras=iyang **baaye=nang cuuci**.*
    raw.rice=ACC good=DAT   wash
    'Wash the rice well!' (K060103rec01)

(9) a. *Luu   **baaye=nang** masà-blaajar **baaye=nang** masà-mnaaji.*
        2S.FAMILIAR good=DAT   must-learn   good=DAT   must-recite
        'You have to learn well and you have to recite well.'

    b. *Lu=ppe umma-baapa=nang **baaye=nang** masà-kaasi thaangang.*
        2S=POSS mother-father=DAT good=DAT   must-give hand
        'You must lend a hand to your parents.' (K060116sng01)

(10) *Itthu vakthu=ka hathu bìssar beecek caaya Buruan **mlaarath=ka***
      DIST time=LOC INDEF big   brown colour bear   difficulty=LOC
      uuthang=dering luvar=nang su-dhaathang.
      forest=ABL   outside=DAT PAST-come
      'Then, a big brown bear came out of the woods with difficulties.'
      (K070000wrt04)

(11) *Aajuth thaakuth=**ka** su-naangis,*
      dwarf fear=LOC   PAST-weep
      'The dwarf wept in fear.' (K070000wrt04)

(12) *Itthu pada [sraathus binthan pada arà-kiilap]_{CLS}=**ke** su-kiilap.*
      DIST PL 100   star   PL   SIMULT-shine=SIMIL PAST-shine
      'They shone like a hundred stars.' (K070000wrt04,K081105eli02)

(13) *Kancil **lompath˜lompath** arà-laari.*
      rabbit jump˜RED   NON.PAST-run
      'The rabbit runs away jumping.' (K081104eli06)

The full structure of the verbal predicate phrase, combining (1) and (6) in then as follows:

(14) $\left[\left\{ \begin{array}{l} \text{ADJ} \\ \text{NP} = nang \\ \text{NP} = \text{ka} \\ \text{CLS} = \text{ke} \\ \text{V}_i \sim \text{V}_i \end{array} \right\} \text{TAM-V} \left(\begin{array}{l} \text{V} \\ \text{VectorV} \end{array} \right) \right]_{VPRED}$

Verbal predicates take between 0 and 3 arguments. Zero-place verbs are represented by *uujang* 'rain' in (15). A monovalent verb is shown in (16), a bivalent verb in (17). The trivalent verb *kaasi* 'give' is shown in (18). (19) shows a verb with four participants, but their status as arguments might be debatable. As discussed in ⊡→11, p. 481 and ⊡→12, p. 493, the distinction between arguments and adjuncts is not easy to make in SLM and might not be relevant at all.

(15) Ø *anthi-uujang.*
      IRR-rain
      'It will rain.' (K081104eli06)

(16) *[Mà-poothong oorang pada]_{NP1} su-dhaathang.*
    INF-cut   man   PL   PAST-come
    'The cutters came.' (K051205nar05)

(17) *Itthusubbath, [deram pada]_{NP1} [jaalang]_{NP2} arà-kijja.*
    therefore   3PL   PL   street   NON.PAST-make
    'Therefore, they build the street.' (K051222nar04)

(18) *[Se=ppe baapa]_{NP1} [incayang]_{NP2}=nang [ummas]_{NP3} su-kaasi.*
    1s=POSS father   3s.POLITE=DAT   gold   PAST-give
    'My father gave him gold.' (K070000wrt04)

(19) *[Itthu baathu]_{NP1}=yang [incayang]_{NP2} [Seelong]_{NP3}=dering [laayeng*
    DIST   stone=ACC   3s.POLITE   Ceylon=ABL   other
    nigiri]_{NP4}=nang asà-baapi.
    country=DAT   CP-bring
    'These stones, he brought them from Ceylon to other countries.'
    (K060103nar01)

Verbal predicates are negated by the prefix *thàrà-* in the past (20), postverbal *thraa* in the perfect (21) and the quasi-prefix *thama-* in non-past tenses (22).

(20) *Ikkang Seelon=ka kitham=nang gaaji **thàrà-**sampe.*
    then   Ceylon=LOC 1PL=DAT   salary   NEG.PAST-reach
    'Then, the salary was not enough for us in Sri Lanka.' (N060113nar04)

(21) Invitations *daapath **thraa**.*
    invitations get   NEG
    'The invitations had not been received.' (K060116nar11)

(22) *Kitham=pe aanak pada **thama**-oomong.*
    1PL=POSS child   PL   NEG.NONPAST-speak
    'Our children do not speak.' (G051222nar01)

Verbal predicates can be seen in all major clause types in SLM. All main clause constructions and all subordinate clause constructions can contain a verbal predicate. The most ample research on verbal predications to date is found in (Slomanson 2006).

10.1.2   Verbal predicates with a vector verb

Verbal predicates with a vector verb consist of a full verb and a vector verb ⊡→5.1.5, p. 173. The vector verb occurs always after the full verb. Vector verb and full verb are parsed into two different phonological words, which distinguishes them from compound verbs ⊡→7.1.5, p. 403. The vector verb modifies the full verb and adds aspectual meaning to it.

(23)   *Kanabisan=ka=jo duva oorang=le **anà-thaau ambel** [Andare duva*
        last=LOC=EMPH   two   man=ADDIT PAST-know take   Andare two
        oorang=yang=le asà-enco-kang aada] katha.
        man=ACC=ADDIT CP-fool-CAUS   exist QUOT
        'At the very end, both women understood that Andare had fooled both of them.' (K070000wrt05)

In example (23), the two women were at first not aware of the trick being played on them. Finally, they discover it, and they change from a state of ignorance to a state of knowledge. This aspectual information is encoded by the vector verb *ambel*. Note that *ambel* does not have its normal meaning *take* in this context.

10.1.3   Verbal predicates with two full verbs

Verbal predicates with two full verbs are serial verb constructions where none of the participating verbs is a vector verb. This is a lot less frequent than the construction involving vector verbs. Just like the preceding construction, the verbs in this construction are parsed into two phonological words, which distinguishes them from compound verbs: ⊡→7.1.5, p. 403. All the examples of this construction in the corpus deal with some kind of motion.

The first example involves the neutral verb *pii* 'go'. In this context, we are dealing with one event of walking, which is coded by two words, *jaalang* 'walk', indicating the manner, and *pii* 'go' indicating motion away from the center.

(24)   *Hathu haari, hathu oorang thoppi mà-juval=nang kampong=dering*
        INDEF day   INDEF man   hat   INF-sell=DAT   village=ABL
        *kampong=nang su-**jaalang pii**.*
        village=DAT   PAST-walk go
        'One day, a man went and walked from village to village to sell hats.' (K070000wrt01)

Care must be taken to not confound serial verbs of this type with verbs of motion with a purposive clause, such as (25), where the purpose of the movement is indicated by the infinitve marker *mà-*, and additionally by the dative marker *=nang*.

(25) *Derang [dìkkath=ka aada laapang]=nang **mà**-maayeng=**nang** su-pii.*
    3PL   vicinity=LOC exist ground=DAT   INF-play=DAT   PAST-go.
    'They went to the nearby ground to play.' (K070000wrt04)

An example of full verb serialization with a more specific verb, *laari* 'run' is (26).

(26) *Aanak su-**laari kluuling**.*
    child   PAST-run roam
    'The child went astray.' (K061019sng01)

Motion towards the deictic center is also possible in a serial verb construction, as *dhaathang* 'come' shows in example (27) and (28).

(27) *Oorang pada kapang-**laari dhaathang**, ini   daara sgiithu=le   suusu*
    man   PL   when-run   come   PROX blood that.much=ADDIT milk
    su-jaadi.
    PAST-become
    'When people came running, the blood had turned into milk.' (K051220nar01)

(28) *See=yang asà-**caari dhaathang**=apa.*
    1s=ACC   CP-search come=after
    'he came in search of me and ...' (K051213nar06)

An example with the motion verb hidden in more elaborate semantics are found in (29)(30), where the word *baapi* 'take.to.a.place' implies motion.

(29) *Daalang=ka light=le   mà-**ambel baapi** thàràboole.*
    inside=LOC   light=ADDIT INF-take   bring cannot
    'You cannot take a light either.' (K051206nar02)

(30) *Aajuth=yang buurung mà-angkath baapi   su-diyath.*[4]
    dwarf=ACC   bird   INF-lift   take.away PAST-try
    'The bird tried to carry the dwarf away.' (K070000wrt04)

In all the examples given in this section but one, the left border of the serial verb is indicated by a preceding TAM-prefix. The right border is indicated by

- a postposition (28),

---

[4]It can actually be argued that the *baapi diiyath* in (30) involves three verbs because *baapi* 'bring/take' can be segmented in to *baa(wa)* 'bring' and *pii* 'go'.

- a modal (29),

- or the end of the clause (24), (26), (27), (30).

There are other examples where the right border is an argument. These have not been selected as examples here because of the difficulty to ascertain whether the argument forms part of the serial verb construction or not. With postpositions, modals, or the end of the clause, it is clear that they are not part of the serial verb constructions.

Between the left and the right boundary, only verbs are found. There is no intervening material. This distinguishes serial verbs from clause chains involving *asà-* ⊡→13.3, p. 529.

The meaning of the serial verb constructions involving two full verbs is transparent and compositional as can be seen from the examples above. Three of the examples above are intransitive (24)(26)(27), while the others are transitive. For the transitives, both component verbs assign agent to the same referent, but the patient assigned by the transitive component verb is not present in *dhaathang* 'come' in (28). In (29), both verbs *ambel* and *baapi* agree in assigning the agent to a generic entity and the patient to a lamp. There are thus no mismatches between the assignments of the semantic roles to the component verbs. An analogous analysis is possible for *angkath* 'lift' and *baapi* 'take.away' in (30).

Given these facts, full verb serialization in SLM can be analyzed as an instance of 'nuclear juncture' (Foley et al. 1985, Van Valin & Foley 1997). In nuclear serialization, the conjoined verbs take a single set of actor and undergoer arguments and cannot be modified or negated independently, whereas in 'core juncture' (not found in SLM), each verb selects its arguments independently, and the verbs can individually be negated or modified. A final type, 'clause juncture', is distinct from the two types discussed in that identity of core arguments is not required. This type is instantiated by clause chains in SLM ⊡→13.3, p. 529.

10.2 Existential predicate

10.2.1 Standard existential predicate

The second type of predicate is the existential predicate, in which one of the existential verbs ⊡→5.1.2, p. 164 *aada* or *duuduk* is used.

(31) $\boxed{TAM - \left\{ \begin{array}{c} aada \\ duuduk \end{array} \right\}}_{VPRED_{exist}}$

Example (32) illustrates this pattern.

(32) Problem *pada=le **aada**.*
      problem PL=ADDIT exist
      'There are also problems.' (N060113nar04)

Aada cannot take the progressive prefix *arà-*, while *duuduk* can and normally does. Both can be inflected for the remaining tenses. The negation of both is normally simply *thraa* (33)(34)(35).

(33) Malay political party *atthu **thraa**.*
      Malay political party INDEF NEG
      'There is no Malay political party.' (K051206nar12)

(34) *Se=ppe umma **thraa**.*
      1s=POSS mother NEG
      'My mother has passed away.' (B060115prs03)

(35) *kàthaama su-aada,   karang **thraa**.*
      before   PAST-exist, now   NEG
      'Before there was rugby, now there is no rugby' (B060115cvs01)

Duuduk in its existential reading, on the other hand, cannot be negated by *thàrà-duuduk*, because this forces a reading as 'stay' or 'sit'. Simple *thraa* has to be used instead (cf. (34)). So, the interrogative existential sentence in (36) can be formed with *duuduk*, but a negative sentence with *duuduk*, like (37), does not express non-existence, but non-location, i.e. the addressee did indeed happen to exist at that point in time, during the fasting period, but he did not stay at the relevant place. If non-existence has to be expressed, *thraa* has to be used, as in (34), rather than a negation of *duuduk*.

(36) *Sudaara sudaari pada arà-**duuduk**=si?*
      brother sister PL   NON.PAST-exist.ANIM=INTERR
      'Do there exist (=Do you have) any brothers or sisters?' (B060115cvs03)

(37) *Puaasa muusing thàrà-**duuduk**=si,* fasting period=*ka?*
      fasting season NEG.PAST=INTERR   fasting period=LOC
      'You were not here during the fasting period, were you?' (B060115cvs03)

This restriction only holds for the existential reading. For the locational reading, both negations (*thàraduuduk* and *thraa* are possible).

The existential predicate is used for five different, but interrelated functions:

- presentationals (38)

- existence (39)

- availability (40)

- location as a special type of existence (41)

- possession as a special type of existence (42)

(38) *Hathu muusing=ka ... hathu kiccil ruuma su-**aada***
       INDEF time=LOC   ... INDEF small house PAST-exist
       'Once upon a time, there was a small house.' (K07000wrt04)

(39) *Karang Kluum̆bu=ka   mlaayu pada **aada**.*
       now   Colombo=LOC Malay PL   exist.ANIM
       'Now, there are many Malays in Colombo.' (N060113nar01)

(40) *Itthu blaakang ini   karang **santham**   **aada**, bukang.*
       DIST after   PROX now   coconut.milk exist TAG
       'Now there is this coconut milk, you know.' (B060115rcp02)

(41) *Se=ppe dhaatha=pe   thiiga aanak=le   **Dubai=ka** arà-**duuduk**.*
       1s=POSS elder.sister=POSS three child=ADDIT Dubai=LOC NON.PAST-exist.ANIM
       'My sister's three children also live in Dubai.' (B060115prs21)

(42) *Thiiga klaaki aade=le,   atthu pompang aade=le*
       three male   younger.sibling one   female   younger.sibling=ADDIT
       ***se=dang arà-duuduk**.*
       1s.DAT   NON.PAST-exist.ANIM
       'I have three younger brothers and one younger sister, too.' (K060108nar01)

10.2.2   Possessive predicate

The possessive predicate construction consists of an existential predicate with additional indication 'to whom' the entity referred to exists, i.e. who possesses it ⊙→15.11, p. 668. Two types of possessive relationship are distinguished: permanent and temporary. In case of permanent possession, the possessor is marked with =*nang* (43) (Ansaldo 2005b:26), in case of temporary possession, the possessor is marked with =*ka* (44). The choice of the existential verb (*aada/duuduk*) is the same as for the general existential construction.

(43) ***Se=dang*** *liima anak klaaki pada **aada**.*
      1s.DAT   five   child male   PL   exist
      'I have five sons.' (permanent) (K060108nar02)

(44) ***Incayang=ka*** *... bìssar beecek caaya hathu* bag ***su-aada**.*
      3s.POLITE=LOC ... big   mud   colour INDEF bag PAST-exist
      'He had a big brown bag with him.' (temporary) (K070000wrt04)

This predicate type can be used in both main clauses and subordinate clauses, as the following example shows, where the location of the money in the bank is expressed in a subordinate clause, and the possession of the money by Mr Yusuf is expressed in the main clause, both clauses making use of *aada*.

(45) *Mr. Yusuf karang incayang=nang$_i$ ini   [[[private bank]=ka$_j$ aada$_j$]*
      Mr Yusuf now   3s.POLITE=DAT   PROX private   bank=LOC exist
      duvith]$_{i,j}$ pada aada$_i$, bukang
      money   PL   exist, tag
      'Now, Mr Yusuf, he had the money that was at that private bank, isn't it?'
      (K060116nar09)

10.3 Modal predicate

A third predicate type is the modal predicate. It consists of a nominal argument or a clause in the infinitive followed by one of the four modal particles *boole* 'can', *thàrboole* 'cannot', *(ka)mau(van)* 'want' and *thussa* 'NEG.want'. The entity for which the modal predicate holds is in the dative case, marked by =*nang*.

(46) $\left\{ \begin{array}{c} \text{NP} \\ \left\{ \begin{array}{c} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * m\grave{a} - \text{V} \end{array} \right\} \left\{ \begin{array}{c} boole \\ th\grave{a}rboole \\ (th\grave{a}r)(ka)mau(van) \\ thussa \end{array} \right\}$ *MODPRED*

(47) shows the use of this predicate type with a nominal complement, (48) with a clausal complement.

(47) *Deran=**nang** [thumpath]$_N$ **maau**.*
      3PL=DAT   place   want
      'They wanted land.' (N060113nar01)

(48) *Derang pada=**nang** [atthu=le   **mà**-kijja]*CLS*=nang **thàràboole**.*
      3pl   PL=DAT   one=ADDIT INF-make=DAT   cannot
      'They could not do a single thing.' (N060113nar01)

This predicate type can be used in main clauses as above, but also in relative clauses as in (49).

(49) *[Kitha=nam boole] mosthor=jo.*
      1PL=DAT   can   way=EMPH
      'The way we can do it' (B060115prs17)

The proclitic variants of the modal particles (⊡→5.12, p. 246) are not modal predicates; they are treated as verbal inflection and are discussed with the verbal predicates ⊡→10.1, p. 461.[5]

10.4   Nominal predicates

The fourth type of predicate is the nominal predicate. Two subtypes can be distinguished, which differ in the information structure context in which they are used: the ascriptive nominal predicate ⊡→10.4.1, p. 472 and the identificational/equational nominal predicate ⊡→10.4.2, p. 474 (cf. Hengeveld 1992:104f).[6]

10.4.1   Ascriptive nominal predicate

The ascriptive nominal predicate (ANP) consist of a predicate NP and an argument NP. It ascribes membership in the class of the predicate NPs to the argument NP. Normally, argument and predicate are simply juxtaposed, but it is also possible to use the copula *asdhaathang(pa)*.

If the argument NP is in the singular, the use of the indefiniteness marker *atthu* is obligatory for the predicate NP (51)-(53). This is not the case for the plural.

(50) ┌─────────────────────────────┐
      │ (COPULA) (*hatthu*) NP*pred* │ *NOMPRED*ascr*
      └─────────────────────────────┘

---

[5]See Noonan (1985:128) for a discussion of general difficulties in establishing the predicate status of modals.

[6]I use "equational" here instead of "equative" used by Hengeveld, which I reserve for comparative contexts.

(51) *[Andare katha arà-biilang ∅] raaja mliiga=ka [**hathu oorang***
Andare QUOT NON.PAST-say] king palace=LOC INDEF   man
***koocak]**pred.*
joke
'The man called Andare was king's jester at the royal palace.'
(K070000wrt05)

(52) *Itthu [**bannyak laama hathu ruuma**pred].*
DIST very   old   INDEF house
'That one was a very old house.' (K070000wrt04)

(53) *Se asdhaathang [**hathu butthul** moderate *Muslim atthu*]pred.*
1s COPULA   one   very   moderate Muslim one
'As for me, I am a very moderate Muslim.' (K051206nar18)

In the plural, nominal predicates do not carry markers of indefiniteness.

(54) *Itthukapang se=ppe baapa, se=ppe kaake,   kaake=pe   baapa,*
then   1s=POSS father 1s=POSS grandfather grandfather=POSS father
kithang samma oorang [Seelong=pe oorang pada]pred.
1PL   all   man   Ceylon=POSS man   PL
'Then my father, my grandfather, grandfather's father, all of us are Ceylonese'
(K060108nar02,K081105eli02)

Nominal predicates are negated by *bukang* in all tenses (55). The use of *hatthu* is
optional in negative contexts. Where the affirmative predicate *Muslim* is marked with
hatthu in (56), the negated predicate *mlaayu* 'Malay' is not.

(55) *Incayang (hatthu) doktor bukang.*
3s.POLITE INDEF   doctor NEG.NONV
'He is/was not a doctor.' (K081105eli02)

(56) *Sindbad* the sailor *hatthu Muslim, mlaayu bukang.*
Sindbad the sailor INDEF   Muslim, Malay NEG.NONV
'Sindbad the Sailor was a Moor, he was not a Malay.' (K060103nar01)

There is no grammaticalized way to mark TAM on nominal predicates ⊙→15.5,
p. 613. Either a lexical solution must be employed, as in (57), or a construction in-
volving *jaadi* 'become' (58).

(57) *Itthu duuva bergaada=jo **kàthaama** oorang ikkang.*
      DIST two   group=EMPH earlier   man   fish
      'These two groups were fishermen.' (K060108nar02)

(58) *Ini   kittham=pe nigiri   su-**jaadi**.*
      PROX 1PL=POSS   country PAST-become
      'This (country) became our country.' (K0512222nar04)

Negated nominal predicates can express temporal reference lexically, like *innam blaakang* 'from now onwards' in (59). The negator is still *bukang* and does not change with regard to the normal negation of nominal predicates.

(59) *See innam   blaakang, hatthu aanak bukang.*
      1s PROX.DAT after   INDEF child NEG.NONV
      'I will never be a child again.' (K081106eli01)

10.4.2   Equational predicate

The equational predicate asserts the identity of the referents designated by two NPs (X = Y). Both NPs are definite. This means that the indefiniteness marker *hatthu* is never employed in this predicate type. In return, one of the referents is very often marked with the emphatic clitic =*jo* (60). Also, the copula *asàdhaathang* is often present (61). This predicate type is often found with kin.

(60) *Suda [itthu kaake=pe   aade=pe   aanak]$_X$=**jo** [baapa]$_Y$.*
      thus DIST grandfather=POSS younger.sibling=POSS child=EMPH father
      'So that grandfather's younger sister's child is my father.' (K051205nar05)

(61) *[Baapa=pe umma]$_X$ **asàdhaathang** [kaake=pe   aade]$_Y$.*
      father=POSS mother COPULA   grandfather=POSS younger.sibling
      'My paternal grandmother was my grandfather's younger sister.'
      (K051205nar05)

10.5   Circumstantial predicate

10.5.1   Standard circumstantial predicate

The fifth predicate type is the circumstantial predicate. It consists of a postpositional phrase.

(62) $\boxed{\text{PP}}$ *CIRCPRED*

It asserts that the two referents are in a relation expressed by the postposition. Very frequent are local postpositional phrases with =*dering* 'from' (63), =*ka* 'at' (64), and =*nang* 'to' (66), but other circumstantials are also possible. The locational type with =*ka* will be discussed more extensively below.

(63)   *Duuva thiiga Kluum̆bu=deri.*
        two   three   Colombo=ABL
        'Two or three are from Colombo.' (K051206nar13)

(64)   a. *Hatthu aade   luvar   nigiri=ka.*
            one   younger.sibling outside country=LOC
            'One younger sibling is abroad,'

        b. *Hatthu aade   Suisse=ka.*
            one   younger.sibiling Suisse=LOC
            'one younger sibling is at Hotel Suisse,'

        c. *Hatthu aade   HNB=ka   arà-bagijja.*
            one   younger.sibling HNB=LOC NON.PAST-work
            'one younger sibling works at Hatton National Bank.' (K061019prs01)

Ansaldo (2008:27) has a nice example of this type (his orthography).

(65)   *Ini buk go-riŋ lu-daŋ.*
        this book I-instr you-DAT
        'This book is from me to you.' (Ansaldo 2008:27(14))

While there is no instance of a circumstantial predicate with the dative in the corpus, Ansaldo's example would be perfectly fine in the Upcountry dialect.

(66)   *Ini   buk   see=dering lorang=nang.*
        PROX book 1s=ABL   2PL=DAT
        'This book is from me to you.' (K081118eli01)

Besides the circumstantial predicates with a spatial meaning, other postpositions can also be used, like the possessive =*pe* in (67).

(67)   *Itthu muusing bannyak* teacher *pada [**Jaapna=pe**]*pred.
        DIST time   many   teacher PL   Jaffna=POSS
        'Back then, many teachers were from Jaffna.' (K051213nar03)

The position of NP and PP can be inverted, as the following example shows. This has consequences for information structure in that (68) is about the town of Gampola and provides new information about it; it is not about Tamils. This contrasts with (63), which is not about the town, but about the people.

(68)   *Gampola=ka bannyak mulbar.*
       Gampola=LOC much   Tamil
       'There are a lot of Hindus in Gampola.' (G051222nar04)

A special case is the use of a relator noun in a circumstantial predicate, as given in (69).

(69)   *Spaaman anà-niiňğgal thumpath=**nang**=le Passara katha arà-biilang*
       3s   PAST-die   place=DAT=ADDIT   Passara QUOT NON.PAST-say
       *nigiri=**nang**=le   dìkkath.*
       country=DAT=ADDIT vicinity
       'The place where he died is close to the village called Passara.'
       (B060115nar05,K081105eli02)

10.5.2   Locational predicate

The locational predicate is the most frequent subtype of the circumstantial predicate. It indicates that the argument is located at the place designated by the predicate. The predicate can be either a noun marked with the locative postposition =*ka* or an adverb, which may or may not bear =*ka*.

(70) $\boxed{\text{NP=}ka}$ *LOCPRED*

(71) $\boxed{\text{DEIC(=}ka\text{)}}$ *LOCPRED*

Example (72) and (73) show a locational predicate with an NP and the locative marker =*ka*. (74) shows the use of an adverb *siini* with and without =*ka*.

(72)   *Kithang **Kandy=ka**=jo.*
       1PL   Kandy=LOC=foc
       'We are from Kandy.' (K051222nar04)

(73)   *Se=ppe kaake   **hathu** estate=**ka**.*
       1s=POSS grandfather INDEF   estate=LOC
       'My grandfather was on an estate.' (K051205nar05)

(74) *Kitham pada **siini**(=**ka**) (arà-duuduk).*
    1PL   PL   here=LOC NON.PAST-exist.ANIM
    'We are here.' (K081104eli06)

Temporal reference need not be expressed on the locational predicate. The following example refers to the past, but this is not coded morphosyntactically.

(75) *Kithang=pe oorang thuuva pada samma **Seelong=ka**.*
    1PL=POSS   man   old   PL   all   Ceylon=LOC
    'Our forefathers were all in Ceylon' (K060108nar02)

This construction cannot only be used for space ⊙→15.4.1, p. 607 but also for time ⊙→15.5.1, p. 614.

(76) *Inni dhaathampa **1987=ka**.*
    PROX COPULA   1987=LOC
    'That happened in 1987.' (K060116nar06)

The negation of this predicate is done just like the negation of the existential predicate, but there are slight differences in meaning which can lead to confusion. In (77), the existential predicate is negated, leading to an interpretation of non-existence, while in (78), the locational predicate is negated, which does not imply that the person is not alive, but rather that she happened to be absent.

(77) *Se=ppe umma **thraa**.*
    1s=POSS mother NEG
    'My mother has passed away.' (B060115prs03)

(78) *Ithu=kapang wife=le   **thraa**.*
    DIST=when   wife=ADDIT NEG
    'At that point in time, my wife was not at home.' (K060116nar04)

10.6 Adjectival predicate

The sixth and final type of predicate is the adjectival predicate. It consists of an adjective, which can be modified by an adverb or a comparative.

(79) $$\left[\left[STD = nang \left\{ \begin{array}{c} libbi \\ liivath \end{array} \right\} \right]_{comp} (ADV)\ ADJ \right]_{ADJPRED}$$

This type has to be distinguished from the use of adjectives as verbal predicates and from the use of adjectives as nominal predicates ⊡→7.4, p. 407. Verbal predicates can be distinguished by their TAM-morphology, while nominal predicates carry *hatthu* in the singular. Nominal predicates in the plural cannot be distinguished from adjectival predicates.

Examples (80)-(83) show typical uses of the adjectival predicate.

(80)   *Dee buthul **jahhath**.*
       3   very   wicked
       'He was very wicked.' (K051205nar02)

(81)   *Samma oorang **baaye**.*
       all   man   good
       'All men are good.' (B060115cvs13)

(82)   *Goppe   aanak pada samma **baaye**.*
       1s.FAMILIAR=POSS child PL   all   good
       'My children are all good.' (B060115cvs13)

(83)   *Skarang biini arà-iingath   puthri **thuuli** katha; Puthri arà-iingath   biini*
       now   wife NON.PAST-think queen deaf   QUOT   queen NON.PAST-think wife
       ***thuuli** katha.*
       deaf   QUOT
       'Now the wife thought the queen was deaf, and the queen thought the wife was deaf.' (K070000wrt05)

(84)   *Suda, inni kaving   **bannyak**_ADV **panthas**_ADJ.*
       thus   PROX wedding much   beautiful
       'So this wedding was very beautiful.' (K060116nar04)

(85)   *Seppe kaaka   se=dang libbi   thiiňǧgi.*
       1s=POSS elder.brother 1s.DAT   remain high
       'My brother is taller than me.' (K081104eli06)

Negation of adjectival predicates depends on the lexeme. Most adjectives are negated by postposed *thraa*, regardless of time reference (86)(87). Some other adjectives are negated by preposed *thàrà-*, again regardless of time reference (88)(89).[7]

---

[7]As for frequencies, this resemble the position of adjectives in the French NP, where most adjectives are postnominal, while there are a few exceptions which are prenominal.

(86)   *Se kaaya thraa.*
       1s rich   NEG
       'I am not rich.' (K081104eli06)

(87)   *Se pàrthaama kaaya thraa.*
       1s earlier   rich   NEG
       'I was not rich.' (K081104eli06)

(88)   *Se=dang **thàrà-sìggar**.*
       1s.DAT   NEG[adj]-healthy
       'I am/was not healthy.' (K081104eli06)

(89)   *Incayang=pe dudukan   hathiyan thaaun=ka=le   laile **thàrà-baae**.*
       3s.POLITE=POSS behaviour other   year=LOC=ADDIT still NEG[adj]-good
       'His behaviour will still be bad even next year.' (K081104eli06)

Adjectives of the second class can only be negated by *thàrà-*. Adjectives of the first class or normally negated by *thraa*, but can be negated by *thàrà-* if they are converted to verbs. In that case, they get dynamic semantics, and do no longer denote a state, but an achievement or an accomplishment. When adjectives of the first class (*thraa*-class) are negated with *thàrà-*, they do get past time reference, and additionally a dynamic reading, which indicates that they are adjectives converted to verbs. Example (86) shows the normal negation of *kaaya* 'rich', with *thraa*, negating a state. In (90), *thàrà-* is used, and it is the event of becoming rich which is negated, not the state (pragmatic implicatures notwithstanding). The event has necessarily temporal reference to the past.

(90)   *Se thàrà-kaaya.*
       1s NEG.PAST[verbal]-rich
       'I did not become rich?*I was not rich/*I am not rich.' (K081104eli06)

The sentence in (90) is thus parallel to (91), with a negated verbal predicate.

(91)   *Incayang nyaari thàrà-dhaathang.*
       3s.POLITE today NEG.PAST[verbal]-come
       'He didn't come today/*He is not coming today.' (K081104eli06)

For negation of adjectival predications referring to the future, the adjective is converted to a verb, and the irrealis verbal negator *thama-* is used.

(92) *Inni pukuran=yang mà-gijja thamau-gampang*
      PROX work=ACC   INF-make NEG.IRR-easy
      'To do that kind of work will never be(come) easy.' (K081106eli01)

Chapter 11

Valency frames and alignment

The different types of predicates discussed above can occur with different numbers of arguments which can be marked for various cases. In the following, we will deal with the valency structure (Bickel 2004), or case frames (Tsunoda 2004), of predicates taking between 1 and 3 arguments.[1]

For predicates taking from one to three arguments, we will see whether their arguments are zero-marked, or whether the predicate assigns a particular case expressed by a certain postposition to its arguments. The postpositions of particular interest here are: *=yang* 'accusative', *=nang* 'dative', and *=dering* 'instrumental'. The instrumental is somewhat peripheral as it can only be used with institutional actors like the police or the government.[2] In this respect, it resembles English institutional plurals (*The police are investigating*) and Dutch feminine agreement for institutions (*Het kabinet en haar beleid* 'The cabinet[neuter] and her[female] policy'). The use of *=dering* on core arguments is therefore limited to a small semantic class of entities and does not form part of SLM core grammar, so that its importance for the SLM alignment system should not be overestimated.[3]

Another issue is the dative marking found on predicates with a modal proclitic on the verb. For instance, we get zero marking in (1), but dative marking in (2), where the modal proclitic *bole=* is present.

(1) *Tony naasi   arà-maakang.*
     Tony cooked.rice NON.PAST-eat
     'Tony eats rice.'

---

[1]There seems to be no clear way to tell arguments from adjuncts in SLM, a property common to other Austronesian languages (Himmelmann 2005:158).

[2]The same holds for Sinhala, see Gair (1991b:14).

[3]Again, see Gair (1991b:14) for a similar argumentation for Sinhala.

(2) *Tony=**nang** naasi   **bole**=maakang.*
     Tony=DAT   cooked.rice can=eat
     'Tony can eat rice.'

The dative case on *Tony* is thus not assigned by the verb *maakang* 'eat', but by the
proclitic *bole*= 'can'.[4] Given the tight integration of the proclitic with its host, it is
difficult to argue that we are dealing with two clauses here. Such a biclausal analysis
is possible for the alternative form with a free word *boole* and the infinitive *mà-* given
in (3).

(3) *Tony=nang naasi   **mà**-maakang boole.*
     Tony=DAT   cooked.rice INF-eat   can
     'Tony can eat rice.'

In (3), we are dealing with a modal predicate *boole*, which assigns dative to *Tony*
and infinitive to *maakang*. In the former case with the proclitic *bole=maakang*, such
a biclausal analysis is problematic. When *bole=* is used, arguments of every predicate
type can take dative marking. We will list these predicates, but as with =*dering*, the
importance of dative assigned by modal proclitics should not be overestimated when
determining the SLM alignment system.

11.1   One-place predicates

One-place predicates are predicates with only one argument. Very often, the argument
does not carry any case marking because the relation between it and the predicate is
clear and there is no need for disambiguation, another argument not being present. The
following two examples show the lack of marking for an actor (4) and an undergoer
(5).

(4) *Itthukapang Tony Hassan=∅ su-pii.*
     then   Tony Hassan   PAST-go
     'Then Tony Hassan left.' (K060116nar09)

(5) *Dee=∅ su-thiidor baava=ka.*
     3s   PAST-sleep down=LOC
     'He slept downstairs.' (K051205nar05)

Psychological and physiological predicates can assign the dative, as in (6), where
the predicate *thàràsìggar* 'sick' assigns dative case to the experiencer of the sickness,

---

[4]See Noonan (1985:128) for general discussion of the problematic case assignment by modals.

go, yielding *godang* (cf. Ansaldo 2005b, 2008, 2009b). The marking of this type of predicate with the dative is typical for South Asia (Masica 1976:159ff).

(6)   *Go=**dang***   *karang bannyak thàràsìggar.*
      1s.familiar=dat now   very   sick
      'I am now very sick.' (B060115nar04)

Institutions take the instrumental =*dering*. This is also possible with one-place predicates.[5]

(7)   Police=***dering** su-dhaathang.*
      police=instr   PAST-come
      'The police came.' (K081105eli02)

When dealing with modal predicates, one place predicates can also mark dative on their argument. This is shown in (8). Normally, *duuduk* 'stay' would not assign dative case to its argument, but when prefixed with *bole=* it does.

(8)   *Kithang=**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)

In some instances, the accusative marker =*yang* can be found as the argument of a one-place predicate.[6]

(9)   Titanic *kappal=**yang** su-thìnggalam.*
      Titanic ship=ACC   PAST-sink
      'The ship "Titanic" sank.' (K081104eli05)

In those cases, care must be taken to check whether we are dealing with a true one-place predicate, or with a two-place predicate where the actor argument is dropped. For instance, the verb *braanak* is used to refer to child birth, and sentences like (10) can be found, where there is only one argument, which is marked with =*yang*.

---

[5]See Gair (2003:791) for institutional actors being marked with the instrumental in Sinhala.

[6]Gair (2003:791) states for Sinhala that 'accusative subjects' imply lack of volition and "that some external force is responsible," which is indeed the case for the Titanic in example (8). In SLM, it is possible to mark inanimate arguments with the accusative in this construction, which is the case of the Titanic. In Sinhala, this is impossible (Gair 1991b:15), but this is due to the general impossibility to mark inanimates with the accusative.

(10) *Itthu=subbath=jo incayang=**yang** siithu anà-braanak.*
      therefore=ᴇᴍᴘʜ   3s.ᴘᴏʟɪᴛᴇ=ᴀᴄᴄ   there ᴘᴀꜱᴛ-braanak
      'Because of that, he was born here'

However, the verb *braanak* is a two-place predicate meaning 'to give birth', as is clear from (11), where both arguments are realized. In (10), and most other cases, the mother is not mentioned, giving the impression that *braanak* takes only one argument.

(11) *Umma see=**yang** su-braanak.*
      mother 1s=ᴀᴄᴄ   past-give.birth
      'My mother gave birth to me.' (K081104eli05)

To return to example (9) about the Titanic, it is fortunately possible to show that the verb *thìnggalam* 'sink' is intransitive because it is one of the rare verbs where the transitive form is suppletive, *cullop*. If an actor causes the sinking of an undergoer, *thìnggalam* cannot be used, but *cullop* must be used instead. This entails that the sentence *Titanic=yang su-thìnggalam* has indeed accusative marking on the only argument of the predicate.

Another instance where we find accusative marking on one-place predicates are existentials. This is not very frequent, and the exact circumstances for this are unclear. Specificity of the theme seems to favour the occurrence of =*yang*.[7] In example (12) =*yang* is optional, but the informants feel that the sentence is better with =*yang*. Note that the hat is definite in (12), indicated by the deictic *itthu*.

(12) *Itthu thoppi=(**yang**) siini aada.*
      ᴅɪꜱᴛ   hat=ᴀᴄᴄ   here exist
      'That hat is here.' (K081106eli01)

Another instance where we find =*yang* is with some adjectival predicates (13)(14). As of now, there is no good explanation for these cases.

(13) *Seelon=pe   duppang muusing=**yang** bannyak panthas anthi-aada.*
      Ceylon=ᴘᴏꜱꜱ future   time=ᴀᴄᴄ   much   beautiful ɪʀʀ-exist
      'Sri Lanka's future will be bright.' (K081104eli05)

---

[7]Specificity is also seen as an important factor in Ansaldo (2005b), but only as far as the occurrence of =*yang* on 'objects' is concerned. The examples given below would probably not qualify as 'objects', but specificity still plays a role in the occurrence of =*yang*.

(14) *Itthu* bag=***yang*** *lorang=nang bannyak bìrrath (liivath).*
     DIST bag=ACC 2PL=DAT   much   heavy more
     'That bag is too heavy for you.' (K08114eli01)

When questioning the location of a specific referent, =*yang* is obligatory, as shown in (15).

(15) *Itthu thoppi=*(**yang**) aapa=ka   aada.*
     DIST hat=ACC   what=inside exist
     'Inside of what is that hat located?' (K081106eli01)

A naturalistic example of =*yang* occurring on the argument of a locational predicate is (16).

(16) Three miles *cara jaau=ka aada [dee   anà-sbuuni duuduk* cave*]=**yang***.
     three miles way far=LOC exist 3s.FAMILIAR PAST-hide   stay   cave=ACC
     'Three miles away from here is the cave where he stayed hidden.'
     (K051206nar02)

To sum up, arguments of monovalent predicates can be marked with zero or the dative. In exceptional cases, they can also be marked with the instrumental or the accusative. If modal proclitics are considered, the proportion of dative marking increases.

11.2 Two-place predicates

With two place predicates, we have to distinguish the actor argument and the undergoer argument. Case marking of these arguments is optional. (17) and (18) give sentences where both arguments lack a postposition.

(17) *Sebastian=∅ saathe=∅ asà-maakang aada=si?*
     Sebastian   satay   CP-eat   exist=INTERR
     'Have you eaten satay, Sebastian?' (B060115cvs02)

(18) *baapa=∅ [...] hatthu pohong=∅ nya-poothong.*
     father   INDEF tree   PAST-cut
     'My father cut a tree in their garden.' (K051205nar05)

While *poothong* 'cut', above, was construed with two non-marked arguments, it is also possible to use the accusative on the undergoer, as shown in (19).

(19) *Ithukapang lorang=pe leher=**yang** kithang=∅ athi-poothong.*
      then   2PL=POSS neck=ACC 1PL   IRR-cut
      'Then we will cut your neck.' (K051213nar06)

Some verbs assign dative case to the undergoer argument, such as *banthu* 'help'
in (20) or *puukul* 'hit' in (21).

(20) *Derang pada=∅ arà-banthu   cinggala raaja=**nang**.*
      3PL   PL   NON.PAST-help Sinhala   king=DAT
      'They help the Sinhalese king.' (K051206nar03)

(21) *Rose-red=∅ buurung=**nang** su-puukul.*
      Rose-red   bird=DAT   PAST-hit
      'Rose-red hit the bird.' (K070000wrt04)

(21) shows one possibility to assign zero and dative to two arguments. The actor
is ∅-marked while the undergoer receives dative marking. Another possibility, which
exists for experiencer verbs, is to mark the experiencer with the dative =*nang* and the
stimulus with zero. This is shown in (22) (and a common South Asian construction
(Masica 1976:159ff)).

(22) *[svaara hatthu]=∅ derang=**nang** su-dìnngar.*
      noise   INDEF   3PL=DAT   PAST-hear
      'They heard a noise.' (K070000wrt04)

Institutional actors take the instrumental as usual. The other argument can be
marked with =*yang* (23), or bear no marking (24).

(23) *See=**yang** police**=dering** nya-preksa.*
      1s=ACC   police=ABL   PAST-enquire
      'I was questioned by the police.' (K051213nar01)

(24) British *Government=**dering*** Malaysia *Indonesia, [inni nigiri   pada]=∅*
      British Government=ABL   Malaysia Indonesia PROX country PL
      samma peegang.
      all   catch
      'The British Government captured Malaysia and Indonesia, those countries.'
      (K051213nar06)

When modal proclitics are considered, the actor can receive dative marking. In

(25), *kithang* 'we' receives dative marking and the theme of reading, *mulbar* 'Tamil', is zero-marked.

(25)   *Kithang=**nang** baaye=nang mulbar=∅ bole=baaca.*
        1PL=DAT   good=DAT   Tamil   can=read
        'We can read Tamil well.' (K051222nar06)

Accusative marking of the undergoer is still possible when a modal proclitic is used. This is found for instance in (26).

(26)   *aathi=**yang** sajja hatthu oorang=**nang** bole=ambel.*
        liver=ACC   only one   man=DAT   can-take
        'Only one person can take the liver.'

When using a verb which normally assigns the dative case to the undergoer a modal proclitic results in both arguments being marked with the dative.

(27)   *Se=**dang** Farook=**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, while the undergoer is the other argument. When using pointing gestures, this can be over-ruled, as in (28).

(28)   *Ini   kaaka=nang   itthu kaaka=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)

To sum up, two-place predicates normally have zero-marked actors, and undergoers are either marked for accusative or dative. In special cases, actors can be marked for instrumental or dative. All case markers can optionally be dropped (Ansaldo 2005b, 2008, 2009b).

11.3   Three-place predicates

With three place arguments, we distinguish an actor argument, an undergoer argument, and a recipient (or goal) argument. The recipient argument is always marked for dative. Dropping of the case marker is not possible there. The actor is normally zero-marked, while the undergoer either takes =*yang* or also zero.

Zero-marking of both actor and undergoer (theme) is shown in (29).

(29) *Se=ppe baapa=∅ incayang=**nang** ummas=∅ su-kaasi.*
      1s=POSS father   3s.POLITE=DAT gold   PAST-give
      'My father gave him gold.' (K070000wrt04)

A more complicated example is (30) with a coordinated noun phrase *jaithanle jaarong pukurjanle* 'sewing and needle work', which functions here as one referent. Again, both actor and theme are zero-marked, but the recipient/beneficiary is marked with the dative.

(30) *Derang=pe umma=∅ derang=**nang** [jaithan=∅=le, jaarong pukurjan=∅=le]*
      3PL=POSS   mother 3PL   sewing=ADDIT needle   work=ADDIT
      su-aajar.
      PAST-teach
      'Their mother taught them sewing and needle work.' (K070000wrt04)

Marking of the undergoer argument with the accusative is shown in (31), where the recipient argument is still marked with the dative.

(31) *Itthu vakthu=ka=jo   Mr Samath=∅=le go=**dang** nya-introduce=king*
      DIST time=LOC=EMPH Mr Samath=ADDIT 1s=DAT   PAST-introduce-CAUS
      *Sebastian=**yang**.*
      Sebastian=ACC
      'At that point in time, Mr Samath introduced Sebastian to me, too.'
      (B060115cvs01)

As before, institutional actors take instrumental marking (32) and modal proclitics assign the dative (33).

(32) Police=***dering*** *see=**yang*** remand=*nang su-kiiring.*
      police=instr   1s=ACC   remand=DAT   PAST-send
      'The police sent me into custody.' (K081105eli02)

(33) *Kithang=**nang** miskin pada=nang duvith bole=kaasi.*
      1PL=DAT   poor PL=DAT   money can=give
      'We can give money to the poor.' (K081104eli05)

To sum up, three place predicates normally mark the actor with zero, the undergoer with the accusative, and the recipient with the dative. As usual, institutional actors take the instrumental, and modal proclitics assign the dative. All case markers can be

dropped with the exception of the dative marking of the recipient.

11.4   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:[8]

- The dative marker can be found on R and P. Additionally, it can be found on S and A if they are experiencers. Furthermore, modal proclitics can assign the dative to S or A

- The accusative marker can be found on P and in rare instances on S

- The instrumental marker can be found on S and A when they are institutional

- Zero can be found on S, A and P. Zero is never found on R.

This distribution can be illustrated as in Figure 11.1.

If we interpret this in terms of a theory of alignment, we find characteristics of accusative alignment and split-S alignment, as well as neutral alignment.[9] The role of the dative cannot be fit into any one of the standard types of alignment. Discussing these similarities in turn, we can say that the SLM system resembles accusative alignment if we discard the peripheral instances of institutional instrumentals and the rare cases of =*yang* attaching to S. In this case, S and A are always zero, while P is marked with =*yang*. We would thus be dealing with a conflation of S and A, a nominative-accusative system. This ignores the dative marking for modals and experiencers, but this need not be a problem, since other languages with 'dative subjects', like German for instance, are also commonly called nominative-accusative languages.

On the other hand, if we think that the institutional actors, and especially the instances of =*yang* found on S are indeed relevant, we find resemblances to a Split-S system. S is marked like A (zero or instrumental) if it is actor-like; if S is undergoer-like, it is marked like P (accusative).

When we turn to the most common case, the dative, we find resemblances to the neutral system in the sense that every argument *can* be marked with the dative. In a neutral system, there is no distinction between the arguments. However, in SLM, dative marking is normally not done simultaneously on several arguments. It is rare that more than one argument is marked with the dative, so that it is possible to distinguish the semantic roles of the arguments, which is not what we find in a neutral system.

To sum up, the SLM system shows resemblances to a number of alignment systems, but is not a clear instance of any of them. This might have to do with the fact that the traditional alignment systems only take into account a mapping of three roles

---

[8]The Sinhala facts are very similar to this. This was found out only after writing this chapter. See Gair (1976a, 1991b), Henadeerage (2002).

[9]See a similar discussion in Gair (1991b:19f) for Sinhala.
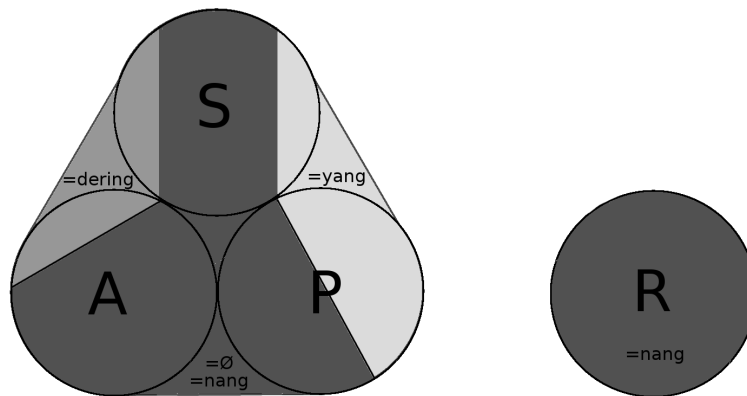
Figure 11.1: The coding of semantic roles in SLM. The instrumental =*dering* can code S and A, but is marginal, as seen by the small portion of the circles it covers. The accusative =*yang* can be used for P, where it is common, and for S, where it is marginal. This difference in frequency is represented by the different sizes of the covered surface. The dative marker =*nang* can be used for S, A and P, as can zero. For expository reasons, dative and zero are not distinguished in the illustration. Zero is much more frequent than the dative in all three roles. R, finally, can only be marked by the dative.

on two cases (the tripartite system being somehow only of theoretical relevance). In SLM, we have more than two markings that are relevant: accusative, dative, zero and possibly the instrumental. These do not fit nicely in the classical S-A-P triangle (cf. Figure 11.1), but rather cut across all of S, A and P. It appears that the three-circles model is thus not very appropriate for SLM. Rather, SLM seems to have a special kind of *semantic alignment* (Donohue & Wichmann 2008) or *role based alignment* (Hengeveld & Mackenzie 2008:328), which goes beyond the classical active-stative-split (Klimov 1974).

One interesting thing to note about the dative is that it normally occurs in contexts with low volitionality (experiencer, modality, recipient). There might be a possibility to interpret the accusative as [+affected], the instrumental as [+active] and the dative as [-volitional] (Ansaldo 2005b). Arguments are marked for this if disambiguation requires it; otherwise, zero can be used. Crucially, the number of arguments of a predicate does not seem to influence the choice of a marker. The syntactic relevance of valency of predicates is thus limited, and it can indeed be argued that in SLM, predicates do not subcategorize for the number of arguments they take. The speaker is free to add as many referents to the predicate as he sees fit, and to suppress as many referents as he thinks will not impede understanding. Any referent can be dropped ⊡→12.4, p. 498, and the semantic role of any referent can be indicated by appropriate postpositions. In that sense, the four referents in (34) are all of the same status. There is not one which would enjoy a privileged status over another in the way in which arguments are privileged over adjuncts.

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

Chapter 12

Grammatical relations and alignment

While in English and many other languages, there is one argument which is more prominent than the other ones, the subject, this is not the case in SLM (cf. Ansaldo 2005b:16). It is very difficult to point out criteria to distinguish the syntactic prominence of the arguments in a clause.[1] Commonly used tests like agreement, word order, case marking, relative clauses or realization of pronouns do not yield any systematic differences. Furthermore, there is no operation which would change the status of an argument from more prominent to less prominent, or from less prominent to more prominent, as would be the case for passivization in English.[2] After English passivization, the formerly more prominent argument is less prominent, and the less prominent argument is more prominent. This lack of a difference in prominence between the arguments is easy to formulate from a descriptive point of view, but might be more difficult to incorporate into some current linguistic theories. This grammatical description being theory-neutral, the theoretical preferences of some formalism do not matter in principle, yet given the importance of 'subject' in most of linguistic theorizing, it is important to provide the empirical groundings on which the absence of this category was arrived at.

Himmelmann (2005:152) lists the following criteria commonly used in Western Austronesian languages to establish grammatical relations (also see (Keenan 1976, Schachter 1996)).

- case marking ⊡→12.1, p. 495

- agreement ⊡→12.2, p. 495

---

[1] This is common in the adstrates as well, as documented at least for Sinhala in Gair (1976a, 1991b) and Henadeerage (2002).

[2] Compare Gair (1991b:19)for the similar Sinhala facts.

- word order ⊡→12.3, p. 495

- pro-drop ⊡→12.4, p. 498

- conjunction reduction ⊡→12.5, p. 500

- control ⊡→12.6, p. 501

- resumptive pronouns ⊡→12.7, p. 502

- passivization ⊡→12.8, p. 502

- binding ⊡→12.9, p. 503

- raising ⊡→12.10, p. 503

- obligatoriness ⊡→12.11, p. 504

- relativization ⊡→12.12, p. 504

Following Van Valin & Foley (1997:250ff), 'subjects'[3] must be instances of 'restricted neutralizations'. 'Neutralization' means that two or more argument receive an identical treatment in syntax although they have different semantic roles. An example is agreement in English. English verbs agree with the actor of transitive verbs, and the only argument of intransitive verbs. The semantic distinctions between an intransitive undergoer and a transitive actor are ignored as far as agreement is concerned. This ignoring of semantic differences is what Van Valin & Foley (1997) call 'Neutralization'. 'Restriction' means that the neutralization must not be generalized: some things must be excluded from the neutralization. If all actors and all undergoers in all transitive and all intransitive sentences were treated alike, we would indeed deal with neutralization, but it would not be restricted. The agreement pattern of English is restricted, since undergoers of transitive verbs are excluded from triggering agreement.[4]

As far as SLM is concerned, for the majority of the domains named above, we are dealing with unrestricted neutralization, i.e. all semantic roles are treated alike. Cases in point are agreement, word order, pro-drop, conjunction reduction, obligatoriness, and relativization. Some other domains cannot be applied because the relevant morphosyntactic operation does not exist in SLM. This applies to passivization, binding, and raising. As argued for in Section 11, case marking does not neutralize semantic distinctions either. The only domain where we might possibly dealing with restrictive neutralization is 'control', but even there, the case is far from clear. The different domains will now be discussed in turn.

---

[3]Researchers like Dixon (1979, 1994) would argue that the topic of this discussion is 'pivot' rather than 'subject'. Following Dixon's definition, this is indeed the case. However, Dixon's definition of subject is so narrow that it is difficult to say anything about 'subjects' which is not trivial, making 'subject' a nearly useless category in language description in general. This does not seem to be a consensus position; many researchers (e.g. Comrie 1981:104ff) keep on using the term 'subject' for what Dixon would call 'pivot', and so will I. I add the caveat that where I use 'subject', Dixon's 'pivot' is intended, except for my discussion of control in SLM, which covers Dixon's 'subject' and, possibly, 'pivot'.

[4]This can be changed through operations like passivization, though.

12.1 Case marking

In many languages, the subject of a transitive sentence is case-marked identically to the only argument of an intransitive sentence. In German, the A argument and the S argument are both marked for nominative, in Basque, the P argument and the S argument are both marked for absolute. As discussed in ☐→11.4, p. 489, SLM cases are assigned on a semantic basis, and the syntactic nature and valency of the arguments' predicate matters little. The only argument of an intransitive predicate (S) can be marked for dative, instrumental, accusative, or zero. It is thus not possible to decide which one of the two arguments of a transitive predication is case-marked like S. Under a strict view, none of the arguments of the transitive clause has the exactly the same case-marking distribution as S, since A is never marked for accusative, and P is never marked for instrumental. Under a more loose view, both A and P are marked like S in that they can take dative and zero marking. Case marking thus does not provide a good criterion to decide which one of the two arguments of a transitive clause is more like the only argument of an intransitive clause, in other words, 'subject' cannot be determined based on the criterion of case marking.

12.2 Agreement

Another feature which can be used to determine the subject is the agreement pattern. If only one argument triggers agreement on the verb, this is the subject. In SLM, there is no generalized agreement, and the cases of agreement found so far ☐→13.9, p. 539 are very weak and only found in some idiolects. For the time being, it seems safe to assume that SLM has no agreement as such, so that this feature cannot be used to determine subjecthood of arguments.

12.3 Word order

In some languages, word order gives cues as to the subjecthood of the argument, e.g. in English, where the subject must precede the verb. Two prominent positions are to be investigated: the position immediately preceding the verb and the position of the first argument in the clause. Furthermore, the distribution has to be controlled for the status of the arguments as nouns or pronouns, since these might behave differently with regard to placement. We will limit the discussion to the position of A. The following example shows that A can be found in a position which is neither the first one, nor the closest to the the verb. In this example, A is pronominal.

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

Example (1) shows that a likely argument for subjecthood, *incayang* '3s.POLITE',
is not found in initial position. It is not adjacent to the verb either. This suggests that
incayang, the topical agent, does not have to occur in any particular position, unlike
English, where *he* would have to occur before *brought*, there is no other possibility in
English.

   A similar situation obtains in (2), where again the pronominal agent is sandwiched
between the other arguments.

(2)   *[Itthu thumpath=yang]ₚ [**incayang**]ₐ [giithu]=jo   [avuliya=nang]=jo*
      DIST   place=ACC   3s.POLITE   like.that=EMPH saint=DAT=EMPH
      su-kaasi.
      PAST-give
      'Like that, he gave that place to the saint.' (K051220nar01)

The above examples treated pronominal arguments. The 'sandwich' position of a
nominal argument is given in (3).

(3)   *[Kithang=pe oorang thuuva pada=yang]ₚ [**Dutch**]ₐ [Seelong=nang]*
      1PL=POSS   man   old   PL=ACC   Dutch   Ceylon=DAT
      anà-aaji.baa.
      PAST-bring.ANIM
      'The Dutch brought our forefathers to Ceylon.' (K060108nar02)

The three examples (1)-(3) show that the A-argument does not have to occur in a
privileged position. This is true of both nominal and pronominal arguments. These
examples also show, that a nominal P-argument can occur in the first position of the
clause. We still have to show, that a pronominal P-argument can occur in the first
position of a clause and that P-arguments can occur in other positions in the clause
than the initial one. Example (4) shows the use of a pronominal P in initial position.

(4)   *[Incayang=yang]ₚ [siaanu]ₐ asà-buunung thaaro=apa.*
      3s.POLITE=ACC   3s.prox   CP-kill   put=after
      'This one has killed him.' (K051220nar01,K081103eli02)

Example (5) shows the use of a nominal P-argument sandwiched between two other arguments, and thus in no privileged position. Example (6) shows the same for a pronominal argument.

(5) *[Aanak pompang duuva=le, derang=pe umma=le]$_A$ [**Buruan=yang**]$_P$*
     child   female   two   3PL=POSS   mother=ADDIT bear=ACC
     [ruuma daalang=nang] su-panggel.
     house inside=DAT   PAST-call
     'The two girls and their mother invited the bear to come inside the house.'
     (K070000wrt04)

(6) *Boole lìkkas=ka [see]$_A$ [**lorang=yang**]$_P$ [mliige=nang] anthi-panggel].*
     can   quick=LOC 1s   2PL=ACC   palace=DAT   IRR=call
     'As soon as possible I will call (the two of) you (girls) to the palace.'
     (K070000wrt04)

Up to now, we have seen that there is no special relative order between the A argument and the P argument. We now turn to arguments marked with the dative. The following two examples show that the relative order of the zero-marked argument and the dative-marked argument is free. (7) has the zero-marked argument in first position, while (8) has the dative occurring first.

(7) *[Se=ppe baapa$_\emptyset$] [incayang=nang]$_{DAT}$ ummas$_\emptyset$ su-kaasi.*
     1s=POSS father   3s.POLITE=DAT   gold   PAST-give
     'My father gave him gold.' (K070000wrt04)

(8) *[Se=dang]$_{DAT}$ laiskali [se=ppe bìnnar mosthor]$_\emptyset$ anà-jaadi.*
     1s=DAT   again 1s=POSS real   way   PAST-become
     'I became my old self again [because the curse was broken].' (K070000wrt04)

Given that the distribution of zero/dative on pronoun/noun is the same in (7) and (8), but the order is different, no privileged position seems to exist.

To sum up, word order does not seem to give a clue as to the relative prominence of arguments in 2+-place-predicates. Word order in SLM is thus pragmatically conditioned as common in South Asia (Bickel 2004).[5]

---

[5]See Bayer (2004:62) and references therein for a more extensive list of preferences influencing word order in languages with free word order.

12.4   Pro-drop

In some languages, the possibility to drop certain arguments but not others can be used as a test for subjecthood. In Standard Spanish, for instance, only subjects can be dropped, but objects can not. If the realization of an argument is optional in Spanish, this is a clear indicator that that argument must be the subject.

SLM differs from Spanish in that any argument can be dropped (Smith et al. 2004).[6] (9) shows the dropping of an agent.

(9) *Itthu=nang blaakang, $\emptyset_A$ dee=yang   su-peegang siini Seelong=ka.*
     DIST=DAT   after   3S.IMPOLITE=ACC PAST-catch here Ceylon=LOC
     'After that (they) caught him here in Ceylon.' (K051206nar02)

The dropping of the agent and the dropping of postpositions are independent of each other. The following example shows a dropped agent and a zero-marked patient/theme, which could easily lead to confusion.

(10) $\emptyset_A$ *Derang=$\emptyset$=le su-baava.*
      3PL= =ADDIT   PAST-bring
      'They were also brought (not: They also brought someone).' (K051213nar06)

From the context, it is clear that the persons referred to as *derang* 'they' are undergoing the action of bringing rather than performing it.

While we have seen dropping of the agent above, the following examples show the dropping of patient or theme. Example (11) is about the speaker and his parents, and the event of being sent to an uncle's house for education. It is clear from cultural knowledge that it is necessarily the parents who were sending the child to the uncle, and not the other way round. This means that the parents are agent and the speaker in his childhood is theme, but it is the speaker that is dropped.

(11) *Hathu thaaun [se=ppe umma-baapa] $\emptyset_{theme}$ [se=ppe maama=pe*
      INDEF   year   1s=POSS mother-father   1s=POSS uncle=POSS
      ruuma=nang] su-kiiring.
      house=DAT   PAST-send
      'One year my parents sent me to my uncle's house.' (K051213nar02)

This speaker's educational career continues and he is admitted to college in (12). Again, the agent is mentioned (*derang* 'they') but not the patient/theme.

---

[6]This is common in both western Austronesian (Himmelmann 2005:171), Sinhala (Gair 2003:813) and Tamil (Lehmann 1989:367).

(12) *S-riibu   sbiilan raathus ùmpathpulu thuuju thaaun, derang $\emptyset_{pat/theme}$*
one-thousand nine   hundred four-ty   seven year   3PL
*Badulla=ka   Dharmadhuutha College=nang anà-*admit-*kang.*
Badulla=LOC Dharmadhuutha College=DAT   PAST-admit-CAUS
'They admitted me to Dharmadhuutha college in Badulla in 1947.'
(K051213nar02)

Interestingly, the same speaker also uttered a sentence (13) where the agent is dropped and the theme realized, but the theme is not case marked (cf. discussion of (10) above).

(13) *[Mà-blaajar=nang] $\emptyset_A$ see=$\emptyset_{theme}$ anà-kiiring [se=ppe maama hatthu=pe*
INF-learn=DAT   1s   PAST-send   1s=POSS uncle   INDEF=POSS
ruuma=nang].
house=DAT
'I was sent to an uncle of mine's to study.' (K051213nar02)

Simultaneous dropping of A and P is also possible, as shown in (14), where both *kithang* 'we' and *aanak pada* 'children' are dropped in the final clause.

(14) a. *Luvar nigiri   [kithang]$_i$=nang mà-pii thàrà-suuka.*
       outside country 1PL=DAT   INF-go NEG-like
       'We do not want to go abroad.'
   b. *Nni nigiri=ka=jo   [kitham=pe aanak buva pada]$_j$=yang*
       PROX country=LOC=emph 1PL=POSS   child fruit PL=ACC
       asà-simpang,
       CP-keep
       'We have raised our children in this country and'
   c. *inni schools pada=nang $\emptyset_i$ $\emptyset_j$ asà-kiiring.*
       PROX schools PL=DAT   CP-send
       'We have send them to the schools here and ... .' (K051222nar04)

Finally, the following example shows the dropping of all three arguments of the verb *juuval* 'sell', the agent *Sindbad*, the recipient *soojor* 'Europeans' and the theme *lakuan bathu* 'gems'.

(15)   a. *Incayang_i [ini Seelong=ka anà-aada lakuan baathu]_j asà-caari,*
            3s.POLITE PROX Seelon=LOC PAST-exist wealth stone   CP-find
            'He was looking for the gems present in Ceylon and'

      b. *[soojor   pada]_k [incayang=sasaama]_i [Seelong=nang]*
            European PL   3s.POLITE=COMIT   Ceylon=DAT
            asà-dhaathang,
            CP-come
            'the Europeans came to Sri Lanka together with him and'

      c. ∅_i ∅_j ∅_k *blaangan arga=nang   anà-juuval.*
                  amount   expensive=DAT PAST-sell
            '(he) sold (them) (the stones) for an expensive price' (K060103nar01)

To sum up, dropping of arguments is not influenced by their syntactic status. Dropping cannot be used to distinguish between more and less prominent arguments.

12.5 Conjunction reduction

Related to the topic of pro-drop is the phenomenon of conjunction reduction. If in the non-initial part of coordinated sentences, an argument can be dropped it must be the subject. This test can be used in English for example, where the sentence *Mary hit John and ∅ ran away* can only mean that Mary ran away, hence *Mary* is the subject.

The test of conjunction reduction works well with English because English does not allow pro-drop in the first place. As discussed in the preceding section, pro-drop of any argument is wide-spread in SLM, and it is difficult to decide whether a given sentence should be analyzed as pro-drop or conjunction reduction (cf. Himmelmann 2005:17). The following examples show instances of what in English would be conjuction reduction, but it might as well be analyzed as simple pro-drop in SLM, parallel to the examples in the preceding section.

The subordinate clause of example (16) introduces both an agent (*kithang* 'we') and a patient (*mlaayu makanan* 'Malay food'). In the main clause, neither of them is repeated, both are dropped/reduced. Note that in the English translation *it* referring to the food is obligatory while this is not the case in SLM.

(16)   a. *Kithang_i puaasa vakthu=ka [mlaayu makan-an]_j asà-gijja,*
            1PL   fasting time=LOC Malay eat-NMLZR CP-make
            'We make Malay food during the fasting period and' (K061019nar01)

      b. ∅_i ∅_j *ruuma pada=nang arà-kaasi.*
              house PL=DAT   NON.PAST-give
            'give (it) to our neighbours.' (K061019nar01)

A similar case is found in (17), where the fourth and final clause consists of the predicate *arà-juuval,* which takes two arguments, one seller and one produce. The seller *incayang* 'he' is introduced in the first clause, the produce *itthu samma* 'all that' (referring to vegetables mentioned earlier) is introduced in the third clause. Both of them are not overtly realized in the final clause (Note again the obligatory *it* in the English translation).

(17)   a. *Incayang=jo$_i$ asà-pii*
           3S.POLITE=EMPH CP-go
           'He goes and'

       b. *paasar=dang asà-pii*
           shop=DAT   CP-go
           'goes to the shop and'

       c. *itthu samma$_j$ asà-baa*
           DIST all   CP-bring
           'brings all that and'

       d. $\emptyset_i$ $\emptyset_j$ *arà-juuval.*
                NON.PAST-sell
           'sells (it).' (B060115cvs07)

To sum up, conjunction reduction, if it is indeed a useful concept in SLM at all, could not be used to distinguish between A and P and hence does not allow to judge the relative prominence of these arguments.

12.6 Control

The only instance where we can find a glimpse of a privileged argument in SLM are control structures. Whereas, generally, in SLM, a non-expressed argument can have any role, in control structures with *kamauvan* 'want/need', it must be the agent. As a test, the verb *ciong* 'kiss' was taken, because desire can be expressed to be the acting or the undergoing part in the act of kissing, which is more difficult for other verbs, like eating or killing. In a predication 'X wants to be part in the act of kissing', will X be associated with the kisser, the kissed, or either? To test this the context of the fairy tale "The Frog Prince" was taken, where a frog wants to be kissed to regain his former self, a prince. The pragmatic context thus forces an association of the wanter with the undergoer. However, this is not what we find. The sentence as in (18) can only mean that the frog wants to be the actor in the act of kissing.

(18) *Se=dang mà-ciong kamauvan.*
      1s=DAT   INF-kiss   want
      'I want to kiss.' (K081104eli05)

If the pragmatically more appropriate reading of the frog wanting to be the under-goer needs to be expressed, (19) must be used, where an indefinite pronoun *saapake* 'somebody' instantiates the actor. This entails that the frog cannot be the actor any-more, and hence is the undergoer.

(19) *Se=dang$_i$ [saapa=ke $\emptyset_i$ mà-ciong] maau.*
      1s=DAT   who=SIMIL   INF-kiss   want
      'I want that someone kisses me/I want to be kissed.' (K081104eli05)

This shows that *kamauvan* 'want' is an agent-controller in SLM. The wanter is automatically assigned the agent role in the complement clause. This is the case even if the the wanter is assigned dative case in the main clause, and the undergoer (as in the case of *ciong* 'kiss') would also be assigned dative case (*se=dang*). Agent-control is thus a stronger force than the desire to associate arguments which take the same case marking. This control relation need not be syntactic ('subject-control', 'syntactic pivot'), however. It is also possible that we are dealing with an 'actor control' relation ('semantic pivot') (cf. Van Valin & Foley 1997:257).[7]

12.7   Resumptive pronouns

In some languages, resumptive pronouns can be used to identify whether an argument is subject or not. SLM uses no resumptive pronouns, so that this test cannot be applied.

12.8   Passivization

Passivization can be used in some languages to identify arguments. This test is useful if an argument cannot be used for a certain operation before passivization, but after passivization, this is possible. That argument has then been promoted to subject status (Foley & Van Valin 1985:306). As an example, *Mary$_i$ hit John$_j$ and $\emptyset_i$ ran away* can only mean that the hitter (*Mary* in this case) ran away. If we want the victim to run away, we can use passivization to promote the former patient to subject: *John$_i$ was hit by Mary$_j$ and $\emptyset_i$ ran away*.

---

[7]Conflation of S and A in 'want' clauses is actually claimed to be a universal feature of human language (Dixon 1979).

SLM has no such operation. There is the vector verb *kìnna* and the (probably etymologically related prefix) *kànà-*, but they do not change the syntactic status of the arguments. What they do is to add an element of adversity and surprise, but this is semantic and has no syntactic repercussions.

The following two examples show that the argument structure of verbs does not change through the use of *kìnna*. In (20), *kìnna* is present and we find the patient marked with zero and the agent/recipient marked with the dative.

(20) *Sdiikith thaaun=nang duppang see ini   Aajuth=nang su-kìnna   daapath.*
     few   year=DAT   before   1s PROX dwarf=DAT   PAST-kìnna get
     'Some years before, I was captured by this dwarf.' (K070000wrt04)

Marking of arguments is exactly identical in (21), but *kìnna* is optional there, because entering a relationship with a man is not necessarily surprising or undesirable, whereas being captured is.

(21) *Se=ppe mavol   ini   oorang=nang su-(kìnna) daapath.*
     1s=POSS daughter PROX man=DAT   PAST-kìnna get
     'My daughter got entangled to this man.' (K081105eli02)

This shows that *kìnna* does not change the syntactic status of arguments, but instead adds a semantic element. This can be obligatory if the expressed proposition forces an adversative reading, but if the proposition is neutral with regard to its being undesirable, *kìnna* is optional. If the proposition is clearly not a case of an adversative event, *kìnna* cannot be used. To conclude, *kìnna* is not a passivization operation which changes the prominence of arguments. It can therefore not be used to test for subjecthood.

12.9   Binding

Binding of reflexive anaphora can be used to establish subjects in some languages. SLM has no reflexive pronouns, so that this test cannot be used.

12.10   Raising

There are no raising verbs in SLM, so that this test cannot be used. The function fulfilled by raising verbs in English is coded by the enclitic *=ke* or the evidential marker *kiyang*, which are located in the same clause as the predicate.

12.11   Obligatoriness

In some languages, a subject must be syntactically present even if there is none semantically. Cases in point are meteorological verbs like *it is raining* in English, where *it* is a dummy subject. In SLM, clauses without arguments are perfectly fine, like ∅ *aràuujang* 'it is raining', so that obligatoriness cannot be used to establish the category of subject in SLM.

12.12   Relativization

In some languages, only certain arguments can be relativized on. In SLM, any argument can be relativized on, so that this test is not useful either. See ⊡→13.2, p. 516 for an extensive discussion of relativization.

12.13   Conclusion

It is very difficult to establish grammatical relations in SLM. The only instance where the notion of 'subject' seems to matter is in control structure with *kamauvan* 'want'. For the following discussion of clause types, the notion of 'subject' is not relevant[8] (nor is the notion of 'object'). In discussing the structure of clauses, I will only refer to *NP*, without implying a particular syntactic role. This means that for instance word order in the main clause is described as *NP NP V* (or *NP* * *V*), rather than *SOV*.

---

[8]Cf. Schachter (1996:51f) for the lack of relevance of this category in grammatical descriptions.

Chapter 13

Clauses

We have discussed how morphemes can be combined into words and how words can be combined into NPs, PPs and predicates. We will now address the next higher unit and discuss how several phrases can be combined to form a clause.

A clause must obligatorily have a predicate (PRED), and it can have one or more nominal phrases (NPs) which encode the arguments of the predicate. These NPs often carry postpositions (POSTP). Additional material can also be put into the clause in the form of adjuncts. All clauses are normally predicate-final (with some exceptions for relative clauses), so that the normal structure of the clause is as follows (The Kleene star * marks that an element can be present 0, 1, or more times).

(1)
$$\left[\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \text{PRED} \right]_{CLS}$$

13.1 Main clauses

The main clause normally consists of a predicate in final position, and a number of adjuncts and arguments to the left. The order of those arguments and adjuncts is free.[1] As argued in ⊡→12, p. 493, SLM grammar has no privileged argument, like 'subject' or 'object'. Rather, all the arguments have the same prominence. This can be represented in the following Tree:[2]

---

[1] This freedom of argument order in the preverbal field is an areal feature (Butt et al. 1994).

[2] The trees follow Givón's (2001) argument for the usefulness of trees based in the style of Chomsky (1957, 1965) to represent constituency, hierarchy, category labels, and linear order, in an elegant way. Generative grammar has used, and modified, trees and tree structure in a number of ways since then, but these more recent types of trees would not serve the illustrative purpose they are intended for here any better than their more traditional counterparts. This holds true regardless of whether one finds the theoretical arguments for the more modern trees compelling or not.

(2)

```
                              S
                     ┌──────┬──┼───────┐
                    NP     NP  NP    ...   PRED
```
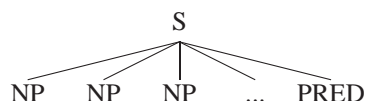
This tree actually resembles Hale's analysis of Japanese clause structure[3] (Hale 1982), in which he denies the existence of a VP in Japanese and argues that Japanese is non-configurational. Mohanan's (1982) analysis of Malayalam, a sister language of Tamil, comes to the same conclusion. Lehmann (1989:177) finally states for Tamil:[4]

> We assume that the verb and all its argument NPs, as well as adverbial adjuncts, are immediate constituents of the sentence. This means that in Tamil there is no syntactic bond [...] which affects the verb and the object NPs but not the subject NP. That is to say, there is no verb phrase (VP) constituent in Tamil. Subjects, objects, and the verb are all immediate constituents of S.

Just like Japanese, Malayalam, or Tamil, SLM seems to lack rigid constituency within the clause, with the exception of the position of the verb, which must be the last or second-to-last element (see below) in the clause.

In SLM, it is very common for arguments to be dropped if they are inferable from context. It appears that Sri Lanka Malay is similar to Late Archaic Chinese in this respect, for which Li (1997) said: "What demands explanation is not zero-anaphora [...] but the appearance of a referential expression, whether a pronoun or a full fledged noun phrase." While this chapter will discuss many clauses with overt arguments, it should be borne in mind that the normal case for Sri Lanka Malay is to have most of the arguments dropped. This is discussed in more detail in Chapter ⊙→16, p. 677.

As for clauses, we can distinguish a number of subtypes, discussed in the sections given:

- Declarative clause ⊡→13.1.1, p. 506

- Copular clause ⊡→13.1.2, p. 510

- Two types of interrogative clauses ⊡→13.1.3, p. 512, ⊡→13.1.4, p. 513

- Imperative clause ⊡→13.1.5, p. 514.

13.1.1 Declarative clause

The declarative clause is the most simple type. It consists of between zero and three arguments followed by a predicative phrase. See the section on predicate types for examples ⊡→10, p. 461.

---

[3]This analogy between SLM and Japanese is also found in the domain of relative clauses ⊡→13.2, p. 516.
[4]Non-configurational analysis for Sinhala have also been proposed by Gair (1998 [1983]) and Henadeerage (2002).

(3) $\left\{\begin{array}{l}\text{NP}\\\text{PP}\\\text{ADJCT}\end{array}\right\}$ * PRED $\Big|_{CLS}$

The order of the arguments is free, but the leftmost argument is normally the topic.[5] An example of this is given in (4). The case marker of the topic is occasionally dropped as in (5)(6).

(4) *[[Ini kaaving=nang aada] haarath-saarath pada]$_{top}$ kitham=pe bannyak*
    PROX wedding=DAT exist traditions   PL   1p=POSS many
    oram   pada arà-kijja.
    people PL   NON.PAST-make
    'The traditions that were at this wedding, many of our people follow them.'
    (K061122nar01)

(5) *[Luvar nigiri]$_{top}$=∅ kithang=nang mà-pii thàrà-suuka.*
    outside country   1PL=DAT   INF-go NEG-like
    'We do not want to go abroad.' (K051222nar04)

(6) *[Samma thumpath]$_{top}$=∅ mlaayu aada.*
    every   place   Malay exist
    'In every place, there are Malays.' (K051222nar04)

Occasionally, material can be found to the right of the predicate (cf. Slomanson (2006:137ff.), Ansaldo (2005b:14)). This is frequently done for arguments encoding location and goal.

(7) NP(=POSTP)$_{TOPIC}$ $\left\{\begin{array}{l}\text{NP}\\\text{PP}\\\text{ADJCT}\end{array}\right\}$ * PRED PP$_{loc}$ $\Big|_{CLS}$

Example (8) shows a local argument following the predicate *anà-laaher* 'was born'.

(8) *see anà-laaher   Navalapitiya=ka$_{loc}$*
    1s PAST-be.born Nawalapitiya=LOC
    'I was born in Nawalapitiya.' (K051201nar01)

---

[5]See Bayer (2004:62) and references therein for a more extensive list of preferred orders in languages with free word order.

When right dislocation of non-spatial arguments occurs, these are often in focus.[6]

(9) $\text{NP(=POSTP)}_{TOPIC} \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \text{PRED} \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\}_{foc}$

In the following three examples, the argument following the predicate is in contrastive focus to other referents mentioned before in discourse (other ethnic groups, other food, other languages).

(10)   *Nni   Peradeniya jaalang=ka samma n-aada   **mlaayu**.*
       PROX Peradeniya road=LOC   all   PAST-exist Malay
       'Everybody in this Peradeniya Rd was Malay.' (K051222nar04)

(11)   *Hindu arà-maakang **kambing**.*
       Hindu NON.PAST-eat   goat
       'Hindus eat GOAT.' (K060112nar01)

(12)   *Itthukapang kitham arà-blaajar   **mulbar**.*
       then   1PL   NON.PAST-learn Tamil
       'Then we learn TAMIL.' (K051213nar02)

Also, very heavy NPs can be dislocated to the right:

(13)   *Kitha=nang_{NP} maau_{PRED} [**kitham=pe mlaayu lorang blaajar lorang=pe***
       1PL=DAT   want   1PL=POSS   Malay 2PL   learn 2PL=POSS
       ***mlaayu kitham blaajar**]_{NP}.*
       Malay 1PL   learn
       'We want that you learn our Malay and that we learn your Malay.'
       (K060116nar02)

But also arguments fulfilling none of the conditions mentioned above can be found after the verb, as shown in the following four examples.

(14)   *Suda hathu week=nang duuva skali arà-dhaathang **daughter**.*
       so   one   week=DAT two   time NON.PAST-come daughter
       'Thus my daughter comes twice a week.' (K051201nar01)

---

[6]This is a common focalization device in Sinhala and Tamil, which entails additional morphological changes in these languages, not mirrored in SLM (Gair 1985a). Right dislocation without morphological changes is less common but does exist in Tamil (Herring 1994) and Sinhala (Paolillo 1997).

(15) *Derang pada arà-mintha* **nigiri**.
    3   PL   NON.PAST-ask country
    'They are asking for a country of their own.' (K051206nar12)

(16) *Itthu=nang blaakang su-dhaathang* **Hambanthota mlaayu pada**.
    DIST=DAT   after   PAST-come   Hambantota   Malay   PL
    'After that came the Hambantota Malays.' (K051206nar14)

(17) *Itthu muusing Islam igaama nya-aajar kaasi* **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)

It could be possible to analyze the last three examples as sentences similar to an English pseudo-cleft. The first part would be a headless relative clause ('What they are asking for', 'Who came then', 'Who taught Islam'), and the second part the instantiation thereof ('a country', 'Hambantota Malaya', 'Jaffna teachers'.). This structure is common in Tamil and could be at the origin of a similar structure in Sinhala as well (Gair 1985a). More research on the intonation contours of these sentences is needed to ascertain whether right dislocation and pseudo-clefts can be distinguished by their intonation.

If we summarize our findings about argument positions we can say that an arbitrary number of arguments can occur before the verb, but only one can occur after the verb.

(18)
$$\left[\left\{\begin{array}{l}\text{NP}\\\text{PP}\\\text{ADJCT}\end{array}\right\} * \text{PRED}\left(\left\{\begin{array}{l}\text{NP}\\\text{PP}\\\text{ADJCT}\end{array}\right\}\right)\right]_{CLS}$$

This resembles very much the structure of the NP as given in ▱→8.1.17, p. 444. The amount and order of elements preceding the head is quite free, while following the head, there is only one position, which has some restrictions to it. As for the noun phrase, these are more absolute (no deictic, possessive or relative clause can ever be used after the noun), while for the clause, they are more lax, more like tendencies (non-spatial arguments tend not to occur after the verb). Similar to the representation of the NP (175) in ▱→8.9, p. 455, we can represent the SLM declarative clause as in (19).

(19)

```
                          S
              _____|_____
             |         |        |
            pre       pred     post
           __|___              |
          |   |  |            NP
         NP  NP ...
```
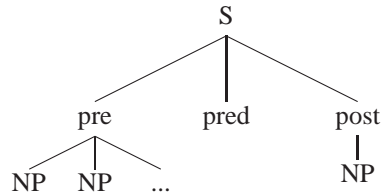
13.1.2 Copular clause

Some predicates can be supported by the copula *(asà)dhaathang(apa)* ☐→5.5, p. 218. This is most often done for predicates of naming (22) or of class-membership, either in an ethnic group (23), a profession (24) or kin (25). English material in the sentence seems to favour the use of the copula.

(20) | NP COPULA name |

(21) | NP COPULA class |

(22) *Se=ppe baapa dhaathangapa **Jinaan Samath**.*
     1s=POSS father COPULA   Jinaan Samath
     'My father was Jinaan Samath.' (N060113nar03)

(23) a. *Se=ppe daughter-in-law=pe   mother asàdhaathang **bingaali**.*
        1s=POSS daughter-in-law=POSS mother COPULA   Bengali
        'My daughter-in-law's mother is Bengali.'

     b. *Ithukapang daughter-in-law=pe   father asàdhaathang **mlaayu**.*
        then   daughter-in-law=POSS father COPULA   Malay
        'Then my daughter-in-law's father is Malay.' (K051206nar08)

(24) *Umma=pe   baapa dhaathangapa hathu **inspector of police**.*
     mother=POSS father COPULA   INDEF inspector of police
     'My mother's father was an inspector of police.' (N060113nar03)

(25) *Baapa=pe umma asàdhaathang **kaake=pe**   **aade**.*
     father=POSS mother COPULA   grandfather=POSS younger.sibling
     'My paternal grandmother was my grandfather's younger sister.'
     (K051205nar05)

Note that the use of the copula is optional in all these cases; it is more common to hear sentences without the copula. The following examples give sentences contrasting with the examples above in the absence of the copula.

(26) *Sudaara sudaari se=ppe naama **Wahida Jamaldiin***.
brother   sister   1s=poss name   Wahida Jamaldeen
'Brothers and sisters, my name is Wahida Jamaldeen.' (B060115prs05)

(27) *Sindbad* the Sailor ***hatthu Muslim***.
Sindbad the Sailor indef   Muslim
'Sindbad the Sailor was a Moor, he was not a Malay.' (K060103nar01)

(28) ∅ *karang* **Dialog GSM**=*ka* **junior executive** *hatthu*.
   now   Dialog GSM=loc junior executive indef
'She is now junior executive at Dialog GSM [phone company].'

(29) *Se=ppe neene   itthu **kaake=pe**   aade*.
1s=poss grandmother dist   grandfather=poss younger.sibling
'My grandmother is that grandfather's younger sibling.' (K051205nar05)

There are occasionally other semantic classes introduced by the copula (see ⊡→5.5, p. 218 for a discussion).

Copular clauses are often used for equational predicates. This is especially true for the naming use, where we assert that referent X is the same as referent Y. We do not assert of any predicate that it is true of my father; rather, we assert that two referents are identical.

Because equation is a symmetric predications (*My father is John* and *John is my father* have the same truth values), inversion of the positions of X and Y does not change the truth value. However, information structure changes, as in (30) and (31). In both examples, the referent first mentioned is topical, while the other one is the comment on the topic.

(30) *[Se=ppe kaake]   asàdhaathang [*estate tea factory officer*]*.
1s=poss grandfather copula   estate tea factory officer
'My grandfather was the estate tea factory officer.' (K060108nar02)

(31) *[Estate=pe* field officer*] asàdhaathangapa [kithang=pe kaake]*.
      estate=POSS field officer COPULA   1PL=POSS   grandfather
      'The estate field officer was our grandfather.' (N060113nar03)

On very rare occasions, the order of the copula and the two NPs can be mixed up,
as in (32).

(32) *asàdhaathangapa incayang* army captain.
      COPULA   3s.POLITE army captain
      'He was an army captain.' (B060115cvs04)

13.1.3   Interrogative clause, clitic

A third clause type is formed by adding the interrogative clitic *=si* to the portion of a
declarative clause one wishes to question, normally the predicate. This is shown for
verbal predicates in (35) and for a nominal predicate in (36). Questioning an argument is shown in (37). Right dislocation does not seem to be possible in interrogative
clauses.

(33) $\boxed{\text{CLS=si}}_{CLS}$

(34) $\left.\begin{cases} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{cases} \text{NP=si PRED}\right|_{CLS}$

.

(35) *Se=pe   uumur masà-biilan=si?*
      1=POSS age   must-tell=INTERR
      '(Do I) have to tell my age?' (B060115prs01)

(36) *Lorang=nang see=yang ingath-an=si?*
      2PL=DAT   1s=ACC   think-NMLZR=INTERR
      'Do you have thoughts on me/are you thinking of me?' (K070000wrt04)

(37) ***Daging baabi=si** anà-bìlli?*
      pork=INTERR   PAST-buy
      'Did you buy PORK?' (K081105eli02)

This is also the Sinhala and Tamil way of forming interrogative clauses.

13.1.4 Interrogative clause, WH

A fourth clause type is the interrogative clause involving an interrogative pronoun. The interrogative pronoun substitutes the queried element. It is normally found *in situ*, but since the word order to the left of the verb is quite free anyway, this also means that *in situ* position cannot be distinguished from initial position. Technically speaking, this is just a special case of the declarative clause, where one or several NPs are instantiated by interrogative pronouns, instead of nouns, pronouns, etc.

$$
(38) \quad \left[\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \text{WH(=POSTP)} \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \text{PRED} \right]_{CLS}
$$

The following three examples show the use of a WH-pronoun in initial position.

(39) **Mana** *nigiri=ka   arà-duuduk?*
which country=LOC NON.PAST-stay
'In which country do you live?' (B060115cvs16)

(40) **Saapa** *m-maati?*
who   PAST-die
'Who died?' (K051213nar07)

(41) **Aapa** *arà-biilang   itthu?*
what NON.PAST-say DIST
'So, what does that mean?' (K051206nar04)

Non-initial position of the WH-pronoun is found in the following two examples.

(42) *Lorang naama kapang-biilang, [baapa **saapa** umma **saapa**] katha.*
2PL   name when-say   father who mother who   QUOT
'When you tell your name, (also tell) who (are) (your) parents.' (B060115nar04)

(43) *[Incayang=pe naama **aapa**],* sir=*pe   naama?*
3S.POLITE=POSS name   what   sir=POSS name
'What's his name, the teacher's name?' (K060103cvs01)

The interrogative pronoun can be doubled. This indicates that an exhaustive list is expected as the answer. A single item will not do. Examples (44) and (45) illustrate

this.

(44) *Aapa anà-bìlli?*
      what PAST-buy
      'What did you buy?' (K081103eli02)

(45) *Aapa˜aapa anà-bìlli.*
      what˜RED   PAST-buy
      'What all did you buy.' (K081103eli02)

13.1.5   Imperative clause

The fifth main clause type is the imperative clause. It consists of between 0 and 2 expressed arguments, none of them agent, and a verb at the right edge. The verb optionally carries the particle *mari* to its left or the imperative suffixes *-la* or *-de* to its right, or both. The verb cannot carry any further TAM marking.

(46) $\left\{\begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array}\right\}$ * (*mari*) V $\left(\begin{array}{l} -la \\ -de \end{array}\right)$

Examples (47) and (48) show the use of a bare verb in the imperative clause, (49) shows the use of *mari*. Examples (50) and (51) show the use of the suffixes while (52) has both *mari* and *-la*.

(47) *Aajuth thaakuth=ka su-naangis, "See=yang ∅-luppas-∅".*
      dwarf fear=LOC   PAST-cry   1s=ACC   leave
      'The dwarf screamed in fear: "Leave me!" ' (K070000wrt04)

(48) *Binthan auntie=ka ∅-caanya-∅, binthan auntie=yang konnyong*
      Binthan auntie=LOC ask,   Binthan auntie=ACC   few
      ∅-panggel-∅.
      call
      'Ask auntie Binthan, call Binthan auntie' (K060116nar06)

(49) **Mari** *maakang.*
     come eat
     'Eat!' (B060115rcp02)

(50) *Allah **diyath-la** inni pompang pada dhaathang aada.*
     Allah watch-IMP PROX female PL come exist
     'Almighty, see, these women have come!' (K061019nar02)

(51) *Pii! ... pii!! ... pii-**de**!!!*
     go   go   go-IMP.IMPOLITE
     'Go! Go now! Bugger off!!!' (not on recordings)

(52) a. *Saayang se=ppe thuan **mari** laari-la.*
       love   1s=POSS sir come.IMP run-IMP
       'Come my beloved gentleman, come here.'

     b. *See=samma kumpul **mari** thaaňdak-la.*
       1s=COMIT   gather come.IMP dance-IMP
       'Come and dance with me.' (N061124sng01)

The imperative prefixes *marà-* 'ADHORT' and *jamà-* 'NEG.IMP' (cf. Slomanson 2008a) are used in a different construction. In this construction, mentioning of the agent (normally *kithang* 'we') is possible.[7]

(53) $(kithang) \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \begin{array}{l} \textit{marà–} \\ \textit{jamà–} \end{array} \text{V}$

(54) *Kitham **marà**-maayeng.*
     1PL   ADHORT-play
     'Let's play.' (K081104eli06)

---

[7]The modal particle *thussa* can be used for prohibitions as well, but this is not an imperative clause syntactically speaking. The speech act of prohibitions is discussed in ⊙→17.1.5, p. 699.

(55) *Hatthu=le **jamà**-gijja baapa ruuma=ka duuduk.*
      INDEF=ADDIT NEG.IMP-do father house=LOC stay
      'Don't do anything, daddy, stay at home!' (B060115nar04)

The adhortative combined with the interrogative particle =*si* forms an adhortative like English *shall we?*

(56) $\left\{\begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array}\right\}$ * *marà*-V=*si*

(57) *Marà-pii=si.*
      ADHORT-go=INTERR
      'Shall we go?' (K081105eli02)

13.2 Relative clause

The relative clause has the same word order as the declarative main clause and precedes its head. Both head-internal and head-external analysis are possible for the SLM relative clause. Given that material can intervene between the relative clause and the head noun ⊡→8.1.14, p. 427, a head-external analysis is preferred here.

(58) ... $\left[\left\{\begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array}\right\} * PRED\right]_{RELC}$ NP(=POSTP)$_{main}$ ... PRED$_{main}$

Relative clauses are only indicated by position.[8] Any declarative main clause can be turned into a relative clause by putting it before a nominal. That nominal can have any semantic role.[9] This simple principle covers all there is to say about relative clauses, still the different possibilities will be discussed in detail.

The Relative Clause Construction is used for relativization properly speaking (*The photo that John left surprised me*), but also for fact clauses (*The fact that John left surprised me*). These two constructions are semantically different in that in the former, we are dealing with a first order entity (an individual x, *photo*) modified by a proposition, whereas in the latter, we are dealing with a third order entity (a proposition

---

[8]Smith et al. (2004) note that relative clauses "are headed by a verbal adjective", yet none of the examples they give has a form glossed as 'verbal adjective'; it appears that all verbs in the examples they give are simply finite verbs.

[9]This is thus very different from other Western Austronesian languages, where only subjects can be relativized (Himmelmann 2005:161).

X, *fact* (Hengeveld 1992:8), whose content is given by the fact clause (cf. Lehmann 1984:46). While these two types are semantically different, this semantic difference is not mirrored in SLM syntax: both use the construction given in (58).[10]

We will first discuss the occurrence of different TAM-markers in relative clauses, to show that they are fully finite[11] ⊡→13.2.1, p. 517. We will then turn to different predicate types in relative clauses ⊡→13.2.2, p. 520 and finally discuss the different semantic roles on which one can relativize ⊡→13.2.3, p. 523. Relative clauses can also occur without a head they modify. These headless relative clauses have been treated in ⊡→8.8.2, p. 453.

13.2.1 TAM in the relative clause

There are no restrictions on the TAM markers which can appear in the relative clause. The following sections give examples of relative clauses in the different tenses.

13.2.1.1 Past *anà-*

The past tense prefix *anà-* can be found in relative clauses.

(59) *[Incayang=pe kàpaala=ka **anà**-aada] thoppi=dering moonyeth pada=nang*
       3s.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
       su-buvang puukul.
       PAST-throw hit
       'He took the hat from his head and violently threw it at the monkeys.'
       (K070000wrt01)

13.2.1.2 Past *su-*

The past tense prefix *su-* can be found in relative clauses.

(60) *[Ruuma duuva subala=ka **su**-aada   rooja pohong komplok duuva]=yang*
       house two   side=LOC PAST-exist rose tree   bush   two=ACC
       asà-baa=apa mliige=pe duuva subla=ka su-thaanàm.
       CP-bring=after palace=POSS two   side=LOC PAST-plant
       'The rose bushes that stood at both sides of the house were brought and planted on both sides of the palace.' (K070000wrt04)

---

[10]See Matsumoto (1997) for a comparable analysis of Japanese relative clauses and fact clauses.
[11]Unlike Sinhala and Tamil, which have non-finite relative clauses.

13.2.1.3 Perfect with *aada*

The perfect tense can be found in the relative clause.

(61) *[Seelon=nang dhaathang **aada**] mlaayu oorang ikkang.*
      Ceylon-DAT come   exist Malay man   fish
      'The Malays who came to Sri Lanka were fishermen.' (K060108nar02)

(62) *Itthu asàdhaathang [baaye=nang vaasil-king   **aada**] dagger hatthu.*
      DIST COPULA   good=DAT   blessed-CAUS exist dagger INDEF
      'That was a well blessed dagger' (K051206nar02,K081105eli02)

13.2.1.4 Non-past *arà*

Arà- in its non-past reading can be found in the relative clause.

(63) *thoppi arà-daagang   oorang*
      hat   NON.PAST-trade man
      'The hat seller' (K070000wrt01)

13.2.1.5 Simultaneous *arà-*

The simultaneous reading of *arà* is also found.

(64) *Suda [puthri=le   biini=le   **arà**-caaňda haari]=le su-dhaathang.*
      so   queen=ADDIT wife=ADDIT SIMULT-meet day=ADDIT PAST-come
      'So then the day came when the wife and the queen were to meet.'
      (K070000wrt05)

This event is located in the past, yet *arà-* is used, which can then not have the non-past meaning, but rather the meaning of 'simultaneous to the time of the matrix sentence' i.e. the coming of the day and the meeting are treated as referring to the same time frame.

The same thing can be said about the next example, where the hearing and the knocking are conceived of as simultaneous. It is not possible for the knocking to refer to non-past, since it is impossible to have heard something in the past which had not yet occurred. Therefore, the non-past reading of *arà-* is not an option here.

(65) *[Kìrras pinthu=nang **arà**-thatti   hathu svaara] su-dìnngar.*
     strong door=DAT   SIMULT-hammer INDEF noise] PAST-hear
     'They heard a noise of hard hammering at the door.' (K070000wrt04)

13.2.1.6 Irrealis *anthi-*

The irrealis is not found often in relative clauses. The following is the only example, where this prefix is used in its epistemic reading.

(66) *[Kithang=nang duppang dhaathang **athi**-aada exiles pada]=jo anà-baa.*
     1PL   before   come   IRR-exist exiles PL=EMPH PAST-bring
     'The exiles, who would have come before us, were brought.' (K060108nar02)

13.2.1.7 Conjunctive participle *asà-*

The conjunctive participle prefix can be used on its own in relative clauses when it expresses perfect tense, as in (67). In this sentence, the existential *aada*, which expresses perfect tense together with *asà-*, is dropped. *Asà-* can also be used on non-final verbs if there are more verbs in the relative clause (68).

(67) *[Tony Hassan uncle=nang **asà**-kaasi (aada) duvith] athi-mintha ambel=si?*
     Tony Hassan uncle=DAT CP-give (exist) money IRR-ask   take=INTERR

     'Shall I ask for the money you gave to uncle Tony Hassan?' (K071011eml01)

(68) *[Banthu-an **asà**-mintha arà-naangis] svaara hatthu derang=nang*
     help-NMLZR CP-beg   SIMULT-cry   sound INDEF 3PL=DAT
     su-dìnngar.
     PAST-hear
     'They heard a sound of crying and begging for help.' (K070000wrt04)

13.2.1.8 Infinitive/purposive *mà-*

The infinitive prefix *mà-* can be used for relative clauses of purpose (69)(70). The dative marker =*nang* can optionally be present in the relative clause (71).

(69) *Lorang se=dang [**mà**-hiidop] thumpath kala-kaasi, ...*
      2PL   1s=DAT INF-stay   place   if-give,   ...
      'If you give me a place to stay, ...' (K070000wrt04)

(70) *Kithang lorang=nang baaye mliiga athi-kaasi, [**mà**-kaaving] panthas*
      1PL   2PL=DAT   good palace IRR-give   INF-marry   beautiful
      pompang pada athi-kaasi, duvith athi-kaasi.
      female   PL   IRR-give   money IRR-give
      'We will give you nice palaces, we will give you beautiful girls to marry, we
      will give you money.' (K051213nar06)

(71) *Itthu=nang aada [divorce **mà**-kijja=nang] hatthu prentha oorang.*
      DIST=DAT   exist divorce INF-make=DAT   INDEF law   man
      'For that there is a lawyer to make the divorce.' (K061122nar01)

13.2.1.9   Negative past *thàrà-*

The negated past tense prefix *thàra-* can also be used in relative clauses.

(72) *Derang pada panggel=nang blaakang [**thàrà**-dhaathang oorang*
      3PL   PL   call=DAT   after   NEG.PAST-come   man
      pada]=nang nya-force-kang kiyang
      PL=DAT   PAST-force-CAUS EVID
      'After they had called (them), (they) apparently forced the people who had not
      come (to join).' (K051206nar07)

13.2.2   Predicate types in the relative clause

All predicate types can be found in relative clauses.

13.2.2.1   Verbal predicate

A simple example of a verbal relative clause is (73).

(73) *Itthu [se arà-**kirijja**]   mosthor=jo.*
    DIST 1s NON.PAST-make manner=EMPH
    'That's the way I do it.' (B060115rcp01)

13.2.2.2 Modal predicate

Modal predicates can be used in relative clauses. This is especially frequent for *boole* 'can', as given in (74).

(74) *[Deram pada [. . .] baae=nang pìrrang mà-kijja **boole**] oorang.*
    3PL PL   good=DAT war INF-make can man
    'They were men who were able to fight well.' (K051213nar06)

Normally, the head noun takes the case the matrix clause requires, as in (74) (where it is zero), but it also occurs that the head noun takes the case required by the relative clause instead of the one required by the matrix clause. In (75), the matrix clause would not assign any case to its only argument. *Boole* in the relative clause normally assigns dative, which is then also marked on the head noun by the postposition =*nang*.

(75) *[Itthu baaye mosthor=nang **bole**=kirja oorang mlaayu]=nang sajja=jo.*
    DIST good manner=DAT can-make man Malay=DAT   only=EMPH
    'The Malays are the only ones who can do it in the right way.' (K061026rcp01)

The negation *thàrboole* can also be used in a relative clause.

(76) a. *[Boole oorang pada] samma dhaathang.*
       can   man PL   all   come
       'The people who could came.'
   b. *[**Thàràboole** oorang pada] su-biilang: kithang=nang mà-dhaathang*
       cannot   man PL PAST-say 1PL=DAT   INF-come
       thàràboole.
       cannot
       'The people who could not come said: "We cannot come".'
       (K051206nar07)

Besides *boole*, *maau* has also been found occurring in a relative clause.

(77) *[Derang=nang **maau** mosthor] baalas katha nya-biilang.*
     3PL=DAT   want manner answer QUOT PAST-say
     I gave them the answer in the way they wanted ' (K051213nar01)

13.2.2.3   Nominal and adjectival predicates

Relative clauses based on nominal predicate clauses and adjectival predicate clauses
are formally indistinguishable from nouns modified by a bare noun (78) or a bare
adjective (79) instead of a a clause.

(78) *[Moonyeth]$_{N/RELC}$ hathu kavanan su-aada.*
     monkey   INDEF group   PAST-exist
     'There was a monkey group.'
     'There was a group which consisted of monkeys.' (K070000wrt01)

(79) *Ini   [laama]$_{ADJ/RELC}$ car pada=jo kithang arà-baapi.*
     PROX old   car PL=EMPH 1PL   NON.PAST-bring
     'It is these old cars we take [to Iraq].'
     'It is these cars which are old that we take [to Iraq].' (K051206nar19)

13.2.2.4   Circumstantial predicates

Circumstantial predicates are normally found with overt marking of the existential, as
in the following two examples.

(80) *[Incayang=pe kàpaala=ka **anà-aada**] thoppi=dering moonyeth pada=nang*
     3S.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey   PL=DAT
     su-buvang puukul.
     PAST-throw hit
     'He took the hat which was on his head and violently threw it at the monkeys.'
     (K070000wrt01)

(81) *Derang [dìkkath=ka **aada**] laapang=nang mà-maayeng=nang su-pii.*
     3PL   vicinity=LOC exist ground=DAT   INF-play=DAT   PAST-go
     'They went to play on the ground which was nearby.' (K070000wrt04)

13.2.2.5 Relative clauses based on utterances

Besides clauses, it is also possible to use utterances to premodify a noun. The following examples show this for a short reported string (82)(83) and a long reported string giving the content of *habbar* 'news' (84).

(82) *Baava=ka   [Kaasim katha] hatthu* family.
      bottom=LOC Kaasim QUOT   INDEF family
      'Below there is a family called "Kaasim".' (K051206nar06)

(83) *Itthukang anà-aada [Mr Janson katha] hathu oorang.*
      then   PAST-exist Mr Janson QUOT   one   man.
      'Then there was a certain Mr Janson.' (K051206nar04)

(84) *Se=dang habbar, [[ini laama* car *pada samma inni* suicide bombers
      1s=DAT news   PROX old   cat PL   all   PROX suicide bombers
      asà-dhaathang car=*yang, aapa, arà*-paavicci   *katha] habbar].*
      CP-come   car=ACC   what NON.PAST-use(Sinh.) QUOT   news
      'I have information, information "these suicide bombers come and take all these old cars and, what, and use them." ' (K051206nar19)

13.2.3   Semantic roles in the relative clause

There is no restriction on the semantic roles that can be relativized on. This will be shown for the different semantic roles in the individual sections below.

13.2.3.1 Agent

The verb *dhaathang* 'come' subcategorizes for Agent.

(85) *[Seelon=nang dhaathang aada] mlaayu oorang ikkang.*
      Ceylon-DAT   come   exist Malay man   fish
      'The Malays who came to Sri Lanka were fishermen.' (K060108nar02)

13.2.3.2 Patient

Patients of both intransitive clauses (*mniiňğgal* 'die' in (86)) and transitive clauses (*baa* 'bring' in (87)(88)) can be relativized on.

(86) *[North-pe   pìrrang-ka mniiňğgal mlaayu pada]=nang bìssar hatthu*
    north=ᴘᴏss war=ᴘᴏss die   Malay ᴘʟ=ᴅᴀᴛ   big   ɪɴᴅᴇꜰ
    religious function.
    religious function
    'A religious function for the Malays who had died in the Northern war.'
    (K060116nar11)

(87) a. *Ka-duuva anà-dhaathang* slaves *pada,*
        ᴏʀᴅ-two ᴘᴀsᴛ-come   slaves ᴘʟ
        'The second to come were slaves,'

      b. *[soldier pada na-baa   oorang pada].*
        soldier ᴘʟ   ᴘᴀsᴛ-bring man   ᴘʟ
        'men brought by soldiers.'

(88) *[[Itthu mà-jaaga=nang] anà-baa   mlaayu]=dring satthu oorang=jo se.*
    ᴅɪsᴛ   ɪɴꜰ-protect=ᴅᴀᴛ ᴘᴀsᴛ-bring Malay]=ᴀʙʟ   one   man=ᴇᴍᴘʜ 1s
    'One of the Malay men brought to protect that (man) is me.' (K060108nar02)

13.2.3.3   Theme

For the semantic role of theme, relativization is possible for all the different shades of
meaning, be they an item located at some place as in (89), a stimulus as in (90), or a
concept known (91).

(89) *Incayang [ini   Seelong=ka anà-aada lakuan baathu] asà-caari.*
    3s.ᴘᴏʟɪᴛᴇ ᴘʀᴏx Seelon=ʟᴏᴄ ᴘᴀsᴛ-exist wealth stone   ᴄᴘ-find
    'He   was   looking   for   the   gems   present   in   Ceylon.'
    (K060103nar01,K081105eli02)

(90) *[Kìrras pinthu=nang arà-thatti   hathu svaara] su-dìnngar.*
    strong door=ᴅᴀᴛ   sɪᴍᴜʟᴛ-hammer ɪɴᴅᴇꜰ noise] ᴘᴀsᴛ-hear
    'They heard a noise of hard hammering at the door.' (K070000wrt04)

(91) *[Se=dang thaau mosthor]=nang karang=nang ka-dhlaapan* generation
   1s=DAT   know way=DAT   now=DAT   ORD-eight   generation
   arà-pii.
   NON.PAST-go
   'As far as I know, we are now in the eighth generation.' (K060108nar02)

13.2.3.4 Experiencer

The person experiencing the ability is relativized on in (92).

(92) *Deram pada ... [baae=nang pìrrang mà-kijja boole oorang]*
   3PL   PL good=DAT war   INF-make can   man
   'They were men who were able to fight well.' (K051213nar06)

13.2.3.5 Recipient

The recipient of the verb *kaasi* 'give' is relativized on in (93).

(93) *[Se   duvith anà-kaasi oorang] su-iilang.*
   1s=DAT money PAST-give man   PAST-disappear
   'The man I gave money to disappeared.' (K081105eli02)

13.2.3.6 Possessor

The husband formerly 'possessed' by the woman is relativized on in (94).

(94) *[Laaki   anà-mniiňğgal hathu pompang].*
   husband PAST-die   INDEF woman
   'A woman whose husband had died.' (K070000wrt04)

13.2.3.7 Location

Locations can be relativized on. In (95), the location of staying is at the same time the location of falling. The locative argument *thumpath* 'place' is found as the semantic role relativized on. The same is true of *thumpath* in (96).

(95) *Siithu [nya-duuduk thumpath]=ka baapa su-jaatho.*
      there PAST-stay   place=LOC   father PAST-fall
      'There, at the place he was staying, my father fell.' (K051205nar05)

(96) *[Spaaman anà-niiŋ̌gal thumpath]=nang=le [Passara katha arà-biilang*
      3s   PAST-die   place=DAT=ADDIT   Passara QUOT NON.PAST-say
      nigiri]=nang=le   dìkkath.
      country=DAT=ADDIT vicinity
      'The place where he died is close to the village called Passara.'
      (B060115nar05)

Locations can also be relativized on in headless relative clauses ⊡→8.8.2, p. 453, as in (97), where the place of graduating (*passing out* in Sri Lanka) is relativized upon.

(97) *[See anà-pass.out ∅] abbisdhaathang* University *of Peradeniya=ka*
      1s PAST-graduate   COPULA   University of Peradeniya=LOC
      'Where I graduated was the University of Peradeniya.' (K061026prs01)

13.2.3.8   Time

Points in time can be relativized on, such as *haari* 'day' in (98), or *thaaun* 'year' in (99).

(98) *Suda [puthri=le   biini=le   arà-caaŋda haari]=le su-dhaathang.*
      so   queen=ADDIT wife=ADDIT SIMULT-meet day=ADDIT PAST-come
      'So then the day came when the wife and the prince were to meet.'
      (K070000wrt05)

(99) *Itthu thaaun=jo [Mahathma Gandhi arà-buunu   thaaun].*
      dist year=EMPH Mahathma Gandhi NON.PAST-kill year
      'That year was the year that Mahathma Gandhi was killed.' (K051213nar02)

13.2.3.9   Instrument

The following sentences shows that the instrument for collecting can be found in a relative clause.

(100) *Incayang=ka [[bannyak panthas ummas baarang pada=le bathu inthan*
       3s=LOC   many   beautiful gold   good   PL=ADDIT stone value
       pada=le anà-punnu-kang]_{RELC} bìssar beecek caaya hathu bag_{headnoun}*]*
       PL=ADDIT PAST-lot-CAUS   big   mud   colour INDEF bag
       su-aada.
       PAST-exist
       'With him, he had a big brown bag with which he collected very beautiful
       golden items and jewellery.' (K070000wrt04)

13.2.3.10 Manner

The relativized argument can have the role of Manner.

(101) *Itthu [se arà-kirija mosthor]=jo.*
       DIST 1s NON.PAST-make manner=EMPH
       'That's the way I do it.' (B060115rcp01)

13.2.3.11 Purpose

If the relativized role is Purpose, the relative clause will be in the infinitive.

(102) *Derang pada=nang [itthu mà-kumpul athu mosthor] thraa.*
       3PL   PL=DAT   DIST INF-add   INDEF way   NEG
       'There is no way for them to collect all this.' (G051222nar01)

(103) *Itthu=nang aada [divorce mà-kijja=nang hatthu prentha oorang].*
       DIST=DAT   exist divorce INF-make=DAT INDEF law   man
       'Then there is a lawyer to make the divorce.' (K061122nar01)

13.2.4 The Relative Clause Construction used as a complement

Besides the use of the Relative Clause Construction to modify referents, as illustrated above, this construction can also be used to indicate the complement of nouns. In (104), the Relative Clause Construction is used for the complement of *habbar* 'news'.

(104)   *Se=dang habbar, [ini   laama* car *pada samma inni* suicide *bombers*
        1s=DAT news   PROX old   cat PL   all   PROX suicide bombers
        asà-dhaathang car=*yang, aapa, arà-paavicci katha habbar].*
        CP-come   car=ACC   what NON.PAST-use QUOT news
        'I have information, information that these suicide bombers come and take all
        these old cars and, what, and use them.' (K051206nar19)

The semantic distinction between, say, (103) and (104) is that in (103) the sub-
ordinate clause modifies and restricts the reference of *prentha oorang* 'lawyer'. In
(104), the subordinate clause gives the *content* of the news; the subordinate clause
does not restrict the reference of *habbar* 'news' to those news which take cars and
use them. Note that in modifying Relative Clause Construction, there is one argu-
ment which subordinate clause and main clause share, e.g. *prentha oorang* 'lawyer'
in (103). This is not the case for *habbar* 'news' in (104), which is not an argument
of the subordinate clause. Furthermore, all semantic arguments of the subordinate
clause are syntactically present in (104), (agent *suicide bombers*, theme *car*), while in
(103), the argument *prentha oorang* 'lawyer' is semantically present in both clauses,
but syntactically, only in the main clause.

Next to finite clauses as (104), clauses containing the infinitive marker *mà-* can
also be used as complement clauses.

(105)   *[Inni mlaayu pada Sri Lanka=nang **mà**-dhaathang=nang* reason] *aapa*
        PROX Malay PL   Sri Lanka=DAT   INF-come=DAT   what
        katha arà-biilang.
        QUOT NON.PAST-tell
        '(I) will tell (you now) what is the reason why these Malays came to Sri
        Lanka.' (N060113nar01)

13.2.5   Discussion

Typologically speaking, SLM is of the Japanese type (Lehmann 1984:50f), which
means that it has preposed relative clauses, and uses the same verb form in relative
clause and main clause. Interestingly, SLM is not of the Dravidian type (Lehmann
1984:70f) like Sinhala or Tamil, which do prepose relative clause, but use a participle
to distinguish the relative clause from the main clause. SLM has emulated many of
the grammatical categories of Sinhala and/or Tamil, but it has not developed a relative
participle as of now.

Slomanson (2006), based on Saldin (1996) (first edition of Saldin (2001)), assumes
a relative pronoun *nya*. The sentence provided for exemplification, however, suggests
that we are not dealing with a relative pronoun, but with an allomorph of the past tense

marker *anà-*. This sentence is given below:

(106) *[[Jalan ka nya lari] ayaŋ]] baisikal atu ka tərbuntur su mati-mati*
      road   P REL run chicken bicyle DET P struck   PAST die
      'The chicken which ran along the road was struck by a bicycle and died.'
      (Slomanson 2006:153) (original orthography)

There are a number of oddities in this sentences, e.g. the reduplication of the final verb, which I will not go into here. What is important is that the string *nya lari ayaŋ* could as well be analyzed as *nya-lari ayaŋ*, where *nya-* would be an allomorph of the past tense prefix *anà-*, as discussed in ⊡→6.2.1.2, p. 272. This is thus a normal finite past tense prefix, and not a relative pronoun.

13.3 Conjunctive participle clause

For sequence of events, a hypotactical construction exists, which consists in using the conjunctive participle (infinite) in all but the last clause. The finite clause has to be the last event.[12]

(107) $$\left[\begin{bmatrix} \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \left\{ \begin{array}{l} asà- \\ jamà- \end{array} \right\} - V \end{bmatrix} * \text{MAIN CLAUSE} \right]$$

The subordinate clauses may only have verbal predicates. This is normally no drawback since all [+dynamic] states-of-affairs are coded by verbs, and the occurrence of [-dynamic] states-of-affairs in a sequence is limited. If another state-of-affair is to be used in a sequence, it will get a dynamic reading (see ⊙→15.2.2, p. 601 for strategies how to do this).

Example (108) shows a typical instance of this construction. There are three clauses, of which the first two contain a verb marked by *asà-* , while the last one contains a verb in another tense. Note that in this case, the referent *oorang pada* 'the people' is introduced in the first clause, while in English, it would be introduced in the matrix clause, in this case the last one.

---

[12]These clause chains (Longacre 1985) are a typical feature of South Asia (Emeneau 1971).

(108)   a. *Oorang pada **asà-pìrrang,***
            man   PL   CP-wage.war
            'After having waged war'

        b. *derang=nang **asà-banthu,***
            3PL=DAT   CP-help
            'and after having helped them'

        c. *siini=jo   se-ciiňǧgal.*
            here=EMPH PAST-settle
            'the people settled down right here.' (K051222nar03)

If one of the subordinate clauses is in the negative, *jamà-* ⊡→6.2.1.6, p. 280 is used instead of *asà-*.

(109)   a. *Liivath aayer **jamà**-jaadi=**nang***
            much   water NEG.NONFIN-become=DAT
            'Without putting too much water (=having not put too much water)'

        b. *itthu aayer=yang hathu blaangan=nang luppas.*
            DIST water=ACC   INDEF amount=DAT   leave
            'leave that water for a while.' (K060103rec01)

There are no restrictions on the roles that the arguments of the conjoined clauses may have. Other languages allow conjunctive participle constructions only if the subject (in those languages) are identical. This is the case for most South Asian languages (Masica 1976:108), but not for SLM as (110) shows:[13]

---

[13]Gair (1976a:58) shows that this is not the case for Sinhala either, although Gair et al. (1989) return to speaking of 'subjects' in Sinhala, relevant for this construction. However, the examples used in the latter paper include dative marking on the purported subjects, so that 'topical argument' seems to be a more relevant characterization than morphosyntactic subject.

(110) a. *Go   **asà**-niiňǧgal,*
           1S.FAMILIAR CP-die
           'I having died'

       b. *Alla go=nya   **asà**-dhaathang,*
           Allah 1S.FAMILIAR=DAT CP-come
           "Allah having come towards me"

       c. *Kuburan **asà**-gaali,*
           grave   CP-dig
           'The grave having been dug'

       d. *Go=nya   kubur-king!*
           1S.FAMILIAR=ACC bury-CAUS
           Bury me!'
           'I die and Allah comes for me and the grave will be dug and they will have
           me buried.' (B060115nar05)

The only argument of the first clause is *go* 'I', while the agent/'subject' of the second one is *Allah*. The arguments of the first and the second clause are not identical. One could argue that the restriction on coordination only holds for zero-anaphora. This is also not the case in SLM, as the third and fourth clause show. If this restriction held, *Allah* would be the subject of these clauses, but this is obviously not the case.

Another example of reference changing between the *asà*-clauses and the main clause is (111), where the first two lines are about the parents and the third one is about the children. Nevertheless, all but the last clause are marked by *asà-*.

(111) a. *Nni nigiri=ka=jo   ∅$_{i(ag)}$ kitham=pe [aanak buva pada]$_j$=yang*
           PROX country=LOC=EMPH   1PL=POSS child   fruit PL=ACC
           asà-simpang.
           CP-keep
           '**We** have raised our children in this country and'

       b. *∅$_i$ ∅$_j$ inni schools pada=nang asà-kiiring, samma asà-kirja.*
               PROX schools PL=DAT   CP-send   all   CP-make
           '**We** have send them to the schools here and done all that and'

       c. *Karang ∅$_j$ asà-blaajar, ∅$_j$ pukurjan asà-kirja ambel.*
           now   CP-learn   work   CP-make take
           'now **they** have learned and **they** have taken up work'

       d. *Skarang ∅$_{i,j}$ siini=jo   arà-duuduk.*
           now   here=EMPH NON.PAST-live
           'and now **we** live here.' (K051222nar04)

13.4   Purposive clauses

The purposive clause resembles the main clause, with the exception that the verb is
in the infinitive and that the purposive clause cannot stand on its own. It must have
a matrix clause it attaches to. Purposive clauses are often additionally marked with
=*nang*, but this is optional (cf. Slomanson 2006:139f).

(112)
$$\left[\left[\left\{\begin{matrix} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{matrix}\right\} * \left\{\begin{matrix} m\grave{a}- \\ jam\grave{a}- \end{matrix}\right\} V\right](=nang)\ \text{MAINCLAUSE}\right.$$

(113)
$$\text{MAINCLAUSE}\left[\left\{\begin{matrix} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{matrix}\right\} * \left\{\begin{matrix} m\grave{a}- \\ jam\grave{a}- \end{matrix}\right\} V\right](=nang)$$

The purposive clause can be center-embedded in the main clause as in (114) or
follow the predicate of the main clause as in (115).

(114) *Derang [dìkkath=ka aada laapang]=nang [mà-maayeng]$_{purp}$=nang su-pii.*
       3PL   vicinity=LOC exist ground=DAT   INF-play=DAT   PAST-go
       'They went to play on the ground nearby.' (K070000wrt04)

(115) *Itthu* cave=*nang kithang*=*le pii aada [mà-liyath]$_{purp}$*=nang.*
       DIST cave=DAT 1PL=ADDIT go exist INF-look=DAT
       'We have also gone to that cave to have a look.' (K051206nar02)

13.5   Subordinate interrogative clauses

Interrogative pronouns can replace NPs in subordinates just as they can in main clauses.
This is done if the speaker reports a question as in (116)(117), or if he confesses his
ignorance as in (118)(119), or if he indicates who knows the answer (120).

(116) *Andare raaja=ka su-caanya [inni mà-kìrring simpang aada **aapa=yang**]*
       Andare king=LOC PAST-ask   PROX INF-dry   keep   exist what=ACC
       katha.
       QUOT
       'Andare inquired from the King what was it that was left [on the mat] to dry.'
       (K070000wrt02)

(117) *Incayang arà-caari* *[inni avuliya **aapa mosthor**] katha.*
3s NON.PAST-search PROX saint what manner QUOT
'He is looking for how that saint was.' (B060115cvs04)

(118) *[Ithu **maana˜maana** thumpath] katha kithang=nang buthul=nang*
DIST which˜RED place QUOT 1PL=DAT correct=DAT
mà-biilang thàrboole.
INF-say cannot
'We cannot tell you exactly which places those were.' (K060108nar02)

(119) *Se=ppe oorang pada [see **saapa**] katha thàrà-thaau subbath see=yang*
1s=POSS man PL 1s who QUOT NEG-know because 1s=ACC
su-uubar.
PAST-chase
'Because my folks did not know who I was, they chased me.' (K070000wrt04)

(120) *Incayang=nang thaau itthu=pe mosthor=atthas punnu, cinggala **aapacara***
3s.POLITE=DAT know DIST=POSS way=about much Sinhala how
anà-banthu katha.
PAST-help QUOT
'He know a lot about that, about how the Sinhalese helped the Malays.'
(K051206nar04)

The five examples above contain content questions, which are marked by *katha*.
Katha can be missing, as in the following example.

(121) *Derang thàrà-thaau [**aapacara** anà-jaadi] ∅.*
3PL NEG-know how PAST-happen
'They did not know how it happened.' (K051213nar01)

It is also possible to have truly embedded questions, which are marked by *=so*.
This is possible for content questions as in as in (122) and for polar questions (123).

(122)   a. *Itthu see arà-gijja   mosthor=jo.*
           DIST 1S NON.PAST-make way=EMPH
           'That's the way I do it.'

         b. *See ini   arà-biilang.*
            1S   PROX NON.PAST-say
            'That's what I am saying.'

         c. *[Laayeng oorang pada **aapacara arà-gijja=so**]   thàrà-thaau.*
            other   man   PL   how   NON.PAST-make=UNDET NEG-know
            'I do not know how other people do it.' (B060115rcp02)

(123)   *[Inni samma anthi-oomong=**so**] thàrà-thaau.*
         PROX all   IRR-say=UNDET   NEG-know
         'I do not know what they say about all that.' (G051222nar02)

In this function, =*so* competes with =*si*, which is found in (124) and (125).

(124)   *[Aashik=nang hathu soldier mà-jaadi   suuka]=**si** **katha** arà-caanya.*
         Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT   NON.PAST-ask
         'He asks if you want to become a soldier, Ashik.' (B060115prs10)

(125)   *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
         Hopefully other know brother sister   PL=LOC can-ask   take
         *[[nya-gijja lai   saapa=kee aada]=**si**   **katha**].*
         PAST-make other who=SIMIL exist=INTERR QUOT
         'Hopefully, you can enquire from another person you know whether there is
         someone else who did something.' (N061031nar01)

Note that in these two examples the quotative *katha* is found, so that the clause
containing =*si* can be analyzed as a reported main clause. In (122) and (123), on the
other hand, *katha* is not found, so that we are not dealing with a reported string, but
rather with a true subordinate clause. The utterance containing =*si* thus has its own
illocution (question), which is reported in the matrix clause with assertive illocution.
There are thus two utterances, and two illocutions.

When =*so* is used on the other hand, we are not dealing with an utterance, but
only with a clause, which cannot have illocutionary force. This can also be seen from
the absence of *katha*, which can only attach to utterances, but not to clauses. The =*si*-
construction thus contains two utterances, whereas the =*so*-construction only contains

one.

The difference can be captured in the following two patterns.

(126) $\left[\left[\left\{\begin{matrix}\text{NP}\\\text{PP}\\\text{ADJCT}\end{matrix}\right\}*\left[\left\{\begin{matrix}\text{NP}\\\text{PP}\\\text{ADJCT}\end{matrix}\right\}*\text{V}_{cls}{=}si\right]_{utt}\ katha\ \text{V}_{cls}\right]_{utt}\right]$

(127) $\left[\left\{\begin{matrix}\text{NP}\\\text{PP}\\\text{ADJCT}\end{matrix}\right\}*\left[\left\{\begin{matrix}\text{NP}\\\text{PP}\\\text{ADJCT}\end{matrix}\right\}*\text{V}{=}so_{cls}\ \text{V}_{cls}\right]_{utt}\right]$
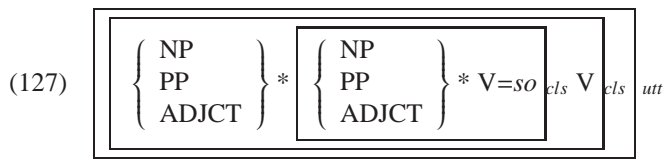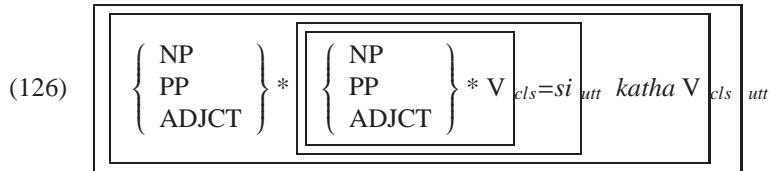
13.6 Supraordination

An interesting feature of SLM syntax is the possibility of the matrix clause occurring *within* the embedded clause. I will call this structure 'supra-ordination'. This is very often done with exlamations of surprise with the aim to make the adressee aware of the exceptional nature of the content (mirativity). An answer is normally not expected. A typographical means to render this meaning in English is the combined use of the question mark and the exclamation mark as in *Do you know what my boss just said !?!?*. In the source language, I will indicate supraordination by inverted brackets]...[as in the following examples.

(128) *Suda inni moonyeth pada]aapa thaau=si[anà-gijja !*
so PROX monkey PL what know=INTERR PAST-do
'Do you know what these monkeys then did!?' (K070000wrt01)

(129) *Suda derang pada siini derang itthu oorang]aapa thaau=si[anà-gijja.*
thus 3PL PL here 3PL DIST man what know=INTERR PAST-make
'Do you know what these men did!?' (K051206nar07)

(130) *Itthu=nang blaakang derang pada]aapa thaau[anà-gijja?*
DIST=DAT after 3PL PL what know PAST-make
'After that do you know what they did?' (K051206nar15)

In the preceding examples, we are dealing with rhetorical questions where the answer is not provided yet. These sentences have a verb in the final position. In

(131), things are a little bit different in that at the end of the utterance all the content is provided, i.e. the hearer does not have to try to find the answer to the rhetorical question because the answer (*jiimath* 'talisman') is already provided.

(131)   *Kithang arà-thaaro   ]aapa thaau=si[jiimath.*
        1PL   NON.PAST-put what   know=INTERR talisman
        'We put you know what? A talisman!' (K051206nar02)

13.7   The position of adjuncts

Adjuncts are formed either by adverbs or by postpositional phrases. They can occur anywhere before, between or after NPs and predicates. The following schema illustrates this for a main clause with right-dislocation.

(132) | (ADJCT) NP (ADJCT) NP (ADJCT) PRED (ADJCT) NP (ADJCT)) |

The following three examples show the use of the adjunct *karang* 'now' in initial, medial and final position.

(133)   ***Karang*** *ini   kitham=pe nigiri   su-jaadi.*
        now   PROX 1PL=POSS   country PAST-become
        'This (country) has now became our country.' (K0512222nar04)

(134)   *Go=dang   **karang** konnyong thàràsìggar.*
        1S.FAMILIAR=dat now   little   ill
        'I am a bit sick these days.' (B060115nar04)

(135)   Associations *pada Bahasa Indonesia Bahasa Malaysia=le*
        associations PL   Bahasa Indonesia Bahasa Malaysia=ADDIT
        introduce-*kang aada **karang**.*
        introduce-CAUS exist now
        'Associations have now introduced Bahasa Indonesia and Bahasa Malaysia.'
        (G051222nar03)

13.8   Reported speech

Reported speech is indicated by the particle *katha* which is put at the end of the reported string. *Katha* is used regardless of the reported string being a clause, a sub-

clausal unit, or an utterance. Example (136) shows a reported clause, (137) shows the use of *katha* on an interjection, which does not meet the criteria to be a clause. See ⊡→6.4.9.1, p. 389 for more discussion and examples.

(136) *Se=ppe oorang thuuva pada anà-biilang [kitham pada* Malaysia=*dering*
       1s=POSS man   old   PL   PAST-say   1PL   PL   Malaysia=ABL
       anà-dhaathang]$_{CLS}$ katha.
       PAST-come   QUOT
       'My elders told me that we had come from Malaysia.' (K060108nar02)

(137) *[Yes]$_{INTERJ}$ katha m-biilang.*
       Yes   QUOT PAST-say
       '(X) said "yes".' (K060116nar11)

The reported utterance can precede (140) or follow (141)(142) the clause containing the verb of utterance or cognizance.

(138) $\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \left\{ \begin{array}{l} \text{V}_{say} \\ \text{V}_{cognize} \end{array} \right\}$ [UTTERANCE *katha*]

(139) $\left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} *$ [UTTERANCE *katha*] (NP) $\left\{ \begin{array}{l} \text{V}_{say} \\ \text{V}_{cognize} \end{array} \right\}$

(140) *[Luu=nya   jadi-kang   rabbu saapa lu=ppe nabi   pada saapa*
       2s.FAMILIAR=ACC become-CAUS prophet who 2s=POSS prophet PL   who
       katha] biilang.
       QUOT   say
       'Say who the prophet is who made you, who are your prophets.'

(141) *Biilang [∅ se=ppe] katha.*
       say   1s=POSS QUOT
       'Say that you are mine.' (K061123sng03)

(142) *Spaaru oorang pada arà-biilang [∅ Seelong=nang* English *anà-aaji   baa]*
       some   man   PL   NON.PAST-say   Ceylon=DAT   English PAST-bring bring
       katha.
       QUOT
       'Some people say that it was the English who brought (us, the Malays,) to Sri
       Lanka.' (K060108nar02)

While *katha* operates on the level of the utterance, and not on the level of the
clause, deictic reference is still readjusted to the present speech situation. This means
that tenses and pronouns change between the original string and the reported string.
For instance, in (141), a fragment taken from a song, the loved one is asked to say
that he belongs to the speaker, as conveyed by the translation 'Say that you are **mine**'.
Here, the deictic reference is readjusted since the literal string would have been 'I am
yours.' as in 'Say: "I am **yours**." ' Another example is (143), where in a double-
embedded utterance, the child is supposed to say 'Your mother has died', but the
intended reading is of course that the child's mother had died, not the father's. This
is obvious from the use of the familiar second person pronoun *luu*, which a father
can use when addressing his child, but which would be impossible if the child was to
address the father.

(143)   a. *Andare aanak=nang su-biilang:.*
             Andare child=DAT   PAST-say
             'Andare said to his child:'

         b. *Aanak.*
             child
             ' "Son, " '

         c. *'lu=ppe umma   su-maathi' katha bitharak=apa asà-naangis mari.*
             2s=POSS mother PAST-die   QUOT scream=after CP-weep   come.IMP
             ' "come and cry and weep 'Your (=my) mother has died!' " (to fool the
             king)'

The deictic center is also readjusted in the temporal domain. In example (144) the
speaker reports expressing her being pleased. The literal string would have been 'We
are pleased' but in reported speech, this is changed to 'We were pleased', because at
the time of speaking, the event has already passed.

(144) *[Suda kithang=le su-suuka] katha anà-biilang.*
       thus   1PL=ADDIT PAST-like QUOT PAST-say
       'So we said that we also liked it.' (K051222nar01)

However, this is not always the case. In (145), a narrative about past events (identifiable by the past marker *su-*) contains an embedded quotation in non-past tense (*arà-*). Still, *arà-* in this sentence could also be analyzed as simultaneous tense ⊡→6.2.2.1, p. 287, so that it is difficult to decide whether temporal readjustment does indeed not take place.

(145)   *Andare ruuma=nang asà-pii biini=nang **su**-biilang [puthri=nang kuuping*
        Andare house=DAT   cp-go wife=DAT PAST-say queen=DAT   ear
        ***arà**-dìnngar kuurang katha].*
        SIMULT-hear little   QUOT
        'Andare went home and said to his wife that the queen did not hear well.'
        (K070000wrt05)

Furthermore, imperative clauses in direct speech are rendered as infinitival clauses in reported speech.

(146)   *Oorang padanang **mà-**dhaathang katha asà-biilang.*
        ma   PL=DAT   INF-come   QUOT cp-say
        'Having told the people to come ...' (B060115cvs01)

Sinhala and Tamil use their quotative constructions in very similar situations, but in both languages, the quotative is homophonous to the conjunctive participle of the word meaning 'to say'. This is not the case in SLM, where *asà-biilang* 'cp-say' is different from *katha*.

13.9   Agreement

Although SLM does not show agreement on the verb as such, we are presently witnessing an incipient grammaticalization process which uses resumptive pronouns just before or after the verb if the antecedent is not immediately adjacent.[14] This might very well evolve into an agreement system.[15] As of now, however, there is no general agreement system to be found in SLM.

(147)   ***Spaaman** avuliya su-jaadi=**spaaman**.*
        3s.POLITE saint   PAST-become=3s.POLITE
        'He became a saint.' (B060115nar05)

---

[14]Whether this already counts as agreement is subject to debate and theoretical orientation (cf. Corbett 2006:99f.).

[15]The mechanism has been described by Givón (1976).

The source for this pattern is possibly Tamil, where the agreement suffix is very often very similar to the pronoun (*niinga*ḷ *pooringa*ḷ 'You(pl.) are going.') and etymologically related. However, some speakers use resumptive pronouns before the verb (148)-(150), not after the verb as Tamil influence would predict.

(148)   ***Dr Draaman*** *duuva thaaun=nang blaakang* **incayang***=su-mniiňǧgal.*
          Dr Draaman two   year=DAT   after   3s.POLITE=PAST-die
          'After two years, Dr Draman died.' (K051213nar08)

(149)   *Itthunam* **samma baae oorang pada** *inni   nigiri=ka=jo*
          therefore all   good man   PL   PROX country=LOC=EMPH
          deram*=sa-duuduk.*
          3PL=PAST-stay
          'It is because of that that all good men settled down in this country.'
          (N060113nar01)

(150)   *Suda* **deram** *inni* political promise *hatthu* **derang***=eng-kaasi* 1958*=ka.*
          so   3PL   PROX political promise INDEF 3PL=PAST-give   1958=LOC
          'So they made this political promise in 1958.' (N060113nar02)

As of now, the instances of this pattern are isolated, so that we cannot speak of SLM having a generalized agreement system.

Chapter 14

Beyond the clause

When morphemes are combined into words, words into phrases and phrases into clauses, this is still not enough to produce meaningful discourse. The relation between the different clauses must also be indicated. This is done in SLM with linking adverbials ⊡→14.1, p. 541. Another construction, tail-head-linkage, also assures discourse continuity ⊡→14.2, p. 543. Finally, coordinating construction can be used to link elements on different levels ⊡→14.3, p. 544.

14.1 Linking adverbials

SLM does not have a category of conjunctions. Instead, a number of adverbs or lexicalized combinations of deictic+postposition are used. The resulting structure is paratactic rather than hypotactic.

(1)　　CLAUSE $\left\{ \begin{array}{c} \text{ADV} \\ \text{DEIC + CLT} \end{array} \right\}$ CLAUSE

The following sentences show the use of the adverbial *suda* 'thus' (2), or the deictic+X-structures *itthule* 'but' (3) and *ikang* 'then' (4).

(2)   a. *Se=ppe blaakang arà-dhaathang se=ppe* fifth generation *se=ppe anak*
           1s=poss back   non.past-come 1s=poss fifth generation 1s=poss child
           klaaki pada.
           boy   pl
           'After me is coming the fifth generation, my sons.'

       b. **Suda** *se=ppe thuuva anak klaaki asàdhaathang dhlapan-blas thaaun.*
           thus 1s=poss old   child boy   copula   eight-teen   year
           'Thus, my eldest son is eighteen.' (K060108nar02)

(3)   *Spaaru mlaayu pada arà-oomong   **itthule** mulbar arà-oomong.*
       some   Malay pl   non.past-speak but   Tamil non.past-speak
       '(Only) few Malays speak (Malay), but they speak Tamil.' (G051222nar04)

(4)   a. *Kandi=ka* Malay mosque=*pe* *blaakang=ka incayang=pe* *zihaarath*
           Kandi=loc Malay mosque=poss back=loc   3s.polite=poss shrine
           aada.
           exist
           'In Kandy behind the Malay Mosque, there is his shrine.'

       b. **Ikang** *derang=pe sudara pompang=jo aada Hanthane=ka.*
           then 3s=poss   sibling sister=emph exist Hantane=loc
           'Then, his sister is in Hantane.' (K060108nar02)

Both the 'adverbial strategy' and the 'deictic+clitic' strategy are found in the small
fragment below.

(5) a. ***Itthule*** *kithang=pe muusing=ka=jo anà-baa.*
     but 1PL=POSS time=LOC=EMPH PAST-bring
     'But it was in our time that they were brought.'

  b. *Itthu kithang hathu vakthu=nang kithang=nang duppang dhaathang*
     DIST 1PL INDEF time=DAT 1PL=DAT before come
     athi-aada.
     IRR-exist
     'So they, em, at that time, they would have come before us.'

  c. Exiles *pada=jo anà-baa.*
     exiles PL=EMPH PAST-bring
     'Exiles were brought.'

  d. ***Itthu=nang=jo*** *see anà-biilang*
     DIST=DAT=EMPH 1S PAST-say
     'That's why I am saying'

  e. *oorang soojor anà-baa katha.*
     man European PAST-bring QUOT
     'that the Europeans brought them.'

  f. *Iiya.*
     yes
     'Yeah.'

  g. ***Suda*** *itthu oorang pada=le Seelong=ka arà-duuduk.*
     thus DIST man PL=ADDIT Ceylon=LOC NON.PAST-stay
     'So those people stay in Ceylon.' (K060108nar02)

14.2 Tail-head linkage

SLM makes extensive use of tail-head linkage. This means that the end of a clause is repeated at the beginning of the following clause (Thompson 1985). This is normally accomplished with a conjunctive participle.[1] Example (6) shows this pattern over several clauses.

---

[1]This pattern is common in Dravidian languages (Emeneau 1971).

(6)   a. *Asà-girijja incayang, su-**pass**ᵢ.*
          cp-do   3s.ᴘᴏʟɪᴛᴇ   ᴘᴀsᴛ-pass
          'Having done that, he passed (the exam).'

      b. *Asà-**pass**ᵢ, kaarang,* Advanced Level *arà-**girijja**ⱼ   **skaarang**.*
          cp-pass   now   Advanced Level ɴᴏɴ.ᴘᴀsᴛ-make now
          'Having passed the exam, then, he will do the Advance Level then.'

      c. ***Karang**,* Advanced Level *asà-**girijja**ⱼ, 1978=ka,* Late Exam *anthi-girijja.*
          now   Advanced Level cp-make,   1978=loc late exam ɪʀʀ-make
          'Then, having done the Advanced Level exams, for 1978ers, he will do
          the late exam.' (K051222nar08)

Example (7) shows another instance of this.

(7)   a. *Mìntha daaging=yang **cuuci**ᵢ.*
          raw   beef=ᴀᴄᴄ   wash
          'Wash the raw beef.'

      b. *Asà-**cuuci**ᵢ, laada=le   gaaram=le bathu giling-an=ka **giiling**ⱼ.*
          cp-wash   pepper=ᴀᴅᴅɪᴛ salt=ᴀᴅᴅɪᴛ stone grind=ɴᴍʟᴢʀ grind
          'Having washed it, grind salt and pepper in a grinding stone.'

      c. *Asà-**giiling**ⱼ, daaging mùntha=yang baathu=ka asà-thaaro, giccak.*
          cp-grind   beef   raw=ᴀᴅᴅɪᴛ   stone=ʟᴏᴄ cp-put   smash
          'Having ground and having put the raw beef on a stone, smash it.'
          (K060103rec02)

14.3   Coordinating constructions

Coordination can be done on a variety of levels of linguistic analysis: words, phrases,
and clauses. In SLM, either zero-coordination ⊡→14.3.1, p. 544 or clitic coordination
⊡→14.3.2, p. 548 is used for nearly all these cases.

14.3.1   Unmarked coordination

Two items can be simply juxtaposed to express coordination.

(8)   ┌─────────────┐
       │ X=∅ Y=∅ │
       └─────────────┘

The coordinated items can either be NPs, PPs, predicates, modifiers or clauses.

14.3.1.1 Unmarked NP coordination

Unmarked coordination of NPs can be found in PPs which contain more than one noun to which the postposition applies. In the following example, the postposition =*dering* applies to the three mentioned countries at the same time.

(9) *Mlaayu pada anà-dhaathang [[nigiri mlaayu]_{NP} [Malaysia]_{NP}*
    Malay PL PAST-come   country Malay   Malaysia
    [Indonesia]_{NP} [Singapore]_{NP}=dering]_{PP}.
    Indonesia   Singapore=ABL
    'The Malays came from the Malay countries Malaysia, Indonesia, Singapore.'
    (K051213nar06)

While in (9), it is the postposition which indicates that the NPs are coordinated, in the following example, we find the countries of origin in coordinated NPs without a postposition.

(10) British government *[Malaysia]_{NP} [Indonesia]_{NP} [ini nigiri pada]_{NP}*
    British government Malaysia   Indonesia   PROX country PL
    samma anà-peegang ambel.
    all   PAST-catch   take
    'The British government captured Malaysia, Indonesia, all these countries.'
    (K051213nar06)

Human referents can also be coordinated by such a construction, as the following example shows.

(11) *Itthukapang [se=ppe baapa]_{NP}, [se=ppe kaake]_{NP}, [kaake=pe*
    then   1s=POSS father   1s=POSS grandfather grandfather=POSS
    baapa]_{NP}, [kithang samma oorang]_{NP} Seelong=pe oorang pada.
    father   1PL   all   man   Ceylon=POSS man   PL
    'Then my father, my grandfather, grandfather's father, all of us are Ceylonese'
    (K060108nar02)

The next example is very similar to the preceding one. An enumeration of NPs referring to different ethnic groups are strung together by unmarked coordination.

(12) School=*ka* *[kithang]*_{NP} *[mulbar aanak pada]*_{NP} *[cinggala aanak pada]*_{NP}
      School=LOC 1PL   Tamil   child PL   Sinhala child PL
      *[sraani pada]*_{NP} *[mlaayu pada]*_{NP}, *samma hatthu=nang=jo anà-duuduk.*
      Burgher PL   Malay PL   all   one=DAT=EMPH PAST-stay
      'In school, the Tamil children, the Sinhalese children, Burghers and Malays,
      we all stayed together.' (K051213nar03)

14.3.1.2 Unmarked PP coordination

Just like NPs, PPs can be strung together by unmarked coordination.

(13) *Kamauvan vakthu=nang, kithang=nang [itthu mosthor=nang]*_{NP}*,*
      want   time=DAT   1PL=DAT   DIST manner=DAT
      *[Malaysian hathu mosthor=nang]*_{NP} *kithang=nang bole=duuduk.*
      Malaysian INDEF manner=DAT   1PL   can-exist.ANIM
      'When we want, we can be like this or in a Malaysian way' (K060108nar02)

14.3.1.3 Unmarked predicate coordination

It might be the case that there is unmarked predicate coordination in SLM, but this can-
not be established. Given the possibility to drop referents retrievable for the hearer, it
is impossible to decide whether there is unmarked predicate coordination or whether
we are rather dealing with clause coordination where all arguments of the second
clause are not overtly realized. This unclear status is indicated by ∅ between paren-
theses. The question whether the element is there, but not overtly realized (clause
coordination), or whether the element is not there (predicate coordination) cannot be
decided.[2]

(14) *Kithang=nang baaye=nang mulbar bole=baaca (∅_{agent}) (∅_{theme})*
      1PL=DAT   good=DAT   Tamil can=read
      bole=thuulis.
      can=write
      'We can read and write Tamil well.' (K051222nar06)

   This analytical problem arises again in (15), where it cannot be decided whether
the assertions are coordinated, or whether we are dealing with clause coordination
with dropped agent.

---

[2]Karunatillake (2004:206) explicitly states for the related Sinhala construction: "Verbs that go in pairs
like [read/write] can be coordinated by simple juxtaposition."

(15) *Kithang lorang=nang baaye mliiga athi-kaasi, (∅) mà-kaaving panthas*
     1PL   2PL=DAT   good palace IRR-give   INF-marry beautiful
     pompang pada athi-kaasi, (∅) duvith athi-kaasi.
     female PL   IRR-give   money IRR-give
     'We will give you nice palaces, we will give you beautiful girls to marry, we
     will give you money.' (K051213nar06)

14.3.1.4 Unmarked modifier coordination

While unmarked coordination of NPs and PPs is common, unmarked coordination of modifiers is found less often. Example (16) restricts the set of referents for *kuvathan* 'strength' to the strength of the arms and the legs, by unmarked coordination.

(16) *Kaaki*$_{mod}$ *thaangan*$_{mod}$ *kuvathan*$_{head}$ *thraada.*
     foot   hand   strong-NMLZR NEG.exist
     'You have no strength in your arms or legs.' (K060116sng01)

Example (17) coordinates two towns of origin which restrict the reference of *oorang* 'man' to people hailing from these towns. The two towns do not carry an overt mark of coordination.

(17) *Thraa, kithang samma* Kandy$_{mod}$ Navalapitiya$_{mod}$ *oorang*$_{head}$.
     NEG   1PL   all   Kandy   Nawalapitiya   man
     'No, we are all Kandy people or Nawalapitiya people.' (K051205nar05)

A complex modification is found in (18), where the coordination is actually not of the same type as above, since the example describes a joint Sinhala-Tamil problem, and not a problem which the Sinhalese and the Tamils happen to share.

(18) *Duuva* week *didaalang cinggala*$_{mod}$ *mulbar*$_{mod}$ *reepoth*$_{head}$.
     two   week within   Sinhala   Tamil   problem
     'Within two weeks, there was the Sinhala-Tamil problem.' (K051213nar01)

This contrasts with (17), where the referents comprise all people with origin Kandy and all people with origin Nawalapitiya. (18) does not refer to the sum of all Sinhala problems and all Tamil problems, but it refers to precisely the one problem which is characterized by Sinhalese and Tamils playing a role in it.

14.3.1.5   Unmarked clause coordination

A chain of independent clauses can be formed without any material indicating co-ordination. In (19), four existential clauses are joined without overt indication of coordination.

(19)   *[Cinggala aada]*_{CLAUSE1} *[mlaayu aada]*_{CLAUSE2} [Moor *aada]*_{CLAUSE3} *[mulbar*
      Sinhala   exist   Malay   exist   Moor exist   Tamil
      *aada]*_{CLAUSE4}.
      exist
      'There are Sinhalese, there are Malays, there are Moors, there are Hindus.'
      (G051222nar04)

Five adjectival predicate clauses are conjoined without overt marking in (20).

(20)   a.   *Kithang=pe nigiri=ka   bedahan   aada.*
        1PL=POSS   country=LOC difference exist
        'There are differences in our country.'

      b.   *[Mulbar laayeng]*_{CLAUSE1} *[cinggala laayeng]*_{CLAUSE2} [Moor
        Tamil   different   Sinhala   different   Moor
        *laayeng]*_{CLAUSE3} *[mlaayu laayeng]*_{CLAUSE4} *[sraani pada*
        different   Malay   different   burgher PL
        *laayeng]*_{CLAUSE5}.
        different
        'Hindus, Sinhalese, Moors, Malays, Burghers, all are different.'
        (K061127nar03)

14.3.2   Coordination with clitics

SLM makes use of coordinating clitics ⊡→6.4.6, p. 359 for all types of coordination. These attach to all of the coordinated items, which can be arguments, adjuncts or predicates. Coordination of clauses with this construction seems to be possible, but is very seldom found. This is expressed by parentheses in (21).

    The use of an 'A-and B-and structure' (Haspelmath 2004a) is typically South Asian.

(21)   ⎧ NP   ⎫   ⎧ NP   ⎫
      ⎪ PRED   ⎪ =COORD. CLITIC ⎪ PRED   ⎪ =COORD. CLITIC
      ⎨ ADJ   ⎬   ⎨ ADJ   ⎬
      ⎩ (CLAUSE)⎭   ⎩ (CLAUSE)⎭

14.3.2.1 Argument coordination

Argument coordination is a regular instance of coordination. The following examples show the use of conjunctive and disjunctive coordination of arguments.

(22) shows the use of overt coordination on NPs unmarked for case, whereas (23) shows the use of overt coordination on PPs.

(22) a. *Thiiga klaaki aade=le*   *hatthu pompang aade=le*
  three male younger.sibling=ADDIT one female younger.sibling
  se=dang arà-duuduk.
  1s=DAT NON.PAST-exist.ANIM
  'I have three sons and one daughter.' (K060108nar01)

(23) *Snow-white=nang=le Rose-red=nang=le ini hatthu=ke*
  Snow.white=DAT=ADDIT Rose.Red=DAT=ADDIT PROX INDEF=SIMIL
  thàrà-mirthi.
  NEG.PAST-understand
  'Snow White and Rose Red did not understand a thing.' (K070000wrt04)

An utterance with two instances of coordination is (24).

(24) a. *See anà-blaajar **mulbar=le** **English=le** itthu muusing.*
  1s PAST-learn Tamil=ADDIT English=ADDIT DIST time
  'I studied in Tamil and English back then.'

  b. *Derang karang arà-blaajar **cinggala=le** **English=le**.*
  3PL now NON.PAST-learn Sinhala=ADDIT English=ADDIT
  'They study in Sinhala and English now.' (K051222nar06)

Disjunctive coordination of arguments is given in (25).

(25) *Ketham pada makanan pada**=si** pakeyan pada**=si** su-baavang.*
  1PL PL food PL=disj clothing PL=disj PAST-bring
  'We brought food or clothing .' (B060115nar02)

14.3.2.2 Adjunct coordination

The distinction between arguments and adjuncts is not clear-cut in SLM ⊡→11, p. 481, but if such a distinction is assumed, example (26) shows that local adjuncts can be overtly coordinated.

(26) *Matale=ka=le   Kluum̌bu=ka=le   su-aada.*
      Matale=LOC=ADDIT Colombo=LOC=ADDIT PAST-exist
      'It was in Matale and in Colombo.' (G051222nar01)

14.3.2.3   Predicate coordination

Predicate coordination is normally not done by clitics, but rather by the anterior
⊡→6.2.1.4, p. 276 or simultaneous ⊡→7.5.4, p. 410 conjunctive participles or zero-
coordination ⊡→14.3.1.3, p. 546.

14.3.2.4   Modifier coordination

There is no instance of overt modifier coordination.

14.3.2.5   Clause coordination

Clause coordination differs from the other kinds of coordination in that the coordinat-
ing clitic normally does not attach to the right edge of the items, but rather to the right
edge of one of the arguments of the clause. In example (27), it is clear that it is not
the mouths which are coordinated, but the clauses containing the predications since
the predicate *paasir* 'sand' is present twice. Still, the coordinating clitic *=le* does not
attach to the predicate, but rather to *muuluthka* 'in the mouth' in both of the clauses.
There is thus a mismatch between semantic scope (clause) and morphological locus
(PP).

(27) *Lu=ppe   **muuluth=ka=le** paasir, se=ppe **muuluth=ka=le***
      2S.FAMILIAR=POSS mouth=LOC=ADDIT sand   1S=POSS mouth=LOC=ADDIT
      paasir.
      sand
      'There is sand in your mouth and there is sand in my mouth.' (K070000wrt02)

A similar observation can be made about (28), where the coordination concerns
two acts of hel-ping, one towards the Sinhalese king and one towards the British.
But the coordinating clitics do not attach to the predicate *mà-fight=nang* or the right
edges of the clauses (*saama*), but to the argument expressing the beneficiary. True, the
crucial difference between the two coordinands is the identity of the beneficiary, but
these beneficiaries participate in different propositions. It is not about helping (A and
B) to do X, but rather about helping A to do X and B to do Y. Again, the morphological
marking of coordination does not reflect its semantic scope.

(28) a. *Cinggala raaja=nang=le na-banthu.*
         Sinhala   king=DAT=ADDIT PAST-help
         '(They) helped the Sinhalese king'

      b. *mà-fight=nang British=saama.*
         INF-fight=DAT   British=with
         'to fight with the British.'

      c. British *oorang pada=nang=le na-banthu.*
         British man   PL=DAT=ADDIT PAST-help
         'and they helped the British men'

      d. *mà-fight=nang cinggala raaja=saama.*
         INF-fight=DAT   Sinhala king=with
         'to fight with the Sinhala king.' (K051206nar04)

There seems to be a tendency to avoid coordinating clitics on the edge of clauses,[3] as exemplified above. This tendency, however, is not universal, as (29) shows, where the two predicates *niiňğgal* 'die' both carry the additive clitic *=le*. This example is highly unusual.

(29) *Ithukang [bini-yang **niiňğgal=le**], [umma-yang **niiňğgal=le**], [aanak pada*
      Then   wife-3.KIN die=ADDIT   mother-3.KIN die=ADDIT   child   PL
      caari   dhaathang].
      search come
      'So then the wife died and the mother died and the children came to look'
      (K051220nar01)

---

[3] Also found in Sinhala (Gair & Paolillo (1997:46), Gair (2003:808)) and Tamil (Lehmann 1989:162), but see Karunatillake (2004:70).

Part IV

Function

In the morphosyntactic part of the book, we discussed how SLM can construct well-formed sentences. For every morpheme and every construction, it was shortly described what functions it could fulfill. That is, we adopted a form-to-function ('semasiological') approach In this chapter we will switch viewpoint to a function-to-form ('onomasiological') approach, and investigate what means SLM uses to encode universal functional domains (Givón 2001a,b).[1] Since form and function are closely interrelated, there will also be a lot of references to the formal chapter. These references are indicated by a boxed arrow ⊡→. References to other functional domains are indicated by a circled arrow ⊙→.

We will first address how propositional content relating to, among other things, participants, states-of-affairs, space, and time can be encoded ⊙→15, p. 557. Then we will turn to the encoding of information structure ⊡→16, p. 677, and to speech acts ⊡→17, p. 697.

---

[1]The ordering of sections is loosely inspired by Functional Discourse Grammar (Hengeveld & Mackenzie 2008).

Chapter 15

Encoding propositional content

In this chapter, we will deal with the factual semantic content of the message that speakers want to transmit to the hearer, i.e. by what means different aspects of the proposition (participants, space, time) are encoded in SLM. The next chapters will deal with the packaging of said information ⊙→16, p. 677.

To return to the semantic content, which is the topic of this chapter: We will first take a look at how participants in propositions are encoded in SLM ⊙→15.1, p. 557. Then we will turn to predication ⊙→15.2, p. 599, and modification ⊙→15.3, p. 606. After these nuclear elements, we will discuss the encoding of space ⊙→15.4, p. 606 and time ⊙→15.5, p. 613. A number of functional domains which have received some typological interests are discussed in the remainder of this chapter. These are

- quantity ⊙→15.6, p. 639,

- modality ⊙→15.7, p. 650,

- conditionals ⊙→15.8, p. 660,

- gradation ⊙→15.9, p. 662,

- comparison ⊙→15.10, p. 664,

- possession ⊙→15.11, p. 668,

- negation ⊙→15.12, p. 671, and

- kin ⊙→15.13, p. 674.

15.1 Particpants

The participants of a proposition are the entities which participate in the state or event. In the sentence *Peter gives Mary the book*, *Peter*, *Mary* and *the book* are the three

participating entities. Entities can be of different orders, which will be discussed in ⊙→15.1.1, p. 558. These participants can fulfill different roles in the proposition, like AGENT, PATIENT or INSTRUMENT. These will be discussed in ⊙→15.1.2, p. 561. Normally, the number of participants and semantic roles are equal and every participant has one role. Sometimes this is not the case, for instance in *Peter cut himself*, there is only one participant, *Peter*, but there are two roles, Agent and Patient. These cases are discussed in ⊙→15.1.3, p. 582. In yet other cases, one wants to include participants in the proposition whose reference is not established, like *someone* or *whoever*. These cases are discussed in ⊙→15.1.4, p. 588. Finally, participants can be modified in a number of ways. These modifications are discussed in ⊙→15.1.5, p. 594.

15.1.1   Participants of different entity orders

We can distinguish between the following entity orders (Lyons 1977, Hengeveld 1992, 2004b, Keizer 1992, Dik 1997, Hengeveld & Mackenzie 2008).

- properties f

- individuals x

- states-of-affairs e

- propositional content p

- utterances U

In SLM, all of these can be participants. We will treat them in turn.

15.1.1.1   Properties

Properties are distinguished from individuals by not having an existence of their own. They are furthermore distinguished from states-of-affairs by not being located in space and time.

Properties can be used as participants. An example is the property 'amount', which participates in the state of ignoring, *thàràthaau*, in (1).

(1)   *Se=dang kalu [**blaangang**] thàrà-thaau.*
       1s=DAT   if   number   NEG-know
       'As for me, I do not know the number (of the descendants).' (K051205nar05)

Often, properties are derived with the nominalizer *-an* ▱→6.2.5.1, p. 305 when they are to be used as participants, as in (2).

(2) *Tsunaami anà-jaadi* **cupath-an**=*nang kithang=nang thàrà-thaau.*
    tsunami   PAST-become quick-NMLZR   1s=DAT   NEG-know
    'We did not know about the speed with which the tsunami came.'
    (B060115nar02)

Note that in (2), the property is modified by *tsunaami anà-jaadi* 'with which the tsunami came'. While the coming of the tsunami is of course localizable in space and time, the property *cupathan* 'speed' is not. It is impossible to say *the speed was at the coast at 7 a.m.*.

15.1.1.2   Individuals

The prototypical participant is an individual. Participants of the individual type are encoded as a noun phrase, which is headed by a noun or a pronoun, as the case may be. In example (3), there are two individual participants, a thief, who is encoded by a noun phrase headed by the third person pronoun *dee* and the victims of the thief, who are also head of a noun phrase.

(3) *[**Dee**] [**oorang pada**]=nang arà-cuuri.*
    3   man   PL=DAT   NON.PAST-steal
    'He steals from the people.'

15.1.1.3   States-of-affairs

States of affairs are second-order entities, which have temporal extension. The morphosyntactic expression of a state-of-affairs as a referent depends on its realization status. If the state-of-affairs is realized, a finite subordinate clause is used, as in (4), where the writing of the letter has already happened. The past tense prefix *anà-* indicates that the subordinate clause is finite.

(4) *[[Lorang suurath=yang mlaayu=dering **anà**-thuulis]=nang bannyak*
    2PL   letter=ACC   Malay=ABL   past=write=DAT   much
    arà-suuka].
    simult-like
    'He liked very much that you wrote the letter in Malay.' (Letter 26.06.2007)

If the state-of-affairs is not realized yet, the subordinate clause will not be finite but be marked with the infinitive *mà-*.[1] In example (5), the predicate *suuka* 'like' has two participants, a liker and a likee. The liker is an individual, which is encoded by the first person pronoun *se(dang)*, while the likee is a state-of-affairs, meeting the wife.

---

[1]Or the negative infinitive *jamà-*.

This state-of-affairs is not realized yet. As a consequence, the verbal predicate carries the infinitive marker *mà-* (cf. Noonan 1985:95ff).

(5) *Se=dang Andare=pe   biini=yang **mà-caaňda** suuka.*
     1s=DAT Andare=POSS wife=ACC INF-meet   like
     'I would like to meet your wife, Andare.' (K070000wrt05)

If the state-of-affairs is the topic, it can receive the postposition =*yang* ⊡→6.4.4.1, p. 328 (6).

(6) *[Se arà-maakang]cls=**yang** lorang=nang atthu creeveth=si?*
     1s NON.PAST-eat=ACC   2s=DAT   INDEF problem=INTERR
     'Do you have a problem with my eating?' (K081103eli04)

For some conventionalized states-of-affairs a lexical solution (i.e. a noun like *pìr-rang* 'war') is available.

15.1.1.4 Propositional content

The difference between states-of-affairs and propositions is that the former are located in space in time, but have no truth value, while the latter are not located in space and time, but do have a truth value. Propositional content can be true or false, and asserted and denied; this is not possible for states-of-affairs. Propositional content is normally indicated by the quotative *katha*, as in the following example, where people discover that they like politics. However, their liking of politics could be an illusion, and the positive truth value asserted here could turn out to be negative. We are thus dealing with propositional content here, and not with a state-of-affairs.

(7) *Itthukapang=jo derang nya-thaau ambel [derang pada* politic=*nang suuka]*
     then=EMPH   3PL   PAST-know take 3PL   PL   politic=DAT like
     katha.
     QUOT
     'Only then will they come to know that they like politics' (K051206nar12)

The quotative marker is not necessary, as is shown in the following example, where the speaker asserts that the president had not sent the subscription money, but again, he could be mistaken and the negative truth value he assigns to the proposition could turn out to be false.

(8) *See su-diya [kithang=pe* president subscription=*yang thàrà-kiiring]* ∅.
     1s PAST-see 1s=POSS   president subscription=ACC NEG.PAST-send
     'I saw that our president had not sent the subscription.' (K060116nar10)

15.1.1.5 Utterances

Utterances can also play the role of participants in SLM. While propositional content can be true or false, utterances do not have a truth value. On the other hand, utterances have illocutionary force and can be judged as to their pragmatic adequacy and felicity, which is not possible for propositional content. In SLM, utterances are marked by *katha* in most of the cases. In example (9), the content of the dwarf's complaint is reported. This content can be judged as to whether it was appropriate or not, but it cannot be evaluated for its truth value.

(9) *[Incayang=pe jeeňğgoth=yang asà-thaarek=apa incayang=nang*
     3s=POSS   beard=ACC   CP-pull=after   3s=DAT
     su-sakith-kang katha]utt anà-maaki.
     PAST-pain-CAUS QUOT   PAST-scold
     'The dwarf complained that they had pulled his beard and hurt him.'
     (K070000wrt04)

Also non-declarative utterances can be used as participants, an example is a reported question in (10). In this example, the interrogative illocution is overtly marked by =*si*.

(10) *[Aashik=nang hathu* soldier *mà-jaadi   suuka=si   katha]utt arà-caanya.*
      Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT   NON.PAST-ask
      'He asks if you want to become a soldier, Ashik.' (B060115prs10)

We see that *katha* is used both for propositional content and for utterances used as entities. The distinction between these two levels does therefore not seem to be very relevant in SLM grammar.

15.1.2 Participant roles

The participants of a proposition can have different semantic roles. The sentence *Mary cut the bread with the knife on the balcony* has four participants, *Mary, the bread, the knife* and *the balcony*. In this example, Mary has the semantic role of agent, the bread is patient, the knife instrument, and the balcony, location. In the following, we will discuss a number of common semantic roles and how they are expressed in Sri Lanka Malay.

With the exceptions of propositional content and utterances, all entities discussed in the previous section can be marked for different semantic roles. The two exceptions are always of the semantic role THEME, which is never marked morphosyntactically in their case.

Participant marking is done by means of postpositions. If the role of a participant is inferable from context or general knowledge of the world, the marking can be left

out (Ansaldo (2005b:26), Ansaldo (2008:31)). Leaving out the participant role is very often done for topical participants in initial position, like in (11). In this case, the speaker trusts the hearer that the latter will be able to retrieve the participant role of *thumpath* 'place(s)' from context, LOCATION for this lexeme. This is not too difficult here, since *thumpath* 'place(s)' is unlikely to have any other role.

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

Sometimes, marking of participant roles is less straightforward. In the following example, the undergoers of *thiikam* 'stab' are not marked for semantic role while the undergoers of *theeṁbak* 'shoot' are.

(12)   a. *Itthukapang oorang pada=∅ thiikam=apa,*
           then   man   PL   stab=after
           'Then people were stabbed'

      b. *oorang pada=**nang** theeṁbak=apa,*
           man   PL=DAT   shoot=after
           'and people were shot'

      c. *se=dang bannyak creeveth pada su-aada.*
           1s=DAT much   trouble PL   PAST-exist
           'and I got a lot of trouble.' (K051213nar01)

15.1.2.1 Agent

The agent is normally ∅-marked (Ansaldo 2005b:19). An example of this is *kithang* '1pl' in (13).

(13)   *Ithu=kapang lorang=pe leher=yang kithang=∅$_{ag}$ athi-poothong.*
      DIST=when   2PL=POSS neck=ACC 1PL   IRR=cut
      'Then we will cut your neck.' (K051213nar06)

If the agent is not a natural person, but an institution, the instrumental marker =*dering* is found, as in the following two examples, where the Government and the police are said to be agents of capturing and interrogating.[2]

---

[2]Instrumental marking of institutional actors is also found in Sinhala (Gair & Paolillo (1997:31), Gair (2003:791)).

(14) British Government=***dering*** Malaysia *Indonesia, inni nigiri   pada samma*
      British Government=ABL   Malaysia Indonesia PROX country PL   all
      peegang.
      catch
      'The British Government captured Malaysia and Indonesia, those countries.'
      (K051213nar06)

(15) *See=yang* police=***dering*** *nya-preksa.*
      1s=ACC   police=ABL   PAST-enquire
      'I was questioned by the police.' (K051213nar01)

15.1.2.2 Patient

The patient of transitive clauses can be marked either with =*yang* or with =*nang*,
but zero-marking is also found (Ansaldo 2008). The semantics of the verb seems to
influence the choice of the postposition, but lexical subcategorization seems to play a
part as well. The verb *puukul* 'hit' subcategorizes for =*nang*,[3] as do *theeṁbak* 'shoot'
(17) and *thiikam* 'stab' (18).

(16) Dutch=***nang*** *mà-**puukul**=jo cinggala raaja pada pii aada.*
      Dutch=DAT INF-hit=EMPH Sinhala king PL   go exist
      ' The Sinhalese kings went to fight the Dutch.' (K051206nar04)

(17) *Oorang pada**=nang theeṁbak**=apa se=dang bannyak creeveth pada*
      man   PL=DAT   shoot=after 1s=DAT much   trouble PL
      su-aada.
      PAST-exist
      'People were shot and and I got a lot of trouble.' (K051213nar01) '

(18) *Incayang=**nang** su-**thiikam**.*
      3=DAT   PAST-stab
      'They stabbed him.' (K051220nar01)

   The verb *poothong* 'cut', on the other hand, governs the accusative marker =*yang*.

---

[3]Its Sinhala counterpart, *gahanawaa* also subcategorizes for the dative (Gair 1991b:18f). See Garus-
inghe (1962:36,92) for a list of Sinhala verbs with dative marking for patients, which partly overlaps with
the SLM verbs.

=yang		=nang
kuthumung 'see'	*biilang* 'say'	*puukul* 'hit'
encokang 'fool'	*baa* 'bring'	*theeṁbak* 'shoot'
diyath 'see'	*thaaro* 'put'	*thiikam* 'stab'
poothong 'cut'	*muuji* 'venerate'	
admit-kang 'admit'	*angkath* 'lift'	
ambel 'take'	*caaṅda* 'meet'	
panggel 'call'	*buunung* 'kill'	
uubar 'chase'	*kiiring* 'send'	
thaanàm 'plant'	*preksa* 'interrogate'	
iingath 'think'	*salbakang* 'save'	
sakithkang 'hurt'	*caari* 'search/find'	
luppas 'let go'	*rubbus* 'boil'	
abbisking 'finish(trs)'		

Table 15.1: Transitive verbs which govern the accusative =*yang* or the dative =*nang* for the undergoer argument.

(19)   *Ithu=kapang lorang=pe leher=**yang** kithang athi-**poothong**.*
    DIST=when   2PL=POSS   neck=ACC   1PL   IRR=cut
    'Then we will cut your neck.' (K051213nar06)

(20)   *Derang hathu papaaya=**yang** asà-**poothong**.*
    3   INDEF   papaya=ACC   CP-cut
    'They cut open a papaya.' (K051220nar01)

It appears that dative marking is more likely for activities affecting humans or other animate entities, like hitting, stabbing and shooting, whereas more general-purpose verbs like *ambel* 'take' and *thaaro* 'put' govern =*yang*. One could also speculate that there is a shade of recipient meaning assigned by these verbs (*receive a blow, shot, stab (wound)*), which might condition the dative marking. Table 15.1 gives an overview over different verbs.

If there is no risk of confounding agent and patient, the patient is commonly not marked by a postposition. This is most notably the case

- if the patient is much lower on the animacy hierarchy than the agent,

- if the patient has plural reference,

- if the action performed on the patient does not affect him much (see Theme below),

-yang	+anim	+pl	+aff	+indiv	+top	+exp	+sent
	(23)	(22)	(21)	(21)	(25)	(23)	(27)
				(23)		(26)	
				(26)		(27)	
-yang	-ANIM	-PL	-aff	-indiv	-top	-exp	-sent
	(21)	(21)	(22)	(22)	(21)	(21)	(21)
	(22)	(23)	(23)	(24)	(22)	(22)	(22)
	(24)	(24)	(24)	(25)	(23)		(23)
	(25)	(25)	(26)		(24)		(26)
	(26)	(26)			(26)		

Table 15.2: Zero-marking of the patient as a function of a number of factors. Numbers refer to the examples below.

- if we are dealing with a sentential complement,

- if the patient is not individuated,

- if the patient is not topical,

- if the other argument is experiencer rather than agent,

Very often, the factors given above conspire, and the occurrence of =*yang* in a sentence can be argued for or against based on more than one of these factors. Table 15.2 gives an overview of which examples can be used to illustrate the factors given above.

(21) *Itthu=kapang baapa derang=pe kubbong=ka hatthu pohong]*_{PAT}=∅

    DIST=when   father 3PL=POSS estate=LOC INDEF tree

    nya-poothong.

    PAST-cut

    'The father felled a tree in their estate.' (K051205nar05)

(22) *[Oorang anà-baava samma thoppi=pada]*_{PAT}=∅ *asà-ambel.*

    man   PAST-bring all   hat=PL   CP-take

    '(The monkeys) took all the hats the man had brought.' (K070000wrt01)

(23) *Ithukapang oorang pada [hathu oorang]*_{PAT}=∅ *asà-kuthumung.*

    then   man PL INDEF man   CP-see

    'Then the men saw a man.' (K051220nar01)

(24) *Baapa=le   aanak=le [guula]_{PAT}=∅ su-maakang.*
      father=ADDIT son=ADDIT sugar   PAST-eat
      'Father and son ate sugar.' (K070000wrt02)

(25) a. *Itthu=nang [aayer]_{PAT}=∅ asà-thaaro=apa*
          DIST=DAT   water   CP-put=after
          'Having added water to it'

      b. *[itthu aayer]_{PAT}=yang baaye=nang arà-boil-kang.*
          DIST   water=ACC   good=DAT   NON.PAST-boil-caus
          'boil the water well.'

(26) *[Kìrras pinthu=nang arà-thatti   hathu svaara]_{PAT}=∅ su-dìnngar.*
      strong door=DAT   SIMULT-hammer INDEF noise]   PAST-hear
      'They heard a noise of hard hammering at the door.' (K070000wrt04)

(27) *Blaakang=jo incayang anà-kuthumung [moonyeth pada thoppi asà-ambel*
      after=EMPH   3S.POLITE PAST-see   monkey   PL   hat   CP-take
      pohong atthas=ka arà-maayeng]_{PAT}=∅.
      tree   top=LOC   SIMULT-play
      'Then only he saw that the monkeys had taken his hats and were playing on
      the top of the trees.' (K070000wrt01)

To illustrate several factors conspiring, *thoppi* in example (22) is at the same time
unindividuated, plural in reference, low on the animacy hierarchy and relatively unaf-
fected (cf. Table 15.2).

If on the other hand, there is a risk of confounding agent and patient, =*yang* is
used more often. This is notably the case under the following conditions.

- the agent is less animate than the patient (28)

- the patient is individuated (29).

(28) *Aajuth=**yang** buurung mà-angkath baapi   su-diyath.*
      dwarf=ACC   bird   INF-lift   take.away PAST-try
      'The bird tried to carry the dwarf away.' (K070000wrt04)

In this example, the dwarf is comparatively little affected, but he is individuated,
and, more importantly, the normal order of more animate entities acting upon less
animate ones is inverted here. Again these factors conspire very often, and in example

(29) the speaker and his persecutors are a) close on the animacy hierarchy, and b) the speaker is individuated. Both these conditions favour the use of =*yang*.

(29) *Se=ppe oorang pada [see saapa] katha thàrà-thaau subbath* **see=yang**
    1s=POSS man   PL   1s   who   QUOT NEG-know   because 1s=ACC
    su-uubar.
    PAST-chase
    'Because my folks did not know who I was, they chased me.' (K070000wrt04)

It should be noted that no hard and fast rule for the occurrence of =*yang* can be given; its occurrence depends on multiple factors, and even then there are unexplainable sentences. An examples is (30) where two things are bought, but one is marked with =*yang* and the other one is not.

(30) *Lai   se* computer=*nang baaru* optical mouse *atthu=∅=le* Encarta2006
    other 1s computer=DAT   new   optical mouse INDEF=ADDIT Encarta2006
    software=**yang**=*le*   *su-bìlli.*
    software=ACC=ADDIT PAST-buy
    'Then I also bought a new optical mouse and the Encarta 2006 software for the computer.' (Letter 26.06.2007)

The semantic factors sketched above have an influence, but also idiolectal and possibly style factors. For every simple determinant factor for the occurrence of =*yang*, counterexamples can be found, but a combination of several of the criteria mentioned above could possibly yield good predictions. Such a multivariate analysis is beyond the scope of this grammatical description.

The affectedness of a patient can additionally be highlighted by the vector verb *thaaro* ⊡→5.1.5.4, p. 185, as in the following examples.

(31) *Incayang=yang siaanu asà-buunung* **thaaro**=*apa.*
    3s.POLITE=ACC   3s.prox CP-kill   put=after
    'This one has killed him.' (K051220nar01)

(32) *See=yang dhaathang* remand=*ka mà-thaarek* **thaaro**=*nang thàràboole*
    1s=ACC   come   remand=LOC INF-pull   put=DAT   cannot
    su-jaadi.
    PAST-become
    'It became impossible to remand me.' (K061122nar03)

Smith et al. (2004) claim that accusative (patient) and dative (recipient) are conflated in SLM, but the base for their claim is unclear (Ansaldo 2005b). The forms =*yang* and =*nang* are clearly distinct (with the exception of a very infrequent =*nya*,

which might be an allomorph of either, and for whose occurrence no explanation can be given as of now). The verbs also clearly subcategorize either for the one or the other, so that there are neither morphophonological nor distributional reasons to assume a conflation of patient marking and recipient marking.

15.1.2.3   Theme

Themes are either marked by ∅ or by *=yang*.

(33)   *Se=ppe baapa incayang=nang ummas=∅ su-kaasi.*
       1s=POSS father 3s.POLITE=DAT   gold   PAST-give
       'My father gave him gold.' (K070000wrt04)

(34)   *Derang=pe pàrhaal pada=**yang** mà-thuulis ambel=nang.*
       3PL=POSS   problem PL=ACC   INF-write   take=DAT
       'To write down their problems.' (K051213nar01)

Since themes are by definition less affected than patients, marking with *=yang* is less common. Themes are never marked by *=nang*. Themes of utterance verbs can also be marked by *atthas* 'about', as in (35).

(35)   *Se=ppe **atthas** laskalli masà-biilang=si?*
       1s=POSS about   again   must-say=INTERR
       'Do I have to tell about myself again?' (B060115prs05)

An interesting case is (36), where the theme is marked by *=yang*, possibly to underscore the individuated nature and anthropomorphic image of the cow venerated by Hindus. I have tried to render this by capitalization in English.

(36)   *Hindu pada sampi=**yang** arà-muuji.*
       Hindu PL   cow=ACC   NON.PAST-venerate
       'Hindus venerate the Cow' (K060112nar01)

15.1.2.4   Recipient

Recipients are always coded with the postposition *=nang*. (37) gives an example.

(37)   *Se=ppe baapa incayang=**nang** ummas su-kaasi.*
       1s=POSS father 3s.POLITE=DAT   gold   PAST-give
       'My father gave him gold.' (K070000wrt04)

15.1.2.5 Experiencer

The semantic role of experiencer is indicated by the dative marker =*nang* (Ansaldo 2005b, 2008) ⬚→6.4.4.2, p. 334, as common in South Asia (Masica 1976, Sridhar 1976a,b, 1979, Verma & Mohanan 1990, Abbi 1994, Bhaskararao & Subbarao 2004a,b). The following sentences give examples for experiencers of mental predicates (38)-(41), sensory predicates (42)(43), and bodily predicates (44).

(38) Tailoring go**dang** *baaye=nang* **thaau**$_{mental}$.
      tailoring 1s=DAT good=DAT   know
      'I know tailoring very well.' (B060115nar04)

(39) *Inni oorang=**nang** itthu **thàrà-thaau**$_{mental}$.*
      PROX man=DAT   DIST NEG-know
      'This man did not know that.' (K070000wrt01)

(40) *Derang pada=**nang** karang Malay arà-**luupa**$_{mental}$.*
      3PL   PL=DAT   now   Malay NON.PAST-forget
      'They are forgetting (their) Malay now.' (G051222nar02)

(41) *Suda Andare=**nang**=le   buthul **suuka**$_{mental}$.*
      thus Andare=DAT=ADDIT very   like
      'So Andare also liked it a lot.' (K070000wrt05)

(42) *Derang=**nang** byaasa svaara hatthu su-**dìnngar**$_{sensory}$.*
      3PL=DAT   habit noise INDEF PAST-hear
      'They heard a familiar voice.' (K070000wrt04)

(43) *Itthu haari=ka=jo   aanak pompang duuva=**nang** ... kiccil jillek Aajuth*
      DIST day=LOC=EMPH child female   two=DAT   ... small ugly dwarf
      *hatthu=yang su-**kuthumung**$_{sensory}$*
      INDEF=ACC   PAST-see
      'On that very day, the two girls perceived a small ugly dwarf.' (K070000wrt04)

(44) *Go=**dang**   karang bannyak **thàràsìggar**$_{bodily}$.*
      1s.FAMILIAR=DAT now   lot   sick
      'I am very sick now.' (B060115nar04)

The fact that the modals *maau* 'want', *thussa* 'NEG.want', *boole* 'can' and *thàr-*

boole 'cannot' mark the argument can also be explained by the fact that the person having the desire or ability (or lack thereof) is normally not actively responsible for that in the moment of speaking (i.e. lacks control, cf. Ansaldo (2005b).). True, a person who can drive a car has actively acquired that knowledge at some point in time, but at the time of speaking, this person is passive.

15.1.2.6 Source

Source can be either marked by *=dering* (45) ⊡→6.4.4.4, p. 342 or the conjuctive participle of the animate existential *(a)s(à)duuduk* ⊡→5.1.2.2, p. 168 (46).

(45) a. *Kitham=pe oorang thuuva pada bannyak dhaathang aada*
          1PL=POSS man old PL many come exist
          *Malaysia=**dering**;*
          Malaysia=ABL
          'Many of our ancestors came from Malaysia;'

      b. *Spaaru Indonesia=**dering** dhaathang aada.*
          some Indonesia=ABL come exist
          'some came from Indonesia.' (K060108nar02)

(46) *Suda see* Trinco=*ka*   **asàduuduk** *Kluum̆bu=nang su-dhaathang.*
      So 1s Trincomalee=LOC from   Colombo=DAT PAST-come
      'So I went from Trincomalee to Colombo.' (K051206nar20)

Since *asàduuduk* is the conjunctive participle of the animate existential verb *duuduk*, it can only be used to indicate the source of animate entities. The stones in (45) are not animate, and *asàduuduk* cannot be used there to indicate the origin of the stones.

The following examples show some more uses of *=dering* to indicate source

(47) *Hathu haari, hathu oorang thoppi mà-juval=nang kampong=**dering***
      INDEF day INDEF man hat INF-sell=DAT village=ABL
      kampong=nang su-jaalang pii.
      village=DAT PAST-walk go
      'One day, a man walked from village to village to sell hats.' (K070000wrt01)

(48) *See asà-retire aada police=**dering**.*
      1s CP-retire after police=ABL
      'Having retired from the police'

(49) Kandy Malay Association=***dering*** *hatthu hatthu oorang pada arà-lompath*
Kandy Malay association=ABL   INDEF INDEF man   PL   NON.PAST-jump
Hill Country=*nang.*
hill country=DAT
'More and more people stepped over from the KMA to the Hill Country Malay Association.' (K060116nar07)

The use of *(a)s(à)duuduk* is exemplified by the sentences below.

(50) *Kithan nya-pii Anuradhapura=dang, Katunaayaka **sduuduk**.*
1PL   PAST-go Anuradhapura=DAT   Katunayaka   from
'We went to Anuradhapura, from Katunayaka [airport].' (K051206nar16)

(51) *Katugastota **asàduuduk** St.Anthony's=nang arà-jaalang*
Katugastota from   St.Anthony's=DAT   NON.PAST-walk
'I walk from Katugastota (train station) to St. Anthony's (school).' (K051201nar02)

15.1.2.7 Goal

Goal is normally coded by =*nang* ⊡→6.4.4.2, p. 334 (Ansaldo (2005b:19,22),Ansaldo (2008:24)), which attaches to the place name (52), the spatial entity (53) or a relator noun (54), but deictics do not take this marker when used as goal (55).

(52) *Soojer   pada incayang=sàsaama Seelon=**nang** asà-dhaathang.*
European PL   3s.POLITE=COMIT   Ceylon=DAT CP-come
'The Europeans came to Sri Lanka together with him.' (K060103nar01)

(53) *Suda lorang=yang ruuma=**nang** anthi-aaji.baapi*
thus 2PL=ACC   house=DAT   IRR=bring.ANIM
'So I will take you to my house.' (K070000wrt04)

(54) *Ithukang ithu bambu giithu=jo   luvar=**nang** arà-dhaathang.*
then   this bamboo like.that=EMPH outside=DAT NON.PAST-come
'Then the bamboo comes out like that.' (K061026rcp04)

(55) *Se=ppe* profession=*subbath se=dang siini=∅ mà-pii su-jaadi.*
      1s=POSS profession=because 1s=DAT here   INF-go PAST-become
      'I had to come here because of my profession.' (G051222nar01)

If the goal of motion is human, the relator noun *dìkkath* is obligatory, as shown in
(56) (cf. Smith et al. 2004).

(56) *Kithang=nang hatthu* job *hatthu mà-ambel=nang kithang=nang hatthu*
      1PL=DAT   INDEF job INDEF INF-take=DAT   1PL=DAT   INDEF
      application *mà-sign kamauvan vakthu=nang=jo kithang arà-pii   inni*
      application INF-sign want   time=DAT=EMPH 1PL   NON.PAST-go PROX
      politicians *pada* **dìkkath=nang**.
      politicians PL   vicinity=DAT
      'When we want to take a job, when we want to sign an application, we ap-
      proach these politicians' (K051206nar12)

=*nang* can attach to proper nouns referring to non-human entities, even if they are
not spatial in the strict sense of the term, such as in (57), where the goal of motion is
an association.

(57) Kandy Malay Association=*dering hatthu hatthu oorang pada arà-lompath*
      Kandy Malay association=ABL   INDEF INDEF man   PL   NON.PAST-jump
      Hill Country=*nang*.
      hill country=DAT
      'More and more people stepped over from the KMA to the Hill Country Malay
      Association.' (K060116nar07)

Goal of motion can also be marked by the locative marker =*ka* (58)(59) (Smith
et al. 2004).

(58) Government=*pe   hathu* **thumpath=*ka*** *asà-pii pukurjan bole=girja*
      government=POSS INDEF place=LOC   CP-go work   can-make
      'They can go to a government place and work there.' (K051222nar05)

(59) *Derang samma oorang hatthu hatthu* **thumpath pada=*ka*** *asà-pii pukurjan*
      3PL   all   man   INDEF INDEF place   PL=LOC   CP-go work
      su-gijja
      PAST-make
      'All those people go to one place or another and work.' (B060115cvs06)

15.1.2.8 Path

Path is coded by =*dering* ⊡→6.4.4.4, p. 342. In the following example, the bird's trajectory has an unknown source and an unknown goal, but leads over the girls' heads. This is indicated by the ablative marker =*dering*.

(60) *Bìssar hathu buurung derang=pe atthas=**dering** su-thìrbang.*
     big   INDEF bird   3PL=POSS top=ABL   PAST-fly
     'A big bird flew over them.' (K070000wrt04)

15.1.2.9 Instrument

Instrument is indicated by =*dering* ⊡→6.4.4.4, p. 342. The instrument can be used by a sentient being, as in (61), but =*dering* can also be used in a wider sense for participants not manipulated by a sentient being as in (62).

(61) *Thaangang=**dering** bukang   kaaki=**dering** masà-maayeng.*
     hand=ABL   NEG.NONV leg=ABL   must-play
     'You must play not with the hands, but with the feet.' (N060113nar05)

(62) ***Daavon=dering*** *thuuthup ada gaaja   hatthu asdhaathang.*
     leaf=**instr**   close   exist elephant INDEF CP-come
     'An elephant, which had been hidden by leaves, appeared.' (B060115nar05)

The semantic role of instrument is normally always marked overtly. An exception to this rule is found in (63), where the ears could be said to be the instruments of hearing, yet they are not marked by =*dering*. Since ears are so prototypically associated with hearing, marking of their semantic role does not seem necessary.

(63) *Andare ruuma=nang asà-pii biini=nang su-biilang puthri=nang **kuuping***
     Andare house=DAT   CP-go wife=DAT PAST-say queen=DAT ear
     arà-dìnngar   kuurang katha.
     NON.PAST-hear little   QUOT
     'Andare went home and said to his wife that the queen did not hear well (lit.: that the queen heard little with her ears).' (K070000wrt05)

15.1.2.10 Beneficiary

Beneficiary is normally coded by =*nang* ⊡→6.4.4.2, p. 334 (64)(65) (Ansaldo 2005b, 2008), but can also be marked with a relator noun *bagiyan* 'behalf'. As with the other

roles, the overt marking can be dropped if the role is clear from context (66). A further possibility is the use of the vector verbs *kaasi* 'give' or *ambel* 'take'.

(64)   *Cinggala raaja=**nang** deram pada banthu aada.*
      Sinhala   king=DAT   3PL   PL   help   exist
      'They have helped the Sinhalese king.' (K051206nar04)

(65)   *Derang=pe umma derang=**nang** jaith-an=le,   jaarong pukurjan=le*
      3PL=POSS   mother 3PL=DAT   sew-NMLZR=ADDIT needle work=ADDIT
      su-aajar.
      PAST-teach
      'Their mother taught them sewing and needle work.' (K070000wrt04)

(66)   *Kettha drampada=∅ bannyak banthu.*
      1PL   3PL   much   help
      'We helped them a lot.' (B060115nar02)

Beneficiary can also be added to another predication which does not require a beneficiary strictly speaking, if the action turned out to be beneficial. This is shown in (67).

(67)   *Giini   duuduk bannyak [**kithang=pe oorang pada**]=nang anà-caape.*
      like.this stay   much   1PL=POSS   man   PL=DAT   PAST-tired
      'Being there, he toiled a lot for our people.' (N061031nar01)

The beneficial nature of an action can additionally be highlighted by the vector verb *kaasi* 'give' as in (68). Note that the benefiting persons are not overtly realized because they have been established as a topic before.

(68)   *Itthu muusing Islam igaama nya-**aajar kaasi** 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)

It is also possible to overtly realize the benefitting participant when *kaasi* is used, as in (69).

(69) *Kithang=pe ini* younger generation=*nang=jo konnyong masà-**biilang***
1PL=POSS PROX younger generation=DAT=EMPH few must-say
***kaasi**, masà-aajar.*
give must-teach
'It is to the younger generation that we must explain it, must teach it.'
(B060115cvs01)

If the action is beneficial to the agent itself, instead of *kaasi* 'give', *ambel* 'take' is used.

(70) *[Tony Hassan* uncle=*nang asà-kaasi duvith] athi-**mintha ambel**=si?*
Tony Hassan uncle=DAT CP-give money IRR-ask take=INTERR
'Shall I ask for the money you gave to uncle Tony Hassan?' (K071011eml01)

In this case, the person performing the action of asking would also profit from it since he has a chance of getting the money.

In rare cases, *ambel* can also be used if the action profits other entities than the agent, as in (71).

(71) *See=yang lorang=susamma diinging muusing sangke-habbis anà-**simpang***
1s=ACC 2PL=COMIT cold season until-finish PAST-keep
***ambel**.*
take
'You have kept me together with you until the cold season was over.'
(K070000wrt04)

In this case of a family providing shelter for a bear, it is the bear who profits, not the family, but still *ambel* is used in this sentence.

The following four examples show the use of *ambel* and *kaasi* for indicating self-benefactive (73) or alter-benefactive (75). Note that the use of the vector verbs implies control, which renders the use of the dative impossible in sentences with these vector verbs.

(72) *Se=dang su-mirthi.*
1s=DAT PAST-understand
'I understood.' (K081106eli01)

(73) *Se su-mirthi ambel.*
1s PAST-understand take
'I understood/I could understand.' (K081106eli01)

(74) *Se su-mirthi-king.*
      1s PAST-understand-CAUS
      'I made him understand.' (K081106eli01)

(75) *Se su-mirthi-king   kaasi.*
      1s PAST-understand-CAUS give
      'I made him understand.' (K081106eli01)

15.1.2.11 Comitative

The comitative is coded by the postposition *(=sà)saama* ⊡→6.4.4.9, p. 352 (76).

(76) *Se=ppe mma-baapa=le=**sàsaama**=jo   arà-duuduk.*
      1s=POSS mother-father=ADDIT=COMIT=EMPH NON.PAST-stay
      'It is together with my parents that I live.' (B060115prs06)

Very often, simple coordination with *=le* ⊡→14.3.2, p. 548 is used instead of the comitative.

15.1.2.12 Purpose

Purpose is expressed by either an infinitive clause ⊡→13.4, p. 532 (77)(78), or the postposition *=nang* ⊡→6.4.4.2, p. 334 (79)(80), or a combination thereof (81)-(83).

(77) *Blaakang Andare [Kandi=ka asduuduk Dikwella arà-pii   jaalang]=ka*
      after   Andare Kandy=LOC from   Dikwella NON.PAST-go road=LOC
      *aayer **mà**-miinong Udamalala kampong=ka su-birthi.*
      water INF-drink   Udamalala village=LOC PAST-stop
      'Then, Andare stopped on the street which leads from Kandy to Dikwella at the hamlet Udamalala to drink water.' (K070000wrt03)

(78) *Arà-blaajar   mlaayu ini   buk **mà**-thuulis kiyang.*
      NON.PAST-learn Malay PROX book INF-write   EVID
      '(Sebastian) is learning Malay to write this book of his, it seems.' (K051222nar07)

(79) *De   laaye hathu nigiri=nang anà-baapi, buunung-king=**nang**.*
      3s.IMPOLITE other INDEF country=DAT PAST-bring kill-CAUS=DAT
      'They brought him to another country to have him executed.' (K051206nar02)

(80) *Karang masiigith=nang arà-pii   liima vakthu sbaayang=**nang**.*
now   mosque=DAT   NON.PAST-go five time pray=DAT
'now he goes to the mosque five times (a day) to pray.' (K051220nar01)

(81) *Freedom=yang **mà**-daapath=**nang** kithang=nang bannyak caape aada.*
freedom=ACC INF-get=DAT   1PL=DAT   much tired exist
'We have worked hard to get our freedom.' (N061031nar01)

(82) *Hathu haari, hathu oorang thoppi **mà**-juval=**nang** kampong=dering*
INDEF day INDEF man   hat   INF-sell=DAT   village=ABL
kampong=nang su-jaalang pii.
village=DAT PAST-walk go
'One day, a man walked from village to village to sell hats.' (K070000wrt01)

(83) *Derang [dìkkathka   aada laapang]=nang **mà**-maayeng=**nang** su-pii.*
3PL   vicinity=LOC exist ground=DAT   INF-play=DAT   PAST-go
'They went to play on the ground nearby.' (K070000wrt04)

These strategies can be found in the same idiolects. The following example shows the use of *mà-* with and without *=nang* in purposive clauses in one stretch of discourse by one speaker.

(84) a. *Kandi=pe   raaja=nang kitham=pe inni banthu-an asà-kamauvan*
       Kandy=POSS king=DAT   1PL=POSS PROX help-NMLZR CP-want
       se-aada.
       PAST-exist
       'The Kandyan king wanted this help of ours'
  b. *Inni raaja=yang **mà**-jaaga=∅.*
       PROX king=ACC   INF-protect
       'to protect this king.'
  c. *[Itthu **mà**-jaaga=**nang** anà-baa   mlaayu]=dering satthu oorang=jo se.*
       DIST INF-protect=DAT PAST-bring Malay]=ABL   one   man=EMPH 1s
       'One of the Malay men brought to protect that (man) is me.'
       (K060108nar02)

While purpose is most often marked on clauses, it is also possible to mark it on NPs, as is the case in the following example, where the noun *thakuthan* 'fear' hosts the postposition *=nang*.

(85) Second world war time=*ka   Kluum̆bu=nang* Japanese *arà*-bomb-*king*
      second world war time=LOC Colombo=DAT   Japanese NON.PAST-bomb-CAUS
      [*thakuth-an*]=***nang***.
      fear-NMLZR=DAT
      'During the second world war, the Japanese bombed Colombo to cause fear.'
      (N060113nar03)

15.1.2.13   Cause and reason

Cause is normally coded by the postposition *subbath* ⊡→6.4.4.6, p. 348.

(86) [*Non-Muslims pada*]=***subbath*** *kithang muuka konnyong arà-cunjikang*
      non-Muslims PL=because   1PL   face   little   NON.PAST-show
      siini.
      here
      'Because of the non-Muslims we do not wear the veil.' (K061026prs01)

(87) *Lorang [see]=**subbath** ithu Aajuth=yang su-salba-king.*
      2PL   1s=because   DIST dwarf=ACC   PAST-safe-CAUS
      'You saved that dwarf because of me.' (K070000wrt04)

(88) [*Ini oorang giini   kapang-jaalang pii caape*]=***subbath*** *jaalang hathu*
      PROX man   this.way then-walk   go tired=because   road   INDEF
      piiňğgir=ka anà-aada   hathu pohong baava=ka su-seeňder.
      border=LOC PAST-exist.inanim INDEF tree   down=LOC PAST-rest
      'Because he was tired from walking then, this man sat down under a tree which
      stood at the side of the street.' (K070000wrt01)

(89) *Suda **itthusubbath**=jo, se laile Marine Engineering asà-kijja ambel*
      so   therefore=EMPH 1s again Marine Engineering CP-make take
      arà-pii.
      NON.PAST-go
      'So, it was because of that that I took up again the Marine Engineering work
      and went (away).' (K051206nar20)

Another possibility is the use of *lanthran* 'because' ⊡→6.4.4.12, p. 356 (90)(91).

(90) *Derang hathu suurath nya-kiiring [see ini   Kandi Mlaayu*
   3PL   INDEF letter   PAST-send 1s   PROX Kandy Malay
   Association=*dering nya-kiisar]=***lanthran**.
   association=ABL   PAST-go.aside=**because**
   'They had written a letter because I had left the Kandy Malay Association.'
   (K061122nar03)

(91) a. *Beeso   luusa   lubaarang arà-dhaathang.*
      tomorrow later.in.the.future festival   NON.PAST-come
      'The day after tomorrow is the festival.'

   b. *Itthu=***lanthran** *kithang=pe ruuma see arà-cuuci.*
      DIST=because 1PL=POSS   house 1s NON.PAST-clean
      'That's why I am cleaning the house.' (K061019prs01)

Finally, a third possibility is =*sikin* ⊡→6.4.4.11, p. 355.

(92) *Inni aari pada=ka kithang=nang* test. *Itthu=***siking**=*jo   see thàrà-kiiring*
   PROX day PL=LOC 1PL=DAT   test DIST=because=EMPH 1s NEG.PAST-send

   'We are having tests these days. That is why I have not sent (anything).'
   (K071203eml01)

15.1.2.14 Value

The value of an item is indicated by =*nang* ⊡→6.4.4.2, p. 334. In (93), the value of the stones indicated earlier in discourse is said to be *baaye lakuvan* 'a good price'. This NP carries =*nang*.

(93) *[Baaye lakuvan]=***nang** *anà-juuval.*
   good   wealth=DAT   PAST-sell
   '(He) sold (them) for a good price.' (060103nar01)

15.1.2.15 Portion

When something is divided into portions, the fraction is also indicated by =*nang* ⊡→6.4.4.2, p. 334, as is the case in dividing meat into three portions in the example below

(94) *Thapi ithu **thiiga=nang** arà-baagi.*
      but   DIST three=DAT   NON.PAST-divide
      'But you divide it into three (parts). (K060112nar01)'

15.1.2.16   Set domain

The superset among which a member is chosen is indicated in English by *among*. In
SLM, this is either indicated by the locative *=ka* ⊡→6.4.4.3, p. 340 (95) or the ablative
=dering ⊡→6.4.4.4, p. 342 (96)-(98).

(95) ***Mlaayu=ka**=jo   bannyak avuliya Seelong=ka aada.*
      Malay=LOC=EMPH many   saint   Ceylon=LOC exist
      'It is among the Malays that there are many saints in Sri Lanka.'
      (K060108nar02)

(96) a. *Karang inni Kandy nigiri=ka aada mlaayu avuliya pada.*
          now   PROX Kandy town=LOC exist Malay saint   PL
          'Now there are Malay saints here in Kandy.'
      b. *Derang=**dering** hatthu avuliya dhaathang aada sini=ka   dìkkath.*
          3PL=ABL   INDEF saint   come   exist here=LOC vicinity
          'Among them there is one saint close to here.' (K051220nar01)

(97) *[Itthu mà-jaaga=nang anà-baa   mlaayu]=**dering** satthu oorang=jo see.*
      DIST INF-protect=DAT PAST-bring Malay=ABL   one   man=EMPH 1s
      'One of the Malays brought to protect him is me' (K060108nar02)

(98) *Seelong=pe makanan=**dering** aapa makanan suuka?*
      Ceylon   food=ABL   what food   like
      'What food do you like within Sri Lankan cuisine?' (B060115cvs02)

15.1.2.17   Temporal domain

The temporal domain in which an event occurs is indicated by *=nang* ⊡→6.4.4.2,
p. 334. In the following examples, this is the week and the day. There is usually some
quantifying element elsewhere in the clause indicating the relation, like *duuva skalli*
'two times' or *ùmpathblas kaayu* 'fourteen miles'.

(99) *Suda **hathu** **week=nang** duuva skali arà-dhaathang* daughter.
 so one week=DAT two time NON.PAST-come daughter
 'Thus my daughter comes twice a week.' (K051201nar01)

(100) *Kithang **hathu** **week=nang** hathu skali duva skaali=ke arà-maakang.*
 1PL one week=DAT one time two time=SIMIL NON.PAST-eat
 'We might eat it once or twice a week.' (K061026rcp04)

(101) *Derang su-jaalang ùmpath-blas kaayu **hathu aari=nang.***
 3PL PAST-walk four-teen mile one day=DAT
 'They walked fourteen miles a day.' (K051213nar03)

15.1.2.18 Duration

Duration is normally not overtly marked. In the following two examples, the time spans *thuuju thaaun* 'seven years' and *spuulu thaaun* 'ten years' do not receive any special marking.

(102) *See [**thuuju thaaun**]=∅ luvar nigiri=ka asà-duuduk karang abbis*
 1s seven year outside country=LOC CP-exist.ANIM now finish
 dhaathang aada.
 coming exist
 'After having stayed abroad seven years, I have now arrived.' (B060115prs13)

(103) *See [**spuulu thaaun**]=∅ siini sri Lanka=ka pukurjan nya-kirja.*
 1s ten year here Sri Lanka=LOC work PAST-do
 'I worked here in Sri Lanka for ten years.' (K061026prs01)

If the duration shall be emphasized, *=dering* ⊡→6.4.4.4, p. 342 can be used.

(104) *[Bannyak aari]=**dering** saapa=yang=ke thàrà-enco-kang.*
 many day=ABL who=ACC=SIMIL NEG.PAST-fool-CAUS
 'For how many days have I not fooled anybody!' (K070000wrt05a)

15.1.2.19 Role

The semantic role of 'role', e.g. for profession or functions, is indicated by the noun *caara* 'way'.

(105)   *Seelong=nang* exiles ***caara*** *nya-kiiring.*
        Ceylon=DAT   exiles way   PAST-send
        'The soldiers were sent to Sri Lanka as exiles.' (K051213nar06)

15.1.2.20   Summary

Table 15.3 gives an overview of the distribution of semantic roles on different morphemes

15.1.3   Mismatches between number of semantic roles and number of syntactic arguments

It is possible that the predicate demands more semantic roles than there are participants in the discourse. This is most notably the case for reflexives ⊙→15.1.3.1, p. 582, self-benefactives ⊙→15.1.3.2, p. 584 and reciprocals ⊙→15.1.3.3, p. 586. The opposite is that a semantic role applies to more than one participant. In that case, the participants are coordinated ⊙→15.1.3.4, p. 588.

15.1.3.1   Reflexive

If one term has more than one role (as is the case of reflexives), there exist three possibilities: Use the vector verb *ambel* ⊡→5.1.5.1, p. 176, use the emphatic clitic =*jo*⊡→6.4.7.1, p. 378, or separate the term into two non-identical subterms. These possibilities will be discussed in turn.

   The first possibility is to use the vector verb *ambel* to express reflexive action, as in (106) and (107). The case assigned by the main verb remains unaffected by this (accusative by *buunung* 'kill', dative by *banthu* 'help'). The emphatic marker =*jo* can be used optionally.

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

(107)   *Incayang incayang=**nang**=(jo) su-banthu **ambel**.*
        3S.POLITE 3S.POLITE=ACC=EMPH PAST-help take
        'He helped himself.' (K081106eli01)

   The second possibility, already alluded to above, is the use of the emphatic clitic =*jo*. This is obligatory in contexts where *ambel* is not possible, for instance in some possessive contexts (108)(109). The emphatic clitic can attach either to the possessor
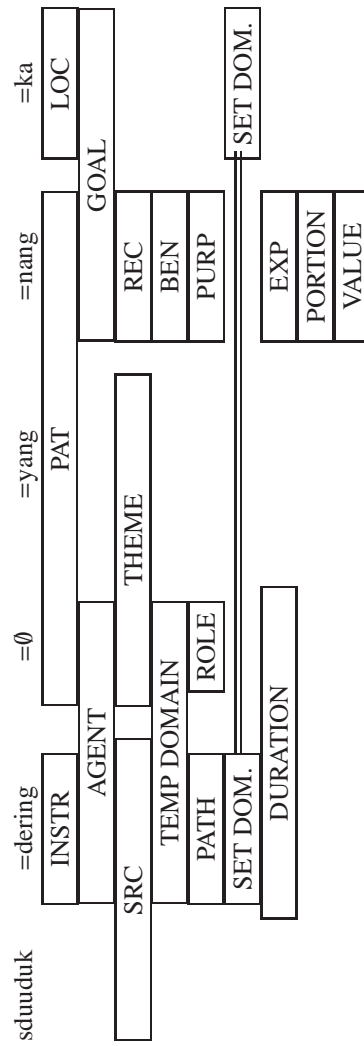
Table 15.3: Repartition of semantic roles on morphemes. =*nang* is used to express eight different semantic roles, =*dering* is used for seven, while the other morphemes have lower numbers of semantic roles they can express. Eight semantic roles can be expressed by more than one morpheme. The two morphemes which can be used for the role SET DOMAIN are linked by a line because for typographical reasons they could not be made contiguous. The semantic roles of COMITATIVE, ROLE, and REASON are left out of the graphic.

or to the possessee, entailing slight changes in meaning, rendered by 'he himself' and 'his own' in English.

(108) ***Incayang=pe=jo***   *ruuma=yang incayang su-ronthok-king.*
      3s.POLITE=POSS=EMPH house=ACC   3s.POLITE PAST-demolished-CAUS
      'He demolished his own house.' (K081106eli01)

(109) *Incayang=pe ruuma=yang **incayang=jo***   *su-ronthok-king.*
      3s.POLITE   house=ACC   3s.POLITE=EMPH PAST-demolished-CAUS
      'He himself demolished his house.' (K081106eli01)

The third possibility is to use an additional term to express the undergoer in a more precise manner, thus making it different from a real reflexive construction. The additional term is normally in a meronymic relationship to the main term, like *diiri* 'body' to the speaker in (110) and (111). The normal case markers are used in this construction as assigned by the verb. Even in this case, *ambel* and *=jo* are normally present.

(110) *Se se=ppe diiri=yang(=jo) su-poothong ambel.*
      1s 1s=POSS body=acc(=foc) past-cut   take
      'I cut myself.' (K081106eli01)

(111) *Se se=ppe diiri=nang(=jo) su-puukul ambel.*
      1s 1s=POSS body=DAT=EMPH PAST-hit   take
      'I hit myself.' (K081106eli01)

15.1.3.2   Self-benefactive

It is possible that an agent performs an action of which he himself benefits. In English, a sentence like *John bought himself a car* comes close to this, since *John bought a car* normally already pragmatically implies, in absence of other context, that John will be the owner of the car. This fact can be highlighted by *himself* in English. In SLM, a different construction is used for actions which profit the agent, namely the vector verb *ambel* ⊡→5.1.5.1, p. 176.

(112) *[Tony Hassan* uncle=*nang asà-kaasi duvith]* athi-mintha ***ambel**=si?*
      Tony Hassan uncle=DAT   CP-give   money IRR-ask   take=INTERR
      'Shall I ask for the money you gave to uncle Tony Hassan (for my own benefit)?' (K071011eml01)

In example (112), the use without *ambel* would be perfectly fine, but the use with

ambel highlights the fact that the speaker would perform an action which would profit himself. It is difficult to argue for this example that it changes the valency of the verb *mintha* 'ask'. The verb is still trivalent (asker, askee, theme), but the emphasis is more on the asker than on the askee or the theme.

Note that overt coding of the self-benefactive is optional, as shown by the following example without *ambel*.

(113) *Derang pada arà-mintha ∅ nigiri.*
     3PL   PL   NON.PAST-ask   country
     'They are asking for a country (of their own).' (K051206nar10)

Ambel can only be used if the beneficiary is actively involved. For *daapath* 'find' in (114), this is not the case. From the context, we do not learn that the speaker was actively involved in getting a new job, hence *ambel* cannot be used. Also note the dative marking on *se* '1s', which indicates lack of control.

(114) *Se=dang baaru hatthu idopan su-daapath (*ambel).*
     1s=DAT new INDEF job   PAST-get   take
     'I got a new job.' (K081106eli01)

If the speaker has found a new job after being actively involved in finding it, *ambel* is possible (115). In this case, nominative (∅) is used instead of the dative, taking away the semantics of lack of control.[4]

(115) *Se=∅ baaru hatthu idopan su-caari   ambel.*
     1s   new INDEF job   PAST-look.for take
     'I found a new job.' (K081106eli01)

On the other hand, if the process of finding a job is not finished yet, *ambel* cannot be used. In (116), we find the non-past marker *arà-*, indicating that at the time of speaking the process of finding/searching is still ongoing. The job has not been found yet, hence no benefit is materialized, and *ambel* cannot be used.

(116) *Se baaru hatthu idopan arà-caari   (*ambel).*
     1s new INDEF job   non.PAST-look.for take
     'I am looking for a new job.' (K081106eli01)

---

[4]The difference between searching and finding is not marked lexically in SLM, but only aspectually. If searching is completed, this implies finding.

15.1.3.3 Reciprocals

In a reciprocal situation, the participants taking part in the reciprocity are both actor and undergoer (or recipient) and their double involvement in the event is indicated by two occurrences of the indefiniteness marker *hatthu* ⬚→6.4.1, p. 319. The following three examples show this for nominal predications.

(117) *Oorang pada **hatthu=nang hatthu** muhabbath.*
   man PL INDEF=DAT INDEF love
   'People like each other.' (K081106eli01)

(118) *Oorang pada **hatthu=nang hatthu** maara.*
   man PL INDEF=DAT INDEF angry
   'People are angry with each other.' (K081106eli01)

(119) *Oorang pada **hatthu=nang hatthu** buthul percaaya.*
   man PL INDEF=DAT INDEF correct trust
   'People (have) trust in each other.' (K081106eli01)

With verbal predications, the vector verb *ambel* can also be used to convey reciprocal semantics (120)(121).

(120) *Ini nigiri=pe oorang pada bannyak arà-buunung **ambel**.*
   PROX country=POSS man PL much NON.PAST-kill take
   'The people in this country kill each other.' (K081106eli01)

(121) *Faarok=le Oomar=le su-buunung **ambel**.*
   Faarok=ADDIT Oomar=ADDIT PAST-kill take
   'Farook and Oomar killed each other.' (K081106eli01)

Both possibilities (*hatthunang hatthu* and *ambel*) can be combined (122).

(122) *Oorang pada **hatthu=nang hatthu** arà-ciong **ambel**.*
   man PL INDEF=DAT INDEF NON.PAST-kiss take
   'People kiss each other.' (K081106eli01)

The verb *ciong* 'kiss' in (122) normally governs the dative, which aligns nicely with the dative case marker =*nang* found on *hatthu* in (122). There are other verbs which govern the accusative (=*yang*), like *buunung* 'kill' in (123). If these verbs are used in the reciprocal construction (124), the =*nang* assigned by reciprocity takes precedence over the =*yang* assigned by the verb. We find *hatthunang hatthu* and not **hatthuyang hatthu*.

(123) *Oorang pompang=**yang** arà-buunung*
      man   female=ACC   NON.PAST-kill
      'The man kills the woman.' (K081106eli01)

(124) *Oorang pada hatthu=**nang** hatthu arà-buunung ambel.*
      man   PL   INDEF=DAT   INDEF NON.PAST-kill take
      'People kill each other.' (K081106eli01)

If through the use of a modal =*nang* is assigned, as in (125), the verb returns to assigning its normal role, accusative in (125).

(125) *Oorang pada=nang hatthu (*nang) hatthu=yang bole-buunung.*
      man   PL=DAT   INDEF DAT   INDEF=ACC   can-kill
      'People can kill each other.' (K081106eli01)

If the reciprocal involves other roles than actor and undergoer, such as recipient in (126), the position of the case marker is on the second item, not on the first as above in (122) and (124).

(126) *Oorang pada hatthu=∅ hatthu oorang=**nang** duvith arà-**kaasi**   **ambel**.*
      man   PL   INDEF   INDEF man=DAT   money NON.PAST-give take
      'People give money to each other.' (K081106eli01)

Note the occurrence of the vector verb *ambel* 'take' together with the full verb *kaasi* 'give' in (126). These normally have opposite semantics, but the semantic content of *ambel* is bleached in this construction, and it fulfills a grammatical function instead of contributing semantic content to the propositions.

If the case marker which would be used in a non-reciprocal clause is neither zero nor =*yang* nor =*nang*, it is omitted in the reciprocal construction. The verb *thaanya* 'ask' normally assigns the locative =*ka* to the askee. This is not possible in the reciprocal construction, and hence there is no case marker found in (127).

(127) *Oorang pada hatthu hatthu arà-thaanya ambel.*
      man   PL   INDEF INDEF NON.PAST-ask take
      'People ask each other (questions).' (K081106eli01)

An exception to this rule might be the ablative case marker =*dering*, which seems to be able to be used prenominally in a reciprocal construction.[5]

---

[5]This is very surprising, given that it is normally a *post*position. This example should be taken with a grain of salt, but might be a worthwhile starting point for future investigations of reciprocity.

(128)   *(dering) oorang pada hatthu hatthu duvith arà-mintha   ambel.*
      ABL   man   PL   INDEF INDEF money NON.PAST-ask take
      'people ask/take money from each other.' (K081106eli01)

The reciprocal must be distinguished from joint reflexive action as in (129).

(129)   *Farook=le   Oomar=le   derang derang=yang su-buunung ambel.*
      Farook=ADDIT Oomar=ADDIT 3   3=ACC   PAST-kill   take
      'Farook and Oomar killed themselves individually.' (K081106eli01)

15.1.3.4   More than one entity in a term

If a term consists of more than one entity, these entities are conjoined by a semantically appropriate coordinating construction ⊡→14.3, p. 544. The zero coordination (130) and the coordination with a clitic (131) are shown below.

(130)   a. *Mlaayu pada duuduk=apa*
        Malay PL   stay=after
        'After the Malays had settled down'
      b. *spaaru mlaayu pada **Singgapur=∅ Indonesia=∅ Malaysia=∅** anà-pii.*
        some Malay PL   Singapore   Indonesia   Malaysia   PAST-go
        '(only) some Malays went (back) to Singapore, Indonesia or Malaysia.'
        (K051213nar07)

(131)   ***Snow-White=nang=le Rose-red=nang=le*** *ini   hatthu=ke*
      Snow.white=DAT=ADDIT Rose.Red=DAT=ADDIT PROX INDEF=SIMIL
      thàrà-mirthi.
      NEG.PAST-understand
      'Snow White and Rose Red did not understand a thing.' (K070000wrt04)

15.1.4   Unknown participants

If a predicate semantically requires an argument, but the precise referential nature of this argument is not known, there are several possibilities: the referent is not important ⊙→15.1.4.1, p. 589, the referent is important and individuated ⊙→15.1.4.2, p. 589, the referent is important and categorial ⊙→15.1.4.3, p. 589, the referent is important and generic ⊙→15.1.4.4, p. 590. A somewhat different possibility is that the speaker wants the hearer to provide the referent, by forming a question ⊙→15.1.4.5, p. 591.

15.1.4.1 Unimportant referent

Unimportant referents which might be required semantically do not have to be expressed in morphosyntax. In example (132), the author of the history book is not known and irrelevant. Therefore, no reference to the author is made.[6]

(132) ∅ *cinggala* history book *atthu thuulis aada.*
    Sinhala history book atthu written exist
    'Someone has written a Sinhala history book (about that)/A Sinhala history book has been written (about that).' (K051213nar06)

15.1.4.2 Individuated referent

If the referent is important and individuated, but further information is not available (like English *somebody, something*), the WH=*so*-construction ⊡→8.4, p. 447 is used to yield an indefinite pronoun.[7]

(133) ***Saapa=so*** Malay exam *arà-girja.*
    who=ADDIT Malay exam NON.PAST-make
    'Someone was taking a Malay exam.' (K060103nar01)

(134) *Thapi **aapacara=so** itthu samma asà-iilang*   *su-aada.*
    But   how=UNDET DIST all   CP-disappeared PAST-exist
    'But somehow it had all disappeared.' (K20070920eml01)

15.1.4.3 Unknown categorial referents

In distinction to individuated referents, which are instantiated by a specific entity in discourse, categorial referents do not refer to a specific entity, but to any member of the indicated category.

In the following example, Andare wants to be dressed like a king. In this case, this does not refer to a specific king, but to any member of the class of kings. This is indicated by the indefinite article *hathu* ⊡→6.4.1, p. 319.

---

[6]It would be possible to read this sentence as 'A Sinhala history book has written something', but since books rarely engage in activities such as writing, no hearer will seriously be tempted by this interpretation.

[7]See Haspelmath (1997:164) for a typological analysis of the combination of an interrogative pronoun with 'or' to yield an indefinite pronoun, which, interestingly, is again a construction shared between SLM and Japanese.

(135)   *[Andare kanabisan=nang anà-mintha] [**hathu** raaja=ke   asà-paake=apa*
        Andare last=DAT   PAST-ask   INDEF king=SIMIL CP-dress=after
        kampong=nang mà-pii maau katha].
        village=DAT   INF-go want QUOT
        'What Andare wanted as a last wish, was to go to the village dressed up as a
        king.' (K070000wrt03)

An example where the categorial nature is emphasized would be English *Do what-
ever you like*, where *whatever* does not refer to an individuated referent, but to any
member of the category of liked things that the hearer wants to instantiate the refer-
ent with. These unknown categorial referents are formed with the WH˜WH ... =*so*
construction in SLM ⊡→8.1.12, p. 424. An example is (136), where a predication is
made about the category of people who have a Malay dress. The precise referents are
unknown to the speakers. They are not individuated nor specific. Any member of the
set of possessors of Malay dresses is invited to wear them at the wedding.

(136)   ***Saapa~saapa**=ka inni mlaayu pakeyan pada aada**=so**, lorang pada ini*
        who~who=LOC   PROX Malay dress   PL   exist=disj 2PL   PL   PROX
        mlaayu pakeyan=samma ini   kaving=nang mà-dhaathang bannyak
        Malay dress=with   PROX wedding=DAT INF-come   much
        uthaama.
        honour
        'Whoever owns such Malay costumes, your coming to the wedding in these
        costumes will be highly respected.' (K060116nar04)

15.1.4.4   Unknown referents, generic

The last possibility for unknown referents is to be generic, i.e. no particular referent
is intended, but the predicate is thought to be true of any referent. An example in
English would be *one* as in *one has to be kind to strangers* or generic *you* as in *you
must obey the law*. In SLM, generic reference is not expressed overtly. An example
is the (137) about the general rules of Sepaktakraw. While in the SLM example, no
referent is expressed, in English the use of *you* is mandatory, alternatively the use of
one or a passive construction.

(137)   ∅ *kaaki=dering masthi-maayeng.*
          leg=ABL   must-play
        'You have to play with your feet.' (N060113nar05)

15.1.4.5 Queried referents

Referents are queried for by replacing them with the interrogative pronoun ⊡→5.7, p. 229 which corresponds to the semantic category of the referent (*aapa* 'what', *mana* 'where', *kaapang* 'when', etc). A postpositions can be added if needed for a more precise indication of semantic role (*aapa=nang* 'for what', *mana=dering* 'from where', etc). The following sections provide more detailed information for specific semantic domains.

Person *Saapa* 'who' is the interrogative pronoun used for persons. It can take the whole array of postpositions to indicate its syntactic and semantic role ⊡→5.7.2, p. 232.

(138) ***Saapa** anà-maathi?*
who PAST-die
'Who died?' (K051213nar07)

Animal There is no special interrogative pronoun to query for animals. One can use *mana binaathang* 'which animal'. This differs from the adstrate Sinhala, where a specialized interrogative pronouns for animals exists, *mokaa* (Garusinghe (1962:29), Karunatillake (2004:259)).

(139) *Mana binaathang lorang=yang anà-giigith?*
which animal 2PL=ACC PAST-bite
'Which animal bit you?' (K051213nar06)

Things *Aapa* 'what' ⊡→5.7.1, p. 231 and *mana* 'which' ⊡→5.7.3, p. 233 are used to query for things in the widest sense. Like *saapa*, they can take any postposition to indicate the syntactic and semantic role.

(140) ***Aapa** anà-jaadi mlaayu pada?*
what PAST-become Malay PL
'What became of the Malays?' (K051213nar06)

(141) ***Mana** nigiri=ka arà-duuduk?*
which country=LOC NON.PAST-stay
'Which country do you live in?' (B060115cvs16)

Quantity   Quantity is asked for with *dhraapa* 'how much'?

(142) ***Dhraapa*** *thaaun **dhraapa** buulang lu   arà-baapi   suusa?*
       how.many year   how.many month   2s.FAMILIAR NON.PAST-bring sad
       'How many years, how man months are you bringing sadness (into my life).'
       (K061123sng01)

(143) *Bìrras   hatthu kilo dhraapa?*
       raw.rice one   kilo how.much
       'How much is one kilo of rice?' (K081106eli01)

Location   Questions for locations are formed with the interrogative pronoun *mana*
'where', which can be used for stative location (144) or goal of motion (145). For
source, *mana=dering* or *manari* 'whence' has to be used (146). An alternative is
mana asduuduk (147).

(144) *Maana(=ka) se=ppe thoppi?*
       where=LOC   1s=POSS hat
       'Where is my hat?' (K081106eli01)

(145) *Maana(=nang) arà-pii?*
       where=DAT   NON.PAST-go
       'Where are you going (to)?' (K081106eli01)

(146) *Maana*(=dering) inni arà-dhaathang?*
       where=ABL   PROX NON.PAST-come
       'Where does this come from?' (K081106eli01)

(147) *Maana asduuduk anà-dhaathang?*
       where from   PAST-come
       'where do you come from?' (K081106eli01)

    If the query is about location in one of a given array of items, *aapa=ka* or can also
be used.

(148) *Itthu thoppi=yang aapa=ka aada?*
       DIST hat=ACC   what=LOC exist
       'On what is that hat?' (K081106eli01)

Time General temporal reference can be queried with *kaapang* 'when'.

(149) *Kaapang loram pada siini arà-dhaathang?*
      when   2PL   PL   here NON.PAST-come
      'When are you coming here?' (K081106eli01)

An amount of time can be queried with *dhraapa laama* 'how long' (Literally *how much while*).

(150) *Ini pukuran=nang dhraapa   laama athi-ambel?*
      PROX work=DAT   how.much while IRR-take
      'How long will it take for this work?' (K081106eli01)

The time can be queried with *pul-dhraapa* 'At what time' (Literally *at how much o'clock*).

(151) *Skaarang vakthu pukul dhraapa=ke   boole=aada?*
      now   time   o'clock how.many=UNDET can=exist
      'What time will it be now?' (K081106eli01)

(152) *Se pukul   dhraapa=nang masà-dhaathang?*
      1s o'clock how.much-DAT must-come
      'At what time should I come.' (K081106eli01)

Manner *Càraapa* is used to query for manner. The inverted form *ap(a)caara* also exists.

(153) *See ini   koolang=yang arà-langka   apacaara/càraaapa?*
      1s PROX river=ACC   NON.PAST-cross how
      'How can I cross this river?' (K081114eli01)

Reason, cause and purpose These questions are formed by *aapa* 'what' followed by the benefactive postposition *=nang* (154). If the cause is a person, *saapa* 'who' can be used instead of *aapa* (155).

(154) *Lorang=nang inni aapa=nang?*
      2PL=DAT   PROX what=DAT
      'Why do you want it?' (K081106eli01)

(155) *Ini saapa=nang arà-baapi?*
      PROX who=DAT   NON.PAST-take.away
      'For whom are you taking this?' (K081106eli01)

A specialized interrogative pronoun to query for reason only is *kànaapa*.

(156) *Kànaapa itthu anà-bìlli?*
      why   DIST PAST-buy
      'Why did you buy that?' (K081106eli01)

Other   Other questions are formed by combining the interrogative pronouns with the relevant postpositions ⊡→5.7, p. 229.

(157) *Ini peena saapa=pe?*
      PROX pen   who=POSS
      'Whose pen is this?' (K081106eli01)

(158) *Farook saapa=ke?*
      Farook who=SIMIL
      'Who does Farook look like?' (K081106eli01)

15.1.5   Modifying participants

In the preceding sections we have seen how participants are encoded in SLM. In this section, we will see how participants can be modified. This can be done in various number of ways: modifications pertaining to quality, such as size or colour, quantitative modifications, indicating a possessor, a location, or time. Temporal modification of a participant is also possible, an English example would be *ex-wife*.

We will discuss these different types of modifications in turn in before we close with a discussion of restrictive and characterizing modifications..

15.1.5.1   Quality

Indication of quality can be done by either a noun or an adjective, which can be either pre- or postposed, or a relative clause, which can only be preposed ⊡→8.1.17, p. 444. Examples (159) and (160) show modification by pre- and postposed nouns, while examples (161) and (162) show the same for pre- and postposed adjectives.

(159) *The⌢ pohong.*
      tea   tree
      'tea tree.'

(160) *Orang ⌣ikkang.*
      man   fish
      'fisherman'

(161) **Baaru⌢ oorang** *pada masà-thaaro.*
      new   man   PL   must-put
      '(We) must put new people.' (K060116nar11)

(162) *Se=ppe* **oorang** ⌢*thuuva pada anà-biilang [kitham pada* Malaysia=*dering*
      1s=POSS man   old   PL   PAST-say   1PL   PL   Malaysia=ABL
      anà-dhaathang katha].
      PAST-come   QUOT
      'My elders told me that we had come from Malaysia.' (K060108nar02)

The adjectival predications may be internally complex, as the negated word for 'good' in (163).

(163) *Go* **thàrà-baaye** *pukujan thama-gijja.*
      1s   NEG-good   work   NEG.IRR-do
      'I do not do bad work.' (B060115nar04)

Dimension is expressed by the adjectives *bìssar* 'big' and *kiccil* 'small'. There are no diminutives or augmentatives.

(164) a. **Bìssar** *atthu kumpulan thraa.*
         good   one   association NEG
         'There is no big association.'

      b. **Kiccil** *kumpulan pada=jo.*
         small association PL=EMPH
         'The associations are small indeed.' (N060113nar01)

Colour is expressed by a colour adjective, which is often found combined with the noun *caaya* 'colour'.[8] Table 15.4 gives the terms for common colours.

---

[8]This is an influence from the adstrates, where colour terms need to be supported by the noun for 'colour', cf. the words for 'blue' in Sinhala (*nil paaṭa*) and Tamil (*niil niṟam*), where *ni(i)l* means 'blue' and the other word means 'colour'.

SLM	gloss	original meaning	SLM	gloss	original meaning
iitham	black	black	iijong	green	green
puuthi	white	white	kuunyith	yellow	turmeric[9]
meera	red	red	beecek	brown	mud
niila	blue	blue			

Table 15.4: Colour terms in SLM, with their original meaning if different.

(165)   *Hatthu komplok bannyak=jo **puuthi caaya**, hathyeng=yang **meera**=jo*
        INDEF   bush   much=EMPH white colour, other=ACC   red=EMPH
        ***meera caaya**.*
        red   colour
        'One bush was very white, the other one was of the reddest red.'
        (K070000wrt04)

   Some colour terms are not adjectives, but nouns, as is the case for the word for
'brown' which is derived from the word *beecek* 'mud'.

(166)   *Sithu=ka, hathu bìssar **beecek caaya** Buruan su-duuduk.*
        there=LOC INDEF big   mud   colour bear   PAST-stay
        'There was a big brown bear.' (K070000wrt04)

15.1.5.2   Quantity

Next to qualitative modifications, referents can also be modified for quantity. This is
discussed in detail in ⊙→15.6, p. 639.

15.1.5.3   Possession

Terms based on nouns can get assigned a possessor. This possessor is then marked
with =*pe* ⊡→6.4.4.5, p. 345. The possessor is always preposed when it is used as a
modifier. Possession can also be expressed predicatively. This is discussed in section
⊙→15.11, p. 668.

---

[9]The original word *kuuning* can also be heard, but because of the phonetic resemblance of *kuuning* and
kuunyith 'turmeric', the latter makes inroads into the domain of colours, also because Sinhala (*kaha*) and
Tamil (*maɲcaḷ*) use the word for 'turmeric' to refer to 'yellow' as well (cf. **Paauw** 2004).

(167) *Kithang sama oorang **lorang=pe suurath**=yang daapath vakthu=ka kithang*
     1PL   all   man   you=POSS letter=ACC   get   time=LOC 1PL
     sama oorang bannyak su-suuka.
     all   man   much   PAST-like
     'All of us were very happy when we received your letter.' (Letter 26.06.2007)

In example (167), it is not because of any letter that the speakers are happy, but the set of letters they are happy about is restricted to the one written by the addressee.

15.1.5.4 Location

A term can be modified to entities at a certain location. This is done with a relative clause containing a locational predicate with the existential *aada*. The existential is obligatory.

(168) *Meeja=ka *(aada) maaňǧga=yang kaasi.*
     table=LOC exist   mango=ACC   give
     'Give me the mango (which is) on the table' (K081118eli01)

15.1.5.5 Time

Terms may be marked as referring to a past referent by means of the adjective *laama* 'former'.

(169) *Se=ppe laama ruuma.*
     1s=POSS old   house
     'My old house.' (K081106eli01)

Furthermore, the relevant period may be used as a possessive modifier with =*pe*. Both adverbs like *karang* 'now' and temporal nouns like *muusing* 'time' are possible here.

(170) ***Karam=pe*** *mosthor=nang, mpapulu aari=ka=jo   sunnath=le*
     now=POSS manner=DAT   forty   day=LOC=EMPH circumcision=ADDIT
     arà-kijja.
     NON.PAST-make
     'For today's way of doing (it), it is on the fortieth day that they also do the circumcision.' (K061122nar01)

(171) **Dovulu=pe** *oorang pada.*
 before=POSS man   PL
 'People in former times.' (K061026rcp04)

(172) *Giithu=jo*   *derang pada,* **itthu muusing=pe** *mlaayu pada.*
 like.that=EMPH 3PL   PL   DIST time=POSS   Malay PL
 'They were like that, the Malays of former times.' (K051206nar04)

15.1.5.6 Restriction and characterization

Modifications can have two different readings, restrictive and characterizing. Restrictive modifications reduce the number of entities the term refers to. *Rich people are unhappy* refers to less entities than *People are unhappy*. The referents are restricted to those people who also have the characteristic of being rich.

Characterization, on the other hand, does not reduce the number of entities the term refers to, but gives additional information. An example would be *loving* in *my loving father*, which characterizes my father, it does not restrict reference. In SLM, all modifications can be used to restrict or characterize terms based on NPs. Terms based on pronouns can only be characterized. The use of restrictive modification has been exemplified amply in the preceding sections, so that only the less common characterization will be discussed in this section.

In (173), the dwarf has been talked about before in the story, and this time he reappears in the claws of a bird. There is only one dwarf in the story, and it is absolutely clear that it is this dwarf that is being talked about. In (173), the dwarf is nevertheless modified by a preposed relative clause, which gives more information about his location. This information does not restrict the set of all possible dwarfs to those being held by birds, it rather gives additional information about the dwarf already identified before: it characterizes him.

(173) *[Kaaki=ka gaaňdas-kang ambel anà-duuduk]characterizing Aajuth=yang*
 leg=LOC   tie-CAUS   take   PAST-stay   dwarf=ACC
 sangke=luppas hathu pollu=dering Rose-red buurung=nang su-puukul.
 until=leave   INDEF stick=ABL   Rose-red bird=DAT   PAST-hit
 'Rose red hit the bird with a stick until it let go of the dwarf, whom he had taken in his claws.' (K070000wrt04)

Another example is the characterization of the Almighty God as the creator of the addressee in (174). This is clearly not restrictive but rather characterizing.

(174) *[Luu   nya-laher-kang]*$_{characterizing}$ *Allah-thaala=nang liima*
      2s.FAMILIAR=ACC be.born-CAUS   Allah-almighty=DAT five
      vakthu mà-sbaayang=nang.
      time   INF-pray=DAT
      'To pray five times to the Almighty God, who created you.' (K060116sng01)

15.2   Predication

Different things can be said about participants. One can say about them that they are in a certain state, that they take part in a certain event, that they are a member of a certain semantic class or that they possess things.

These different types of semantic predicates are encoded by different constructions in SLM. Following Hengeveld (1992:106), I distinguish the following semantic types of predicates:[10]

- states

    - property assignment ⊙→15.2.1.1, p. 599
    - status assignment ⊙→15.2.1.2, p. 600
    - identity ⊙→15.2.1.3, p. 600
    - localization ⊙→15.2.1.4, p. 601
    - possession ⊙→15.2.1.5, p. 601

- events ⊙→15.2.2, p. 601

A related, yet different functional domain is causation ⊙→15.2.3, p. 606.

15.2.1   States

15.2.1.1   Property assignment

Stative properties can be assigned in SLM by adjectival predicates ⊡→10.6, p. 477 as in (175) or by verbal predicates ⊡→10.1, p. 461 as in (176). Some states like colours are expressed by nominal predicates ⊡→10.4, p. 472 (177).

(175) *Samma oorang **baae**$_{ADJ}$.*
      all   man   good
      'All men are good.' (B060115cvs13)

---

[10]The list is reordered and somewhat simplified here for expository reasons. Namely the predicate types of "classification", "instantiation", "factuality", and "interpretation" are left out.

(176) *Itthu pada [sraathus binthan pada arà-kiilap]=ke   su-**kiilap**ᵥ.*
       DIST PL   1000   star   PL   SIMULT-shin=SIMIL PAST-shine
       'They shone like a thousand stars shine.' (K070000wrt04)

(177) *Hatthu komplok bannyak=jo puuthi caaya, hathyeng=yang meera=jo*
       INDEF bush   much=EMPH white colour, other=ACC   red=EMPH
       *[**meera caaya**]ₙ.*
       red   colour
       'One bush was very white, the other one was of the reddest red.'
       (K070000wrt04)

The defective verbs ⊡→5.1.3, p. 169 *suuka* 'like' and *thaau* 'know' also encode states.

15.2.1.2   Status assignment

Status assignment means membership of an entity in a class of "established functional, professional or ideological group" (Dik 1980:76). This is expressed by nominal predicates ⊡→10.4, p. 472 in SLM.

(178) ***Sindbad** the Sailor **hatthu Muslim**, mlaayu bukang.*
       Sindbad the Sailor INDEF   Muslim, Malay NEG.NONV
       'Sindbad the sailor was a Moor, he was not a Malay.' (K060103nar01)

Optionally, the copula *(asà)dhaathang(apa)* ⊡→5.5, p. 218 can be used.

(179) *Se **asdhaathang** hatthu butthul moderate Muslim atthu.*
       1s COPULA   one   very   moderate Muslim one
       'As for me, I am a very moderate Muslim.' (K051206nar18)

15.2.1.3   Identity

The predication of identity of two referents is done by the equational construction ⊡→10.4.2, p. 474, which juxtaposes two NPs. This construction often uses the copula *asàdhaathang* (180) or the emphatic clitic *=jo* (181).

(180) *Baapa=pe umma **asàdhaathang** kaake=pe   aade.*
       father=POSS mother COPULA   grandfather=POSS younger.sibling
       'My paternal grandmother was my grandfather's younger sister.'
       (K051205nar05)

(181) *Suda [itthu kaake=pe   aade=pe   aanak]=**jo**   baapa.*
      thus   DIST   grandfather=POSS younger.sibling=POSS child=EMPH father
      'So that grandfather's younger sister's child is my father.' (K051205nar05)

15.2.1.4   Localization

Localization is expressed by a locational predicate ⊡→10.5.2, p. 476 with *=ka*.

(182) *Kithang=pe oorang thuuva pada samma Seelong**=ka**.*
      1PL=POSS   man   old   PL   all   Ceylon=LOC
      'Our forefathers were all in Ceylon' (K060108nar02)

     Additionally, localization can be expressed by an existential construction with specification of the semantic role of LOCATION on one participant, in the following example *Dubai*. The existential predicate differs from the pure locational predicate by the existence of an existential verb, in this case *duuduk*.

(183) *Se=ppe dhaatha=pe   thiiga aanak=le   Dubai=ka arà-**duuduk**.*
      1s=POSS elder.sister=POSS three   child=ADDIT Dubai=LOC NON.PAST=stay
      'My sister's three children also live in Dubai.' (B060115prs21)

15.2.1.5   Possession

Possessive predicates are expressed by a possessive predicate ⊡→10.2.2, p. 470 with either the dative, used for permanent possession, or the locative, used for temporary possession. For more details on possession, see ⊙→15.11, p. 668.

(184) *Se=**dang** liima anak klaaki pada **aada**.*
      1s=DAT   five   child male   PL   exist
      'I have five sons.' (K060108nar02)

(185) *Incayang**=ka** ... bìssar beecek caaya hathu bag su-**aada**.*
      3s.POLITE=LOC ... big   mud   colour INDEF bag PAST-exist
      'He had a big brown bag with him.' (K070000wrt04)

15.2.2   Events

Events are characterized by being [+dynamic]. They are always encoded as verbs ⊡→10.1, p. 461. An example is given in (186).

(186)   *Baapa=le   aanak=le   guula su-maakang.*
        father=ADDIT child=ADDIT sugar PAST-eat
        'Father and son ate sugar.' (K070000wrt02)

It is possible to use lexemes from the adjective class to denote events and not states. In this case, they undergo conversion to verbs, take verbal morphology and refer to the process of the state denoted by the adjective coming into being. The adjective *bìssar* 'big' in (187) is used with verbal morphology and then denotes not the state of being big, but the event of becoming big, i.e. growing up.

(187)   *Itthu=nam blaakang=jo, kitham pada **anà-bìssar**.*
        DIST   after=EMPH 1PL   PL   PAST-big
        'After that, we grew up.' (K060108nar02)

The use of an adjective in a verbal frame changes its aktionsart from static to dynamic. It is important to distinguish aktionsart from grammatical aspect in this regard. Aktionsart does not interfere with the marking of aspect. The following examples show that dynamic aktionsart can combine with the past tense (with in this case perfective semantics) (188), the conjunctive participle (189) also giving a perfective reading, but also with *arà-* in (190), giving an imperfective reading.

(188)   *hatthu spuulu liimablas thaaun=nang jaalang blaakang, inni kumpulan*
        INDEF ten   fifteen   year=DAT   go   after   PROX association
        ***sa**-mampus*
        PAST-dead
        'About ten, fifteen years after that, the association became defunct.'
        (K060116nar01)

(189)   *Aanak pada **asà**-bìssar, skuul=nang anà-pii.*
        child PL   CP-big   school=DAT PAST-go
        'The children grew up and went to school.' (K051222nar04)

(190)   *Ruuma **arà**-kiccil.*
        house NON.PAST-small
        'The houses are getting small.' (K051222nar04)

The above examples show that the use of adjectives in a verbal frame is not a case of grammatical aspect (inchoative or ingressive), but a change in lexical aspect from [-dynamic] to [+dynamic]. This can further be demonstrated by the fact that adjectives in a verbal frame can be marked for completive aspect. If the verbal frame did indeed indicate inchoative/ingressive aspect, this should be incompatible with the marking of

completive. Example (191) shows the use of the adjective *thuuňduk* 'bent' in a verbal frame, yielding 'becoming bent'. This is marked with the vector verb *abbis*, which indicates completive aspect. The final meaning is then 'having finished becoming bent', which is indeed what one must do to enter a small room like a cave. This example demonstrates that the verbal reading of adjectives has to do with lexical aspect, but not with grammatical aspect, given that grammatical aspect is already expressed elsewhere (*abbis*), without leading to ungrammaticality.

(191)   *Giithu=jo*   *thuuňduk **abbis**=jo*   *masà-pii.*
       like.that=EMPH bent   finish=EMPH must-go
       'You must enter there completely crouched.' (K051206nar02)

The change of aktionsart from static to dynamic through the use in a verbal frame is also possible with loan words, as shown in (192)

(192)   *Se=dang **arà**-**late***   *bukang, see arà-dhaathang.*
       1s=DAT   NON.PAST-late TAG   1s   NON.PAST-come
       'I am getting late, goodbye (goodbye=I will [go and] come).' (B060115cvs08)

Example (193) is more involved. In this example, the adjective *lummas* 'soft' is first converted to a verb 'become soft', which in turn is the only constituent of the NP to which the postpositions =*nang* attaches. Example (193) thus shows two instances of conversion, one from adjective to verb on the morphological level, and one from clause to NP on the syntactic level.

(193)   *[[[[∅ Lummas_{ADJ}] −∅]_V]_{CLS} =∅]_{NP}*   =*nang blaakang minnyak*
            soft   -VRBLZR =NMLZR =DAT after   coconut.oil
       klaapa=ka   *inni=yang gooreng.*
       coconut=LOC PROX=ACC fry
       'After it has become tender, fry this in coconut oil.' (K060103rec02)

While adjectives can undergo this change in lexical aspect by zero-derivation, periphrases have to be used for states denoted by other word classes. States denoted by nouns have to take *jaadi* 'become' to get a dynamic reading, which allows for the application of perfective aspect. This is the case in (194), where the past tense marker *su-* conveys a perfective reading. The speaker in (194) informs us that in former times, the country was not theirs, but a change of state took place, and now the country is theirs. This change of state implies dynamic aktionsart, which is not a possibility for nouns. Therefore, the periphrasis with the verb *jaadi* 'become' is used, which allows for the expression of dynamic aktionsart.

(194) *Ini   kitham=pe **nigiri**   su-**jaadi**.*
      prox 1pl:poss   country PAST-become
      'This became our country.' (K051222nar04)

Things are similar with the following example, where at a time t_0, the person was not a cancer patient, but at a time t_1 had become a cancer patient. This again has to be expressed by *jaadi*.

(195) *Incayang **cancer patient** se-**jaadi**.*
      3s.POLITE cancer patient PAST-become
      'He became a cancer patient.' (K060116nar15)

For the word *enco* 'fooled', the same holds true.

(196) *Thapi=le Andare thàrà-**jaadi**   **enco**.*
      But=ADDIT Andare NEG.PAST-become fooled
      'But Andare did not get fooled.' (K070000wrt02)

This process is also possible for mass nouns, as shown in (197) for the mass nouns *daara* 'blood' and *suusu* 'milk'.

(197) *Oorang pada kapang-laari dhaathang, ini   daara sgiithu=le   **suusu***
      man   PL   when-run   come   PROX blood that.much=ADDIT milk
      *su-**jaadi**.*
      PAST-become
      'When people came running, the blood had turned into milk.' (K051220nar01)

The following sentences give some more examples of this use.

(198) *Aashik=nang **hathu** soldier **mà-jaadi**   suuka=si   katha arà-caanya.*
      Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT NON.PAST-ask
      'He asks if you want to become a soldier, Ashik.' (B060115prs10)

(199) *Baaye **meera caaya** kapang-**jaadi**, thurung-king.*
      good   red   colour when-become descend-CAUS
      'When it turns into a nice red colour, remove it (from the fire).'
      (K060103rec02)

The same construction can be used for modal predications. Modal predications are of static aktionsart, but if for some reason a change in deontic state (from possible to impossible or the other way round) takes place, a verbal periphrasis with *jaadi* must

be employed to give the dynamic reading.[11] In (200), the high commissioner was first able to come, but finally became unavailable. This change of state is once more expressed by *jaadi* 'become'.

(200) *Itthu blaakang=jo, kitham=pe* AGM=*nang duppang,* high commissioner
dist after=EMPH 1PL=POSS AGM=DAT before high commissioner
cultural show=*nang mà-dhaathang=nang* **thàràboole s-jaadi.**
cultural show=DAT INF-come-DAT cannot PAST-become
After that, before our Annual General Meeting, it became impossible for the High Commissioner to attend the cultural show. (K060116nar23)

A similar constellation obtains in (201).

(201) *See=yang dhaathang* remand=*ka mà-thaarek thaaro=nang* **thàràboole**
1s=ACC come remand=LOC INF-pull put=DAT cannot
su-jaadi.
PAST-become
'It became impossible to remand me.' (K061122nar03)

Nominal predicates construed with experiencers can mark change of state with the verb *pii* 'go', as in (202). In this example, the experiencer *incayang* is marked with the dative marker =*nang* which has 'allative' as an additional meaning. This makes it possible to use it with the verb *pii* 'go', entailing that the semantic role changes from experiencer to goal.

(202) *Incayang=**nang** baaye=nang maara su-**pii**.*
3s.POLITE=DAT good=DAT angry PAST-go
'He became really angry (Anger went upon him).' (K070000wrt01)

Finallly, verbs construed with experiencers, like *thàthaava* 'laugh' can also use this periphrasis with *pii* to emphasize the change of state.[12]

(203) *Itthu vakthu=ka Andare asà-maathi anà-duuduk mosthor kuthumung=apa*
DIST time=LOC Andare CP-dead PAST-stay manner see=after
*raaja=nang **thàthaava** su-**pii**.*
king=DAT laugh PAST-come
'When he then saw the way that Andare had died and lay there, the king started to laugh.' (K070000wrt03)

---

[11]See Karunatillake (2004:224) for a related construction in Sinhala.

[12]Note however that no physical object changes places towards the experiencers/goals in (202)(203); if that was the case, the relator noun *dìkkath* would have to be used on the human referent ⊙→15.1.2.7, p. 571.

There is one example where a word likely to be a noun is used without *jaadi*, but rather in the verbal frame discussed above, marked by a TAM-prefix. This is the loan word *pension* in (204), but it could actually be argued that *pension* in SLM does in fact denote the property 'retired', and therefore is an adjective, which can undergo conversion to become a verb for dynamic reading, as usual.

(204) *Derang pada samma konnyong aari pukurjan asà-gijja,* ***su*-pension***.
        3PL   PL   all   few   day work   CP-make PAST-pension
        'They worked some time and then got pensioned.' (K051222nar06)

15.2.3 Causation

There are two ways to indicate causation: the causative morpheme *-king* ⊡→6.2.5.2, p. 307 and a construction using a verb of saying like *biilang* 'say'.

(205) *Baaye meera caaya kapang-jaadi,* ***thurung-king***.
        good red   colour when-become, descend-CAUS
        'When [the food] has turned to a nice rose colour, remove (it) [from the fire].'
        (K060103rec02)

Verbally causing someone to do something can be expressed with an utterance verb in the main clause and an infinitive clause with the action ordered.

(206) *Oorang pada=nang **mà-dhaathang katha** asà-**biilang***.
        man   PL=DAT   INF-come   QUOT CP-tell
        'He told the people to come.' (B060115cvs01)

15.3 Modification

Different modification strategies exist in SLM. These depend fully on the lexical category of the head noun, regardless of its semantic class. The different possibilities are discussed in ⊡→8.1.17, p. 444 for referential phrases (NPs) and ⊡→10.1, p. 461 for predicate phrases.

15.4 Space

Participants, states and events are always located in space. This section discusses the different possibilities to give spatial information in SLM. The next section will deal with the counterpart of space: time. Some concepts, like figure and ground are important for both domains.

We can distinguish three different kinds of spatial reference: absolute (non-deictic) reference (*in India*) ⊙→15.4.1, p. 607, deictic reference with reference to the speaker (*here*) ⊙→15.4.2, p. 607 and figure-ground relations, which involve several entities (*between a rock and a hard place*) ⊙→15.4.3, p. 609. Finally, events can also have an inherent directionality, like *come and bring* or *go and take* in English ⊙→15.4.4, p. 612.

15.4.1   Giving the non-deictic reference space

Non-deictic reference space is indicated by the locative marker =*ka* ▣→6.4.4.3, p. 340, which is attached to an NP with local reference. This can be a common noun as in (207) or a proper noun as in (208). In rare cases, the locative marker is not present on proper nouns, as in (209).

(207)   *Se m-blaajar* **estate=*ka***.
       1s PAST-learn estate=loc
       'I learned on the estate.' (K051213nar05)

(208)   *[**Kandi=ka** arà-duuduk]* military *rejimen   mlaayu.*
       Kandy=LOC NON.PAST-stay military regiment Malay
       'The Military regiment Malays who stayed in Kandy.' (K060108nar02)

(209)   *Seelon samma thumpath=∅ mlaayu aada.*
       Ceylon all   place   Malay exist
       'There are Malays all over Sri Lanka.' (K051222nar04)

15.4.2   Giving the deictic reference space with regard to speech act participants

The spatial location of referents with regard to the speaker can be given by the deictics *ini* 'proximal' and *itthu* 'distal'.

(210)   ***Inni*** *sudaari=pe femili=ka   bannyak oorang tsunami=da s-puukul.*
       PROX sister=POSS family=LOC many   man   tsunami=DAT CP-hit
       'Of this sister's family, many members were hit by the Tsunami.' (B060115nar02)

(211) **Itthu** ports=*ka laama kar asà-baapi.*
   DIST ports=LOC old   car CP-bring
   'We brought the old cars to those ports.' (K051206nar19)

The spatial location of events with regard to the speaker can be indicated by the spatial adverbs *siini* and *siithu*. These can optionally bear a locative clitic as in (212)(213), or be used without one, as in (214)(215).

(212) **Siini=ka** *settle daapath, itthu=nam blaakang bannyak oorang pada* **siini**
   here=LOC s.   get,   DIST=DAT after   many   man   PL   here
   se-duuduk.
   PAST-stay
   'They settled down here, after that many people came to stay.'
   (G051222nar03)

(213) **Siithu=ka** *se cinggala oorang=pe cinggala em-blaajar.*
   there=LOC 1s Sinhala man=POSS Sinhala PAST-learn
   'It was there that I learned the Sinhalas' Sinhala.' (K051213nar02)

(214) *Spaaru* **siini**=∅ *su-duuduk.*
   few   here   PAST-stay
   'Few stayed here.' (K051205nar04)

(215) *Incayang* **siithu**=∅ *asà-kaaving.*
   3s.POLITE there   CP-marry
   'He married there.' (K051206nar18)

While marking of the locative is optional, marking of the allative is impossible for deictics (216) and the marking of the ablative is obligatory (217).

(216) a. **siini=nang*
   b. **siithu=nang*

(217) *See arà-sumpa   paanas muusing dhaathang=thingka see siini=***dering**
   1s NON.PAST-promise hot   time   come=middle   1s here=ABL
   arà-pii.
   NON.PAST-go
   'I promise, I will leave when summer will have come.' (K070000wrt04)

The spatial character can be highlighted by the used of *subla* 'side', as is the case in the following example.

(218) a. *Biini itthu **subla**=dering arà-bitharak;*
           wife DIST side=ABL   NON.PAST-scream
           '(Andare's) wife screamed from that side;'

     b. *puthri inni **subla**=dering arà-bitharak.*
           queen DIST side=ABL   NON.PAST-scream
           'the queen screamed from this side.' (K070000wrt05)

Spatial location close to the hearer can be indicated by *sanaka*, but this is rarely done. Normally *siini* is used, if the hearer is thought to be close to the speaker, or *siithu* if the hearer happens to be farther away.

Among third person pronouns, *siaanu* can only be used for persons close to the speaker.

Deictic and non-deictic spatial reference can be combined, as in (219), where we find non-deictic *Sri Lankaka* and deictic *siini*.

(219) *See spuulu thaaun **siini Sri Lanka=ka** pukurjan nya-kirja.*
        1s ten   years here Sri Lanka=LOC work   PAST-make
        'I worked here in Sri Lanka for ten years.' (K061026prs01)

More precise indications of spatial location with regard to speech act participants can be given by using one or more of them as ground in a figure-ground relation. This will be explained in the next section.

15.4.3 Figure-ground relations

Besides the non-deictic reference space and the reference space with regard to speech act participants, it is possible to give the spatial relation between two or more arbitrary entities. The locative marker =*ka* can be used to convey a generic figure-ground relation, where nothing about the precise disposition of figure and ground in implied. More precise constellations can be obtained by using relator nouns ⊡→5.2.2, p. 199 (cf. Adelaar 1991). These relator nouns indicate stative information (top, bottom, front, back). Lative information (from the inside, to the middle) can be obtained by combining them with the relevant postpositions =*ka* 'essive', =*nang* 'allative' =*dering* 'ablative'.

The most common case is to combine two nouns in such a figure ground relation. The following examples show a number of relator nouns

(220) *Ini pohong_ground* **atthas=ka** *[moonyeth hathu kavanan]_figure su-aada.*
       PROX tree   top=LOC   monkey   INDEF group   PAST-exist
       'On top of this tree was a group of monkeys.' (K070000wrt01)

(221) *Andare_figure [hathu pohong]_ground=pe* **baava=ka**   *kapang-duuduk.*
       Andare   INDEF tree=POSS   bottom=LOC when-sit
       'When Andare sat down under a tree.' ((K070000wrt03))

(222) *Soore=ka,   [Snow-white=le   Rose-red=le]_figure derang=pe*
       Evening=LOC Snow.white=ADDIT Rose.Red=ADDIT 3PL=POSS
       umma=samma appi_ground **dìkkath=ka** *arà-duuduk ambel.*
       mother=COMIT fire   vicinity=LOC SIMULT-sit   take
       'In the evening, Snow White and Rose Red used to sit down next to the fire
       with their mother.' (K070000wrt04)

   Pronouns can be used in this construction as well (223).

(223) *[Incayang]_ground=pe baa=ka   [spaaru]_figure aada.*
       3s.POLITE=POSS   down=LOC some   exist
       'There are some [Malays] down the hill from where he lives.' (K051213nar05)

   The interpretation of the relator noun need not be literal. In (224), the neighbour
from above does not live directly on top of the speakers, but rather on the next floor.
Still, the use of *ruuma* 'house' or *thattu* 'roof' is not possible, and the simple pronoun
kitham is used instead.

(224) *Marian kitham=pe (*ruuma) atthas (*thattu)=ka arà-duuduk.*
       Marian 1PL=POSS house   top   roof=LOC   NON.PAST-live
       'Marian lives above us.' (K081106eli01)

   Relative indications of location can be expressed by *thangang naasi subla* 'hand'
+ 'rice' + 'subla'='right hand side' and *thangang kiiri subla* 'left hand side' (The
latter has no transparent meaning). The following four examples show this, as well as
using *thìnnga* for 'in the middle' and *duppang* for 'opposite'.

(225) *Izi se=ppe* **thangang naasi** *subla=ka arà-duuduk.*
       Izi 1s=POSS hand   rice   side=LOC NON.PAST-sit
       'Izi is sitting at my right hand side.' (K081106eli01)

(226) *Sebastian se=ppe **thangang kiiri** subla=ka arà-duuduk.*
Sebastian 1s=poss hand   left   side=loc non.past-sit
'Sebastian is sitting at my left hand side.' (K081106eli01)

(227) *See **thìnnga**=ka arà-duuduk.*
1s   middle=loc non.past-sit
'I am sitting in the middle.' (K081106eli01)

(228) *Sebastian se=ppe **duppang**=ka arà-duuduk.*
Sebastian 1s=poss front=loc   non.past-sit
'Sebastian is sitting in front of me.' (K081106eli01)

The essive has been exemplified by the sentences above. Allative (229) and ablative (230) are exemplified by the following examples.

(229) *Kithang arà-pii   inni politicians pada dìkkath=**nang**.*
1PL   non.past-go PROX politicians PL   vicinity=DAT
'We approach these politicians.' (K051206nar12)

(230) a. *Biini itthu subla=**dering** arà-bitharak.*
wife DIST side=ABL   non.past-scream
'(Andare's) wife screamed from that side;'
   b. *Puthri inni subla=**dering** arà-bitharak.*
queen DIST side=ABL   non.past-scream
'the queen screamed from this side.' (K070000wrt05)

The perlative is also formed with =*dering*, as shown in the following example.

(231) *Bìssar hathu buurung derang=pe atthas=**dering** su-thìrbang.*
big   INDEF bird   3PL=POSS top=ABL   PAST-fly
'A big bird flew over them.' (K070000wrt04)

Reciprocal grounding is also possible as in (232), where the two places are close to each other. This is highlighted by the use of the dative marker =*nang* on both, as well as the use of two additive markers =*le*.

on	atthas
under	baava
in front	duppang
behind	blaakang
next to	dìkkath
in	daalang
out	luvar
in the middle	thìnnga
thangang naasi	right
thangang kiiri	left

Table 15.5: Spatial orientation

(232)   *[Spaaman anà-niiňǧgal **thumpath=nang=le**] [Passara katha arà-biilang*
      3s   PAST-die   place=DAT=ADDIT   Passara QUOT NON.PAST-say
      ***nigiri=nang=le]**   dìkkath*.
      country=DAT=ADDIT vicinity
      'The place where he died and the village called Passara are close to each other.'
      (B060115nar05)

15.4.4   Indicating the spatial orientation of an event

The spatial orientation of an event is indicated by attaching the clitics for the semantic roles of source, goal and location ⊙→15.1.2.6, p. 570 ⊙→15.1.2.7, p. 571 to the NP expressing the place.

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

The deictic adverbs *siini* and *siithu* are never combined with =*nang* to indicate goal. They can be combined with the ablative marker =*dering*, as shown in (236).

(234)   Second world war *vatthu siini-∅ dhaathang aada atthu kappal.*
      second world war time here   come   exist INDEF ship
      'During the second world war a ship came here.' (K051206nar07)

(235) *[Boole oorang pada]=nang siithu=∅ boole=pii.*
can man PL=DAT there can=go
'All men who are able to go may go.' (B060115cvs01)

(236) *Oorang mlaayu siithu=**dering** dhaathang=apa cinggala raaja=nang=le*
man Malay there=ABL come=after Sinhala king=DAT=ADDIT
anà-banthu.
PAST-help
'The Malays came from there and helped the Sinhala king.' (K051206nar04)

For the spatial orientation of motion with regard to a figure, the lexical solutions *pii* 'go', *dhaathang* 'come' and *baalek* 'return' are available. If a container is involved, *maasok* 'enter' and *kuluvar* 'exit' are used. For vertical motion, the verbs *naaek* 'ascend' and *thuurung* 'descend' are available. The verb *bavung* 'get up' also includes a spatial component. (237) gives an example of the verb *thuurung* 'descend'.

(237) a. *Ithu oorang=nang baaye=nang nanthok pii=nang blaakang.*
DIST=DAT man=DAT good=DAT sleep go=DAT after
'After the man had well fallen asleep'

b. *Pohong=dering baava=nang asà-**thuurung**.*
tree=ABL down=DAT CP-descend
'(the monkeys) climbed down from the tree and'

c. *[Oorang anà-baava] samma thoppi=pada asà-ambel.*
man PAST-bring all hat=PL CP-take
'took all the hats the man had brought and'

d. *Mà-maayeng=nang su-mulain.*
INF-play=DAT PAST-start
'started to play.' (K070000wrt01)

15.5 Time

Just like space, time can be anchored in a non-deictic way. The temporal expression *in 1984* corresponds to the spatial expression *in India* ⊙→15.5.1, p. 614.[13] As a second possibility, time can be indicated relative to the speech act ⊙→15.5.2, p. 615. We can distinguish the periods before, during and after the speech act. The temporal relation between a figure and a ground can be expressed in very much the same fashion as for space, as in *before the war*, which corresponds to *before the church* ⊙→15.5.3, p. 625.

---

[13]It is true that *in 1984* is actually anchored in a figure-ground relation with the conventional birth date of Jesus Christ, but speakers normally do not conceptualize this birth when they use such a time reference.

Taking into account the start time, end time, duration and overlap of the two events, we can distinguish a number of situations, such as precedence, subsequence, point coincidence, etc.

The location of time of an event contrasts with its internal temporal structure. We can distinguish phasal information (start, progression, end) ⊙→15.5.4, p. 635 from aspectual information (perfective, imperfective) ⊙→15.5.5, p. 638. A discussion of the encoding of temporal frequency concludes this section ⊙→15.6.3, p. 649.

15.5.1 Giving non-deictic reference time

The means SLM uses to indicate non-deictic time resemble the means used for absolute space. Temporal reference is marked with the postposition =*ka* (238). This postposition is frequently dropped in rapid speech (239).

(238) *Soore=**ka**, Snow-white=le Rose-red=le derang=pe umma=samma*
Evening=LOC Snow.white=ADDIT Rose.Red=ADDIT 3PL=POSS mother=COMIT
appi dìkkath=ka arà-duuduk ambel.
fire vicinity=LOC SIMULT-sit take
'In the evening, Snow White and Rose Red used to sit down next to the fire with their mother.' (K070000wrt04)

(239) *Kithang anà-pii pugi week=∅.*
1PL PAST-go last week
'We went last week.' (K051206nar07)

Years are indicated by their number in Common Era. For this, English numbers are normally used (240), but Malay numbers can also be found (241). Also months are indicated by their English names (241).

(240) *Suda **nineteen-ninety-four**=ka se=ppe husband su-nniiňğgal.*
thus nine-teen-ninety-four=LOC 1s=POSS husband PAST-die
'So, my husband died in 1994.' (K051201nar01)

(241) *Limapulunnam=ka, April buulang=dìkka.*
56=LOC april month=vicinity
'In '56, around the month of April.' (K060108nar01)

For the days of the week as given in Table 15.6, both native and English words are used.

The hours of the clock are indicated by preposing *pu(ku)l* to the number.

Monday	(h)ari sinnen
Tuesday	(h)ari slaasa
Wednesday	(h)ari rubbo
Thursday	(h)ari k(h)uumis
Friday	(h)ari jumahath
Saturday	(h)ari satthu
Sunday	(h)ari ahath/ahadh

Table 15.6: Days of the week. The letters in parentheses might be preferred by speakers in the orthography, but are not present phonetically.

itthu muusingka	in former times
kumaareng dovulu	the day before yesterday
kumaareng	yesterday
nyaari	today
beeso(na)	tomorrow
luuso(na)	(the/some) day after tomorrow
pugi[14]X	last X
hathiya X	next X

Table 15.7: Temporal adverbs

(242) *Pukul ìnnam.*
       hit   six
       'Six o'clock.' (B060115cvs09)

Special words exist for 'noon' (*thìngaari*) and 'midnight' (*thìngamaalang*). Fractions of the hour are expressed by *spaaru* 'half' and *kaarthu* 'quarter'.

15.5.2 Giving the reference time with regard to the speech act

Besides non-deictic time reference, events can be situated with regard to the speech act, to wit, before, during and after the speech act. Another possibility is 'general truth', which has no fixed time-reference.

We can distinguish lexical solutions from grammaticalized solutions. (Cf. Table 15.7.)

As for grammaticalized solutions, SLM features a developed TAM system, which is used to convey temporal meaning.

There are no morphosyntactic limitations for lexical expressions of time, but due to the impossibility of having more than one verbal prefix (Slomanson 2006), the

---

[14]This is a development from historical **piggi* 'go', (Adelaar 2005:214).

presence of a non-tense prefix (i.e. aspectual, phasal, modal, infinitive) blocks the expression of tense on the verb. The resulting verb can then have any time reference. In the following example, the verbal prefix *kapang-* 'when' blocks expression of tense on the word, and the resulting clause is ambiguous with regard to time-reference. Furthermore, and unrelated to the use of *kapang-*, *thàràthaau* also is ambiguous with regard to time-reference, so that this sentence consists of two clauses which do not indicate the reference time.

(243) *Incayang **kapang**-dhaathang, cinggala incayang=nang thàrà-thaau.*
      3s.POLITE when-come   Sinhala 3s.POLITE=DAT   NEG-know
      'When he came/comes/will come, he does/did/will not know Sinhala.'
      (K060108nar02)

Both lexical and grammatical solutions will be discussed in more detail for the different logical possibilities.

15.5.2.1   Before speech act

The first logical possible constellation is that the reference time precedes the time of the speech act. This constellation can be expressed by both lexical and grammatical means. Among the lexical means, we number temporal adverbs like *kumaareng* 'yesterday' or less specific *(kà)thaama* 'earlier', as well as constructions involving the adverb *pugi* 'last' or the distal deictic *itthu*.

(244) ***Kumaareng**=le thuuju=so   dhlaapan=so oorang asà-buunung, ...*
      yesterday=ADDIT seven=UNDET eight=UNDET man   CP-kill   ...
      'Again yesterday, seven or eight men were killed .' (K051206nar11)

(245) ***Kàthaama** kithang kiccil muusing=ka inni Peeradheniya jaalang=ka*
      before   1PL   small time=LOC   PROX Peradeniya   road=LOC
      samma an-aada   mlaayu.
      all   PAST-exist Malay
      'Before, when we were children, it was all Malays in the Peradeniya Road.'
      (K051222nar04)

(246) *Kithang anapii **pugi** week.*
      1PL   PAST-go last week
      'We went last week.' (K051206nar07)

(247) ***Itthu muusing=ka*** *cinggala thraa.*
      DIST time=LOC   Sinhala NEG
      'At that time, there was no Sinhala.' (K051222nar06)

(Kà)thaama can also be used to indicate the amount of time that lies between the event and the speech act.

(248) ***S-raathus***   ***limapulu thaaun=nang thaama*** *incayang bannyak*
      one-hundred five-ty   year=DAT   early   3s.POLITE much
      igaama=pe   *atthas sbaayang naaji.*
      religion=POSS about pray   recite
      '150 years ago, he recited a lot about religion,' (K051220nar01)

The absolute use and the relative use of *(kà)thaama* are exemplified at the same time in the following passage.

(249) a. *Se=ppe baapa dhaathangapa* government servant.
         1s=POSS father cop   government servant
         'My father is/was a government servant.'

      b. *Inni* railway department=*ka* head guard *hatthu **kàthaama***.
         PROX railway department=LOC head guard INDEF   before
         'He was a head guard in the railway department before.'

      c. *Itthu=nang **kàthaama** incayang* second world war=*ka* CLI *Seelon*
         DIST=DAT   before   3s.POLITE second world war=LOC CLI Ceylon
         lightning infantry *katha athu* soldier.
         lightning infantry QUOT INDEF soldier
         'Before that he was a soldier in the so-called CLI, the Ceylon lightning infantry, in the second world war.' (G051222nar01)

If time reference is inferable from discourse, non-verbal predicates need not carry overt indication of time reference. In example (250), the preceding sentence had established past time reference, and the nominal predication in (250) need not be marked for time reference.

(250) *Itthu bannyak laama hathu ruuma.*
      DIST very   old   INDEF house
      'That one was a very old house.' (K070000wrt04)

Lexical solutions are available for both verbal and non-verbal sentence types. Grammatical solutions, on the other hand, can only be employed with verbs (or converted adjectives).

Verbal predications with past reference are normally indicated by the prefixes *anà-* ⊡→6.2.1.2, p. 272 and *su-* ⊡→6.2.1.1, p. 271.

(251) 58=*ka*=*jo*   **anà-*mulain.*
       58=LOC=EMPH PAST-start
       'It started in '58.' (B060115prs17)

(252) *Aanak pompang duuva*=*nang slaamath katha* **su-*biilang.*
       child   girl   two=DAT   goodbye QUOT PAST-say
       'He said "Goodbye" to the two girls.' (K070000wrt04)

Location in time before the speech act can also be expressed by the perfect construction ⊡→5.1.6.1, p. 192. In distinction to the English present perfect, this does not imply relevance for the time of speaking. Example (253) shows the use of the perfect tense in a context relevant to the time of speaking, while (254) gives an example of the perfect construction being used in a context without relevance to the speech situation.

(253) *See*=*le*   **pii aada** *dhraapa*=*so*   *duuva thiiga skalli.*
       1s=ADDIT go exist how.many=UNDET two   three time
       'I myself have been there, how many, maybe two or three times.'
       (B060115nar05)

(254) Dutch period=*ka* *derang pada* **dhaathang aada**.
       Dutch period=LOC 3PL   PL   come   exist
       'They came during the Dutch period' (K051206nar05)

Events which are asserted to not have happened in the past take the negation marker *thàrà-* (255) or postverbal *thraa* in the perfect (256). As with the affirmative, this does not entail any relevance for the time of speaking, unlike English present perfect.

(255) *S.* subscription **thàrà-*kiiring*
       S. subscription NEG.PAST-send
       'S. did not send the subscription.' (K060116nar10)

(256) *Kithang baaye mlaayu arà-oomong   katha incayang **biilang thraa**.
       1lp   good Malay NON.PAST-speak QUOT 3s.POLITE say   NEG
       'He has not said that we speak good Malay.' (B060115prs15)

Non-verbal predications do not show any special negative marking for the past. This is exemplified for an existential predication in (257).

(257) *Oorang pada=nang hathu oorang=nang creeveth **thraa**.*
man   PL=DAT   INDEF man=DAT   trouble NEG
'People had no trouble with each other (*also possible:* People have no trouble with each other).' (K051220nar01)

Subordinate clauses which carry an indication of relative tense (simultaneous (*arà-*) or anterior *asà-*) cannot express absolute tense. They then inherit the tense meaning of the matrix clause. In example (258), the verb *kuthumung* 'see' in the matrix clause is in the past tense, indicated by *anà-*. The verb *ambel* 'take' in the subordinate clause has the anterior prefix *asà-*, while *maayeng* 'play' in the subordinate clause has the prefix *arà-*, indicating the simultaneity of seeing and playing and the anteriority of taking with respect to seeing. Neither of the verbs in the subordinate clause is marked for past tense, but the past tense meaning is inherited from the matrix clause.

(258) *Blaakang=jo incayang **anà**-kuthumung [moonyeth pada thoppi*
after=EMPH   3S.POLITE **past**-see   monkey PL   hat
***asà**-ambel   pohong atthas=ka **arà**-maayeng].*
anterior-take tree   top=LOC **simult**-play
'Then only he saw that the monkeys had taken his hats and were playing on the top of the trees.' (K070000wrt01)

Another example is (259) where the act of hearing is simultaneous to the act of crying. Note that the conjunctive participle does not imply anteriority in this case, but rather coordination. Further note that we are dealing with past reference.

(259) *[Banthu-an asà-mintha **arà**-naangis] svaara hatthu derang=nang*
help-NMLZR CP-beg   SIMULT-cry   sound INDEF 3PL=DAT
***su**-dìnngar.*
PAST-hear
'They heard a sound of crying and begging for help.' (K070000wrt04)

Durative or habitual actions in the past can also be coded by the progressive marker *arà-*. This overrides the expression of past tense. The following two examples show durative marking (260) and habitual marking (261).

(260) *Itthu kumpulan=dang derang=jo bannyak **arà**-banthu.*
DIST association=DAT 3PL=EMPH much   NON.PAST-help
'It is them who help that association a lot.' (B060115cvs01)

(261) *Itthu kalu, [...] Dubai=ka asà-duuduk laama kar pada kitham **arà**-baapi*
       DIST if   [...] Dubai-LOC CP-stay   old   car PL   1PL   NON.PAST-take
       Iraq ports=nang
       Iraq ports=DAT
       'In that case, we bring old cars from Dubai to the ports in Iraq.'
       (K051206nar19)

The use of *arà-* in past contexts is clearly seen in (262), where we have a non-deictic indication of time (1958), but the verb is not marked with a past tense prefix, but with *arà-*, which marks habitual in this case.

(262) *Muula pàrthaama, Badulla ruuma saakith=ka **s-riibu**   **sbiilan raathus***
       before   first   Badulla house sick=LOC   one-thousand nine   hundred
       ***lima-pulu dhlaapan=ka** pukurjan **arà**-gijja   vakthu.*
       five-ty   eight=LOC   work   SIMULT-make   time
       'Before, when I was working in Badulla in 1958.' (K051213nar01)

Lexical and grammatical indication of time can be combined, as in (263), where we find the lexical marker *kàthaama* and the grammatical marker *nya-*.

(263) *See **kàthaama** pukurjan **nya**-kirja.*
       1s earlier   work   PAST-do
       'I used to work in former times.' (B060115prs01)

15.5.2.2 Simultaneous to speech act

Temporal situation conceived of as simultaneous to the speech act can be expressed by a limited number of lexical means, or by the verb prefix *arà-* ⊡→6.2.2.1, p. 287 (Smith & Paauw 2006:164). Lexical solutions include the use of *(s)kaarang* 'now', *nyaari* 'today' or *ini X* 'this X', where X is a temporal noun like *thaaun* 'year' or *muusing* 'time, period'. This temporal noun should refer to a period which includes the time of speaking. The following sentences are examples of simultaneity to the speech act expressed lexically.

(264) *Suda **skaarang** kitham=pe aanak pada laaeng pukurjan pada arà-girja.*
       thus now   1PL=POSS child PL   other work   PL   NON.PAST-make
       'So our children are doing other jobs now.' (K051222nar05)

(265) *Itthusubbath=jo **nyaari** go   laile arà-duuduk.*
       therefore=EMPH today 1S.FAMILIAR again NON.PAST-stay
       'That's why I still stay (here).' (B060115nar04)

(266) ***Ini   muusing**=dika oorang ikkang Hambanthota=ka arà-duuduk.*
       PROX time=LOC   man   fish   Hambantota=LOC NON.PAST-stay
       'Presently, the fishermen are at Hambantota.' (K081106eli01)

Events taking place in the time frame of the speech act are coded by *arà-* for verbal
predicates (267) ⮕→6.2.2.1, p. 287, and ∅-coded at all for other predicates (268).

(267) *Suda itthu oorang pada=le Seelong=ka **arà-duuduk**.*
       thus DIST man   PL=ADDIT Ceylon=LOC NON.PAST-stay
       'So these people also stay in Ceylon.' (K060108nar02)

(268) *Se asdhaathang hatthu butthul* moderate *Muslim atthu.*
       1S COPULA   one   very   moderate Muslim one
       'As for me, I am/was/will be a very moderate Muslim.' (K051206nar18)

Simultaneity can be conceived in a wide sense (**Smith & Paauw** 2006:164). In
example (269) about a Chinese converted to Islam, it is not clear whether he is going
to the mosque right now. Simultaneity to the time of speaking is conceived in a wider
sense here, meaning that at the time of speaking the convert has the habit of going to
the mosque, without vouching for his going at the present moment.

(269) a. *Ini   ciina oorang Islam=nang asà-dhaathang*
          PROX China man   Islam=DAT cp-come
          'That Chinaman came to Islam and'

       b. *ini   asà-kaaving=apa*
          PROX CP-marry=after
          'married and'

       c. ***karang** masiigith=nang **arà-pii**   liima vakthu sbaayang=nang.*
          now   mosque=DAT   NON.PAST-go five time   pray=DAT
          'now he goes to the mosque five times (a day) to pray.' (K051220nar01)

Negated predicates with time reference simultaneous to the speech act are marked
with *thama-* (270) ⮕→6.2.2.5, p. 298 for verbs. All other predication types take their
standard negation, which is the same regardless of time reference.

(270) *Kitham=pe aanak pada **thama**-oomong.*
1PL=POSS child PL NEG.NONPAST-speak
'Our children do not speak.' (G051222nar01)

15.5.2.3 After speech act

Temporal reference to a point in time after the speech act can be made with some lexical means, and grammatical means for verbal predications. Lexical solutions include *beeso* 'tomorrow' or *(beeso) luusa* 'after tomorrow'.

(271) *Ruuma birsi=nang arà-simpang. **Beeso**=nang kithang arà-mnaaji*
house clean=DAT NON.PAST-keep. Tomorrow=DAT 1PL NON.PAST-pray
'We keep the house clean. Tomorrow, we will pray.' (K061019prs01)

(272) ***Beeso luusa** lubaarang arà-dhaathang.*
tomorrow later.in.the.future festival NON.PAST-come
'The day after tomorrow is the festival.' (K061019prs01)

Additionally, *duppang* 'future' can be used to indicate future temporal reference to after the speech act. Note that the relator noun *duppang* 'before', constructed with =*nang*, is used for past reference, while the full noun *duppang* 'future' is used for future reference. This is similar to German, where the preposition *vor* 'ago' is used to indicate temporal distance in the past (*vor neun Jahren* 'nine years ago'), while the future is said to lie before us *vor uns* as well, whereas the past lies behind (*hinter*).

(273) *Kithang=**nang duppang**$_{RELN}$ lai duuva bàrgaada asà-dhaathang aada.*
1PL=DAT before other two family NON.PAST-come exist
'Before us, there were two other families.' (K060108nar02)

(274) ***Duppang muusing**$_{fullnoun}$=ka=le Dodangwela aapacara=le thama-bìssar.*
before time=LOC=ADDIT Dodangwela how=ADDIT NEG.IRR-big
'Even in the future, Dodangwela [a village close to Kandy] will not be big.'
(K081106eli01)

(275) ***Duppang muusing**$_{fullnoun}$=ka oorang ikkang Negombo=nang anthi-pii.*
future time=LOC man fish Negombo=DAT IRR-go
'In the future, the fishermen will go to Negombo.' (K081106eli01)

As for grammatical solutions, verbal predications which are thought to take place

after the speech act are either marked with *arà-* ⊡→6.2.2.1, p. 287 (276) or *anthi-*
⊡→6.2.2.2, p. 290 (277).

(276) *Paanas muusing dhaathang=thingka see siini=dering **arà**-pii.*
       hot   season come=middle   1s here=ABL   NON.PAST-go
       'When the hot season comes, I will leave from here.' (K070000wrt04)

(277) *Ithu=kapang lorang=pe leher=yang kithang **athi**-poothong.*
       DIST=when   2PL=POSS neck=ACC 1PL   IRR-cut
       'Then we will cut your neck.' (K051213nar06)

Non-verbal predicates technically do not have to be marked for future reference.
However, asserting the future truth of a proposition often implies pragmatically that
it is not true at the time of speaking, so that a construction conveying this change of
state from false to true is preferred. This can be done by using adjectives in a verbal
predication (278), or by using a construction involving *jaadi* 'become' (279) ⊡→5.1.6,
p. 192. *Jaadi* will bear the irrealis marker *anthi-* then.

(278) *Ithukapang gaathal **anthi**-kuurang.*
       then   itching IRR-less
       'Then the itching will become less.' (K060103cvs02)

(279) *Incayang hatthu guru   athi/*arà-jaadi.*
       3S.POLITE INDEF teacher IRR/NON.PAST-become
       'He will become a teacher.' (K081106eli01)

Considering negation for propositions referring to a point in time after the speech
act, *thama-* is employed for verbs (280) and adjectives (281), while other predicates
will use a periphrasis to indicate that the future state will not come into being (282).
In the rare case that the speaker wants to assert that a state is not true at the time of
speaking and will not be true in the future either, this can be done by a normal non-
verbal predication, but lexical material is required to make the time reference clear
(283).

(280) *See lorang=nang **thama-**sakith-kang.*
       1s 2PL=DAT   NEG.NONPAST-sick-CAUS
       'I will not hurt you.' (K070000wrt04)

(281)   *Inni pukuran=yang mà-gijja **thamau**-gampang*
        PROX work=ACC   INF-make NEG.IRR-easy
        'To do that kind of work will never be(come) easy.' (K081106eli01)

(282)   *Se hatthu guru   **thama**-jaadi.*
        1s INDEF   teacher NEG.IRR-become
        'I will not become a teacher.' (K081106eli01)

(283)   *See **innam   blaakang** hatthu aanak **bukang**.*
        1s   PROX.DAT after   INDEF child NEG.NONV
        'I will never be a child again.' (K081106eli01)

15.5.2.4   General truth

General truth independent of time frame is expressed with the progressive marker
arà- ⊡→6.2.2.1, p. 287 (*thama-* in the negative) for verbal predicates and is not coded
overtly for non-verbal predicates.

   In example (284), the speaker's name Tony is true irrespective of the time of speak-
ing. In example (285), the circumcision at the fortieth day is a general truth as well
which is true about circumcisions in the past, present and future alike. Both are coded
by *arà-*.

(284)   *See=yang Tony katha **arà**-panggel.*
        1s=ACC   Tony QUOT NON.PAST-call
        'I am called "Tony".' (K060108nar01)

(285)   *Karam=pe mosthor=nang, mpapulu aari=ka=jo   sunnath=le*
        now=POSS manner=DAT   forty   day=LOC=EMPH circumcision=ADDIT
        ***arà**-kijja.*
        NON.PAST-make
        'For today's way of doing (it), it is on the fortieth day that they also do the
        circumcision.' (K061122nar01)

   General truth can sometimes be difficult to distinguish from habitual. In this gram-
mar, I distinguish general truth, coded by *arà-*, from habitual, which can be coded by
arà- or *anthi-*. See Section 15.6.2 for a discussion of habitual events.

   The normal negation of *arà-* is *thama-*. This prefix is also used to negate propo-
sitions with general time reference. Example (286) shows general truth of a negative
predication: Hindus generally do no eat beef, which again is independent of time

reference to present past or future. This is coded by *thuma-*, an allomorph of *thama-*.

(286) a. *Hindu **arà**-maakang kambing.*
           Hindu NON.PAST-eat   goat
           'Hindus eat goat.'

       b. *samping **thuma**-maakang*
           beef   NEG.NONPAST-eat
           '(they) don't eat beef.' (K060112nar01)

General truth of non-verbal predications does not receive special marking. In example (287) the general truth of the speaker's profession being engineering is conveyed by the copula, which can also be used with other time references.

(287) *Se=ppe percariyan asdhaathang engineering.*
        1s=POSS earning   COPULA   engineering
        'My profession is engineer.' (K061026prs01)

15.5.3 Figure and ground in the temporal domain

Just like two entities can be related to each other in space, this is also possible in time. Depending on start point, end point, duration and overlap of the two events, a certain number of semantic constellations are possible.[15] In SLM, all of them can be expressed lexically, and some can also be expressed grammatically. Very often, relator nouns are employed to indicate the precise relationship.

Any point in time can be referred to anaphorically by the distal deictic *itthu*. The proximal deictic *ini,* on the other hand, specifies the time of the speech act as temporal grounding. The precise kind of relation is then indicated by a relator noun like *blaakang* 'after', with dative marker =*nang* added to *ini*. In the process *ini=nang* is contracted to *innam* (not to be confounded with *ìnnam* 'six').

(288) *See innam   blaakang hatthu aanak bukang.*
        1s PROX.DAT after   INDEF child NEG.NONV
        'I will never (=from now onwards) be a child again.' (K081106eli01)

15.5.3.1 Precedence

Precedence obtains when the time frame of the figure precedes the time frame of the ground.

---

[15]This list is inspired by Givón (2001b:330).

(289)   *fig*   *gr*
         ⊸o   •—

In SLM, it is coded by *duppang* 'before' or *(kà)thaama* 'before'.

Duppang is used as an adverb in this function. The ground is coded with the dative marker =*nang*, while the figure does not receive special marking. The ground can be expressed by a noun, as in (290), or by a verb in the infinitive, as in (291)

(290)   *Itthu blaakang=jo, [**kitham=pe** AGM]*$_{ground}$=***nang** duppang, [high*
        dist after=EMPH   1PL=POSS   AGM=DAT   before   high
        commissioner cultural show=*na mà-dhaathang=nang thàràboole*
        commissioner cultural show=DAT INF-come-DAT   cannot
        s-jaadi]$_{figure}$.
        PAST-become
        After that, before our Annual General Meeting, it became impossible for the
        High Commissioner to attend the cultural show. (K060116nar23)

(291)   *Itthule [**see**=**yang mà-kiiring**=**nang**]*$_{ground}$ *duppang [incayang see=yang*
        but   1s=ACC   INF-send=DAT   before   3s.POLITE 1s=ACC
        hathu Buruan mà-jaadi   su-bale-king]$_{figure}$.
        INDEF bear   INF-become PAST-turn-CAUS
        'But before he sent me back, he turned me into a bear.' (K070000wrt04)

If the ground is not given, the time of the speech situation is taken as a default. In that case, the amount of time separating the event from the speech act can be indicated by =*nang*, as in (292).

(292)   ∅$_{ground}$ **Sdiikith thaaun**=**nang** *duppang, [see ini   Aajuth=nang su-kìnna*
                few   year=DAT   before   1s PROX dwarf=DAT   PAST-patfoc
        daapath]$_{figure}$.
        get
        'Some years ago, I fell prey to this dwarf.' (K070000wrt04)

The situation is similar with *(kà)thaama*. Here as well, the ground is marked by =*nang*. The following example also shows that a deictic can serve as ground.

(293)   *[Inni]*$_{ground}$=*nang kàthaama [see Navalapitiya=ka   duuduk aada]*$_{figure}$.
        PROX=DAT   before   1s Nawalapitiya=LOC stay   exist
        'Before that, I was in Nawalapitiya.' (G051222nar01)

15.5.3.2 Subsequence

Subsequence obtains when the time frame of the ground precedes the time frame of the figure.

(294) $\overset{gr}{\underset{\bullet}{\rule{0pt}{0pt}}}$ $\overset{fig}{\underset{\circ}{\rule{0pt}{0pt}}}$

Subsequence of events need not be coded but can be inferred by the succession of utterances, which is iconic. Subsequence can be coded lexically by *itthuka(apa)ng* 'and then' ⊡→7.6.3, p. 412 or *blaakang* 'after' ⊡→5.2.2.1, p. 201. As for grammatical means, the conjunctive participle *asà-* ⊡→6.2.1.4, p. 276 very often indicates subsequence.

The non-coding of succeeding events is exemplified by (295), where the getting up is followed by the search, but no grammatical or lexical element indicates the temporal relation between the two events. This is then inferred to be subsequence as the default case.

(295)   a. *Oorang su-baavung*
      man   PAST-get.up
      'The man got up'

    b. Ø *thoppi pada yang anà-caari.*
       hat   PL   PAT   PAST-search
      '(and) looked for the hats.' (K070000wrt01)

A lexical solution for subsequence is *blaakang* 'after'. As with precedence, the ground is indicated by =*nang*.

(296)   *[Itthu]ground=nam **blaakang** [bannyak oorang pada siini se-duuduk]figure.*
    DIST=DAT   after   many   man   PL   here PAST-stay
    'After that, many people settled down here.' (G051222nar03)

Just like *duppang*, *blaakang* can also be used without overt grounding, in which case it just emphasizes that the event is subsequent to whatever was there before.

(297)   *[Ø]ground Blaakang, [incayang=nang baaye=nang nanthok su-pii]figure.*
          after   3s.POLITE=DAT   good=DAT   sleepy   PAST-go
    'Then the man became very sleepy.' (K070000wrt01)

The combination of a deictic and a postposition *ithukapang* 'then' and the particle *suda* 'then' are rather discourse markers ⊡→7.6, p. 411, but also convey a meaning of subsequence.

(298) **Ithu=kapang** *umma-baapa su-biilang lorang=nang kaaving thàrboole.*
       DIST=then   mother-father PAST-say 2PL=DAT   marry   cannot
       'Then the parents said, you cannot marry.' (K051220nar01)

(299) **Suda** *derang=nang hathyang muusing=sangke mà-duuduk su-jaadi.*
       then 3PL=DAT   other   time=until   INF-stay PAST-become
       'So they had to wait until another time' (K051220nar01)

There exist three grammaticalized means to express subsequence, the conjunctive participle *asà-* ▣→6.2.1.4, p. 276 (cf. Slomanson 2008a), the postposition =*apa* ▣→6.4.4.10, p. 353 and the (plu)perfect construction ▣→5.1.6.1, p. 192. *Asà-* and =*apa* are frequently combined, but this is not necessary. The following examples show the exclusive use of *asà-* (300) and =*apa* (301), and the combination thereof (302).

(300) a. *Pìrrang=nang anà-dhaathang oorang pada **asà**-pìrrang.*
          war=DAT   PAST-come   man   PL   CP-wage.war
          'The people who had come to the war battled (and)'

       b. *Derang=nang **asà**-banthu.*
          3PL=DAT   CP-help
          'helped him [the king] (and)'

       c. *Siini=jo   se=ciiňǧgal.*
          here=EMPH PAST-settled
          'settled down (here).' (K051222nar03)

(301) a. *Mlaayu pada duuduk=**apa**.*
          Malay PL   stay=after
          'After the Malays had settled down'

       b. *Spaaru mlaayu pada Singapur Indonesia Malaysia anà-pii.*
          some   Malay PL   Singapore Indonesia Malaysia PAST-go
          '(only) some Malays went (back) to Singapore, Indonesia or Malaysia.'
          (K051213nar07)

(302) a. *Thàrà-kalu [ini oorang thoppi arà-kumpul]=yang*
    NEG-if   PROX man hat SIMULT-collect=ACC

    asà-kuthumung=apa,
    CP-see=after
    'Furthermore, when (they) had seen the man collect the hats'

  b. *moonyeth pada=le asà-dhaathang creeveth athi-kaasi katha.*
    monkey PL=ADDIT CP-come   trouble IRR=give QUOT
    'the monkeys would certainly go and cause (some other) trouble.'
    (K070000wrt01)

The pluperfect with *asà-* is given in (303)-(304). The theoretical possibility of forming the pluperfect with =*apa* was not found in the corpus.

(303) *Incayang mliiga=nang kapang-pii, Raaja hathu thiikar=ka guula **asà**-siibar*
    3s.POLITE palace=DAT when-go king INDEF man=LOC sugar CP-spread

    *mà-kìrring simpang **su**-aada.*
    INF-dry   keep   PAST-exist
    'When he was going to the palace, the King had sprinkled sugar in a mat and had left it to dry.'

(304) *Itthu bannyak laama hathu ruuma, itthule ruuma duuva subla=ka panthas*
    DIST very   old   INDEF house, but   house two   side=LOC beautiful

    *rooja kumbang pohong komplok duuva **asà**-jaadi **su**-aada.*
    rose flower   tree   bush   two   CP-become PAST-exist
    'That was a very old house, but still on the two sides of the house, there had grown two beautiful rose bushes.'

Next to the pluperfect, the normal perfect can also be used to refer to the past of the past. The marking of past tense on the existential *aada* is optional, as the following two examples show.

(305) *Kanabisan=ka=jo duva oorang=le anà-thaau ambel [Andare duva*
    last=LOC=EMPH   two man=ADDIT PAST-know take   Andare two

    *oorang=yang=le **asà**-enco-kang 0̸-**aada** katha].*
    man=ACC=ADDIT CP-fool-CAUS   exist QUOT
    'Finally the two women understood that Andare had fooled both of them.'
    (K070000wrt05a)

(306)   a. *Derang pada kathahan thama-thuukar.*
         3PL   PL   word   NEG.NONPAST-change
         'They would not change their word.'

    b. *Karang* British government=*nang derang kathahan **asà**-kaasi ∅-**aada**.*
         now   British Government=DAT 3PL   word   CP-give exist
         'Now, they had given their word to the British government.'
         (K051213nar06)

The following passage of a narrative shows the use of four different strategies to encode subsequence: zero coding, *blaakang, ithukapang* and *asà-*.

(307)   a. *Incayang=nang baaye=nang maara su-pii.*
         3S.POLITE=DAT   good=DAT   anger PAST-go
         'He got very angry.'

    b. ∅ *Incayang=pe kàpaala=ka anà-aada thoppi=dering moonyeth*
          3S.POLITE=POSS head=LOC   PAST-exist hat=ABL   monkey
        pada=nang su-buvang puukul.
        PL=DAT   PAST-throw hit
        'He took the hat from his head and violently threw it at the monkeys.'
        (K070000wrt01)

    c. ***Ithu=kapang** ithu moonyeth pada=le anà-maayeng duuduk   thoppi*
         DIST=then   DIST monkey PL=ADDIT PAST-play   exist.ANIM hat
        pada=dering inni oorang=nang su-bale-king   puukul.
        PL=ABL   PROX man=DAT   PAST-return-CAUS hit
        'Then, the monkeys also threw back to the man the (other) hats with which they had been playing.'

    d. *Itthu=nang **blaakang** inni oorang lìkkas~lìkkas thoppi pada=yang*
         dist   after   PROX man fast~RED   hat   PL=ACC
        ***asà**-kumpul ambel sithu=ka=dering su-pii.*
        CP-collect take   there=LOC=ABL PAST-go
        'After that the man quickly picked up his hats and left that place.'

15.5.3.3 Simultaneity

Simultaneity obtains when two states-of-affairs are true at the same point in time.

(308)
        fig
     — — ——
         gr
     — — ——

Simultaneity is lexically coded by *vatthu* 'time' or *=kapang* 'when' ⊡→5.7.4, p. 234. Both[16] are used as a postposition on an NP, which can be attached on a noun (309), a deictic (310) or a clause (311).

(309) *[*World war*]=**vatthu**.*
      world war=time
      'During the world war.' (K051206nar07)

(310) *[Itthu]=**vatthu** Malaysia=ka anà-duuduk Military Regiment.*
      DIST=time   Malaysia=LOC PAST-stay   Military Regiment
      'During that time there was a Military Regiment from Malaysia.'
      (K060108nar02)

(311) *[Kiccil]=**kapang** kithang sudaara pada samma cricket arà-maayeng.*
      small=when   1PL   siblings PL   all   cricket NON.PAST-play
      'When we were small, us children used to play cricket.' (K051201nar02)

In subordinate clauses, the prefix *arà-* codes simultaneity. In (312), the acts of seeing and playing are simultaneous, which is encoded by *arà-* on the verb *maayeng* 'play'. Taking (*ambel*) precedes the playing, but can be construed as simultaneous to seeing, or as preceding seeing.

(312) *Blaakang=jo incayang **anà**-kuthumung [moonyeth pada thoppi*
      after=EMPH   3S.POLITE PAST-see   monkey   PL   hat
      ***asà**-ambel   pohong atthas=ka **arà**-maayeng].*
      ANTERIOR-take tree   top=LOC   SIMULT-play
      'Then only he saw that the monkeys had taken his hats and were playing on the top of the trees.' (K070000wrt01)

Furthermore, simultaneity of two events can be expressed by putting one of the events in a reduplication construction (313) ⊡→7.5.4, p. 410.

(313) *Kithang nyaanyi˜nyaanyi su-thaaňdak.*
      1PL   sing˜RED   PAST-dance
      'We danced while singing/we sang and danced.' (K081114eli01)

A naturalistic example of this construction being used for simultaneity is (314).

---

[16]*Vatthu* 'time' is seen as a 'real' noun in this grammar, although it shows some behaviour typical of relator nouns. It appears that *vatthu* is acquiring the function of relator noun through grammaticalization. See DeLancey (1997) for a discussion of this path.

(314) *[Lu=ppe   muuluth=ka=le   paasir, se=ppe muuluth=ka=le*
      2S.FAMILIAR=POSS mouth=LOC=ADDIT sand   1S=POSS mouth=LOC=ADDIT

      *paasir katha **biilang˜biilang**] baaye=nang baapa=le   aanak=le   guula*
      sand   QUOT say˜red   good=DAT   father=ADDIT child=ADDIT sugar

      su-maakang.
      PAST-eat

      'Saying "There is sand in your mouth and there is sand in my mouth" both
      father and son ate the sugar.' (K070000wrt02)

15.5.3.4   Point coincidence

Point coincidence obtains when the end point of the first event coincides with the
beginning of the second event.

(315) *gr fig*
      ▬●▬

    It is expressed by the verbal prefix *kam-/ka-/kapang-* ⊡→6.2.1.7, p. 283. In example
(316), the lifting of the bamboo commences exactly at that point in time when the
steam comes out.

(316) *Aavi luvar=nang **kapang**-dhaathang, itthu bambu=yang giini*
      steam outside=DAT when-come   DIST bamboo=ACC like.this

      angkath=apa pullang arà-thoolak.
      lift=after   slow   NON.PAST-push

      'When the steam comes out, lift the bamboo like this and push it slowly.'
      (K061026rcp04)

    In example (317), the habit of gathering is lost as soon as the people take up work.

(317) a. *Inni pukurjan=nang **kam**-pii,*
          PROX work=DAT   when-go
          'When they go to this work,'

       b. *deram pada itthu mà-kumpul hatthu mosthor thraa.*
          3PL   pl   DIST INF-add   INDEF habit   NEG
          'they lack the habit of gathering' (G051222nar01)

    Even if the second state-of-affairs becomes true immediately on the completion
of the first one, *blaakang* 'after' is an alternative to *kapang-*. In (318), the speaker
became an orphan as soon as his parents died, not after that. Still *blaakang* is used.

(318) *Se=ppe umma=le   baapa=le   mniiňǧgal=nang blaakang, se hatthu*
     1s=POSS mother=ADDIT father=ADDIT die=DAT   after   1s INDEF
     yathiim su-jaadi.
     orphan PAST-become
     'After my father and mother had died, I became an orphan.' (K081114eli01)

15.5.3.5 Terminal boundary

The configuration of a ground point in time indicating the terminal boundary of the figure can be schematized as follows.

(319)
$$\begin{array}{c} fig \\ {\longrightarrow}\!\!\!\!\! \\ \uparrow_{gr} \end{array}$$
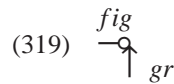
Terminal boundary is expressed by the adposition *sangke* 'until' ⊡→6.4.4.8, p. 350, which can be a postposition (320)-(322) or a preposition (323).

(320) *Suda derang=nang [hathyang muusing]=**sangke** mà-duuduk su-jaadi.*
     thus 3PL=DAT   other   time=until   INF-stay PAST-become
     'So they had to wait until another time' (K051220nar01)

(321) *Incayang [thujupul liima thaaun]=**sangke** incayang anà-iidop.*
     3s.POLITE seventy five year=until   3s.POLITE PAST-live
     'He lived until seventy-five.' (K060108nar02)

(322) *Suda thiga-pulu satthu=ka duudukapa thiga-pulu ìnnam=**sangke**.*
     so   three-ty one=LOC from   three-ty six=until
     'So, from '31 to '36.' (N061031nar01)

(323) *[Kaaki=ka gaaňdas-kang ambel anà-duuduk Aajuth=yang **sangke**=luppas]*
     leg=LOC   tie-CAUS   take PAST-stay   dwarf=ACC until=leave
     hathu pollu=dering Rose-red buurung=nang su-puukul.
     INDEF stick=ABL   Rose-red bird=DAT   PAST-hit
     'Rose red hit the bird with a stick until it let go of the dwarf he had taken in his claws.' (K070000wrt04)

Like the other postpositions, *sangke* can also attach to a clause (324). This is not possible for the prepositional use.

(324) *[Buruan diinging abbis]=**sangke** siithu su-siiňğgal.*
   bead cold finish=until there PAST-stay
   'The bear stayed there until the cold season was over.' (K070000wrt04)

The terminal boundary can coincide with the moment of speaking, but does not imply that the state-of-affairs will not continue afterwards, as in the following example.

(325) *Nyaari=**sangke** se inni ruuma=ka=jo arà-duuduk.*
   today=until 1s PROX house=LOC=EMPH NON.PAST-stay
   'I have been living here until today.' (K060108nar01)

In the example above, the terminal boundary is the time of speaking, but this does not imply that the speaker moves out of his house. It just indicates that for the moment the terminal boundary of the event 'living', which is still progressing, is the time of speaking.

15.5.3.6 Initial boundary

The initial boundary of a ground point in time, after which the period covered by the figure begins, can be schematized as follows.

(326) *fig*
     ⊶—
    ⌐ *gr*

Initial boundary is expressed by means of *(a)s(à)duuduk* ⊡→5.1.2.2, p. 168 (327). Whereas source in a spatial sense can be expressed by both *(a)s(à)duuduk* and *=dering*, the latter is not possible for temporal relations. An alternative realization of *asduuduk* is *duuduk=apa* (328).

(327) *Itthu muusing **asduuduk** sangke=nyaari se pukuran arà-gijja.*
   DIST time from until=today 1s work NON.PAST-make
   'From that time on until today I have been working.' (K060108nar01)

(328) *Suda thiga-pulu satthu=ka **duudukapa** thiga-pulu ìnnam=sangke.*
   so three-ty one=LOC from three-ty six=until
   'So, from '31 to '36.' (N061031nar01)

Initial boundary can also be expressed by *kapang-* 'when', as in (329), or by *vatthu=ka* 'at the time of' as in (330). In these examples, we are not dealing with point coincidence, since the blackness of the speaker did not come about on completion of the process of birth. This contrasts with becoming an orphan (318), which is

only true as soon as both parents have died.

(329) *See **kapang**-laaher, bannyak iitham.*
    1s   when-be.born   much   dark
    'When I was born, I was very dark.' (K081114eli01)

(330) *See=yang anà-braanak **vatthu=ka** asàduuduk, see iitham.*
    1s=ACC   PAST-bear   time=LOC   from   1s   dark
    'Ever since I was born, I have been very black.' (K081114eli01)

15.5.3.7 Intermediacy

(331)

Intermediacy is also expressed by *kapang-* 'when' ⊡→6.2.1.7, p. 283, as in (332), where the time of the earth quake entirely falls within the stay of the speaker in Pakistan.

(332) *Se Pakistan=ka   kapang-duuduk hatthu buumi ginthar-an   anà-jaadi.*
    1s Pakistan=LOC when-stay   INDEF earth shiver-NMLZR PAST-become
    'When I was in Pakistan, there was an earth quake.' (K081106eli01)

15.5.4 Phasal information

Besides being embedded in the time frame of the outside world, events also have an internal temporal structure. For expository reasons, we distinguish phasal information, which is concerned with the beginning, progress and end of an event from aspectual information, which is concerned with boundedness of the event. Depending on theoretical orientation, linguists may or may not agree with this choice, but it makes the presentation more straightforward. At this point, I do not want to endorse any particular theoretical relation between phasal and aspectual information, the separation is purely practical.

Phasal information is concerned with indicating the progress that the event has made in its completion. We distinguish the beginning, the progress and the end.

15.5.4.1 Beginning

To indicate that the action is in its inception, SLM uses the verb *mulain* 'start'. The main verb is marked with the infinitive prefix *mà-* and normally precedes *mulain*.

(333) *[sini=pe   raaja=nang mà-banthu] **anà-mulain***
      here=ᴘᴏss king=ᴅᴀᴛ   ɪɴꜰ-help   ᴘᴀsᴛ-start
      '(They) started to help the local king.' (K060108nar02)

(334) *Siithu=jo   see [mà-blaajar] anà-**mulain**.*
      there=ᴇᴍᴘʜ 1s   ɪɴꜰ-learn   ᴘᴀsᴛ-start
      'It was there that I started to study.' (K051213nar02)

15.5.4.2 Progression

Progression is normally not coded when referring to the present. In the past, progressive in subordinates can be coded by *arà-* ⊡→6.2.2.1, p. 287, used instead of the normal marking for past tense (*su-/anà-*). However, this could be analyzed as simultaneous relative tense as well.

(335) *Blaakang=jo incayang anà-kuthumung [moonyeth pada thoppi asà-ambel*
      after=ᴇᴍᴘʜ 3s.ᴘᴏʟɪᴛᴇ ᴘᴀsᴛ-see   monkey   ᴘʟ   hat   ᴄᴘ-take
      *pohong atthas=ka **arà-maayeng**].*
      tree   top=ʟᴏᴄ   sɪᴍᴜʟᴛ-play
      'Then only he saw that the monkeys had taken his hats and were playing on the top of the trees.' (K070000wrt01)

Progression can be emphasized by the use of the vector verb *duuduk* 'sit', as in the following three examples.

(336) *[Dee   arà-sbuuni **duuduk** cave] asaraathang sini=ka   asàduuduk*
      3s.ɪᴍᴘᴏʟɪᴛᴇ sɪᴍᴜʟᴛ-hide sit   cave] ᴄᴏᴘᴜʟᴀ   here=ʟᴏᴄ from
      hathu three miles cara jaau=ka.
      ɪɴᴅᴇꜰ three miles way far=ʟᴏᴄ
      'The cave where the evildoer remained hidden is about three miles from here.' (K051206nar02)

(337) *Incayang suda [aapa=ke hathu pukurjan] mà-girja arà-diyath **duuduk**.*
      3s.ᴘᴏʟɪᴛᴇ thus what=sɪᴍɪʟ ɪɴᴅᴇꜰ word   ɪɴꜰ-make ɴᴏɴ.ᴘᴀsᴛ-try sit
      'Now he is looking forward to do some kind or other of work.' (K051222nar08,K081104eli06)

(338) *Loram pada anà-dhaathang vakthu=dika se spaathu anà-gijja* **duuduk**.
2PL   PL   NON.PAST-come time=LOC   1s shoe   PAST/IRR-make stay
'When you came, I was making shoes.' (K081104eli06)

15.5.4.3 Continuation

To indicate that an event is still going on, *laile* is used.

(339) a. *Seelong independent state anà-jaadi=nang=apa,*
Ceylon   independent state PAST-become=DAT=after
'After Sri Lanka had become independent,'

b. *kithang=nang independence anà-daapath=nang=apa.*
1PL=DAT   independence PAST-get=DAT=after
'after we had obtained the independence,'

c. **laile** *derang anà-duuduk* under the Commonwealth.
still 3PL   PAST-stay   under the Commonwealth
'they still were under the Commonwealth.' (K051222nar06)

Laile is used in both positive and negative contexts, as the following two examples show.

(340) *Se **laile** lorang=nang saaya.*
1s still 2PL=DAT   love
'I still love you.' (K081104eli01)

(341) *Se=dang **laile** piisang **thàrà**-daapath.*
1s=DAT   still banana NEG.PAST-get
'I still have not got any bananas.' (K081104eli01)

Given that their phonological and semantic similarity leads to confusion, it is important to distinguish *laile* 'still', *lai* 'more', *laskali* 'again' and *laayeng* 'different', as illustrated in the following examples.

(342) *Naasi **laile** asà-maatham thraa.*
rice   still CP-cooked   NEG
'The rice is not cooked yet.' (K081104eli01)

(343)   **Lai** *masà-rubbus.*
        more must-boil
        'Cook it some more.' (K081104eli01)

(344)   **Laskali** *masà-rubbus.*
        again   must-boil
        'Cook it again.' (K081104eli01)

(345)   *Se=dang* **laayeng** *naasi maau.*
        1s=DAT   different rice   want
        ' I want different rice.' (K081104eli01)

15.5.4.4   End

To express that the event has reached its completion *ab(b)is* ⊡→5.1.5.3, p. 183 is used. Completive is not always easy to distinguish from subsequent events ⊙→15.5.3.2, p. 627, since the inception of the second event often implies the completion of the first.

(346)   *Kaaving* **abbis** *derang pada=nang=le aanak aada.*
        marry   finish 3PL   PL=DAT=ADDIT child   exist
        'The wedding finished, they also got children.' (K051206nar07)

15.5.5   Aspectual structure

Phasal information discussed in the last section is used to highlight a portion of an event, be it the beginning, the progress or the end. Aspectual information, which will be discussed now, is concerned with the internal structure of the event, i.e. is it conceived of as including its temporal boundaries or not. Iterative, distributive and habitual are not concerned with the internal structure of an event, but rather with several events; these areas are discussed under Event Quantification ⊙→15.6.2, p. 647.

Perfective and imperfective   Perfectivity and Imperfectivity are normally not expressed in SLM, but subsequent events marked by a chain of *asà*-marked verbs, or events marked for completive with *abis* are normally perfective (Smith & Paauw 2006:171)). All other TAM (quasi-)prefixes are underspecified for (im)perfectivity. It seems to be a reasonable hypothesis that the two past markers *anà-* and *su-* mark some kind of aspectual distinction, but evidence for this hypothesis could not be found.

See ⊡→6.2.1.3, p. 273 for a discussion of the similarities and differences between *su-* and *anà-*.

15.6 Quantification

Quantification can be used for referents (*two book*) or events (*he returned twice.*) Referent quantification will be discussed in ⊙→15.6.1, p. 639; quantification of events follows in ⊙→15.6.2, p. 647.

15.6.1 Quantification of referents

NPs based on nouns and under restricted circumstances NPs based on plural pronouns (and very rarely even clauses) can be modified for quantity. Different types of quantitative semantics are: plurality ⊙→15.6.1.1, p. 639, definite quantity ⊙→15.6.1.2, p. 640, indefinite quantity ⊙→15.6.1.3, p. 641, definite order ⊙→15.6.1.4, p. 646, and indefinite order ⊙→15.6.1.5, p. 646.

15.6.1.1 Plurality

Plurality can be expressed on NPs based on nouns by the plural marker *pada* ⊡→6.4.2, p. 323. This is not obligatory if the plurality can be inferred from context. The following examples show the presence and absence of the plural marker on the word *mlaayu* 'Malay' in very similar contexts.

(347) *Bannyak mlaayu **pada** Hambanthota=ka arà-duuduk.*
much　　Malay　PL　　Hambantota=LOC NON.PAST-exist.ANIM
'There are many Malays in Hambantota.' (B060115nar02)

(348) *Cinggala=∅ aada　mlaayu=∅ aada* Moor=∅ *aada mulbar=∅ aada.*
Sinhala　　　exists Malay　exist Moor　exist Tamil　　exist
'There are Sinhalese, Malays, Moors, and Hindus.' (G051222nar04)

Plurality need not be expressed if a numeral is mentioned in the clause, as in (349), but nothing precludes using *pada* nevertheless, as in (350).

(349) *Kithang hathu* week=*nang hathu skali **duva** skaali=∅=ke　arà-maakang.*
1PL　　one　week=DAT　one　time　two　time=　=SIMIL NON.PAST-eat
'We might eat it one or twice a week.' (K061026rcp04)

(350)   *Se=dang duppang lai*   ***thiiga*** *generation **pada** ini   nigiri=ka*
        1s=DAT before   other three   generation PL   PROX country=LOC
        anà-duuduk
        PAST-exist
        'Before me another three generation existed in this country' (K060108nar02)

Plurality can be emphasized on plural pronouns by adding *pada* as well. In the
following example, the plurality of the word *kitham* '1pl' is emphasized by adding
pada.

(351)   ***Kitham pada**=pe baasa=ka*   *nni* grammar *thraa.*
        1PL   PL=POSS language=LOC PROX grammar NEG
        'There is no grammar in our language.' (G051222nar02)

Unlike the example above, the following example has *kitham* in it without *pada*.

(352)   *Ini   **kitham**=∅=pe nigiri   su-jaadi.*
        PROX 1PL=∅=POSS   country PAST-become
        'This (country) became our country.' (K051222nar04)

Plurality can also be expressed with *pada* on finite headless relative clauses, as in
(353).

(353)   *[[Seelon=nang anà-dhaathang]**pada**] mlaayu pada.*
        Ceylon=DAT   PAST-come   PL   Malay PL
        'Those who had come to Ceylon were the Malays.' (N060113nar01)

What is treated as 'plurality' here might actually rather be an instance of 'collec-
tive nominal aspect' (Rijkhoff 2002). See ⊡→6.4.3, p. 325 for a discussion.

15.6.1.2   Definite quantity

Absolute number can be indicated by cardinal numbers ⊡→5.10, p. 242 on nominal or
pronominal NPs. Numerals precede or follow the noun (354)(355), but always follow
the pronoun (356).

(354) ***Thiiga*** *oorang,* ***thiiga*** *oorang=le,* ***thiiga*** *oorang pada=jo itthu ini*
     three   man,   three man=ADDIT, three man   PL=EMPH DIST PROX
     volleyball *arà-*play*-king=kee.*
     volleyball NON.PAST-play-CAUS=SIMIL
     'Three men, three persons, three people play it, like volleyball.'
     (N060113nar05)

(355) *[panthas rooja kumbang pohong komplok **duuva**] asà-jaadi su-aada.*
     beautiful rose flower   tree   bush   two   CP-grow PAST-exist
     'Two beautiful rose bushes had grown.' (K070000wrt04)

(356) *Mr Sebastian aada, se aada,* ***kitham duuva*** *arà-oomong.*
     Mr Sebastian exist 1s exist 1PL   two   NON.PAST-speak
     'You are here, I am here, the two of us are talking.' (K060116nar05)

The plural particle *pada* is often present when numerals modify nouns, but not obligatory (358). It is not present when the numeral modifies a pronoun (356).

(357) *Thuuju* generation ***pada*** *asà-biilang.*
     seven   generation PL   CP-say
     'Seven generations, they say.' (K060108nar02)

(358) *Mlaayu thigapuluthuuju baasa*   ∅ *aada.*
     Malay 37   language   exist
     'There are 37 Malay languages.' (K060116nar02)

For nominal NPs, the numeral may also follow, but is more often put in front of the noun. *pada* is normally present.

Measures normally do not take *pada*. In example (359), the word *kaayu* 'mile' is used without *pada*.

(359) *Duuva kaayu ∅ kithang masà-pii.*
     two   mile   1PL   must-go
     'We have to walk two miles.' (K051213nar03)

15.6.1.3 Indefinite quantity

In the domain of indefinite quantity, we can distinguish the extreme points of totality and zero quantity, with a number of intermediate levels.

In SLM, quantity of entities can be expressed by quantifiers like *bannyak* 'many'
⊡→5.9, p. 240 or by WH-words combined with boolean clitics like *saapa=so* 'some-
one', *kaapang=ke* 'some day' *kaapang=le* 'always' or *kaapang=pon* 'never' ⊡→8.4,
p. 447.

Normally, there is a set of referents over which the quantity ranges, like *aanak*
in (360), but it is also possible to express general quantities which range over all
imaginable referents (361).

(360)   ***Bannyak aanak*** *pada karang mlaayu thama-oomong.*
        much   child   PL   now   Malay   NEG.NONPAST-speak
        'Many children do not speak Malay now.' (G051222nar02)

(361)   ***Bannyak*** *mà-biilang thàrboole.*
        much   INF-say   cannot
        '(I) can't tell you much.' (K051206nar12)

Zero proportion   A proportion of zero is indicated by a negated predicate, mostly of
the verbal, existential or adjectival type. Zero proportion is mostly not marked by any
other marker than the negation of the predicate. In (362), the verbal negator *thama-*
suffices to imply that none of the fighters changed their words.

(362)   *Derang pada* 0 *kathahan thama-thuukar.*
        3PL   PL   word   NEG.IRR-change
        'They would not change their words.' (K051213nar06)

Similar things can be said about (363), where the absence of any Malay is only
marked by the existential negator *thraa*.

(363)   *Itthuka*   0 *mlaayu **thraa**, bannyak=nang* English=*jo   aada.*
        dist=LOC   Malay NEG   much=DAT   English=EMPH exist
        'There is no Malay over there, it is all English which is there.' (B060115prs15)

If the total absence of applicable reference shall be emphasized, the enclitic =*pon*
can be used. This is the case in (364) where the set to which the negative predicate
applies is stated by *oorang* 'men' and the totality of the absence is indicated by a
following =*pon*. Note that *oorang* is in this case accompanied by the indefinite article
hatthu. In this construction, the indefinite article precedes the noun, whereas in other
constructions, it can also follow the noun.

(364) *Kithang **hatthu**=oorang=**pon** thàrà-iingath.*
  1PL   INDEF=man=any   NEG.PAST-think
  'We cannot think of any person.' (B060115nar02)

Total absence of any referent of any set is applied by adding =*pon* to the indefinite article *atthu* as in (365). The missing mention of domain (cf. *oorang* just above) indicates that the negation ranges over any and all possible referents, there is not a single thing which you can do on the bus.

(365) Bus=*ka* ∅ **hatthu=pon** *mà-kirja thàràboole.*
  bus=LOC INDEF=NEG INF-make cannot
  'You can't do anything on the bus.' (K061125nar01)

More than one entity   Cardinality greater than one need not be specified, but can be expressed by a numeral (366), the plural word *pada* (367), or both, if necessary (368).

(366) ***Duva-pulu ìnnam riibu   ùmpath raathus lima-pulu duuva*** votes ∅
  two-ty   six   thousand four   hundred five-ty   two   votes
  incayang=nang anà-daapath.
  3s.POLITE=DAT   PAST-get
  'He got 26,452 votes.' (N061031nar01)

(367) *Itthu vatthu=ka itthu nigiri **pada**=ka arà-duuduk.*
  DIST time=LOC DIST land PL=LOC   NON.PAST-stay
  'At that time, (they) lived in those countries.' (N060113nar01)

(368) *Se=dang **liima** anak klaaki **pada** aada.*
  1s=DAT five child male PL   exist
  'I have five sons.' (K060108nar02)

Furthermore, the following sections also treat different possibilities to indicate sets of cardinality greater than one.

Low proportion   A low proportion of referents, like English *few,* is indicated by *sdiikith* or *konnyom* and the affirmative or negative predicate, as the case may be.

(369)   *Siithu Dubai=ka Sri Lanka=pe   oorang mlaayu pada **sdiikith***
        there Dubai=LOC Sri Lanka=POSS man   Malay   PL   few
        arà-duuduk.
        NON.PAST-saty
        'There in Dubai, there are few Sri Lankan Malays.' (K061026prs01)

(370)   ***Konnyong** mlaayu=jo   Seelong=ka thiinngal aada.*
        few   Malay=EMPH Ceylon=LOC settle   exist
        'Few Malays have settled down in Sri Lanka.' (K051222nar06)

Unclear proportion   To indicate that the exact proportion is unclear, but not negligible, *spaaru* is used. *Spaaru* indicates a slightly higher proportion than *sdiikith* or *konnyom*. It can refer to about half of the set, which is not possible for *sdiikith* or *konnyom*. In this respect, *spaaru* resembles English *some*, which covers more important proportions than *few*.

(371)   ***Spaaru** oorang pada su-pii, **spaaru** oorang pada su-birthi.*
        some   man   PL   PAST-go some   man   PL   PAST-stop
        'Some people left and some people stayed.' (B060115nar01)

(372)   ***Spaaru** oorang pada thàrà-pii   sindari.*
        some   man   PL   NEG.PAST-go from.here
        'Some men did not go from here.' (K051206nar07)

High proportion   A high proportion is indicated by *bannyak*.

(373)   *Itthu=pe   pada=jo **bannyak** mlaayu pada karang siini aada.*
        DIST=POSS PL=FOC   much   Malay PL   now   here exist
        'It's their folks we get a lot of today here.' (K051205nar04)

(374)   *Oorang pada thiikam apa, oorang pada=nang theeṁbak apa, se=dang*
        man   PL   stab   after man   PL=DAT   shoot   after 1s=DAT
        ***bannyak** creeveth pada su-aada.*
        much   trouble PL   PAST-exist
        'People were stabbed, people were shot, I had a lot of problems.'
        (K051213nar01)

Additionally, a lexical solution like *punnu* 'full' can be used (375)(376). Using *guunung* 'mountain' for this was also overheard, but could not be verified in the corpus and needs additional verification.

(375) *Pake-yan=ka=le*   ***punnu*** *bedahan thraa*
       dress=NMLZR=LOC=ADDIT full   difference NEG
       'Even at our dress there is not much difference.'

(376) ***Punnu*** *mlaayu oorang=nang=le cinggala   mà-blaajar thàrà-suuka=nang;*
       full   Malay man   DAT=ADDIT Sinhala   INF-learn
       derang   laayeng nigiri pada=nang su-pii.
       NEG-like=DAT 3PL   other country   PL=DAT PAST-go
       'Many Malays did not want to learn Sinhala and went to other countries.'
       (K051222nar06)

Totality   To indicate that all entities of the set are included in the predication, either *samma* 'each/every/all' ⊡→5.9, p. 240 or the WH=*le*-construction ⊡→8.1.12, p. 424 is used.

Example (377) shows the use of bare *samma* to indicate the totality of the people knowing the speaker.

(377) Doctors *pada=so* police *ASP=so*   judge=*so,*   ***samma*** *oorang thaau*
       doctors PL=UNDET police ASP=UNDET judge=UNDET all   man   know
       see=yang
       1s=ACC
       'Whether they be doctors, policemen, assistant superintendents of police or judges, all men know me.' (B060115nar04)

More emphasis on the totality can be given by adding =*le* to either the entity itself (378) or *samma* (379).

(378) ***Samma*** *oorang=**le** saanak.*
       all   man=ADDIT relative
       'All people are relatives.' (K051206nar07)

(379) *Suda incayang=pe   aanak pada **samma=le**   musicians pada=jo.*
       thus 3s.POLITE=POSS child PL   every=ADDIT musicians PL=EMPH
       'So all his children are musicians.' (G051222nar01)

Finally, the combination of an appropriate WH-word with again =*le* yields a uni-

versal quantifier as well. In this case, =*le* can be attached to the entity itself as in (380) or to the predicate as in (381)

(380)   *Skarang **maana** aari=**le**   atthu atthu oorang=yang arà-buunung.*
        now   which day=ADDIT one   one   man=ACC   PAST-kill
        'Now people kill each other every day.' (K051206nar11)

(381)   *Lai   **saapa** mlaayu kuthumung=**le** aapa=ke   connection hatthu aada.*
        other who Malay see=ADDIT   what=SIMIL connection INDEF   exist
        'If you see any other Malay, there will always be some kind of connection.'
        (K051206nar07)

15.6.1.4   Definite order

Definite order is expressed by an ordinal derived from the cardinal by *ka-* ⊡→6.2.6.3, p. 316. Ordinal numbers can only precede the noun. They do not seem to be used with pronouns or clauses.

(382)   *Se asdhaathangpa kitham=pe femili=ka   **ka-duuva** aanak*
        1s COPULA   1PL=POSS   family=LOC ORD-two   child
        'I am the second child in our family.' (K060108nar01)

15.6.1.5   Indefinite order

To indicate indefinite order of an entity, the ablative =*dering* is used together with the relator noun *daalang* 'inside'. A precise quantity like *liima* 'five' can be used (383), or a collective noun like *kumpulan* 'group' in (383).

(383)   *Se pàrthaama *(liima) oorang=dering daalang race=yang su-abbis-king.*
        1s first   five   man=ABL   inside race=ACC PAST-finish-CAUS
        'I finished the race among top five.' (K081106eli01)

(384)   *Se pàrthaama kumpulan=dering daalang race=yang su-abbis-king.*
        1s first   group=ABL   inside race=ACC PAST-finish-CAUS
        'I finished the race among the first.' (K081106eli01)

15.6.2 Event quantification

Next to referents, events can also be quantified (Hengeveld & Mackenzie 2008:179f). I will limit myself to events taking place several times in this discussion. We can distinguish habitual event quantification, which indicates that the event takes place regularly, from iterative event quantification, which indicates that the event takes place more than once, but without any commitment to the regular status. Another type of event quantification is the distributive, which indicates that the event was performed by several participants at different points in time, as in *the guests arrived two by two*. There number of events in this case is clearly higher than one, there are several instances of arrivals. Distributive event quantification will be the last type to by discussed.

Habitual   Habitual can be marked by the present tense quasi-prefix *arà-* (Ansaldo 2009b) ⊡→6.2.2.1, p. 287 (385) or the irrealis quasi-prefix *anthi-* ⊡→6.2.2.2, p. 290 (386).

(385)   a. *Hindu **arà**-maakang kambing.*
            Hindu NON.PAST-eat   goat
            'Hindus eat goat.'

        b. *samping **thuma**-maakang*
            beef   NEG.NONPAST   eat
            '(they) don't eat beef.' (K060112nar01)

(386)   *Saudi=ka   ontha **anthi**-kaasi.*
        Saudi.Arabia=LOC camel IRR-give
        'In Saudi Arabia, they give camels.' (K060112nar01)

Irrealis marking for habitual context can also be used with past reference (cf. Givón 1994). The following example has the adverb *kàthaama* 'before', which indicates past reference, but still uses *anthi-* to convey the habitual reading.

(387)   *Punnu mlaayu pada **kàthaama** English=jona **anthi**-oomong.*
        many Malay PL before   English=PHAT IRR-speak
        'Many Malays used to speak English in former times, isn't it?'
        (K051222nar06)

Things are the same with the following example, where an explicit reference to the past ('There once was a time') is combined with the irrealis marker to convey habituality.

(388)   *Samma kithang=pe mlaayu, hathu muusing su-aada,   samma*
      all   1PL=POSS Malay INDEF time   PAST-exist all
      cinggala=dering=jo **athi***-oomong.*
      Sinhalese=ABL=EMPH IRR-talk
      'There once was a time where all Malays (in our home) would speak in Sin-
      hala.' (B060115cvs01)

Iterative   There are no means to indicate iterative or repetitive aspect in the strict
sense. The repetition of an event can be indicated by *laskal(l)i* 'again'.

(389)   *Asà-kumpul, blaakang* **laskalli** *asà-beat-king, inni daalang=ka kithang*
      CP-add   after   again CP-beat-CAUS PROX inside=LOC 1PL
      aayer masà mà-libbi-king.
      water must INF-remain-CAUS
      'Having added (the ingredients), having then beaten them again, we must re-
      serve the water in the middle.' (B060115rcp02)

    Laskalli can be reduplicated as in (390).

(390)   *Se=ppe pake-yan   pada bannyak koothor. Itthu=subbath laskalli laskalli*
      1s=POSS dress-NMLZR PL   much   dirt   DIST=because again   again
      mà-cuuci su-jaadi
      INF-wash PAST-become
      'My clothes were very dirty. That's why I had to wash them over and over
      again.' (K081106eli01)

Another construction which comes close in meaning and involves reduplication of
an infinitive is given in (391).

(391)   *Itthu=nang blaakang karang aari* **mà-pii m-pii**   *karang kitham=pe mlaayu*
      DIST=DAT after   now   day INF-go PAST-go now   1PL=POSS Malay
      arà-mulain.
      NON.PAST-start
      'Now, after that, as days went by, our Malay is beginning.'
      (K051222nar03)(K081106eli01)

Distributive   Distributive is indicated by a reduplicated numeral, a head noun, and
the plural marker.

(392)

$$\text{NUM~NUM N PL} \left\{ \begin{array}{l} \text{NP} \\ \text{PP} \\ \text{ADJCT} \end{array} \right\} * \text{PRED}$$

This is possible for numbers greater than one (393), but also for *hatthu* 'one' (394) (395).

(393) *Duuva duuva oorang pada su-pii.*
two   two   man   PL   PAST-go
'People left two by two.' (K081106eli01)

(394) *Hatthu hatthu oorang pada su-pii.*
one   one   man   PL   PAST-go
'People left one by one.' (K081106eli01)

(395) Kandy Malay Association=*dering* **hatthu hatthu** *oorang* **pada**
Kandy Malay association=ABL   INDEF   INDEF   man   PL
arà-lompath   Hill Country=*nang.*
NON.PAST-jump hill   country=DAT
'More and more people stepped over from the KMA to the Hill Country Malay
Association.' (K060116nar07,K081106eli01)

15.6.3   Temporal frequency

Generally, temporal frequency can be expressed lexically by a temporal noun like *aari* 'day' or *vatthu* 'time' which is modified by a quantifier ⊡→5.9, p. 240, such as *spaaru vatthu* 'some time'. The reader is referred to ⊙→15.6.1.3, p. 641 for more discussion of this pattern involving quantifiers. The temporal domain ⊙→15.1.2.17, p. 580 (like 'per week', 'per year') is indicated by =*nang* ⊡→6.4.4.2, p. 334.

Zero frequency is expressed by a negated predicate. There is no special word for *never*, but the negation can be reinforced by *kaapang* 'when' with the enclitics =*pon* 'any' or =*le* 'additive', as in example (396) and (397).

(396) *Suda itthu kithang=nang* **kaapang=pon** *thama=luupa.*
thus   DIST   1PL=DAT   when=any   NEG.NONPAST=forget
'So, we will neverever forget this.' (B060115nar02)

(397) *Go*   **kaapang=le** *saala   thama-gijja.*
1s.FAMILIAR when=ADDIT wrong NEG.NONPAST-make
'I never do any wrong.' (B060115nar04)

The opposite of this is total frequency (*always*). This can be expressed either lexically by *subbang vatthu* or by again *kaapangle* 'when=ADDIT' (398).

(398)   *Girls'* High School *Kandi=ka   se=dang **kaapang=le** udahan   hatthu*
     girls   high   school   Kandy=LOC   1s=DAT   when=ADDIT   invitation   INDEF
     arà-kiiring.
     NON.PAST-send
     'I am always invited to the Girls' High School in Kandy.' (K061127nar03)

Finally, the WH=*le*-construction with a temporal noun can be used as well.

(399)   *Dee **maana** aari=le   asà-dhaathang, thìngaari vakthu=nang kalthraa*
     3   which   day=ADDIT   CP-come   noon   time=DAT   otherwise
     maalang vakthu=nang ...
     night   time=DAT
     'He came every day, at noon or otherwise during the night (and attacked).'
     (K051206nar02)

15.7   Modality

Speakers do not only exchange absolute truths. They also convey their estimation of the likelihood, necessity and desirability of a situation. This is the domain of modality. Following Hengeveld (2004a), modality can be divided along two axes, the target of evaluation ⊙→15.7.1, p. 650 and the domain of evaluation ⊙→15.7.2, p. 651.

15.7.1   Targets of modality

As for the target of modality, we can distinguish whether it affects an event as in *One must go to school*, where the obligation conveyed by *must* covers the general need to attend school. This event-oriented modality can be contrasted with participant-oriented modality, as in *John must go to school at 7h45*, where the target of the obligation is a participant, John. In SLM, the difference between participant-oriented modality and event-oriented modality is marked by the presence vs. absence of the participant alone.

In SLM, any event-oriented target can be made participant-oriented by adding an NP including the participant. This participant is normally marked with the dative =*nang* ⊞→6.4.4.2, p. 334. Conversely, any participant-oriented modality statement can be changed into an event-oriented one by dropping the participant (Although this is potentially ambiguous due to the high number of dropped topical participants in general, see ⊙→16.1.2, p. 679). In example (400), we are dealing with event modality: no one can go far into that cave. Hence no participant is mentioned. In (401), on the

other hand, we are dealing with participant modality, because it is only the speaker who is unable to go. This is indicated by mentioning *se* '1s' and adding the dative, in this case *=dang*.

(400) *Bannyak jaau mà-pii **thàrboole**, itthu=ka*
     much   far INF-go cannot   DIST=LOC
     'You/one cannot go far there/It is impossible to go far into that cave.'
     (K051206nar02)

(401) ***Se=dang** karang jaau mà-pii **thàràboole**.*
     1s=DAT   now   far INF-go cannot
     'I cannot go far now.' (K061120nar01)

In (402), a recipe, we are dealing with event-oriented modality. Eggs are required for the recipe no matter the identity of the participant.

(402) *Thullor maau.*
     egg   want
     'You need eggs [to make a dessert].' (B060115rcp02)

This contrasts with (403), where the participant for whom the necessity holds is specified, the speaker, expressed by *sedang*. People other than the speaker (who might have different requirements for their jobs) are not under this obligation.

(403) ***Se=dang** se=ppe pukujan pada=nang baaru hatthu kaar **maau***
     1s=DAT 1s=POSS work   PL=DAT   new INDEF car need
     'I need a new car for my work.' (K081114eli01)

A third type is proposition-oriented modality, covering the speaker's views and beliefs, and commitment to the truth of the utterance. This type can be found in the epistemic and evidential domains and will be covered below.[17]

15.7.2 Domains of modality

Hengeveld (2004a) distinguishes 5 domains of modality:

- facultative modality, concerned with intrinsic or acquired capacities

- deontic modality, concerned with what is permissible

- volitive modality, concerned with what is desirable

---

[17]There are some languages where proposition-oriented modality can also be combined in the volitive domain, but this is not the case in SLM.

- epistemic modality, concerned with what is known about the actual world

- evidential modality, concerned with the source of information

These subtypes will be discussed in turn.

15.7.2.1   Facultative modality

Facultative modality deals with the (in)capacities of participants and general facilities which permit performance of certain actions.

Capacity   Capacity is coded by *boole* ⊡→5.12.1, p. 248. There is no difference between physical (404) and mental ability (405).

(404)   *[**Boole** oorang pada]=na siithu boole pii.*
       can   man   PL=DAT   there can go
       'The men who can go may go.' (B060115cvs01)

(405)   *Cinggala bahasa   saapa=nang=le **bole**=bicaara siini.*
       Sinhala   language who=DAT=ADDIT can-talk   here
       'Anybody can talk Sinhala here.' (K051206nar14)

    Example (404) shows the use of *boole* to indicate capacity (first occurrence) and permission (second occurrence). This diverging interpretation of *boole* in the two clauses makes the sentence non-tautological. In return, speakers normally do not utter tautologies, so that the fact that (404) was uttered proves that the two tokens of *boole* must indeed have different interpretations.

Incapacity   Incapacity is expressed by *thàrboole* ⊡→5.12.2, p. 249.

(406)   *Bannyak mà-biilang **thàrboole**.*
       much   INF-say   cannot
       '(I) can't tell you much.' (K051206nar12)

(407)   *Derang pada=nang atthu=le   mà-kijja=nang **thàràboole**=subbath ....*
       3PL   PL=DAT   one=ADDIT INF-do=DAT   cannot=because
       'Because they couldn't do anything.' (N060113nar01)

(408)   *Lorang=nang aayer=ka   appi mà-mnyala-king thàràboole.*
       2PL=DAT   water=LOC fire INF-burn-CAUS   cannot
       'You cannot light fire in water.' (K081114eli01)

15.7.2.2 Deontic modality

Deontic modality covers the external or internal need to (not) perform a certain action.

Permission Permission is coded by the modal particle *boole* ⊡→5.12.1, p. 248.

(409) *Lai sdiikith aari=jo go=dang **bolle**=duuduk.*
other little day=EMPH 1s.FAMILIAR can=stay
'I may stay some more days [on Earth because Allah allows me to].'
(B060115nar04)

Interdiction Deontic interdiction is normally expressed by a periphrasis involving the adjective *thàràbaae* 'not.good'. Another possibility is the use of *thàrboole* 'cannot' ⊡→5.12.2, p. 249. The interdiction is inferred from the primary meaning of incapacity.

(410) Cigarette *mà-miinong **thàràboole**.*
cigarette INF-drink cannot
'It is forbidden to smoke.' (K060116nar04)

(411) *Maalang=nang mà-svara-kang **thàràboole**.*
night=DAT INF-noise-CAUS cannot
'It is forbidden to make noise at night.' (K081114eli01)

The speech act of interdiction is not deontic and is covered under prohibition ⊙→17.1.5, p. 699.

Obligation Obligation is most often expressed by the quasi-prefix *masthi-* ⊡→6.2.2.3, p. 294. This is an exception to the rule in that it does not take the dative, but the nominative. Still, it can be used for participant modality as in (412) or for event modality as in (413).

(412) a. *Luu=∅ baaye=nang **masà**-blaajar baaye=nang **masà**-mnaaji.*
2s.FAMILIAR good=DAT must-learn good=DAT must-recite
'You have to learn well and you have to recite well.'

b. *Lu=ppe umma-baapa=nang baaye=nang **masà**-kaasi thaangang.*
2s=POSS mother-father=DAT good=DAT must-give hand
'You must lend a hand to your parents.' (K060116sng01)

(413) *Hathu oorang kala-pasiith, hathu* chance **masà**-*kaasi ithu oorang=nang=le.*
       INDEF man   if-trouble   INDEF chance must-give   DIST man=DAT=ADDIT
       'When a person has trouble, we must give him a chance.' (K060116nar07)

Periphrastic constructions involving the existential *aada* and *jaadi* 'become' are
also possible ⊡→5.1.6, p. 192.

(414) *Se=ppe* profession=*subbath* **se=dang** *siini* **mà**-*pii su*-**jaadi**.
       1s=POSS profession=because 1s=DAT   here INF-go PAST-become
       'I had to come here because of my profession.' (G051222nar01)

(415) **Se=dang aada** *ini*   army *pada=yang* **mà**-*salba-kang*=**nang**.
       1s=DAT   exist PROX solder PL=ACC   INF-escape-CAUS=DAT
       'I had to save these soldiers.' (K051213nar01)

The construction with *aada* indicates a general obligation, whereas the construc-
tion with *jaadi* indicates an obligation which was caused by a change of circum-
stances. In example (414), the need to move to another town was caused by a change in
professional status, whereas in (415), no such change happened, it was just the general
human obligation to help the soldiers which obtained. The obligation to help people
in need is not brought about by a change in environment, but is a moral value (even if
in this case it is applied to a concrete situation). This general obligation irrespective
of circumstances is indicated by *aada*.

The following examples show the difference between the constructions with *jaadi*
and with *aada*. The need to swim in case of a flood is indeed caused by the circum-
stances, which is why *jaadi* is used. That need is not caused by authority, which is
why it is impossible to use *aada* there.

(416) *Hatthu baanjir anà-dhaathang=siking, se=dang* **mà**-*bìrnang*
       INDEF flood PAST-come=because   1s=DAT INF-swim
       su-**jaadi**/*su*-**aada**.
       PAST-become/PAST-exist
       'Because a flood was coming, I had to swim.' (K081114eli01)

The inverse situation holds in (417), where the need to pray five times a day is
imposed by religious authority and not by the circumstances. This is why *aada* is
possible, and *jaadi* is not.

(417) *Lorang=nang hatthu aari=nang liima skali **mà**-sbaayang*
       2PL=DAT   one   day=DAT five time INF-pray
       aada/*arà-jaadi.
       exist/NON.PAST-become
       'You have to pray five times a day.' (K081114eli01)

Finally, there are some cases where the need to do something is permanent, and not caused by authority or exceptional circumstances. One of these cases is breathing. Then, only the quasi-prefix *masà-* is possible.

(418) *Mà-iidop=nang, naapas **masà**-ambel.*
       INF-live=DAT   breath must-take
       'You must breathe in order to survive.' (K081114eli01)

(419) **mà-iidop=nang naapas **mà**-ambel **aada**.*
       INF-live=DAT   breath INF-take exist
       '(You are compelled to breathe in order to survive).' (K081114eli01)

Necessity   Necessity is expressed by *(ka)mau(van)* ⊡→5.12.3, p. 250. This particle is also used for desire.

(420) *Thullor maau.*
       egg   want
       'You need eggs [to make a dessert].' (B060115rcp02)

(421) *Kithang=nang hathu* application *mà-sign **kamauvan** vakthu=nang=jo,*
       1PL=DAT   INDEF application INF-sign want   time=DAT=EMPH
       kithang arà-pii   inni politicians pada dìkkath=nang.
       1PL   NON.PAST-go PROX politicians PL   vicinity=DAT
       'When we want/have to sign an application, we approach these politicians'
       (K051206nar12)

The borderline between obligation, necessity and desire is generally not clear cut in the Lankan languages. Sri Lanka Malay has a more clear cut distinction between obligation conveyed by *masthi-* and desire conveyed by *(ka)maau(van)*, but this seems to be eroding, and indiscriminate use of *maau* for both desire and obligation is becoming more common, paralleling the Sinhala and Tamil semantics. Example (422) shows the use of *maau* to convey mild obligation. This particle is normally used to express desire (see above).

(422) *Baapa=nang* mosque=*nang mà-pii* **maau**.
　　 father=DAT　mosque=DAT　INF-go　want
　　 'Father, you should go to the mosque. (*literally*: Father, you want to go to the mosque)' (B060115nar04)

Lack of necessity　Lack of necessity has to be distinguished from lack of desire ⊙→15.7.2.3, p. 657. Lack of necessity is coded by *thàr(ka)mauvan* ⊡→5.12.3, p. 250, while lack of desire is coded by *thussa* ⊡→5.12.4, p. 253.[18]

(423) *Se=dang se=ppe　pukujan pada=nang baaru hatthu kaar* **maau**
　　 1s=DAT　1s=POSS work　PL=DAT　　new　INDEF car　need
　　 'I need a new car for my work.' (K081114eli01)

(424) *Loram pukujan=nang baaru hatthu kaar* **thàràmauvan**.
　　 2PL　work=DAT　　new　INDEF car　NEG.need
　　 'You do not need a new car for your work.' (K081114eli01)

15.7.2.3　Volitive modality

Volitive modality is concerned with what is desirable. The desire can be participant-oriented as in *John wants to give money to the poor*, event-oriented *Giving money to the poor is desirable*, or proposition-oriented *Mary wants John to give money to the poor*. The difference is that the person feeling the desire participates in the desired event in the first case and is not mentioned in the second case. In the third case, the person feeling the desire is mentioned but does not participate in the event.

　　We can distinguish presence of desire from absence of desire. These two possibilities will be discussed below.

Desire　Desire is coded by the modal particle *(ka)mau(van)* ⊡→5.12.3, p. 250, which derives from the lexical word *kamauvan* 'desire'. The meaning of *kamauvan* is ambiguous between 'need' and 'desire' (425).[19]

---

[18]My informants furthermore added that *thàràjamauvan* was a 'common mistake', so that that form seems to have a certain frequency as well, even if I have not come across it personally.

[19]This ambiguity only exists in affirmative contexts, in negative contexts, lack of need is coded by *thàrkamauvan* (424), and lack of desire is coded by *thussa* (430).

(425) *Se=dang baaru hatthu kaar (**ka)maau(van)**.*
1s=DAT new INDEF car want
'I want/need a new car.' (K081114eli01)

The following three examples show the use of *maau* for a desired object (426) and a desired state (427). Both examples are participant-oriented in that the person(s) feeling the desire are given.

(426) *Deran=nang thumpath **maau**.*
3PL=dat place want
'They wanted land.' (N060113nar01)

(427) *Itthusubbath=jo incayang=nang ini Sri Lankan Malay mà-blaajar **maau**.*
therefore=EMPH 3s.POLITE=DAT PROX Sri Lankan Malay INF-learn want
'This is why he wants to learn this Sri Lanka Malay.' (B060115prs15)

As for event-oriented volitive modality, it appears that SLM, like the other Sri Lankan languages, does not draw a sharp distinction between what is desirable (volitive) and what is required (by law or morale, deontic). Event-oriented volitive modality seems to be expressed by *masthi-*, otherwise used for obligation, as shown below.

(428) *Minnyak klaapa ini raam̆buth=dering mas-goosok.*
coconut.oil coconut PROX hair=ABL must-rub
'You must rub coconut oil (over the itching) with (human) hair.' (K060103cvs02)

There is only one example of proposition-oriented volitive modality in the corpus, which is by a speaker who frequently drops grammatical morphemes, so that this example must be taken with a grain of salt. It is conjectured that the infinitive prefix *mà-* should surface in the subordinate clause, but this could not be tested.

(429) *Kithang=nang maau kitham=pe mlaayu lorang blaajar, lorang=pe mlaayu*
1PL=DAT want 1PL=POSS Malay 2PL learn 2PL=POSS Malay
kitham blaajar.
1PL learn
'We want that you learn our [Sri Lankan] Malay, and we learn your [Malaysian] Malay.' (K060116nar02)

Lack of desire Lack of desire is expressed by *thussa* ⊡→5.12.4, p. 253. This is thus different from the coding of lack of necessity, which is *thàrkamauvan* ⊡→5.12.3,

p. 250 (424).

(430) *Se=dang baaru hatthu kaar thussa.*
       1s=DAT new   INDEF car NEG.want
       'I do not want a new car.' (K081114eli01)

15.7.2.4 Epistemic modality

Epistemic modality is concerned with the likelihood, probability and faith that speakers have in the truth of the propositions they state. Foley & Van Valin (1984:213) give the following continuum for epistemic modality

(431) real ← necessary – probable – possible → unreal

In SLM, five levels of certainty can be distinguished. This was tested with a small made up setting. Some relatives have left for Badulla, and the question is whether at the time of speaking they have reached Badulla or not. The informants have spontaneously proposed percentages of probability for the different constructions, and this is repeated here.

(432) *Derang Badulla=na sampe*
       they   Badulla=DAT reach

> *aada*   *100%*
> exist
> 'They have arrived at Badulla.'
>
> *anthi=jo aada*   *75%*
> IRR=FOC exist
> 'They will have arrived at Badulla.'
>
> *anthi-aada*   *50%*
> IRR-exist
> 'They might have arrived at Badulla.'
>
> *thama=jo   aada*   *25%*
> NEG.IRR=EMPH exist
> 'They might not have arrived at Badulla yet.'
>
> *thama-aada*   *0%*
> NEG.IRR-exist
> 'They have not arrived at Badulla yet.'

We see that the perfect construction with *aada* conveys certainty, while its negation *thama-aada* conveys certainty of the contrary. The irrealis marker *anthi-* conveys the most uncertain estimation (50%). By attaching the emphatic clitic *=jo* ⊡→6.4.7.1,

p. 378 to either *anthi-* or *thama-,* a somewhat greater likelihood of the event is expressed.

Other means to express epistemic modality, like adverbs or particles do not seem to exist. (433) is the only occurence of the construction given above in the corpus.

(433) *Bìssar aanak asà-dhaathang **anthi-aada** ruuma=nang.*
      big   child CP-come   IRR=exist house=DAT
      'My big child will have come home.' (B060115cvs08)

The following sentence is from an unrelated elicitation session, but illustrates the epistemic use of *thama-aada*.

(434) *Incayang hatthu thookal, itthu=le spaaru vakthu=nang, incayang*
      3s.POLITE INDEF fool,   but   some time=DAT   3s.POLITE
      *Maldives=nang **asà-birnang thama-aada**.*
      Maldives=DAT CP-swim   NEG.IRR-exist
      'He may be a fool, but never would he have tried to swim to the Maldives.'
      (K081114eli01)

15.7.2.5 Evidential modality

Speakers can indicate the source of information: is it first-hand knowledge or hearsay? In SLM, first-hand knowledge does not receive special marking, while non-first-hand information can be indicated by the evidential marker *kiyang* ⊡→6.4.8.1, p. 386 for hearsay. The use of *kiyang* is optional, though, and not very frequent.

(435) *Seelong* Airport=*yang duva-pulu-ùmpath vakthu=le asà-bukka*
      Ceylon Airport=ACC two-ty-four   hour=ADDIT CP-open
      *arà-simpang **kiyang**.*
      NON.PAST-stay EVID
      'The Ceylon Airport will stay open 24h, it seems.' (Letter 26.06.2007)

In example (435), the speaker cannot vouch for the truth of the information he provides about the opening hours of the Colombo airport. He has only second hand knowledge of this, presumably taken from the media. This lack of first-hand knowledge is indicated by *kiyang*.

A lexical solution to convey evidential modality is to use *biilang* 'say'

(436) *Itthu nya-aada katha=le   **arà-biilang**.*
      DIST PAST-exist QUOT=ADDIT NON.PAST-say
      '[About a talisman] It is also said that there was a talisman.' (K051206nar02)

Both possibilities are exemplified below.

(437)   Maldives=*ka   baaru hatthu* president *aada katha* **arà-biilang**.
        Maldives=LOC new   INDEF   president exist QUOT NON.PAST-say
        'It is said that there is a new president in the Maldives.' (K081114eli01)

(438)   Maldives=*ka   baaru hatthu* president *aada* **kiyang**.
        Maldives=LOC new   INDEF   president exist EVID
        'It is said that there is a new president in the Maldives.' (K081114eli01)

Finally, a third possibility involves the 'undetermined' clitic =*so* and indicates that the speaker is not completely sure about the propositional content he is conveying.

(439)   President=*yang asà/su/anà-buunung*=**so   mana**=**so**.
        president=ACC   CP-/PAST-/PAST-kill=UNDET which=UNDET
        'The president was assassinated or something.' (K081114eli01)

15.8   Conditionals

Condition is expressed by the particle *kalu* ⊡→5.14.2, p. 260 (affirmative) or *kalthra* (negative) ⊡→5.14.3, p. 266 in the conditional clause. The consequence is normally marked with the irrealis marker *anthi-* in the affirmative or *thama-* in the negative.[20]

(440)   a. *See lorang=nang* **thama**=*sakith-kang*.
            1s 2PL=DAT   NEG.IRR=pain-CAUS
            'I will do you no harm'
        b. *Lorang see=yang diinging=dering* **kala**-*aapith.*
            2PL   1s=ACC   cold=ABL   if-look.after
            'if you protect me from the cold.' (K070000wrt04)

---

[20]Talking about conditions, especially counterfactuals, is very rare and the aim of elicitation sessions on counterfactuals was not always clear to the informants.

(441) a. *Derang hathu papaaya=yang asà-poothong=apa,*
         3PL   INDEF papaw=ACC   CP-cut=after
         'When they cut open a papaya.'

     b. *Papaaya=ka suum̌bu hatthu **kal**-thaaro,*
         papaw=LOC wick   INDEF when-put
         'and put a wick in the papaya.'

     c. *Aayer=ka **kal**-thaaro **thama**=myaalak.*
         water=LOC when-put NEG.NONPAST=burn
         'and put it into water, it would not burn.'

     d. *Minnyak   **kal**-thaaro **anthi**-myaalak.*
         coconut.oil when-put IRR-burn
         'When they put coconut oil, it would burn.' (K051220nar01)

If a modal prefix or particle is used, this supersedes the use of *anthi-*. This is the case in the following examples, where *masthi-* and *bole=* take the preverbal position and block *anthi-* from occurring there.

(442) *Hathu oorang **kala**-pasiyeth, hathu* chance ***masà**-kaasi ithu*
     INDEF man   if-suffer   INDEF chance must-give DIST
     oorang=nang=le.
     man=DAT=ADDIT
     'When one man suffers, we must give him a chance.' (K060116nar07)

(443) *Siini=jo   incayang=yang **kala**-baava, **bole**=thaau ambel.*
     here=EMPH 3s.POLITE=ACC if-bring   can=know take
     'If you bring him here, he can come to know.' (K061030mix01)

Conditionals of unavailability can be formed by adding *kalthraa* to the NP expressing the lacking substance (444).

(444) *Lorang=ka duvith (*aada) **kal**-thra, kithang anthi-banthu.*
     2PL=LOC   money exist if-NEG   1PL   IRR-help
     'If you have no money available, we will help you.' (K081103eli04)

In the affirmative, the use of the existential *aada*, which is lacking in (444), is obligatory (445).

(445) *Se=dang saayap **kala-aada**, bole=thìrbang.*
       1s=DAT wing if-exist   can=fly
       'If I had wings I could fly.' (K081114eli01)

The conditional marker attaching to a present tense predicate indicates possibility
(446), while its being used on a perfect tense predication indicates counterfactuality
(447).

(446) *Lorang asà-caape **kala-blaajar**, lorang=nang A/L bole=pass.*
       2PL   CP-tired if-learn   2PL=DAT   A/L can=pass
       'If you study hard, you will be able to pass the A/L exams.' (K081114eli01)

(447) *Sampi! Luu   maalas! A/L thàrà-pass.   Lorang baae=nang*
       cow   2s.FAMILIAR lazy   A/L NEG.PAST-pass 2PL   good=DAT
       *asà-caape **asà**-blaajar **kala-aada**, mà-pass=nang su-aada*
       CP-tired CP-learn   if-exist   INF-pass=DAT PAST-exist
       'Lazybones! If you had studied harder, you would have passed your A/L
       exam.' (K081114eli01)

However, if the context is clear, counterfactuals can also be formed with the
present tense. In (448) it is clear that humans will not grow wings, and that we are
dealing with a counterfactual condition.

(448) *Se=dang saayap kala-aada, bole=thìrbang.*
       1s=DAT wing if-exist   can=fly
       'If I had wings I could fly.' (K081114eli01)

Another possibility to express conditionals, which needs further research, is the
use of the 'undetermined' clitic on a past tense verb in the main clause, as given in
(449).

(449) Wicket=*ka su-kìnna=so,*   out.
       wicket=LOC PAST-trike=UNDET out
       'When the ball strikes the wicket, you are out.' (K081105eli02)

15.9 Gradation

In SLM, gradation is done lexically, via *butthul* 'correct' or *bannyak* 'a lot' (Saldin
2001:67), which are all given with their primary meaning here, although in the context
of secondary modification, they would probably all rather be glossed as 'very'. The

following sentences give examples of these words first in their primary meaning and then used as modifiers of modifiers.

The word *but(t)hul* has 'correct' as its primary meaning, this is shown in (450). The intensive reading is shown in (451). Note that the intensive reading in (451) has no geminated stop, which might be indicative of its being not a lexical, but a functional morpheme.

(450) *Lu=ppe nabi   saapa katha **butthul** balas-an   asà-biilang, ...*
　　　 2s=POSS prophet who   QUOT correct answer-NMLZR CP-say
　　　 'Having given the correct answer as to who your prophet is, ...'

(451) *Dee   **buthul** jahhath.*
　　　 3s.IMPOLITE very   wicked
　　　 'He was very wicked.' (K051205nar02)

The second intensifier is *bannyak* 'much', whose literal meaning is given in (452), while the intensifier meaning is shown in (453).

(452) ***Bannyak** Muslim oorang pada arà-duuduk.*
　　　 many   Muslim man   PL   NON.PAST-exist.ANIM
　　　 'There are many Muslims (in the Middle East).' (K061026prs01)

(453) ***Bannyak** thuuva oorang, nya-blaajar oorang.*
　　　 much   old   man   PAST-learn   man
　　　 'A very old man, an educated man.' (K060116nar07)

For positive predications, *baaye* 'good' can also be used as an intensifier.

(454) ***Baaye** meera caaya kapang-jaadi, thurung-king.*
　　　 good   red   colour when-become descend-CAUS
　　　 'When it turns into a nice red colour, remove it (from the fire).' (K060103rec02)

(455) *Dee   arà-cuuri   **baaye** kaaya oorang pada=dering.*
　　　 3s.IMPOLITE NON.PAST-steal good rich   man   PL=ABL
　　　 'He steals from very rich people.'

As for decreasing intensity, it appears that *sdiikith* or *konnyong* are used, both having 'few' as their literal meaning. Unfortunately, there are no instances of *konnyong* and *sdiikith* used for modification of modification in the corpus, only for modification of an adjectival predicate. These are given here for reference.

(456)   ***Konnyong*** *thàràsìggar go=dang.*
      few   sick   1s=DAT
      'I am a little sick.' (B060115nar04)

(457)   *Se=ppe biini **sdiikith** thuuli.*
      1s=POSS wife few   deaf
      'My wife is a little bit deaf.' (K070000wrt05)

15.10   Comparison

Within the domain of comparison, we can distinguish types with an overt indication of the standard of comparison (equation, similarity, superiority, inferiority, superlative) and types where the standard of comparison is not overtly indicated (elative, abundantive, caritive). The aforementioned types typically refer to properties. An additional type, referring to processes, is correlative comparison of the type *the more you X, the more Y*. These types will now be discussed in turn.

15.10.1   Equation

Equation is formed by adding *keejo* to the second element of the equation. The first element is normally marked by *=le*. Neither of the elements of the equation receives special case marking.

(458)   *Incayang*(=le) see=kee=jo   kaaya.*
      3s.POLITE=ADDIT 1s=SIMIL=EMPH rich
      'He is as wealthy as me.' (K081114eli01)

15.10.2   Similarity

Similarity is expressed by the clitic *=ke* ⊡→6.4.6.3, p. 369 on the item the term is similar to.

(459)   *Se=dang **baapa=ke**   soldier mà-jaadi   suuka.*
      1s=DAT   father=SIMIL soldier INF-become like
      'I want to become a soldier like daddy.' (B060115prs10)

15.10.3 Superiority

Superiority is marked either by *liivath* 'more' or *libbi* 'remain'. The standard is marked with the dative.

(460) *Lorang se=**dang libbi** kaaya.*
       2PL   1s=DAT remain rich
       'You are more wealthy than me.' (K081114eli01)

There appear to be some particular semantics with regard to the use of *libbi* and *liivath*. Depending on the adjective, one or the other might be better (461) (462).

(461) *Se=ppe ruuma loram=pe=nang liivath/*libbi bìssar.*
       1s=POSS house 2PL=POSS=DAT   more/remain big
       'My house is bigger than yours.' (K081106eli01)

(462) *Se=ppe ruuma loram=pe=nang *liivath/libbi kiccil.*
       1s=POSS house 2PL=POSS=DAT   more/remain small
       'My house is smaller than yours.' (K081106eli01)

The above examples seem to point to a distinction between positive qualities, graded by *liivath* and negative qualities graded by *libbi*. However, things are more difficult, as the following examples show, where both *mlaarath* 'difficult' and *gampang* 'easy' are graded by *liivath*. This aspect is in need of further research.

(463) Japanese English=*nang liivath mlaarath.*
       Japanese English=DAT   more difficult
       'Japanese is more difficult than English.' (K081106eli01)

(464) Japanese English=*nang liivath gampang.*
       Japanese English=DAT   more easy
       'Japanese is easier than English.' (K081106eli01)

15.10.4 Inferiority

Inferiority is formed like superiority, but *kuurang* 'few, less' is added after the property word.

(465)   *Lorang se=dang libbi   kaaya **kuurang**.*
        2PL   1s=DAT   remain rich   few
        'You are less wealthy than me.' (K081114eli01)

15.10.5   Superlative

The superlative is expressed by *anà-* ⊡→6.2.1.8, p. 285 (466), the emphatic clitic *=jo*
⊡→6.4.7.1, p. 378 (467), or a combination thereof (468).

(466)   *Bill Gates duniya=ka anà-kaaya oorang.*
        Bill Gates world=LOC SUPERL-rich man
        'Bill Gates is the richest man in the world.' (K081103eli04)

(467)   *Anjing oorang=pe baae=jo   thumman.*
        dog   man=POSS good=EMPH friend
        'The dog is man's best friend.' (K081104eli06)

(468)   *Anà-muuda=jo   anak klaaki.*
        SUPERL-young=EMPH child male
        'The youngest one is a son' (K060108nar02,K081103eli04)

15.10.6   Elative

The elative is expressed by *=jo* ⊡→6.4.7.1, p. 378.

(469)   *Hatthu komplok **bannyak=jo** puuthi caaya.*
        one   bush   very=EMPH white colour
        'One bush was very, very white.' (K070000wrt04)

15.10.7   Abundantive

There is no special form for abundantive. The normal intensifiers *bannyak* or *buthul*
can be used.

(470)   *Incayang buthul gummuk.*
        3s.POLITE correct fat
        'He is too/very fat.' (K081114eli01)

The excess can be additionally marked by *liivath* 'more', but this is optional.

(471) *Itthu* bag=*yang lorang=nang bannyak bìrrath (**liivath**).*
DIST bag=ACC 2PL=DAT   much   heavy more
'That bag is too heavy for you.' (K081114eli01)

15.10.8 Caritive

To indicate that the degree is not sufficient as compared to a non-expressed standard, *thàràsampe* 'does not reach' is used. *Thàràsampe* can be used with nouns or with adjectives (472).

(472) *Se=ppe kaake=nang   sìggar/sìgaran thàràsampe.*
1s=POSS grandfather=DAT healthy/health insufficient
'My grandfather is not well enough.' (K081114eli01)

The purpose for which the degree is insufficient can be indicated by an infinitive clause.

(473) ***Pohong mà-seereth=nang** oorang thàrà-sampe.*
tree   INF-drag=DAT   man   NEG.PAST-reach
'There were not enough men to drag the tree.' (K051205nar05)

For adjectives, the normal negation can be used, complemented by *punnu* 'full' and an intensifier.

(474) *Ini   makanan **bannyak/giithu punnu** pùddas **thraa**.*
PROX food   much/like.that full   spicy NEG
'This food is not spicy enough.' (K081114eli01)

15.10.9 Correlative

Correlative comparison is formed by a reduplicated infinitive.

(475) *Caabe **mà-thaaro mà-thaaro**, pùddas liivath.*
chili INF-put   INF-put   spicy more
'The more chili you put, the spicier (the food becomes).' (K081114eli01)

(476)   *Haari **mà-pii mà-pii**, laam̃bath liivath.*
        day   INF-go INF-go delay   more
        'Days pass by and the delay is caused/The more days pass, the more delay is
        caused.' (K081114eli01)

15.11   Possession

Within the realm of possession, we can distinguish three different constellations (assertions in boldface): either the possessee is asserted (*I have **a car***), or the possessor is asserted (*the car is **mine***), or they are both part of the presupposition (***My car broke down***). Furthermore, permanent and temporary possession can be distinguished, which is relevant for SLM, as are differences in animacy, while differences in alienability are not relevant.

There is no verb meaning 'to have', in line with general South Asian (Masica 1976:166) and Austronesian typology (Himmelmann 2005:139). There is a verb *puunya* 'possess', but this is hardly ever used.[21]

15.11.1   Assertion of the possessee

The assertion of the possessee's being in a possessive relationship with the possessor is done with a combination of the postpositions =*nang* ⊡→6.4.4.2, p. 334, =*ka* ⊡→6.4.4.3, p. 340 or =*samma* ⊡→6.4.4.9, p. 352 with the existentials *aada* ⊡→5.1.2.1, p. 165 or *duuduk* ⊡→5.1.2.2, p. 168. =*nang* is used for permanent possession, while =*ka* is used for temporary possession. *Duuduk* can only be used for animate possessees (which are almost always kin), while *aada* can be used for any possessee. =*samma* is only used for temporary possession of animates. The following sections show various combinations of these parameters.

	animate	inanimate	abstract
permanent	=nang+duuduk (477) (478)	=nang+aada	=nang+aada (485)
	=nang+aada (479)	(481)(482)	
temporary	=samma+duuduk (480)	=ka+aada	
		(483)(484)	

15.11.2   Permanent possession of animates

Possessed animates like kin are always construed with the dative marker =*nang* on the possessor. The existential can be either *duuduk* or *aada*.

---

[21]The possessive marker =*pe* has grammaticalized from *pu(u)nya* and is frequent.

(477) *Se=dang hathu maven arà-**duuduk***.
　　　1s=DAT INDEF son NON.PAST-exist.ANIM
　　　'I have a son.' (B060115prs05)

(478) *Se=dang duuva pompang aade=le* hathu klaaki
　　　1s=DAT two female younger.sibling=ADDIT one male
　　　aade=le *anà-**duuduk***.
　　　younger.sibling=ADDIT PAST-exist.ANIM
　　　'I had two younger sisters and a younger brother.'

(479) ***Se=dang** liima anak klaaki pada **aada***.
　　　1s=DAT five child male PL exist
　　　'I have five sons.' (K060108nar02)

15.11.3 Temporary possession of animates

The temporary possession of animates uses *duuduk* as expected, but instead of the locative =*ka*, the comitative =*samma* is used.

(480) *Maaling=samma moonyeth hatthu arà-**duuduk***.
　　　thief=COMIT monkey INDEF NON.PAST-exist.ANIM
　　　'The thief has a monkey (with him).' (K090327eml01)

15.11.4 Permanent possession of inanimates

Inanimates can never be construed with *duuduk*. *Aada* always has to be used.

(481) *Mr. Yusuf, karang, incayang=**nang** [ini* private bank=*ka aada duvith*
　　　Mr. Yusuf now 3s.POLITE=DAT PROX private bank=LOC exist money
　　　pada] ***aada***
　　　PL exist
　　　'Mr. Yusuf owned the money which was deposited in this private bank.'
　　　(K060116nar09)

(482) *Se=dang hatthu ruuma aada*.
　　　1s=DAT INDEF house exist
　　　'I have a house.' (K081114eli01)

15.11.5 Temporary possession of inanimates

Temporary possession of inanimates is construed with the locative =*ka* and the inanimate existential *aada*.

(483) *Incayang=**ka** ... bìssar beecek caaya hathu bag su-**aada**.*
       3s.POLITE=LOC ... big   mud   colour INDEF bag PAST-exist
       'He had a big brown bag with him.' (K070000wrt04)

(484) *See=**ka**=jo   bannyak ini   panthong ∅.*
       1s=LOC=EMPH much   PROX song
       'I possess a lot of these songs [on sheets].' (K060116nar04)

15.11.6 Possession of abstract concepts

Abstract concepts can never be construed with =*ka* or *duuduk*. =*nang* and *aada* have to be used.

(485) *Se=**dang** bannyak creeveth pada su-**aada**.*
       1s=DAT lot   trouble PL   PAST-exist
       'I had a lot of trouble.' (K051213nar01)

15.11.7 Assertion of the possessor

If the possessee has been established before, but not the possessor, an equational construction ⊡→10.4.2, p. 474 is used (cf. Hengeveld 1992:104f). This construction assigns the possessum to the class of items possessed by the possessor. This class is indicated by marking the possessor with the possessive postposition =*pe*.

(486) *Se=ppe saayang jiiva biilang [**se=ppe**] katha.*
       1s=POSS love   life say   1s=POSS QUOT
       'Love of my life, say that you are mine.' (K061123sng03)

(487) *Itthu muusing bannyak teacher pada [**Jaapna=pe**].*
       DIST time   many   teacher PL   Jaffna=POSS
       'Back then, many teachers were from Jaffna.' (K051213nar03)

15.11.8 Presupposition of the possessor and the possessee

If the possessive relation is presupposed, the possessive postposition =*pe* ⊡→6.4.4.5, p. 345 is used (488). If the possessor is expressed by a monosyllabic pronoun, =*ppe* is used (489). Other pronouns take the normal form (490).

(488)   *Kithang=pe **baapa=pe*** *naama Mahamud.*
       1PL=POSS   father=POSS name Mahamud
       'Our father's name is Mahamud.' (B060115nar03)

(489)   *Lu=**ppe*** *muuluth=ka=le paasir, se=**ppe** muuluth=ka=le*
       2S.FAMILIAR=POSS mouth=LOC=ADDIT sand   1S=POSS mouth=LOC=ADDIT
       paasir.
       sand
       'There is sand in your mouth and there is sand in my mouth.' (K070000wrt02)

(490)   ***Kitham=pe*** *ruuma dikkath=ka.*
       1PL=POSS   house vicinity=LOC
       'Close to our house.' (K051220nar01)

15.12 Negation

Negation gives negative truth value to a proposition. In SLM negation patterns vary with predication type, clause type, tense, and information structure. Every different type of predicate (verbal, existential, modal, nominal, adjectival, circumstantial) has a corresponding negation pattern. The verbal predicate has an additional negation type used in subordinate clauses. In the domain of tense, we can distinguish past, perfect, non-past and future as relevant tense domains. As for information structure, predicate negation is different from constituent negation. Table 15.8 gives an overview of the different negation patterns. Please see section 10 and the relevant subsections for more discussion and examples.

Investigating Table 15.8, we observe a number of generalizations. For instance, negation of circumstantial predication and negation of constituents are both done by *bukang*. This could point to circumstantial predicates not being predicates in their own rights, but rather constituents with a 'real' predicate which is not overtly realized.

Another item which occurs frequently is *thraa*, which is used for perfect tense negation of verbs, negation of locational predicates and negation of adjectives. The first two of this can be explained by the presence of *aada* in the affirmative counterpart. Since *thraa* is the negative form of *aada*, the occurrence of *thraa* in perfect

---

[22]This suppletive negation of the existential is an exception to the generalization formulated by Himmelmann (2005:138), that the existential is negated by the common verbal negator.

	past	perfect	present	future
predicate negation				
fin. verbal clause	thàrà-V	V thraa		thama-V
inf. verbal clause	jamà			
existential	thraa			
nominal	bukang			thama-jaadi/bukang
adjectival1	ADJ thraa			thama-ADJ
adjectival2	thàrà-ADJ			thama-ADJ
circumstantial	bukang			
locational	thraa[22]			
animate locational	thàràduuduk	duuduk thraa		thama-duuduk
(ka)maau(van)	thàrkamauvan/thussa			
boole	thàrboole			
constituent neg.	bukang			

Table 15.8: Negation patterns for various predicate types and tenses.

and locational predicates where *aada* is used in the affirmative is not surprising. The negation of adjectives by *thraa* cannot be explained in this manner. Note that speakers differ as to the negation patterns of adjectives. Depending on speaker and individual lexeme, the pattern with *thraa* or the pattern with *thàrà-* can be found.

When *thàrà-* is used to negate adjectives, it can be used in both past and present tenses, while in its other use for negating verbs, it necessarily has past tense reference.

For negation of non-verbal predications referring to the future, some special periphrases exist. All these involve the non-past negative marker *thama-* and an extra verb, like *jaadi* 'become'.

Conjunctive participle clauses and purposive clauses are marked by *jamà-* when negated. Other subordinate clauses like relative clauses or argument clauses take the same negation as main clauses.

Indicating that none of the possible referents would yield a positive truth value (*no student came*) and indicating that for all referents, the truth value is negative (*The students did not come*) is done in a very similar fashion in SLM. The only difference between the two is that in the former case the indefinite article *atthu* is used before the noun and a suitable coordinating clitic like *=le*, *=ke* or *=pon* is used after the NP.

(491) *Kithang=pe **hatthu** oorang=**le** minister jaadi   **thraa**.*
        1PL=POSS   INDEF   man=ADDIT minister become NEG
        'Not one of our has (ever) become a minister.' (N061031nar01)

(492) *Derang=nang Kluum̆bu=pe   samma* association=*le*   support*;*
       3PL=DAT   Colombo=POSS all   association=ADDIT support
       *Kluum̆bu=pe   **hatthu*** association=*le*   *kithang=nang **thàrà**-*support.
       Colombo=POSS INDEF   association=ADDIT 1PL=DAT   NEG.PAST-support
       'All Colombo associations supported them, not one association from Colombo
       supported us' (K060116nar06)

(493) *Kithang **hatthu**=oorang=**pon thàrà**-iingath.*
       1PL   INDEF=man=any   NEG.PAST-think
       'We cannot think of any person.' (B060115nar02)

   Rarely, the indefinite article and the clitic are both found after then noun.

(494) *See pukaran=**hatthu=pon thama**=gijja,   ruuma=ka arà-duuduk.*
       1s   work=INDEF=any   NEG.NONPAST-make house=LOC NON.PAST-stay
       'I don't do any work, I stay at home.' (B060115prs03)

   The clitic attaches after the postposition, as can be seen from the following two
examples, where the postposition =*nang* intervenes between the noun *oorang* 'man'
and the clitic =*le*.

(495) *Incayang=pe   muusing=ka kithang=pe* Malays *pada **atthu***
       3s.POLITE=POSS time=LOC   1PL=POSS   Malays PL   INDEF
       ***oorang=nang=le** [*parliament=*nang mà-dhaathang=nang thumpath*]
       man=DAT=ADDIT parliament=DAT   INF-come=DAT   place
       ***thàrà**-daapath.*
       NEG.PAST-get
       'During his time, no man of our Malays got a place to go to parliament (i.e. a
       seat).' (N061031nar01)

(496) *Incayang=le   kithang=pe mlaayu pada **hatthu** oorang**=nang=le***
       3s.POLITE=ADDIT 1PL=POSS   Malay PL   INDEF   man=DAT=ADDIT
       *thumpath **thàrà**-kaasi.*
       place   NEG.PAST-give
       'He did not not give any position to any of our Malays either.' (N061031nar01)

   It is also possible to use the combination of *atthu* and a clitic without a noun. In
this case, the clitic =*ke* was also found in the corpus, next to =*le* and =*pon*.

(497)   *Laayeng **hatthu=∅=le** thraa.*
        different INDEF=ADDIT   NEG
        'There is nothing else.' (B060115nar04)

(498)   Bus=*ka **hatthu=∅=pon** mà-kirja thàràboole.*
        bus=LOC INDEF=NEG   INF-make cannot
        'You can't do anything on the bus.' (K061125nar01)

(499)   *Snow-white=nang=le   Rose-red=nang=le   ini   **hatthu=∅=ke***
        Snow.white=DAT=ADDIT Rose.Red=DAT=ADDIT PROX INDEF=SIMIL
        ***thàrà**-mirthi.*
        NEG.PAST-understand
        'Snow White and Rose Red did not understand a thing.' (K070000wrt04)

15.13 Kin

The family relations of people show differences in their encoding between the languages of the world. The SLM kinship system distinguishes between males and females, generations, and relative age. Siblings and cousins are not distinguished, but there are many different types of uncles and aunts, depending on whether they are male or female, maternal or paternal, elder or younger. Parents' elder siblings and their spouses are only distinguished for sex. Parents' younger siblings are also distinguished for sex of the parent, and spouses are different from consanguineous relatives. Grandparents and grandchildren are only distinguished by sex, while greatgrandparents and greatgrandchildren are not distinguished by sex either, but the word for greatgrandparents includes the morpheme for the grandparents.

The terms *maven* 'son, nephew, younger man' and *mavol* 'daughter, niece' can also be heard, but it was not clear to which individuals they could refer. When addressing older members of the community, younger speakers use *uncle, auntie* (English) or *kaake, neene* 'grandfather, grandmother', depending on the age. When elder speakers address younger ones, they can use *maven* or *mavol*. Among peers, the sibling terms *kaaka* 'elder male', *dhaatha* 'elder female' or *aade* 'younger sibling' are used. These follow the proper name, so *Imi kaaka* 'Elder brother Imi'.

Kekulawala (1982) first studied the kinship system of SLM (cited in Bichsel-Stettler (1989)) and analyzed it as lineo-bifurcate collateral type.
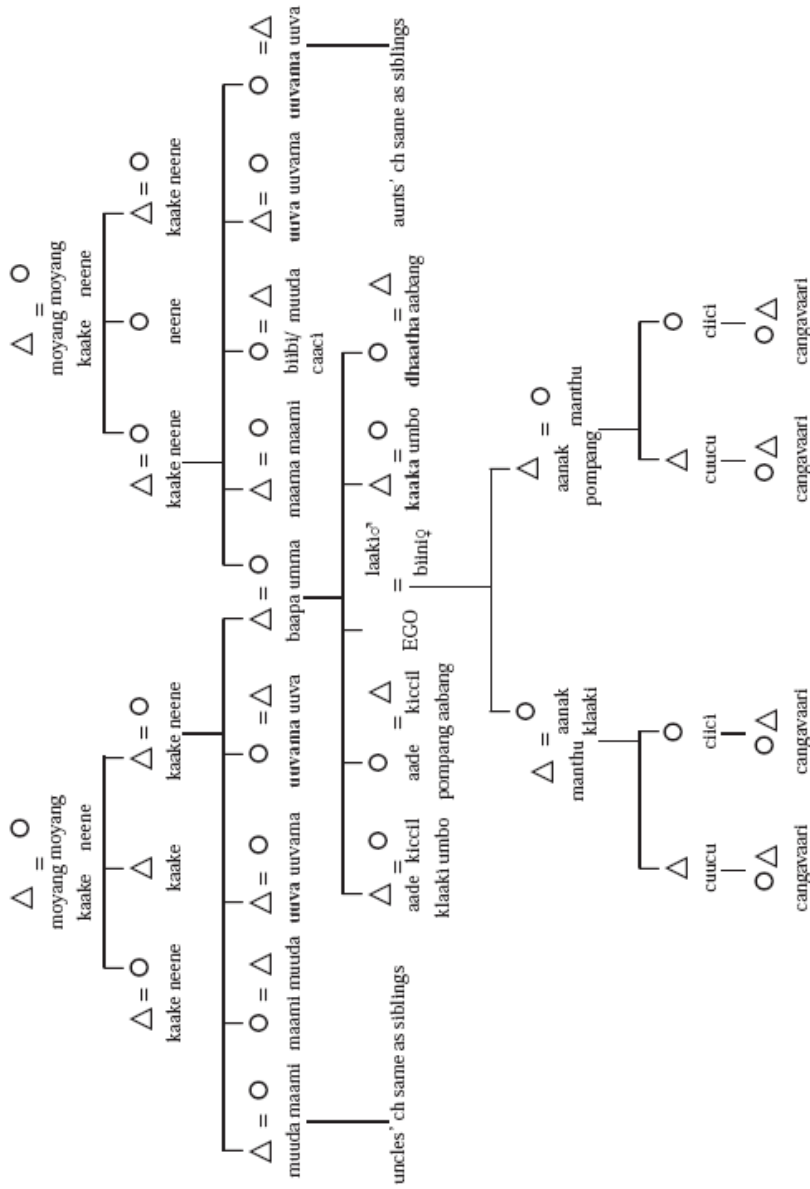
Figure 15.1: Kinship relations in SLM. Bold face denotes elder siblings. Normal font denotes younger siblings, if there is a distinction with bold face in the same generation. Relative age of spouses of parents' siblings is not relevant.

Chapter 16

Information flow

Having discussed how to encode propositional content, we now turn to information flow. Information flow deals with the way in which the speaker structures the information she presents so that the hearer's knowledge about the world is modified in the way the speaker intended. For this end, it is important to track what the speaker is talking about: the referents. This is discussed in Section ⊙→16.1, p. 677. Furthermore, it is important to now which one of the referents is the topic of the message, and what portion is the comment on the topic. This is discussed in Section ⊙→16.2, p. 685. Next to identifying the aboutness of the message, it is also important to tell the presupposition from the assertion. Strategies for this are discussed in ⊙→16.3, p. 689. Section 16.4 on page 692 discusses how to cancel implicatures. While these sections deal with how the speaker packages the information, the last section ⊙→16.5, p. 694 is on parsing and deals with how the hearer can decode the message packaged by the speaker.

16.1   Referents and reference

When several referents are established in discourse, it is important to make clear what is said about which referent. Four types of referents will be distinguished here.[1]

- *Unidentifiable* referents are those referents which have not been talked about before ⊙→16.1.1, p. 678.

- *Active* referents are those referents which are currently topical ⊙→16.1.2, p. 679.

- *Accessible* referents are those referents which are not topical right now, but are still retrievable ⊙→16.1.3, p. 684.

---

[1] This taxonomy is based on Prince (1981), Chafe (1987), Lambrecht (1994), Van Valin & Foley (1997), but slightly simplified.

- *Inactive* referents are those referents which have already been introduced in discourse, but have been replaced by other referents in the focus of consciousness (16.1.4).

16.1.1   Unidentifiable referents

In order to comment on referents, one has to introduce them first. New referents are referents which cannot be assumed to be identifiable by the addressee. In SLM, new referents are normally introduced with existential ⊡→10.2, p. 468 or locative predicates ⊡→10.5.2, p. 476.

Example (1) introduces a new referent *hathu kiccil ruuma* 'a small house', on which the following sentences in the narrative will provide more information. The introduction of that house is done with an existential predicate with *aada.*

(1)   *Hathu muusing=ka ... **hathu** kiccil ruuma su-**aada***
      INDEF time=LOC   ... INDEF small house PAST-exist
      'Once upon a time, there was a small house.' (K07000wrt04)

Very often, the indefinite article is used in presentationals. In following occurrences, the indefinite article is not used anymore. This can be seen in (2), where the prince is introduced with the indefinite article in the first line, but the second occurrence of *aanak raaja* 'prince' does not carry the indefiniteness marker *hatthu.*

(2)   a.   *Derang anà-baalek sajja=jo,   sithu=ka panthas **hathu Aanak raaja***
           3PL   PAST-turn   only=EMPH   there=LOC beautiful INDEF king   child
           su-aada!
           PAST-exist
           'Just when they turned around, there was a beautiful prince.'

      b.   *Buruan=pe kuulith incayang=pe   kaaki baava=ka   jaatho su-aada!*
           bear=POSS skin   3s.POLITE=POSS leg   bottom=LOC fall   PAST-exist
           'The bear's hide had fallen to his feet.'

      a.   ***Aanak raaja** su-biilang: ...*
           child king PAST-say
           'The prince said: ...' (K070000wrt04)

Another method is to introduce new referents as comment on an already existing topic. In this case, the indefinite article is also used. In example (3), a new referent *hatthu pohong* 'a tree' is introduced, which will have some consequences for the further development of the story. This referent is introduced as part of the two-place predicate *poothong* 'cut', and not as an existential as in (1).

(3) *Itthu=kapang baapa derang=pe kubbong=ka **hatthu** pohong nya-poothong.*
     DIST=when   father 3PL=POSS estate=LOC INDEF tree   PAST-cut
     'The father then felled a tree in their estate.' (K051205nar05)

Next to arguments, new referents can also be introduced in adjuncts, as in (4), where a tree is introduced as a locative adjunct (supposing that a difference between arguments and adjuncts can indeed be established, ⊡→11, p. 481).

(4) *Andare **hathu pohong**=pe baava=ka   kapang-duuduk ...*
     Andare INDEF tree=POSS   bottom=LOC when-sit
     'When Andare sat down under a tree ...' ((K070000wrt03))

16.1.2 Active referents

Active referents are those which are in the current focus of consciousness. They are normally not overtly coded in SLM as they are inferable from context (Smith et al. 2004).[2] Active referents may also be realized as pronouns for emphasis. Speaker and hearer are always active, which is why they are very often not overtly expressed.

It is possible to drop the agent of the sentence, as in (5), but arguments with other roles can also be dropped, as in (6)(7). Dropping several roles is also possible (8)(9).[3]

(5) *Hindu arà-maakang kambing, ∅ samping thuma-maakang.*
     Hindu NON.PAST-eat goat   beef   NEG.NONPAST-eat
     'Hindus eat goat, they do not eat beef.' (K060112nar01)

In example (5), the first clause contains an overt reference to *Hindu(s)*. This makes this referent active in discourse. The assertion *aràmaakang kambing* 'eat goat' is then applied to that referent. In the second clause, the assertion *samping thumamaakang* 'do not eat beef' is to be applied to *Hindu*. Since *Hindu* is the last active referent, overt expression/repetition is not required.

Examples (6) and (7) show that roles other than the agent/'subject' can be dropped.

(6) *Dee ∅_{theme} athi-kaasi miskin oorang pada=nang.*
     3   IRR-give poor   man   PL=DAT
     'He would give (it=the money) to the poor.' (K051206nar02)

---

[2]SLM thus obeys to the 'Principle of Absent Information' as formulated by Minegishi (2004:305):'Avoid repetition, overt or psychological, so as not to add unnecessary connotation or contrast.'

[3]This liberal dropping of constituents is called "Super pro drop" in Gair et al. (1989) for Sinhala.

(7) *Hathu thaaun se=ppe umma-baapa Ø_{theme} se=ppe maama=pe*
     INDEF year   1s=POSS mother-father   1s=POSS uncle=POSS
     ruuma=nang su-kiiring.
     house=DAT   PAST-send
     'One year my parents sent (me) to my uncle's house.' (K051213nar02)

The following example shows that more than one referent can be dropped in a
clause, in this case both agent and patient.

(8) *Ø_{pat} mà-mathi-king=nang, Ø_{ag} Ø_{pat} siithu=jo   anà-baapi.*
       INF-dead-CAUS=DAT   there=EMPH PAST-bring
     'It was there that (they) brought (him) to make (him) dead.' (K051206nar02)

The following stretch of discourse shows how overt marking and zero marking
alternate, and how two arguments are dropped in the last clause.

(9) a. *Incayang=jo_i   asà-pii,*
         3s.polite=EMPH CP-go
         'He goes there and'

     b. *Ø_i paasar=dang asà-pii,*
           shop=DAT   CP-go
         'goes to the shop and'

     c. *itthu samma_j asà-baa,*
         DIST all   CP-bring
         'brings everything and'

     d. *Ø_i Ø_j arà-juuval.*
               NON.PAST-sell
         'sells (it).' (B060115cvs07a)

The fourth and final clause of (9) consists of the predicate *arà-juuval*, which takes
two arguments, one seller and one produce. The seller *incayang* 'he' is introduced in
the first clause, and the produce *itthu samma* 'all that' (referring to vegetables men-
tioned earlier) is introduced in the third clause. Both seller and produce are in the
focus of consciousness and can thus be dropped in the last sentence. The referents
need not be overtly realized because there is no risk of confusing the participant roles
of *juuval* 'sell'. Selling requires a vendor and a good. Knowledge of the world tells
us that vendors are normally people, and goods can be vegetables. We also know that
vendors are normally not vegetables, and the entities sold are nowadays rarely people.
This means that the two introduced referents do not compete for the semantic roles
that *juuval* 'sell' licenses. As a consequence, there is no risk of confusion, and neither
referent is overtly expressed/repeated. The role of the undergoer is instantiated with

the active referent *itthu samma* 'all that' for semantic reasons. The predicate requires an additional role, for which no referent is provided in the clause. The hearer thus has to select one of the referents in the focus of consciousness and check which one of them could possibly instantiate the role of actor. Luckily, a suitable accessible referent is available, *incayang* 'he'.

The amount of inference required by the hearer can be quite high. Associating semantic roles and referents is very hard in the following example and requires a lot of work by the hearer.

(10) a. *[Derang pada]ᵢ panggel=nang blaakang, ∅ᵢ [thàrà-dhaathang oorang*
     3PL   PL   call=DAT   after   NEG.PAST-come   man

     *pada]ⱼ=nang nya-*force*-kang kiyang.*
     PL=DAT   PAST-force-CAUS EVID

     'After they had called (them), (they) apparently forced the people who had not come (to join),'

   b. *blaakang ∅ⱼ thàrà-pii   kiyang.*
      after   NEG.PAST-go EVID

      'but still (they) did not go it seems' (K051206nar07)

The context is that a ship came from Malaysia, and the sailors wanted to convince the Sri Lankan Malays to get on that ship. They called them and tried to force them, but the Sri Lankans would not follow. Two referents are retrievable in the first part, the foreigners, referred to by *derang pada* 'they', who are agent of the predicates *panggel* 'call' and *force-kang* 'force', and the Sri Lankan Malays, who are goal and patient of those two predicates, respectively. The final clause in this example contains a negated predicate, *thàràpii* 'did not go', but the argument is dropped. The verb *pii* 'go' requires only one argument, an agent, and it is tempting to associate it with the foreigners, since they have been agents in the two preceding clauses, and seem to be more topical as well. Furthermore, in case of 'subject' playing a role in SLM grammar, we would definitely expect that it is the foreigners who are associated with the role of agent that the verb *pii* 'go' assigns. Our knowledge of the world, however, makes clear that it must be the locals who did not go despite being forced. So the role of agent in the final clause is assigned not to the referent having been agent in the preceding clauses, but to the undergoer referent, which furthermore was less topical. There are no morphosyntactic clues for the resolution of the zero in the final clause. We have to rely on our knowledge of the world, which tells us that after an act of forcing, the persons most likely to take part in an act of leaving are the forced ones, not the forcers. We are thus not interested in whether the foreigners left; what is interesting is whether the locals left. Only after all of these inferences can the reference of the participant of the verbal predicate *thàràpii* 'did not go' be established.

While dropping of arguments is common, realization of pronouns is not ruled out. Active referents may be coded by pronouns or by ∅ as the following sequence

shows, where new referents are introduced by nouns. The next reference to this freshly
introduced referent is normally by the pronoun *incayang*, while references thereafter
are commonly zero.

(11) a. Malay *thiiga aanak pada arà-duuduk.*
          Malay three child PL   NON.PAST-stay
          'There are three Malay children [in this house].'

      b. ***Ka-thaama aanak***$_{new}$ *dhaathampa klaaki.*
          ORD-early child   COPULA   boy
          'The first one is a boy.'

      c. ***Incayang***$_{active}$ *skaarang=nang* Colombo University=*ka arà-blaajar.*
          3S.POLITE   now=DAT   Colombo University=LOC NON.PAST-study
          'He studies at Colombo University.'

      d. ∅$_{active}$ *blaajar=apa*
                study=after
          'Having finished his studies,'

      e. *mareng   dovulu **incayang***$_{active}$=*nang* HSBC=*ka   hatthu pukuran*
          yesterday earlier 3s.polite=DAT   HSBC=LOC INDEF job
          se-daapath.
          PAST-get
          'he got a job a the HSBC bank the day before yesterday.'

      f. ***Ka-duuva* Daughter**$_{new}$ *Swarnamali=ka blaajar=apa*
          ORD-two daughter   Swarnamali=LOC study=after
          'The second child, a daughter, studied at Swarnamali School'

      g. ***incayang***$_{active}$ French lesson *kijja.*
          3S.POLITE   French lesson make
          'and did French.'

      h. ∅$_{active}$ *Karang* Dialog GSM=*ka   junior executive hatthu.*
                now   Dialog GSM=LOC junior executive indef
          'She is now junior executive at Dialog GSM [phone company].'

      i. ***Kanabisan aanak***$_{new}$, *incayang=nang Swarnamali=ka=jo   blaajar.*
          last   child   3s.polite=DAT Swarnamali=LOC=EMPH learn
          'The last child studies at Swarnamali school.'

      j. ***Incayang***$_{active}$ 2007=*ka   A-level exam kijja.*
          3S.POLITE   2007=LOC A-level exam make
          'She will pass the A-level exams in 2007.' (G051222nar01)

In some languages, coordinated sentences require that elements dropped in later
clauses have the same role/grammatical relation as in their last overt realization. An
example is *Mary came and hit John*, which can only mean that Mary hit. If we want

to say that Mary was hit, we have to specify the argument, as in *Mary came and John hit **her***. In SLM, this is not the case. Semantic roles of elements can be adjusted to the new predicate as necessary. An example of this is (12), where the person leaving work and the person getting pension are coreferential and both realized by zero, but in the first clause, the person has the role of agent, and in the second, the role of recipient.

(12) $\emptyset_{ag}$ *Pukurjan=yang su-luppas-kang, karang* $\emptyset_{rec}$ *pension arà-daapath.*
　　　　work=ACC　　PAST-leave-CAUS　now　　　pension NON.PAST-get
　　　'He quit work and now gets his pension.' (K051222nar06)

Another example is (13): a respectable Malay is first given the position of High Commissioner before he is sent to Pakistan. The first clause refers to this person by *incayang* 'he', which is zero-marked but has the semantic role of recipient. In the second clause, this referent is not overtly realized, but is the theme of the action of sending, which is not the same as recipient.

(13) a. *Itthu=nang blaakang incayang=∅ [Sri Lanka=pe　high commissioner*
　　　　　DIST=DAT　after　　3s.POLITE　Sri　Lanka=POSS high commissioner
　　　　　position=yang kaasi]=apa.
　　　　　position=ACC　give=after
　　　　　'After that, they gave him the position of Sri Lankan High Commissioner and'

　　　b. *∅ ∅ Pakistan=nang anà-kiiring.*
　　　　　Pakistan=DAT　PAST-send
　　　　　'they sent him to Pakistan.' (N061031nar01)

Similar things can be said about (14), where the sister is agent in the subordinate clause, where the argument would normally be zero-marked, but possessor in the main clause, which would normally be marked by the dative.

(14) a. *[Hathyang dhaatha]_{ag} asà-kaaving aada.*
　　　　　other　　　elder.sister CP-marry　exist
　　　　　'My other elder sister married.'

　　　b. $\emptyset_{poss}$ *thiiga aanak aada.*
　　　　　three　child　exist
　　　　　'She has three children.' (K061019prs01)

An example with accusative marking is (15), where the speaker is the accusative-marked patient of birth, and the zero-marked theme of being black.

(15) *See=yang_{patient} kapang-braanak, 0_{theme} bannyak iitham.*
      1s=ACC   when-bear   much   black
      'When I was born, I was very dark.' (K08114eli01)

An example, with the postposition *=apa* instead of the conjunctive participle *asà-*, is (16).[4] The first two clauses treat *oorang pada* 'people', while the last sentence, the main clause, does not have an argument coreferential with the aforementioned *oorang* 'people'.

(16) *oorang pada thiikam=apa, oorang pada=nang theeṁbak=apa, se=dang*
      man   PL   stab=after   man   PL=DAT   shoot=after   1s=DAT
      bannyak creeveth pada su-aada.
      much   trouble   PL   PAST-exist
      'People were stabbed, people were shot, I had a lot of problems.'
      (K051213nar01)

16.1.3   Accessible referents

Accessible referents are referents which are "textually, situationally or inferentially available by means of [their] existence in the physical or linguistic context or [their] relation to something in the physical or linguistic context but [are] not yet the current focus of consciousness" (Van Valin & Foley 1997:200).

An example of an accessible referent is (17), where the son had not been talked about before. Still, he is "inferentially available" by being a referent naturally connected to a middle-aged woman. Accessible referents can then be coded as definite NPs in SLM, and can be topical, as seen in (17).

(17) a. *Se=dang karang ruuma=nang masà-pii.*
          1s=DAT now   house=DAT   must-go
          'I have to go home now.'

      b. **Bìssar aanak** *asà-dhaathang anthi-aada*
          big   child   CP-come   IRR-exist
          'My big son will have come home.' (B060115cvs08)

16.1.4   Inactive referents

Inactive referents are those that have been talked about earlier but which have been superseded by other referents. They are in the hearer's long-term memory, but not

---

[4]*=apa* and *asà-* are in an unclear relationship, see ☞→6.2.1.4, p. 276 and ☞→6.4.4.10, p. 353.

in short-term memory. When new reference to these referents is made, the indefinite article *atthu* is not used (unlike for unidentifiable referents), but a noun is required to reestablish reference; a pronoun or zero would not be enough. In the following story, two girls find a dwarf and help him. The first reference to the dwarf is made when he is introduced as the argument of *kuthumung* 'see', where this introduction is supported by the indefinite article *hathu= ... =hatthu*, present twice in the NP. Later, they meet the dwarf again, but reference to him has become inactive in the meantime, so he has to be reintroduced with a full NP. This time, this happens without the indefinite article. The absence of the indefinite article implies definiteness or retrievability in this case (there is no definite article in SLM).

(18) a. *Itthu haari=ka=jo   aanak pompang duuva=nang [. . .] **hathu** kiccil*
          DIST day=LOC=EMPH child girl   two=DAT   INDEF small
          *jillek **aajuth hatthu**=yang su-kuthumung.*
          ugly dwarf INDEF=ACC   PAST-see
          'On that very day, the two girls saw a small ugly dwarf.'

      b. *Derang incayangyang su-salba-king.*
          3PL   3s.polite=ACC PAST-escape-CAUS
          'They saved him.'

      c. [. . .]

      d. *Hathyang aari=ka Snow-white=le   Rose-red=le   [. . .] berry*
          other   day=LOC Snow-white=ADDIT Rose-red=ADDIT   berry
          kapang-picca-kang, bìssar hathu buurung derang=pe atthas=dering
          when-broken-CAUS big   INDEF bird   3PL=POSS top=ABL
          su-thìrbang.
          PAST-fly
          'On another day, when Snow-white and Rose-red were picking berries, they saw a big bird flying over them.'

      e. [. . .]

      f. *Derang su-kuthumung [ithu buurung=pe kuuku=ka **aajuth** asà-sìrrath*
          3PL   PAST-see   DIST bird=POSS   claw=LOC dwarf CP-stuck
          kìnna arà-duuduk].
          strike simult-exist.ANIM
          'They saw the dwarf being held in the claws of that bird.' (K070000wrt04)

16.2 Topic

Topic refers to the element about which the sentence says something (Lambrecht 1994:118). We can distinguish ongoing topics, which are treated above under *active referents* ⊙→16.1.2, p. 679 and new topics. Furthermore, we can distinguish con-

trastive topics and non-contrastive topics.

16.2.1   New topic

It is rare for discourse to start off with a new topic. It is more common to introduce a referent in one of the manner discussed in ⊙→16.1.1, p. 678. The referent so introduced can then be used as a topic in following sentences. The following stretch of discourse provides an example for this. The new information *Irish nuns* goes to the right in the first sentence. In the second sentence, it is old information, and is found in topical position at the beginning.

(19)   a. *Itthu muusing=ka kithang=nang anà-aajar [**Irish nuns pada**]*$_{newinf}$.
        DIST time=LOC   1PL=DAT   PAST-teach Irish   nuns   PL
        'In former times the ones who taught us something were Irish nuns.'

   b. *[**Irish nuns**]*$_{oldinf}$*, derang=pe* English *baaye, derang baaye=nang*
      Irish   nuns   3PL=POSS   English good 3PL   good=DAT
      anà-aajar.
      PAST-teach
      'Irish nuns, their English is good, they taught well.' (K051222nar06)

In the example above, the nuns were first introduced, and became the active referent. As such, they could be used in topic position. If the topic of the sentence is not active in discourse, *kalu* 'as for' (also the conditional marker) can be used to signal this. In example (20), the new topic of the stay in the town of Galle is marked with *kalu*.

(20) ***Galle**=ka **kalu**, se=ppe* cousin brothers *pada=samma see anà-jaalang*
     Galle=LOC if   1s=POSS cousin brothers PL=COMIT   see PAST-walk
     skuul=nang.
     school=DAT
     'As for when in Galle, I used to walked to school with my cousin brothers.'
     (K051201nar02)

New topics are commonly introduced by a linking adverbial ⊡→14.1, p. 541, indicating that the preceding topic has come to an end and that the speaker moves on to a new topic. In the following examples, the topics are

1. the past (a,b), and a digression in (d,e),

2. the present (c),

3. a conclusion (f,g,h), and

4. a result (i)

The start of a new topic is indicated by the adverbials in all cases.

(21)   a. *Kàthaama ikkang, oorang ikkang.*
            first   fish,   man   fish
            "The first were fish.. were fishermen.'

       b. *Kà-duuwa=jo, slaves pada.*
            ORD-two=EMPH slaves PL
            'As for the second, they were slaves.'
            —

       c. **Itthule** *kithang=pe muusing=ka=jo anà-baa.*
            but   1PL=POSS   time=LOC=EMPH PAST-bring
            'But it was in our time that they were brought.'

       d. *Itthu kithang hathu vakthu=nang kithang=nang duppang dhaathang*
            DIST 1PL   INDEF time=DAT   1PL=DAT   before   come
            athi-aada.
            IRR-exist
            'So they, em, at that time, they would have come before us.'

       e. Exiles *pada=jo anà-baa.*
            exiles PL=EMPH PAST-bring
            'Exiles were brought.'
            —

       f. **Itthu=nang=jo** *see anà-biilang*
            DIST=DAT=EMPH 1S PAST-say
            'That's why I am saying'

       g. *oorang soojor   anà-baa   katha.*
            man   European PAST-bring QUOT
            'that the Europeans brought them.'

       h. *Iiya.*
            yes
            'Yeah.'
            —

       i. **Suda** *itthu oorang pada=le Seelong=ka arà-duuduk.*
            thus DIST man   PL=ADDIT Ceylon=LOC NON.PAST-stay
            'So those people stay in Ceylon.' (K060108nar02)

16.2.2   Non-contrastive topic

Topics are normally zero, or at the leftmost position (cf.(19)b). Since the order of
terms before the verb is free, no special morphosyntactic operation (like passivization
or extraction in English) is required to put an element into that position. In example

(22), the undergoer *itthu baathuyang* 'those stones' is found at the leftmost position, and not one of the other referents.

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

In this case, the stones (*itthu baathu=yang*) could be put in any position before the verb, but since this sentence is 'about' the stones, they are put in initial position to highlight their topical role.

If the item chosen to be topic is an active referent, it need not be mentioned at all since it is inferable from context (see above).

16.2.3   Contrastive topic

Another kind of topic is the contrastive topic *As for A ..., (but) as for B ...* (Lambrecht 1994:291), where two different states-of-affairs hold for two different referents. This can be marked by =*jo* ⊡→6.4.7.1, p. 378 in Sri Lanka Malay, although this is optional. In a narrative about different waves of Malay immigration to Sri Lanka (K060108nar02), the introduction of the first and second wave as referents receives no special marking, while the third wave is marked as contrastive topic by =*jo*.

(23)   a.   *Ka-duuva an-dhaathang* slaves *pada,*
            ORD-two PAST-come   slaves PL
            'The second to come were slaves,'

       b.   soldier *pada anà-baa   oorang pada.*
            soldier PL   PAST-bring man   PL
            'people brought by soldiers.'

       c.   *Ka-thiiga=jo   rejiment.*
            ORD-three=EMPH regiment
            'The third, then, were the regiment [Malays].' (K060108nar02)

Another example is (24), where foreign countries and the native country are compared. The second item of the comparison *nni nigirika* 'in this country' receives =*jo*-marking.

(24) a. *Luvar nigiri kithang=nang mà-pii thàrà-suuka.*
outside country 1PL=DAT INF-go NEG-like
'We do not want to go abroad.'

   b. *Nni nigiri=ka=**jo** kitham=pe aanak buva pada=yang*
PROX country=LOC=EMPH 1PL=POSS child fruit PL=ACC
asà-simpang.
CP-keep
'We have raised our children in this country' (K051222nar04)

16.3 Presupposition and assertion

A dichotomy related to topic and comment, yet subtly different, is the opposition between presupposition and assertion, and the notion of focus. Focus is the portion of the utterance which makes the difference between the presupposition and the assertion (Lambrecht 1994). The elements which are in focus are said to be in the focus domain. If the focus domain includes the predicate, we speak of predicate focus. If the focus domain does not include the predicate, but rather an argument, we speak of argument focus. So *John drank tea.* has two different interpretations depending on the focus domain. *John DRANK TEA* has predicate focus. The presupposition was that John did something, the assertion is that John drank tea, and the focus, the difference between presupposition and assertion is DRANK TEA. We learn that John's activity consisted in the ingestion of a liquid rather than going out for a hike or some other activity. *John drank TEA* on the other hand has the presupposition that John drank something, the assertion that John drank tea, and the focus domain is TEA alone. We learn that the beverage John consumed is tea, rather than coffee. In the first case, the focus domain includes the predicate 'drink'; in the second one it does not. There are two additional minor types of foci, namely sentence focus, where the focus domain covers the whole assertion (and the presupposition is void), and contrastive focus (Lambrecht 1994:233,286).

In the following sections, I will discuss predicate focus ⊙→16.3.1, p. 689, argument focus ⊙→16.3.2, p. 689, sentence focus ⊙→16.3.3, p. 691, and contrastive focus ⊙→16.3.4, p. 691 are encoded in SLM.

16.3.1 Predicate focus

Predicate focus is the default case. There is no special marking for it in SLM.

16.3.2 Argument focus

An argument is put into focus by attaching the clitic *=jo* ⊡→6.4.7.1, p. 378.

(25) *[Sri Lanka=ka]=**jo***   *kaaving.*
      Sri Lanka=LOC=EMPH marry
      'Marry in Sri Lanka!' (B060115cvs03)

The focus domain of example (25) is *Sri Lanka*, and does not include *kaaving*
'marry'. This is clear from the context of the utterance, a dialogue where marriage is
discussed. The presupposition is that the addressee will marry, the question is only
where. The presupposition is not what the addressee should do in general, marry in
Sri Lanka, work in Dubai or drive fancy cars in Monaco. The focus domain restricted
to the argument is marked by =*jo* in example (25).

A similar example is (26), where the focus domain does not include the act of
saying, but only the content of Mahinda's utterance.

(26) *[Itthu katha]=jo Mahindha arà-biilang.*
      DIST   QUOT=EMPH Mahindha NON.PAST-say
      'That's what Mahinda [Rajapaksa, President of Sri Lanka] is saying.'
      (K051206nar11)

The focused element does not have to appear in initial position, as the following
example shows.

(27) *Itthu kumpulan=dang **derang=jo** bannyak arà-banthu.*
      DIST association=DAT 3PL=EMPH much   NON.PAST-help
      'It is they who help the association a lot.' (B060115cvs01)

The following example also clearly states the presupposition in the first part, but
provides the surprising assertion in the second part.[5]

---

[5]For focalization purposes, one speaker has a kind of pseudo-cleft construction involving the copula
a(bbi)sdhaathang, which is shown in the following two examples. This construction seems to be idiolectal.

(1) *[See anà-*pass.out*] abbisdhaathang* University of Peradeniya=*ka*
     1s   PAST-graduate COPULA   University of Peradeniya=LOC
     'Where I graduated was the University of Peradeniya.' (K061026prs01,K081104eli05)

(2) a. *[Itthu arà-kirja]   abbisdhaathang.*
         DIST NON.PAST-make COPULA
         'How you make it is'

     b. *thullor asà-ambel=apa baaye=nang asà-puukul=apa ...*
         egg   CP-take=after   good=DAT CP-hit=after
         'you take eggs and beat them well and ...' (K061026rcp01)

(28) *Suda kanabisan=ka raaja Andare=yang mà-enco-king asà-pii,* **raaja=jo**
thus last=LOC   king Andare=ACC INF-fool-CAUS CP-go king=EMPH
su-jaadi   enco.
PAST-become fool
'So finally the king had tried to make a fool out of Andare, but it was the king
who turned out to be the fool.' (K070000wrt02)

A second possibility is extraposition of the focussed element to the right ⊡→13.1.1,
p. 506. This is given in (29).

(29) *Itthu=nang blaakang su-dhaathang [***Hambanthota mlaayu pada***].*
DIST=DAT   after   PAST-come   Hambantota   Malay PL
'After that came the Hambantota Malays.' (K051206nar14)

16.3.3   Sentence focus

Sentence focus, used for thetic statements, does not seem to be distinguished from
predicate focus by segmental or intonational material in SLM. What does distinguish
sentence focus from predicate focus sentences is that all relevant NPs must be realized,
which is hardly ever the case for predicate focus sentences.

16.3.4   Contrastive focus

Contrastive focus indicates that a state-of-affairs contrasting with the presupposition
holds, as in *John is not a teacher, he is a doctor*. In SLM, this is also indicated by
=jo ⊡→6.4.7.1, p. 378. In example (30), two contrastive propositions about the size
of Malay Associations in Sri Lanka are presented. The first one asserts that the idea
that there are big associations is wrong, while the second one asserts that the opposite
idea, small associations, is correct. This contrast is marked by *=jo* in the second
proposition.

(30)   a. **Bìssar** *atthu kumpulan thraa.*
           big   INDEF association NEG
           'There is not one big association.'

    b. **Kiccil** *kumpulan pada=**jo**.*
           small association PL=EMPH
           'there are SMALL ASSOCIATIONS.' (N060113nar01)

A similar situation obtains in (31), where two languages are compared.

(31) *Itthu=ka **mlaayu** thraa, bannyak=nang* **English=jo**   *aada.*
      DIST=LOC Malay   NEG   much=DAT   English=EMPH exist
      'There is no Malay. What there is, is a lot is English.' (B060115prs15)

16.4   Canceling implicatures

The knowledge of the world leads the speakers to make inferences based on the com-
municative content and what it implies. If the speaker assumes that such inferences
have incorrectly been made, he can signal this, like English *but* as in *He is a sports
star, **but** he is not rich.* In this case, the conventional implicature drawn from our
knowledge of the world (that sports stars are normally wealthy) is overtly canceled by
but.

   In SLM, adversative strategies like the one above are a lot less common than in
English. If implicatures need to be canceled, this is normally done by using =*le*, which
is also used for normal coordination (cf. Haspelmath 2004b).

   The following examples show the use of *itthule* ⬚→7.6.4, p. 412. The implicature
of paying the ransom was that the speaker would be sent back as he is, but instead and
unexpectedly, he is turned into a bear.

(32) a. *Se=ppe baapa incayang=nang ummas su-kaasi.*
          1s=POSS father 3s.POLITE-DAT   gold   PAST-give
          'My father gave him the gold.'

      b. ***Itthule*** *see=yang mà-kiiring=nang duppang incayang see=yang hathu*
          But   1s=ACC   INF-send=DAT   before   3s.POLITE 1s=ACC   INDEF
          Buruan mà-jaadi   su-bale-king.
          bear   INF-become PAST-turn-CAUS
          'But before he send me back, he turned me into a bear.' (K070000wrt04)

   Another example of an implicature cancelled is (33), where despite the sizeable
amount of votes obtained, the person was not elected.

(33) a. *Duva-pulu ìnnam riibu   ùmpath raathus lima-pulu duuva* votes
          two-ty   six   thousand four   hundred five-ty   two   votes
          incayang=nang anà-daapath.
          3s.polite=DAT   PAST-get
          'He got 26,452 votes.'

      b. ***Itthu=nang=le*** *incayang=nang=le inni thumpath thàrà-daapath.*
          DIST=DAT=ADDIT 3s.polite=DAT=ADDIT PROX place   NEG.PAST-get
          'In spite of that, he did not get the seat either.' (N061031nar01)

Normally, *itthu* is used as an anaphora for the content whose implicature is to be cancelled. In rare cases, *=le* can be found without *itthu* in this function. An example where *=le* attaches to a noun is given in (34).

(34) *Itthukapang, kumpul-an=nang=**le**, punnu mlaayu pada lai   Seelong=ka*
      then   mix-NMLZR=DAT=ADDIT full   Malay PL   more Seelon=ADDIT
      aada
      aada
      'Then, and in spite of all this mixing, there are still a good lot of Malays in Sri Lanka.' (K060108nar02)

An example where *=le* attaches to the verb *kaala* 'lose' is (35).

(35) a. *Sudaara TB Jayah inni state council election pada=nang duuduk*
          Brother   TB Jayah PROX state council election PL=DAT   exist.ANIM
          aada.
          exist
          'Brother TB Jayah was in that state council election.'
      b. *Thiga-pulu ìnnam=ka incayang itthu=dering su-kaala.*
          three-ty   six=LOC 3s.POLITE DIST=ABL   PAST-lose
          'He lost in the '36 elections.'
      c. *Kaala=**le**   thàrà=na=apa incayang=nang appointed member=pe*
          lose=ADDIT NEG=DAT=after 3s.polite=DAT   appointed member=POSS
          hathu thumpath=yang government=ka anà-kaasi.
          INDEF seat=ACC   government=LOC PAST-give
          'In spite of having lost, he was given a seat as appointed member in the government.' (N061031nar01)

If *=le* is used together with *kala-* 'if', the meaning is irrealis 'even if' instead of realis 'although' (36).

(36) *Incayang pukuran kala-gijja=le, laile hatthu miskin.*
      3s.POLITE work   if-make=ADDIT still INDEF poor
      'Even if he works, he will still remain a pauper.' (K081104eli05)

It is also possible to combine *=le* in its adversative meaning with interrogative pronouns (37). Note the similarity to the WH=CLT-construction discussed in ⊡→8.1.12, p. 424.

(37)   *Incayang **dhraapa**   pukuran gijja=**le**,   laile hatthu miskin.*
       3s.POLITE how.much work   make=ADDIT still INDEF   poor
       'However much he might work, he will still remain a pauper.' (K081104eli05)

The arbitrariness of the degree can be reinforced by *boole* 'can' (38).

(38)   *Itthu oorang dhraapa   kaaya bole=jaadi=le,   see dee=yang*
       DIST man   how.much rich   can=become=ADDIT 1s   3s.IMPOLITE=ACC
       thama-kaaving.
       NEG.IRR-marry
       'However rich that man might be, I will never marry that bastard.'
       (K081104eli05)

16.5   Parsing

One important step in structuring information is to indicate what belongs together and what does not. The coherence of constituents can be marked by intonation, pausing, and positional information in SLM.

16.5.1   Identifying phonological constituents

Intonational cues present in SLM are falling pitch for the end of a declarative sentence and rising pitch at the end of interrogative and subordinate sentences ⊡→4.8, p. 140. By paying attention to these cues, the hearer can segment the speech signal in a number of presuppositions and assertions.

Identification of individual words is aided by the bimoraic foot structure with its consequences on vowel length and consonant gemination ⊡→4.5, p. 120. On hearing a long vowel or a geminated consonant, the hearer can immediately assume that he is dealing with the penultimate syllable of a lexeme (with the exception of the very limited number of monosyllables with a long vowel.) In this way, a word boundary after the final syllable can be retrieved.

Given the positional restriction of a number of phonemes which cannot occur in final position, such as palatals or voiced stops, the hearer has negative evidence that a word boundary is not present at that place. The negative cues for the beginning of a word are much weaker. Only /ŋ/ cannot occur in this position and hence precludes a word boundary where it occurs.

16.5.2   Identifying syntactic constituents

The assertive contour ⊡→4.8.1, p. 142 can be used to identify the predicate phrase, while the presuppositional contour ⊡→4.8.3, p. 145 can be used to identify NPs and PPs. The further parsing of VPs can be done by identifying a verbal prefix. What follows the prefix must be part of the verb, but what precedes cannot. NPs and PPs can be chopped starting from the right by identifying first clitics and then postpositions. As soon as the segmental material can no longer be interpreted as a clitic or a postposition, the stem has been reached. Further cues are given by a long vowel or a geminate consonant, which indicate the penultimate syllable of a lexeme.

16.5.3   Identifying syntactic constituency

Syntactic constituency can be indicated in the languages of the world by head marking, dependent marking, and positional information (Nichols 1986). SLM does not make use of head marking, but uses dependent marking extensively to indicate the relationship between two words. This is mostly done with postpositions, which indicate the semantic roles of the arguments of a predicate, as in (39).

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

A postposition on the dependent is also used to indicate possessors in the noun phrase: *=pe*. The following example contains five possessive relationships. In all of these, the dependent (possessor) is overtly marked, while the head is not.

(40)   *Itthu=kaapang [se=**ppe** baapa] [se=**ppe** kaake]   [[se=**ppe***
        DIST=when   1s=POSS father   1s=POSS grandfather 1s=POSS
        *kaakee=**pe**]   baapa] kithang samma oorang [Seelon=**pe**   oorang]*
        grandfather=POSS father, 1PL   all   man   Ceylon=POSS man
        pada.
        PL
        'Then my father and my grandfather and my grandfather's father, all of us people became Ceylon people.' (K060108nar02)

Furthermore, a number of subordinate clauses indicate their dependent status on the verb. This is the case for clauses with *asà-* and *mà-* (41).

(41)   a.   *[Pohong=dering baava=nang **asà**-thuurung],*
            tree=ABL   down=DAT   CP-descend
            'Tthe monkeys) climbed down from the tree and'

       b.   *[oorang anà-baava samma thoppi=pada **asà**-ambel],*
            man   PAST-bring all   hat=PL   CP-take
            'took all the hats the man had brought and'

       c.   *[**mà**-maayeng]=nang su-mulain.*
            INF-play=DAT   PAST-start
            'started to play.' (K070000wrt01)

However, not all subordinate clauses are overtly marked for their dependent status. (42) consists of a main clause and an argument clause, but the embedded status of the argument clause is not signalled morphosyntactically.

(42)   *Blaakang=jo incayang anà-kuthumung [moonyeth pada pohong atthas=ka*
       after=EMPH   3S.POLITE PAST-see   monkey   PL   tree   top=LOC
       *arà-maayeng]*arg.
       SIMULT-play
       'Then only he saw that the monkeys were playing in the top of the tree.'
       (K070000wrt01)

Besides segmental information of the dependent in clauses, NPs and clause combinations, positional information is also available. Dependents tend to precede their head in all areas of SLM grammar. Nominal modifiers precede nouns ⊡→8.1.17, p. 444, arguments precede verbs, subordinate clauses precede main clauses ⊡→13.1, p. 505.

Within the phonological phrases identified by intonation contours, the morphosyntactic head is normally to the right, and modifiers are to the left. This means that the predicate (head) of a sentence can easily be identified. The arguments are then the remaining elements. Similar things can be said about NPs, where the rightmost element is normally the head noun. True, there are some cases where the head is not the rightmost element. This is not very often the case for clauses, but much more common for NPs. This entails that for the hearer, it is much more difficult to establish the constituency of the noun phrase as compared to the constituency of the clause. If one assumes that NPs have little constituency and are more appositional in nature ⊡→8.9, p. 455, this need not be a problem.

Chapter 17

Pragmatics

After we have seen which are the elements and constructions of SLM grammar, and how different meanings can be expressed by them, we now turn to the different communicative goals that speakers may want to achieve through the use of these units. We will first discuss different types of speech acts (requesting confirmation, requesting action, answering, etc.) ⊙→17.1, p. 697 and then turn to the ways how speech acts can be nuanced with regard to force and politeness ⊙→17.2, p. 708.

17.1   Speech acts

17.1.1   Providing information

If a speaker wants to provide information, he uses an assertion. Assertions have been dealt with extensively in the preceding chapters, and all relevant information can be taken from there.

Speakers do not only want to provide information. Sometimes, they also want the hearers to do something, e.g. provide missing information or perform a certain action (Givón 2001b:319). They might also want to express their agreement with the hearer or someone else to perform a certain action. This will be dealt with in the following sections.

17.1.2   Requesting information

Requesting information can be divided into two subtypes: requesting information about the truth of a certain proposition, and requesting information about the content of a certain proposition. The former type is called 'polar question' and the latter, 'content question'.

Content questions are formed in SLM by means of an interrogative clause for content ⊡→13.1.4, p. 513, which resembles a normal clause, but replaces the NP or PP

queried for by an interrogative pronoun like *saapa* 'who'.

A special type of content question is concerned with metalinguistic information about a translation of a word into another language. For this, the dative marker =*nang* is used. The interrogative pronoun is either *aapa* as in (1) or *aapayang* as in (2).

(1) "Simple"=***nang aapa*** *arà-biilang.*
     simple=DAT   what NON.PAST-say
     'What do you say for "simple"?' (B060115cvs01)

(2) "Terrorist"=***nang aapayang*** *arà-biilang, mlaayu=dering.*
     terrorist=DAT   what   NON.PAST-say Malay=ABL
     'How do you say "terrorist" in Malay?' (K051206nar02)

The second possibility are polar questions, which are formed by a polar interrogative clause ⊡→13.1.3, p. 512, or by a declarative clause followed by the tag *bukang* 'isn't it?'

The clitic =*si* is neutral as to the expected answer. In example (3), the speaker is not sure whether she has the obligation to tell her age or not. If a positive answer is expected, the declarative sentence followed by *bukang* 'isn't it' is used (4)(5).

(3) *Se=pe uumur masà-biilan=si?*
     1=POSS age   must-tell=INTERR
     'Do I have to tell my age?' (B060115prs01)

(4) *Itthu saala, bukang?*
     DIST wrong NEG.NONV
     'That is wrong, isn't it?' (K060116nar11)

(5) *Kandi=ka   ithu Thuan Skiilan aada, bukang?*
     Kandy=LOC DIST Thuan Skiilan exist TAG
     'Now there is this Thuan Skiilan in Kandy, you know?' (B060115cvs04)

When expecting a negative answer, a negative declarative sentence is used, followed by =*si*.

(6) *Puaasa muusing thàrà-duuduk=si?*
     fasting season   NEG-stay=INTERR
     'You were not here during the fasting season, were you?' (B060115cvs03)

Questions are also marked by a particular intonation contour (Section 4.8.4, p. 147). If that contour is clear, the segmental marking (=*si*) can be left out from polar

questions.

17.1.3 Requesting confirmation

Requesting confirmation of the expected truth value has been discussed above in the preceding section. Another kind of confirmation which can be requested is whether the hearer is actually following the words of the speaker. This can be done by *thaau=si* 'you know?'. The hearer can then confirm verbally or non-verbally, and the speaker can continue his turn.

(7) *Bannyak pukurjan anà-kirja soojor   pada-saama,* **thaau=si**   *soojor?*
    much   work   PAST-do European PL=COMIT   know=INTERR European
    'He worked together with "soojors". You know "soojor"?' (K061026prs01)

Another possibility to request confirmation is *=jona* ⊡→6.4.7.2, p. 385.

(8) *Punnu mlaayu pada kàthaama* English=***jona*** *anthi-oomong.*
    many Malay PL   earlier   English=PHAT IRR=speak
    'Many Malays would speak English, y'know, in former times.' (K051222nar06)

17.1.4 Requesting permission

Permission is requested by adding the interrogative clitic *=si* ⊡→6.4.6.1, p. 360 to a verb marked by either the infinitive *mà-* ⊡→6.2.1.5, p. 279 or the adhortative *marà-* ⊡→6.2.2.6, p. 301.

(9) *See* **mà**-*maakang=***si***?*
    1s   INF-eat=INTERR
    'Shall I eat?' (K081106eli01)

(10) *Kithang* **marà**-*maakang=***si***?*
    1PL   ADHORT-eat=INTERR
    'Shall we eat?' (K081106eli01)

17.1.5 Requesting action

Next to providing and requesting information, speakers also sometimes want other people to perform certain actions, or to refrain from performing certain actions. The

most straightforward way to request action is an imperative construction ⊡→13.1.5, p. 514.

(11)   *Binthan* auntie=*ka   caanya, Binthan* auntie=*yang konnyong panggel.*
       Binthan auntie=LOC ask,   Binthan auntie=ACC few   call
       'Ask auntie Binthan, call Binthan auntie' (K060116nar06)

Milder than the preceding utterance is the use of *maau* 'want' ⊡→5.12.3, p. 250, as in the following examples.

(12)   *Baapa=nang* mosque=*nang mà-pii* **maau**.
       father=DAT   mosque=DAT   INF-go want
       'Father, you should go to the mosque/ Father, you want to go to the mosque.'
       (B060115nar04)

(13)   *Manis-an=nang   mà-thaaro guula* **maau** *gula   paasir konnyong* **maau**.
       sweet-NMLZR=DAT INF-put   sugar want sugar sand few   want
       'In order to make it sweet, you need sugar, you need a bit of crystal sugar.'
       (B060115rcp02)

The negation of *maau*, *thussa* ⊡→5.12.4, p. 253, can be used for negative commands, as in (14).

(14)   *'Thussa mà-thaakuth', Buruan su-biilang.*
       NEG.IMP INF-fear   bear   PAST-say
       ' "Don't be afraid," said Bear.'/' "I don't want you to be afraid", said Bear.'
       (K070000wrt04)

Another more subtle way is to use *lorangnang ... (thàr)boole* 'you can(not) ...' ⊡→5.12.2, p. 249 (15), or the verbal prefix *masthi* 'must' ⊡→6.2.2.3, p. 294 with an unexpressed referent (16). Both imply that it would be preferable that the action be (not) performed, but lack coercive power (in the case of *(thàr)boole*) or the direct expression of the addressee (in the case of *masthi*).

(15)   *Ithu-kapang umma-baapa su-biilang lorang=nang kaaving* **thàrboole**.
       DIST=when mother-father PAST-say 2PL=DAT   marry   cannot
       'Then the parents said: "You cannot marry/Don't marry." ' (K051220nar01)

(16) *Minnyak klaapa ini raam̌buth=dering **masà**-goosok.*
coconut.oil coconut PROX hair=ABL must-rub
'You must rub coconut oil (over the itching) with (human) hair.'
(K060103cvs02)

Adhortatives can be formed with *marà-* ⊡→6.2.2.6, p. 301.

(17) *Kitham **marà**-maayeng.*
1PL ADHORT-play
'Let's play.' (K081104eli06)

17.1.6 Answering

Answering affirmatively is done by *iiya* (18) or *saya*. Negative answers are given by *thraa*, regardless of predicate type (19).

(18)   a.   *SN: Butthul* sportsman*, bukang?*
correct sportsman, TAG
'They are good sportsmen, aren't they?'

b.   *SLM: **iiya iiya**,* sportsman.
Yes, yes, sportsman
'Yes indeed, they are good sportsmen.'
[...]

c.   *SN:* Football *aada?*
Football exist
'Is there football being played in Badulla.'

d.   *SLM: **Iiya*** football *aada.*
yes football exist
'Yes, there is football.' (B060115cvs01)

It is also possible to repeat the predicate of the question for affirmative purposes, or to repeat a verbal predicate and add a negative prefix (*thàrà-/thama-*) as a negative answer. This can be seen for example in (19), where the predicate *kaaving* 'marry' is repeated with the negative prefix *thama-* in the answer (Why the non-past prefix is chosen here is unclear, though).

(19)   a. *Q: Sebastian su-**kaaving**=si?*
            Sebastian PAST-marry=INTERR?
            'Is Sebastian you married?'

        b. *A: **Thraa thraa**, Sebastian thama-**kaaving**.*
            no   no   Sebastian NEG.NONPAST-marry
            'No, no, Sebastian is not married.'[1](B060115cvs03)

If the answer to a question reasserts a known referent, this referent is often redu-
plicated. This is the case in (20), where the string *dìkkath* is present in the question
and hence active in discourse. In the answer, this string is present twice, indicating
that the speculation of the hearer as to the vicinity of the school has been understood
and is indeed correct.

(20)   a. *Q: Skuul   dìkkath=si?*
            school vicinity=INTERR
            'Is the school close (expecting yes).'

        b. *A: Dìkkath dìkkath.*
            vicinity vicinity
            'Oh yeah, it is close indeed.' (K081104eli06)

17.1.7   Allowing

Allowing is normally done with the modal *boole* 'can' ⊡→5.12.1, p. 248.

(21)   a. *Dadaa, se=dang sinema mà-liiyath=nang bolle=pii=si?*
            Daddy 1s=DAT   cinema INF-see=DAT   can=go-INTERR
            'Daddy, can I go to the cinema?'

        b. *Iiya, Izi=nang bolle=pii.*
            yes, Izi=DAT   can=go
            'Yes, you can go, Izi.' (K081104eli06)

The lexical verb for 'to allow' is *luppas*, which also means 'leave (behind)'.

(22)   *See Izi=yang arà-luppas   sinema mà-liyath=nang mà-pii.*
        1s Izi=yang NON.PAST-leave cinema INF-watch=DAT   INF-go
        'I allow you/Izi to go to the cinema.' (K081104eli06)

---

[1]A second native speaker answers instead of the researcher.

17.1.8 Wishing

Wishes can be made by invoking Allah. Interestingly, the complement of wishes is marked with the pseudo-prefix *masà-* 'must', instead of an irrealis marker or an infinitive, which would have been expected.

(23) *Yaa, Allah, se arà-mintha   se=dang itthu pukuran masà-daapath katha.*
Yeah, Allah, 1s NON.PAST-wish 1s=DAT DIST job   must-get   QUOT
'Allah, I wish I could get that job.' (K081104eli06)

Wishes can also be indicated by *incalla* 'if God wills'.

(24) *Incalla   [lai   thaau sudaara sudaari pada]=ka bole=caanya ambel*
Hopefully other know brother sister PL=LOC   can-ask   take
[nya-gijja lai   saapa=kee aada]=si   katha.
PAST-make other who=SIMIL exist=INTERR QUOT
'Hopefully, you can enquire from another person you know whether there is someone else who did something.' (N061031nar01)

17.1.9 Mirativity

Mirativity can be expressed by the insertion of *thaausi?* 'you know what!?'.

(25) *Suda inni moonyeth pada]aapa thaau=si[anà-gijja !*
SO   PROX monkey PL   what know=INTERR PAST-do
'Do you know what these monkeys then did!?' (K070000wrt01)

Another possibility involves the use of the irrealis marker *anthi-.* shown in (26).

(26) *Itthu saapa **anthi-aada**!*
DIST who IRR-exist
'Who could/might that be!' (K070000wrt04)

17.1.10 Promises

Promises are made with the irrealis marker *anthi-* ⊡→6.2.2.2, p. 290 (27) or with the non-past marker *arà-* ⊡→6.2.2.1, p. 287 (28).

(27)   *Kithang lorang=nang baaye mliiga **athi**-kaasi, mà-kaaving panthas*
       1PL   2PL=DAT   good palace IRR-give   INF-marry beautiful
       *pompang pada **athi**-kaasi, duvith **athi**-kaasi.*
       girl   PL   IRR-give   money IRR-give
       'We will give you beautiful palaces and women to marry and money.'
       (K051213nar06)

(28)   *See arà-**sumpa**   paanas muusing dhaathang=thingka see siini=dering*
       1s NON.PAST-promise hot   time   come=middle   1s here=ABL
       ***arà**-pii.*
       NON.PAST-go
       'I promise, I will leave when summer will have come.' (K070000wrt04)

17.1.11   Explanations

Explanations are often structured along the pattern of statement, rhetorical question, repetition of statement. The following stretch of discourse by one speaker is in English, but the pattern is also found in SLM even though it is not in the corpus.

(29)   a. Two miles we have to walk.

       b. How many miles?

       c. Two mile *duuva kaayu kithang masà-pii.*
           two mile two   mile 1PL   must-go
           'Two miles we had to walk.' (K051213nar03)

17.1.12   Greeting people

While the other native ethnic groups in Sri Lanka do normally not express greetings verbally, this is different for the Malays.[2] The most common greeting is *slaamath*, which can be uttered at arrival or departure, and also on the telephone. It can be complemented by the period of the day *slaamath paagi* 'good morning', *slaamath soore* 'good afternoon/evening', and *slaamath maalang* 'good night'. More informal greetings do not exist, but English *hello* can be used for such purposes.

   More formal is *slaamath dhaathang* 'welcome', which can be uttered at the same occasions as its English counterpart.

---

[2]Tamil *vanakkam* and Sinhala *Aayu boovan* are only used for special occasions. One Sinhala informant reported that he could not remember when he had last used the greeting. Things are a bit different for Moors, who use *Salam aleikum* on a regular basis.

17.1.13 Taking leave

When leaving, one can use *slaamath* as described above, or one can use *spi dhaathang* 'go and come'. This is modelled on an analogous construction in Sinhala and Tamil. The rationale for this is that it is considered unfortunate to say *I leave* because that might mean departure from this world. Therefore, one has to specify that one will return, and this is done by adding the *dhaathang* 'come' part. Optionally, the *spi*-part can be left out, which leaves the curious situation that you can say *I am coming* to indicate you are leaving, as in the following example.

(30) *Se=dang arà*-late,   *bukang,* **see arà-dhaathang**.
      1s=DAT   NON.PAST-late TAG   1s   NON.PAST-come
      'I am getting late, aren't I, goodbye.' (B060115cvs08)

17.1.14 Thanking

Just as with greetings, Sinhalese and Tamils do not use verbal means for this function, but Malays do. The normal way of thanking is *thriima kaasi* 'thank give'.

(31) *(Bannyak) thriima kaasi.*
      much   thank   give
      'Thank you very much.' (ubiquitous)

17.1.15 Asking a favour

Asking a favour is normally done with an interrogative construction involving *=si*. Given that the favour is typically beneficial to the asker, the vector verb *kaasi* 'give' is often used then.

(32) *Mamaa, se=dang itthu=yang anthi-ambe **kaasi=si**?*
      mum   1s=DAT   DIST=ACC   IRR-take   give=INTERR
      'Mummy, can you buy that for me?' (K081104eli06)

(33) *Mamaa, se=dang nyaari daaging athi-maasak **kaasi=si**?*
      mum, 1s=DAT today meat   IRR-cooked give=INTERR
      'Mama, can you cook beef for me today?' (K081104eli06)

     The asking of favours involves polite pronouns. It is even possible to use double marked plural forms like *loram pada* with singular reference then.[3]

---

[3]Tamil also uses plural pronouns for polite reference (Lehmann 1989:100).

(34) *Saayang, **loram pada**=sesama nyaari* Galle Face *mà-pii*
      love   2PL   PL=COMIT   today Galle Face INF-go
      arà-dhaathang=si?
      NON.PAST-come=INTERR
      'Darling, shall I come and take you to the Galle Face Green [a favourite hang-
      out for young couples in Colombo]?' (K081104eli06)

   (35) gives a request about a third person.

(35) Auntie, *se arà-mintha incayang se=ppe masà-jaadi   katha.*
      auntie   1s NON.PAST-ask 3s.polite 1s=POSS must-become QUOT
      'Madam, may I ask for the hand of your daughter.' (K081104eli06)

17.1.16   Granting a favour

Granting a favour is done by *iiya* 'yes' or *butthul* 'correct, OK'. (36) is the positive
answer for request (32).

(36) *Iiya, butthul, se athi-ambel kaasi.*
      yes   correct 1s IRR-take   give
      'Yes, I will buy it for you.' (K081104eli06)

   The positive reply to (34) is (37). Note the use of the terms *saayang* 'love' for the
girl in (34) and *jiiva* 'life' for the boy in (37).

(37) *Iiya, jiiva, se thi-dhaathang.*
      yes   life   1s IRR-come
      'Yes, dear, I will come.' (K081104eli06)

17.1.17   Declining a favour

The negative answer to (32) is (38), the negative answer to (34) is (39), the negative
answer to (35) is (40).

(38) *Thraa, se loran=nang thama-ambel kaasi.*
      NEG   1s 2PL=DAT   IRR-take   give
      'No, I won't buy it for you.' (K081104eli06)

(39) *Thookal=si?*
      fool=INTERR
      'Are you crazy?/Get lost!' (K081104eli06)

(40) *Thookal, itthu kalu, thama-jaadi!*
      fool   DIST if   NEG.IRR-become
      'You fool! As for that, that will never happen!' (K081104eli06)

17.1.18   Telling your age

There are two ways to tell your age, one modelled on the European pattern and one on the Lankan pattern. The European pattern is stating the number of your last birthday, while the Lankan way is stating the number of your next birthday with the verb *gìnnap* 'complete'. Both cases put the argument in the dative case, in the examples *sedang*.

(41) *Karang se=dang ìnnam-pulu liima (thaaun) uumur.*
      now   1s=DAT six-ty   (year) five   age
      'Now I'm sixty-five.' (N060113nar03)

(42) *Sangke=nyaari se=dang limapulu thaaun arà-gìnnap.*
      until=today   1s=DAT fifty   year   NON.PAST-complete
      'I am in the course of my fiftieth year on earth (= I am 49).' (K060108nar01)

17.1.19   Introducing yourself

A formal way to introduce yourself is given in (43).

(43) *Sudaara sudaari se=ppe naama Wahida Jamaldiin.*
      brother sister   1s=POSS name Wahida Jamaldeen
      'Brothers and sisters, my name is Wahida Jamaldeen.' (B060115prs05)

17.1.20   Expressing annoyance

Annoyance is not often expressed openly. The following possibility was gathered through elicitation.

(44) *Cumma duuduk-la!*
     idle   stay-IMP
     'Mind your own business!' (K081105eli02)

17.1.21 Ending a conversation

An explicit way to end a conversation is given in (45). The sentences in (46) and (47) are heard very often, but I am not very sure about their pragmatic implications. They might be a polite hint to stop now, or, on the contrary, they might signal desire to continue.

(45) *Skaarang marà-birthi-king, se=dang lai   pukuran aada.*
     now   ADHORT-stop-CAUS, 1s=DAT more work   exist
     'Let's stop now, I still have some work to do.' (K081104eli06)

(46) *Lai   aapa aada?*
     more what exist
     'What else?' (K081104eli06)

(47) *Lai   aapa=ke   (mà-oomong) aada=si?*
     more what=SIMIL INF-talk   exist=INTERR
     'Is there something else to say?' (K081104eli06)

17.1.22 Hesitation

Hesitation is marked by the proximal deictic *inni*, as in (48), when looking for the right word.

(48) *laile itthu=yang   arà-..**inni**..-iingath-kang   inni   police=dering*
     still DEM.DIST=ACC NON.PAST-..HESIT..-think-CAUS DEM.PROX police=ABL
     'Still the police bring into memory this story.' (K051206nar02)

17.2   Blending in the social tissue

This section is mainly concerned with politeness. The main expression of politeness is through personal pronous (see Table 5.5 on page 225). By choosing the polite form, distance is conveyed, whereas the intimate form conveys closeness. As for bound

morphology, the optional imperative suffix *-la* ⊡→6.2.4.1, p. 304 makes a request more polite, while the impolite imperative suffix *-de* ⊡→6.2.4.2, p. 304 yields a very impolite request , but increases the force of the request.

17.2.1 Terms of address

When speaking to relatives, the term for the kinship relation may be used instead of the 2*nd* or 3*rd* person pronoun (Saldin 2001:55). (49) and (50) show the use of kin terms as terms of address.

(49) *Hatthu=le jamà-gijja,* **baapa** *ruuma=ka duuduk.*
INDEF=ADDIT NEG.NONFIN-make father house=LOC stay
'Father, do not do anything and stay at home!' (B060115nar04)

(50) a. *Andare aanak=nang su-biilang:*
Andare child=DAT PAST-say
'Andare said to his child:'

 b. **Aanak**
child
' "Son, " '

 c. *'Lu=ppe umma su-maathi' katha bitharak=apa asà-naangis mari.*
2S=POSS mother PAST-die QUOT scream=after CP-weep come.IMP
' "come and cry and weep 'My mother has died!' " '

The use of kin terms when referring to third person is shown in (51), a stretch of discourse about the father of the speaker, who, despite of being introduced early on, is mostly referred to as *baapa*. The zeroes indicate instances where the referent is not overtly realized. Also note the reference to *aade* 'younger sibling' in a similar sense.

(51)   a. *Itthu blaakang **baapa** sajja anà-duuduk.*
           DIST after   father only PAST-saty
           'The father stayed at home alone.'

       b. ***Baapa** asà-duuduk suda giithu   samma oorang=pe samma aanak*
           father CP-stay   thus like.that all   man=POSS all   child
           *pada=pe ruuma **baapa** arà-pii.*
           PL=POSS house father NON.PAST-go
           'Father stayed there and then like that father went to all folks'', all chil-
           dren's homes.'

       c. *Giithu   asàduuduk   kithang transfer su-dhaathang Kandy=nang.*
           like.that cp.stay/from 1PL   transfer PAST-come   Kandy=DAT
           'From there, we were transfered to Kandy.'

       d. *Kandy=nang dhaathang=jo **baapa**=le   su-dhaathang Kandy=nang.*
           Kandy=DAT come=EMPH   father=ADDIT PAST-come   Kandy=DAT
           'Only when we came to Kandy did father come to Kandy.'

       e. *∅ kithang=pe ruuma=ka=le   anà-duuduk, **aade**=pe*
             1PL=POSS   house=LOC=ADDIT PAST-stay   younger.sibling=POSS
           ruuma=ka=le   anà-duuduk.
           house=LOC=ADDIT PAST-stay
           'He stayed at our house and he stayed at (my) younger sibling's house.'

       f. *∅ itthu samma ruuma=ka asà-duuduk*
             DIST all   house=LOC CP-stay
           'Having stayed in all those houses'

       g. last=*ka   **baapa** Badulla=nang anà-pii*
           last=LOC father Badulla=DAT   PAST-go
           'at last, father returned to Badulla.' (K051205nar05)

   Next to kin terms, proper names can also be used as terms of address (52)(53).

(52)   *Andare=nang asà-panggel anà-biilang: 'Se=dang **Andare**=pe   biini=yang*
       Andare=DAT cp-call   PAST-say   1s=DAT   Andare=POSS wife=ACC
       mà-caañda suuka.'
       INF-meet   want
       'He called Andare and said: "I want to meet Andare's wife".' (K070000wrt05)

(53)   ***Aashik**=nang hathu soldier mà-jaadi   suuka=si   katha arà-caanya.*
       Aashik=DAT   INDEF soldier INF-become like=INTERR QUOT NON.PAST-ask
       'He asks if you want to become a soldier, Ashik.' (B060115prs10)

Elder relatives may also be addressed as or referred to with the English terms *uncle* and *auntie* (54)-(56). This does not imply that the person is actually a sibling of a parent.

(54) **Uncle***, kitham=pe umma anà-biilang giini.*
uncle 1PL=POSS mother PAST-say   like.this
'Uncle, our mother said the following: ... .' (K051220nar01)

(55) *Diya-la,* **uncle** *se=ppe SSC exam mà-ambel=nang duva-pul aari su-aada.*
see-IMP uncle 1s=POSS SSC exam INF-take=DAT   two-ty   day PAST-exist
'See, uncle, there are (were) only twenty days left until my SSC exam.'
(K051220nar01)

(56) a. *Thuan Kuddhuus, Thuan Thungku, Thuan Skiilan, Thuan Idriis,*
        Thuan Kuddhuus Thuan Thungku Thuan Skiilan Thuan Idriis
        'Thuan Kuddhuus, Thuan Thungku, Thuan Skiilan, Thuan Idriis,'

     b. *Thuan –* **uncle***, hathyang naama pada saapa ?*
        Thuan uncle other   name PL   who
        'Thuan ...   — what's the names of the others again, uncle?'
        (K060108nar02)

17.2.2 Distancing

Distancing oneself from the provided content can be done by *sedang kalu* 'as for me', *sedang thaau mosthornang* 'as far as I know' or *boolebiilang* 'you could say that ...'

(57) *Kaake=nang   dhraapa=so   bannyak aanak pada,* ***se=dang kalu***
      grandfather=DAT how.many=UNDET much   child PL   1s=DAT if
      blaangang thàrà-thaau.
      amount   NEG-know
      'Grandfather had several children, many children, as for me, I do not know the number.' (K051205nar05)

(58) ***Se=dang thaau mosthor=nang*** *inni=jo   inni kithang=pe inni*
      1s=DAT   know manner=DAT   PROX=EMPH PROX 1PL=POSS   PROX
      Seelong=nang political news.
      Ceylon=DAT   political news
      'As far as I know, this is the political news for Sri Lanka.' (N061031nar01)

(59) **Bole-biilang** *se=ppe kaake=le*   *asàdhaathang hathu* army officer
      can-say   1s=POSS grandfather=ADDIT COPULA   INDEF army officer
      Badulla=ka.
      Badulla=LOC
      'You can say that my grandfather also was an army officer in Badulla.'
      (B060115cvs05)

Direct attribution of the source is also possible, as in (60), where the source of the
information *spaaru oorang pada* 'some people' is explicitly mentioned.

(60) **Spaaru oorang pada arà-biilang** *Seelong=nang* English *anà-aaji.baa*
      some   man   PL   NON.PAST-say Seelon=DAT   English PAST-bring.ANIM
      katha
      QUOT
      'Some people say it is the English who brought them to Sri Lanka.'
      (K060108nar02)

17.2.3   Reinforcing

Slomanson (2008b) reports the use of the string *apa kata k@m-bilang* (his orthogra-
phy) for reinforcing purposes. The following is an example he gives.

(61) *Apa kata k@m-bilang, April ka jo   e-datang   (aDa)*
      what word when-say   April in FOC ASP-come AUX
      'What I say to you is that it has come in April.' (Slomanson 2008b:7(17a))

With commands, *-la* ⊡→6.2.4.1, p. 304 can be used as a polite reinforcer, while
-de is a very impolite way to reinforce a command ⊡→6.2.4.2, p. 304.

Bibliography

ABBI, A. (1994). *Semantic universals in Indian languages*. Shimla: Indian Institute of Advanced Study.

ABBI, A. (2001). *A Manual of Linguistic Field Work and Structures of Indian Languages*. München: Lincom Europa.

ABBI, A. & D. GOPALAKRISHNAN (1991). "Semantics of explicator compound verbs in south Asian languages". *Language Sciences*, 13(2):161–180.

ABEYASEKERE, S. (1987). *Jakarta – A history*. Singapore: OUP.

ADELAAR, K. A. (1985). *Proto-Malayic; The reconstruction of its phonology and parts of its morphology and lexicon*. Ph.D. thesis, Leiden University.

ADELAAR, K. A. (1991). "Some notes on the origin of Sri Lankan Malay". In Steinhauer (1991), pp. 23–37.

ADELAAR, K. A. (2005). "Structural diversity in the Malayic subgroup". In Adelaar & Himmelmann (2005), pp. 202–226.

ADELAAR, K. A. & N. P. HIMMELMANN (eds.) (2005). *The Austronesian Languages of Asia and Madagascar*. London, New York: Routledge.

ADELAAR, K. A. & D. J. PRENTICE (1996). "Malay: its history, role and spread". In Wurm et al. (1996), pp. 673–693.

AIKHENVALD, A. Y. (2004). *Evidentiality*. Oxford: OUP.

AIKHENVALD, A. Y. (ed.) (2006a). *Serial verb constructions – a cross-linguistic perspective*. New York: OUP.

AIKHENVALD, A. Y. (2006b). "Serial verb constructions in typological perspective". In Aikhenvald (2006a), pp. 1–68.

ALIEVA, N. F., V. D. ARAKIN, A. K. OGLOBIN & Y. H. SRIK (1991). *Bahasa Indonesia: Descripsi dan Teori*. Yogyakarta: Penerbit Kanisius.

ALLEYNE, M. C. (2000). "Opposite processes in creolization". In Neumann-Holzschuh & Schneider (2000), pp. 125–134.

AMEKA, F. K., A. DENCH & N. EVANS (eds.) (2006). *Catching language – The Standing Challenge of Grammar Writing*. Berlin, New York: Mouton de Gruyter.

ANNAMALAI, E. & S. B. STEEVER (1998). "Modern Tamil". In STEEVER, S. B. (ed.) "The Dravidian Languages", London: Routledge. pp. 100–128.

Ansaldo, U. (2005a). "Contact-induced morphologization. Motivations for emergence of case in Kirinda Java". Paper presented at NRG3, Santiago de Compostela, 17-20 July 2005.

Ansaldo, U. (2005b). "Typological admixture in Sri Lanka Malay. The case of Kirinda Java." Unpublished manuscript, available at http://home.medewerker.uva.nl/u.ansaldo/index.html.

Ansaldo, U. (2008). "Sri Lanka Malay revisited: Genesis and classification". In Dwyer, A., D. Harrison & D. Rood (eds.) "A world of many voices: Lessons from documented endangered languages", Studies in Language, Amsterdam, Philadelphia: John Benjamins. pp. 13–42.

Ansaldo, U. (2009a). "Contact language formation in evolutionary terms". In Aboh, E. & N. Smith (eds.) "Complex processes in new languages", Amsterdam: John Benjamins.

Ansaldo, U. (2009b). *Contact languages, Ecology and Evolution in Asia*. Cambridge: Cambridge University Press.

Ansaldo, U. & L. Lim (2006). "Globalisation, empowerment and the periphery – The Malays of Sri Lanka". In Elangayiyan, R., R. McKenna Brown, N. D. M. Ostler & M. K. Verma (eds.) "Vital Voices: Endangered Languages and Multilingualism. Proceedings of the FEL X Conference", Bath, Mysore: Foundation for Endangered Languages, Central Institute of Indian Languages. pp. 39–46.

Ansaldo, U. & S. Matthews (2007). "Deconstructing creole: The rationale". In Ansaldo et al. (2007), pp. 1–18.

Ansaldo, U. & S. Nordhoff (2006). "The role of typological pressure in contact-induced restructuring". Paper presented at the conference Syntax of the World's Languages II, Lancaster, UK.

Ansaldo, U. & S. Nordhoff (2009). "Complexity and the age of languages". In Aboh, E. O. & N. Smith (eds.) "Complex processes in new languages", Amsterdam: John Benjamins.

Ansaldo, U. & S. Nordhoff (submitted). "Typological dynamics in contact language formation. Congruence, frequency and metatypy in Sri Lanka Malay grammar".

Ansaldo, U., S. Matthews & L. Lim (eds.) (2007). *Deconstructing Creole*. Amsterdam: John Benjamins.

Apoussidou, D. & S. Nordhoff (2008). "Feet in Sri Lanka Malay – no stress please". Paper presented at OCP 5, January 2008, Toulouse.

Arden, A. H. (1934). *A progressive grammar of common Tamil*. Christian Literature Society, 4th edn. 1st edn. 1891.

Arends, J. (1989). *Syntactic developments in Sranan: creolization as a gradual process*. Ph.D. thesis, University of Nijmegen.

Arends, J., P. Muysken & N. Smith (eds.) (1995). *Pidgins and Creoles : an introduction*. Amsterdam: John Benjamins.

Arokianathan, S. (1981). *Tamil clitics*. Trivandrum: Dravidian Linguistics Associa-

tion.

ASHER, R. E. (1982). *Tamil*. Amsterdam: North-Holland Publishing Company.

ASHER, R. E. & E. ANNAMALAI (2002). *Colloquial Tamil - The complete course for beginners*. London, New York: Routledge.

BAKKER, P. (1995). "Over convergentie in een sprachbund". In HULS, E. & J. KLATTER-FOLMER (eds.) "Artikelen van de Tweede Sociolinguïstische conferentie", Delft: Eburon. pp. 37–51.

BAKKER, P. (1996). "Language intertwining and convergence: typological aspects of the genesis of mixed languages". *STUF*, 49(1):9–20.

BAKKER, P. (2000a). "Convergence intertwining: an alternative way towards the genesis of mixed languages". In Gilbers et al. (2000), pp. 29–35.

BAKKER, P. (2000b). "Rapid Language Change: creolization, intertwining, convergence". In Renfrew et al. (2000), pp. 586–620.

BAKKER, P. (2006). "The Sri Lanka *Sprachbund*: The newcomers Portuguese and Malay". In MATRAS, Y., A. McMAHON & N. VINCENT (eds.) "Linguistic Areas – Convergence in Historical and Typological Perspective", Houndmills, New York: Palgrave Macmillan. pp. 135–159.

BAKKER, P. & M. PARKVALL (2005). "Reduplication in Pidgins and Creoles". In Hurch & Mattes (2005), pp. 511–531.

BALASUBRAMANIAN, T. (1980). "Timing in Tamil". *Journal of Phonetics*, 8:449–467.

BAYER, J. (2004). "Non-nominative subjects in comparison". In Bhaskararao & Subbarao (2004a), pp. 49–76.

BEYTHAN, H. (1943). *Praktische Grammatik der Tamilsprache*. Wiesbaden: Harrassowitz.

BHASKARARAO, P. & K. V. SUBBARAO (eds.) (2004a). *Non-nominative subjects*, vol. I. Amsterdam and Philadelphia: John Benjamins.

BHASKARARAO, P. & K. V. SUBBARAO (eds.) (2004b). *Non-nominative Subjects*, vol. II. Amsterdam, Philadelphia: John Benjamins.

BICHSEL-STETTLER (1989). *Aspects of the Sri Lanka Malay Community and Its Language*. Master's thesis, Universität Bern.

BICKEL, B. (2004). "The syntax of experiencer verbs in the Himalayas". In Bhaskararao & Subbarao (2004a), pp. 77–112.

BIRDSONG, D. (1989). *Metalinguistic performance and interlinguistic competence*. New York: Springer.

BLOCH, J. (1934). *L'indo-aryen du Véda aux temps modernes*. Paris: Adrien-Maisonneuve.

BONTA, S. (2004). *Negombo Fisherman's Tamil: A Case of Contact-Induced Language Change from Sri Lanka*. Ph.D. thesis, Cornell University.

BONTA, S. (2008). "Negombo Fishermen's Tamil (NFT): A Sinhala Influenced Dialect from a Bilingual Sri Lankan Community". *International Journal of Dravidian Lin-*

guistics, XXXVII(2):133–140.

Booij, G., C. Lehmann, J. Mugdan & S. Skopeteas (eds.) (2004). *Morphologie – Ein internationales Handbuch zur Flexion und Wortbildung*, vol. 2. Berlin: Walter de Gruyter.

Botma, B. (2004). *Phonological Aspects of Nasality – An Element-Based Dependency Approach*. Ph.D. thesis, Universiteit van Amsterdam.

Burah, T. A. (2006). *Saga of The Exiled Royal Javanese Unearthed*. Dehiwala: Sri Devi.

Butt, M. (2003). "The Light Verb Jungle". *Harvard Working Papers in Linguistics*, 9:1–49.

Butt, M., T. H. King & G. Ramchand (eds.) (1994). *Theoretical Perspectives on Word Order in South Asian Languages*. CSLI.

Bybee, J., J. Haiman & S. A. Thompson (eds.) (1997). *Essays on Language Function and Language Type*. Amsterdam: John Benjamins.

Chafe, W. (1976). "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View". In Li (1976), pp. 25–55.

Chafe, W. (1987). "Cognitive constraints on information flow". In Tomlin (1987), pp. 21–52.

Chater, J. (1815). *A grammar of the Cingalese language*. Colombo: Government Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.

Chung, S.-F. (2005). "Kena as a third type of Malay passive". *Oceanic Linguistics*, 44(1):194–214.

Clements, G. N. (1990). "The Role of the Sonority Cycle in Core Syllabification". In Kingston, J. & M. E. Beckman (eds.) "Papers in Laboratory Phonology I", Cambridge: CUP. pp. 283–333.

Clements, G. N. (1999). "Affricates as noncontoured stops". In Fujimura, O., B. D. Joseph & B. Palek (eds.) "Proceedings of LP '98: item order in language and speech", Prague: Karolinum Press. pp. 271–299.

Codrington, H. W. (1926). *A Short History of Ceylon*. London: Macmillan.

Collins, J. T. (1992). "Preliminary notes on Berau Malay". In Martin, P. W. (ed.) "Shifting patterns of language use in Borneo", Williamsburg, Va: Department of Anthropology, College of William and Mary in Virginia. pp. 297–333.

Comrie, B. (1981). *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

Comrie, B. (1985). *Tense*. Cambridge: CUP.

Corbett, G. (2006). *Agreement*. Cambridge: CUP.

Croft, W. (2000). *Explaining language change : an evolutionary approach*. Harlow: Longman.

CROFT, W. (2001). *Radical Construction Grammar*. Oxford: OUP.

DE CASPARIS, J. G. (1961). "New Evidence on Cultural Relations between Java and Ceylon in Ancient Times". In SOPER, A. C. (ed.) "Artibus Asiae", Ascona: Artibus Asiae Publishers. pp. 241–248.

DE HAAN, F. (1935). *Oud Batavia*. Bandoeng: Nix.

DE SILVA, K. M. (1981). *A history of Sri Lanka*. London: Hurst.

DE SILVA, M. W. S. (1972). *Vedda Language of Ceylon. Texts and Lexicon*. München: R. Kitzinger.

DE SILVA, M. W. S. (1974). "Convergence in Diglossia: The Sinhalese situation". In Southwoth & Apte (1974), pp. 60–91.

DE SILVA JAYASURIYA, S. (2002). "A Unique Malay: Sri Lankan Malay Creole". *NUSA*, 50:43–57.

DELANCEY, S. (1997). "Grammaticalization and the gradience of categories: Relator nouns and postpositions in Tibetan and Burmese". In Bybee et al. (1997), pp. 51–69.

DEUMERT, A. & S. DURLEMAN (eds.) (2006). *Structure and Variation in Language Contact*. Amsterdam: John Benjamins.

DIK, S. (1980). *Studies in Functional Grammar*. London: Academic Press.

DIK, S. (1997). *The theory of functional grammar*. Berlin: Mouton de Gruyter.

DISSANAYAKA, J. B. (2003). *Say it in Sinhala*. Pannipitiya: Stamford Lake.

DIXON, R. M. W. (1979). "Ergativity". *Language*, 55(1):59–138.

DIXON, R. M. W. (1994). *Ergativity*. Cambridge: CUP.

DIXON, R. M. W. (1997). *The rise and fall of languages*. Cambridge: CUP.

DIXON, R. M. W. (2006). "Serial verb constructions – conspectus and coda". In Aikhenvald (2006a), pp. 338–350.

DONOHUE, M. & S. WICHMANN (eds.) (2008). *The typology of semantic alignment*. Oxford: OUP.

DRYER, M. (2006). "Descriptive theories, explanatory theories, and Basic Linguistic Theory". In Ameka et al. (2006), pp. 207–234.

DUPONT, P. (1959). "Les Buddha dits d'Amarâvatï en Asie du Sud-Est". *Bulletin de l'École Française d'Extrême-Orient*, XLIX(2):631–636.

DURIE, M. & M. ROSS (eds.) (1996). *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford: OUP.

EMENEAU, M. B. (1956). "India as a linguistic area". *Language*, 32:3–16.

EMENEAU, M. B. (1969). "Onomatopoetics in the Indian Linguistic Area". *Language*, 45(2):1274–99.

EMENEAU, M. B. (1971). "Dravidian and Indo-Aryan: the Indian linguistic area". In SJOBERG, A. F. (ed.) "Symposium on Dravidian Civilization", Austin: Center for Asian Studies, University of Texas at Austin. pp. 33–68.

EMENEAU, M. B. (1974). "The Indian linguistic area revisited". In Southwoth & Apte

(1974), pp. 92–134.

Ewing, M. C. (2005). "Colloquial Indonesian". In Adelaar & Himmelmann (2005), pp. 227–258.

Fairbanks, G. W., J. W. Gair & M. W. S. De Silva (1968). *Colloquial Sinhalese (Sinhala)*. Ithaca: South Asia Programm, Cornell University.

Fedson, V. J. (1993). "Complex verb-verb predicates in Tamil". In Verma (1993), pp. 63–75.

Foley, W. A. (1986). *The Papua Languages of New Guinea*. Cambridge: CUP.

Foley, W. A. & R. D. Van Valin (1984). *Functional syntax and universal grammar*. Cambridge: CUP.

Foley, W. A. & R. D. Van Valin (1985). "Information Packaging in the clause". In Shopen (1985a), pp. 282–364.

Foley, W. A., & M. Olson (1985). "Clausehood and verb serialization". In Nichols, J. & A. C. Woodbury (eds.) "Grammar inside and outside the clause", Cambridge: Cambridge University Press. pp. 17–60.

Gair, J. W. (1967). "Colloquial Sinhala Inflectional Categories and Parts of Speech". *Indian Linguistics*, 27:32–45. Reprinted in Gair (1998), pp. 13–25.

Gair, J. W. (1968). "Sinhalese Diglossia". *Anthropological Linguistics*, 10(8):1–15. Reprinted in Gair (1998), pp. 213–223.

Gair, J. W. (1971). "Action Involvement Categories in Colloquial Sinhala". In Zamorra, M., J. M. Mahar & H. Orenstein (eds.) "Themes in Culture: Essays in Honor of Morris E. Opler", Manila: Kayumanggi Publishers. pp. 238–256. Reprinted in Gair (1998), pp. 25–46.

Gair, J. W. (1976a). "Is Sinhala a Subject Language?" In Verma (1976), pp. 39–63.

Gair, J. W. (1976b). "The Verb in Sinhala, with some Preliminary remarks on Dravidianization". *International Journal on Dravidian Linguistics*, 10(8):259–273.

Gair, J. W. (1982). "Sinhala: an Indo-Aryan isolate". *South Asian Literary Association Journal*, 5(3):51–64. Reprinted in Gair (1998) pp. 1–12.

Gair, J. W. (1985a). "Sinhala Focused Sentences: Naturalization of a Calque". In Krishnamurti et al. (1985), pp. 147–164. Reprinted in Gair (1998), pp. 153–169.

Gair, J. W. (1985b). "Sinhalese Diglossia Revisited, or, Diglossia Dies Hard". In Krishnamurti et al. (1985), pp. 322–336. Reprinted in Gair (1998), pp. 224–136.

Gair, J. W. (1991a). "Discourse Deixis and Situational Deixis in Sinhala". In Balachandran, L. B. & B. R. Reddy (eds.) "Studies in Dravidian and General Linguistics: A Festschrift for Bh. Krishnamurti", Hyderabad: Osmania University. pp. 448–467. Reprinted in Gair (1998), pp. 111–126.

Gair, J. W. (1991b). "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. (1998). *Sinhala and Other South East Asian Languages*. New York,

Oxford: OUP.

GAIR, J. W. (1998 [1983]). "Nonconfigurationality, Movement and Sinhala Focus". In Gair (1998), pp. 50–64. Revised version of a paper presented at the Linguistics Association of Great Britain, Newcastle, September 1983.

GAIR, J. W. (2003). "Sinhala". In CARDONA, G. & D. JAIN (eds.) "The Indo-Aryan Languages", London: Routledge. pp. 766–818.

GAIR, J. W. & W. S. KARUNATILLAKE (1974). *Literary Sinhala*. Ithaca: Cornell University South Asia Program.

GAIR, J. W. & J. C. PAOLILLO (1988). "Sinhala non-verbal sentences and argument structure". *Cornell Working Papers in Linguistics*, 8:39–78. Reprinted in Gair (1998), pp. 87–110.

GAIR, J. W. & J. C. PAOLILLO (1997). *Sinhala*. München: LINCOM.

GAIR, J. W. & S. SUSEENDIRARAJAH (1981). "Aspects of the Jaffna Tamil Verbal System". *International Journal of Dravidian Linguistics*, 10(2):370–384. Partially reprinted in Gair (1998) pp.170–181.

GAIR, J. W., W. SUSEENDIRARAJAH & W. S. KARUNATILLAKE (1979). *Spoken [Sri Lanka Jaffna] Tamil*. s.l.: University of Sri Lanka Press.

GAIR, J. W., B. LUST, L. SUMANGALA & M. RODRIGO (1989). "Acquisition of Null Subjects and Control in Some Sinhala Adverbial Clauses". *Papers and Reports on Child Language Development*, 28:97–106. Reprinted in Gair (1998), pp. 271–285.

GARUSINGHE, D. (1962). *Sinhala, the Spoken Idiom*. München: Max Hueber.

GEIGER, W. (1938). *A Grammar of the Sinhalese Language*. Colombo: Royal Asiatic Society, Ceylon Branch.

GEIGER, W. (1973). *Kleine Schriften zur Indologie und Buddhismuskunde*, *Glasenapp-Stiftung, Veröffentlichungen*, vol. 6. Wiesbaden: Steiner.

GIL, D. (2004). "Riau Indonesian *sama*: Explorations in macrofunctionality". In Haspelmath (2004a), pp. 371–426.

GILBERS, D., J. NERBONNE & J. SCHAEKEN (eds.) (2000). *Languages in Contact*. Amsterdam: Rodopi.

GIVÓN, T. (1976). "Topic, pronoun and grammatical agreement". In Li (1976), pp. 151–188.

GIVÓN, T. (1994). "Irrealis and the subjunctive". *Studies in Language*, 18(2):265–338.

GIVÓN, T. (2001a). *Syntax: A Functional-Typological Introduction*, vol. I. Amsterdam: Benjamins.

GIVÓN, T. (2001b). *Syntax: A Functional-Typological Introduction*, vol. II. Amsterdam: Benjamins.

GOONETILEKE, H. A. I. (1970–1983). *A bibliography of Ceylon*. Zug: Inter Documentation Corp.

GREENBERG, J. H. (1963). "Some universals of grammar with particular reference to the order of meaningful elements". In GREENBERG, J. H. (ed.) "Universals of

language : report of a conference held at Dobbs Ferry, New York, April 13-15, 1961", Cambridge, MA: The M.I.T. Press. pp. 73–113.

GRIJNS, C. D. (1991). *Jakarta Malay*. Ph.D. thesis, Universiteit Leiden.

GUNASEKARA, D. M. (1891). *A comprehensive Grammar of the Sinhalese Language*. Sri Lanka Sahitya Mandalaya.

HALE, K. L. (1982). "Preliminary remarks on configurationality". In PUSTEJOVSKY, J. & P. SELLS (eds.) "Proceedings of the 12th annual meeting of the North-Eastern Linguistics Society", Amherst. pp. 86–96.

HALIM, A. (1974). *Intonation in relation to syntax in Bahasa Indonesia*. Jakarta: Djambatan.

HALLE, M. & J.-R. VERGNAUD (1980). "Three-dimensional phonology". *Journal of Linguistic Research*, 1:83–105.

HASPELMATH, M. (1997). *Indefinite Pronouns*. Oxford: OUP.

HASPELMATH, M. (ed.) (2004a). *Coordinating Constructions*. Amsterdam: John Benjamins.

HASPELMATH, M. (2004b). "Coordinating constructions: an overview". In Haspelmath (2004a), pp. 3–40.

HAWKINS, J. A. (1994). *A performance theory of order and consituency*. Cambridge: CUP.

HELLMANN-RAJANAYAGAM, D. (1990). "The politics of the Tamil past". In Spencer (1990), pp. 107–122.

HENADEERAGE, K. (2002). *Topics in Sinhala Syntax*. Ph.D. thesis, Australian National University.

HENGEVELD, K. (1992). *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.

HENGEVELD, K. (2004a). "Illocution, mood, and modality". In Booij et al. (2004), pp. 1190–1201.

HENGEVELD, K. (2004b). "State-of-affairs concepts". In Booij et al. (2004), pp. 1104–1111.

HENGEVELD, K. & L. MACKENZIE (2008). *Functional Discourse Grammar*. Oxford: Oxford University Press.

HERRING, S. (1994). "Afterthoughts, Antitopics, and Emphasis – The syntacticization of Postverbal Position in Tamil". In Butt et al. (1994), pp. 119–152.

HIMMELMANN, N. P. (2005). "The Austronesian languages of Asia and Madagascar: typological characteristics". In Adelaar & Himmelmann (2005), pp. 110–181.

HOLM, J. (1989). *Pidgins and Creoles*. Cambridge: CUP.

HOOK, P. E. (1974). *The Compound Verb in Hindi*. Ann Arbor: Center for South & Southeast Asian Studies, University of Michigan.

HOPPER, P. & E. TRAUGOTT (2003). *Grammaticalization*. Cambridge: CUP, 2nd edn.

HUALDE, J. I. (2005). *The Sounds of Spanish*. Cambridge: CUP.

Hurch, B. & V. Mattes (eds.) (2005). *Studies on reduplication*. Berlin: Mouton de Gruyter.

Hussainmiya, B. A. (1987). *Lost Cousins: The Malays of Sri Lanka*, *Occasional Paper*, vol. 2. Kebangsaan: Universiti Kebangsaan.

Hussainmiya, B. A. (1990). *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Universiti Kebangsaan.

Hussein, A. (2007). *Sarandib – An Ethnological Study of the Muslims of Sri Lanka*. Dehiwala: Asiff Hussein.

Jakobson, R. C., G. M. Fant & M. Halle (1952). "Preliminaries to speech analysis: The distinctive features and their correlates." Tech. Rep. 13, MIT Acoustics Laboratory.

Jayawardena-Moser, P. (1996). *Aufbauwortschatz Singhalesisch - Deutsch*. Wiesbaden: Harrassowitz, 3rd edn.

Jayawardena-Moser, P. (2004). *Grundwortschatz Singhalesisch - Deutsch*. Wiesbaden: Harrassowitz, 3rd edn.

Jespersen, O. (1904). *Lehrbuch der Phonetik*. Leipzig, Berlin: Teubner.

Kachru, Y. (1993). "Verb serialization in syntax, typology and historical change". In Verma (1993), pp. 115–133.

Kager, R. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: Foris.

Kandiah, T. (1996). "Syntactic "Deletion" in Lankan English: Learning from a new variety of English about —". In Baumgardner, R. J. (ed.) "South Asian English – Structure, Use, and Users", Delhi: OUP. pp. 104–126.

Karunatillake, W. S. (2004). *An Introduction to Spoken Sinhala*. Colombo: Gunasena, 3rd edn.

Keane, E. (2001). *Echo words in Tamil*. Ph.D. thesis, Merton College, Oxford.

Keenan, E. L. (1976). "Towards a Universal Definition of "Subject"". In Li (1976), pp. 303–334.

Kehrein, W. (2002). *Phonological Representation and Phonetic Phasing. Affricates and Laryngeals*. Tübingen: Niemeyer.

Keizer, E. (1992). *Reference, predication and (in)definiteness in functional grammar : a functional approach to English copular sentences*. Ph.D. thesis, Vrije Universiteit Amsterdam.

Kekulawala, S. L. (1982). "Kinship terminology in Sri Lankan Malay: A contribution to the study of language universals". *Kelaniya Journal of Humanities and Social Sciences of the University of Kelaniya*, 1(2):207–224.

Kiparsky, P. (1985). "Some consequences of lexical phonology". *Phonology Yearbook*, 2:83–138.

Klimov, G. A. (1974). "On the Character of Languages of Active Typology". *Linguistics: An Interdisciplinary Journal of the Language Sciences*, 131:11–25.

Krishnamurti, B., C. P. Masica & A. Sinha (eds.) (1985). *South Asian Languages: Structure, Convergence and Diglossia*. Delhi: Motilal Banarsidass.

Ladefoged, P. (1975). *A Course in Phonetics*. New York: Harcourt Brace.

Lambrecht, K. (1994). *Information structure and sentence form – Topic, focus, and the mental representations of discourse referents*. Cambridge: CUP.

Lassen, C. (1847). *Indische Alterthumskunde*. Bonn: Kittler/Williams & Norgate.

Lefebvre, C. (2001). "Relexification in Creole Genesis and Its Effects on the Development of the Creole". In Smith & Veenstra (2001), pp. 9–42.

Lehmann, C. (1984). *Der Relativsatz – Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Gunther Narr Verlag.

Lehmann, T. (1989). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

LePage, R. B. & A. Tabouret-Keller (1985). *Acts of identity : Creole-based approaches to language and ethnicity*. Cambridge: CUP.

Letterman, R. (1997). *The Effects of Word-Internal Prosody in Sinhala: A Constraint-Based Analysis*. Ph.D. thesis, Cornell University.

Levi, S. V. (2008). "Phonemic vs. derived glides". *Lingua*, 118:1956–1978.

Levinson, S. C. & N. Enfield (eds.) (2001). *Field Manual*. Nijmegen: Language & Cognition Group, Max Planck Institute for Psycholinguistics.

Li, C. N. (ed.) (1976). *Subject and topic*. New York, San Francisco, London: Academic Press.

Li, C. N. (1997). "On Zero Anaphora". In Bybee et al. (1997), pp. 275–300.

Lim, L. (2007). "Mergers and acquisitions: on the ages and origins of Singapore English particles". *World Englishes*, 26(4):446–473.

Lim, L. & U. Ansaldo (2006). "Keeping Kirinda vital: The endangerment-empowerment dilemma in the documentation of Sri Lanka Malay". *Amsterdam Centre for Language & Cognition Working Papers*, 1:51–66.

Lim, L. & U. Ansaldo (2007). "Identity alignment in the multilingual space: The Malays of Sri Lanka." In Anchimbe, E. (ed.) "Linguistic identity in multilingual postcolonial spaces", Cambridge: Cambridge Scholars Press. pp. 218–243.

Longacre, R. E. (1985). "Sentences as combinations of clauses". In Shopen (1985b), pp. 235–286.

Ludowyk, E. F. C. (1966). *The Modern History of Ceylon*. London: Weidenfeld and Nicholson.

Lyons, J. (1977). *Semantics*. Cambridge: CUP.

Malten, T. (1989). *Reduplizierte Verbstämme im Tamil*. Wiesbaden: Steiner Verlag.

Marga Institute, Sri Lanka Center for Development Studies (ed.) (1988). *The Muslim Community of Sri Lanka*. Colombo. Reprint of a study carried out in 1984/85 by Giom Kurukulasuriya for the Dr Sheikh Shams Al-Fassi Foundation of Sri Lanka. All data taken over from citations in Bichsel-Stettler (1989).

Masica, C. P. (1976). *Defining a Linguistic Area: South Asia*. Chicago: University of Chicago Press.

Masica, C. P. (1991). *The Indo-Aryan Languages*. Cambridge: CUP.

Matsumoto, Y. (1997). *Noun-modifying constructions in Japanese*. Amsterdam. Philadelphia: John Benjamins.

Matzel, K. (1983). *Einführung in die singhalesische Sprache*. Wiesbaden: Harrassowitz, 2nd edn.

McCarthy, J. J. & A. Prince (1993). "Prosodic Morphology – Constraint Interaction and Satisfaction". Tech. Rep. 3, Rutgers University Center for Cognitive Science.

Milroy, L. & M. Gordon (2003). *Sociolinguistics*. Malden: Blackwell Publishing.

Minegishi, M. (2004). "Southeast-Asian languages". In Bhaskararao & Subbarao (2004a), pp. 301–317.

Mithun, M. (1984). "The Evolution of Noun Incorporation". *Language*, 60:847–894.

Mohanan, K. P. (1982). "Grammatical Relations and Clause Structure in Malayalam". In Bresnan, J. (ed.) "The Mental Representation of Grammatical Relations", Cambridge: MIT Press. pp. 504–89.

Mufwene, S. S. (2001). *The Ecology of Language Evolution*. Cambridge: CUP.

Muysken, P. C. (1981). "Halfway between Quechua and Spanish: the case for relexification". In Highfield, A. & A. Valdman (eds.) "Historicity and variation in creole studies", Ann Arbor: Karoma. pp. 52–78.

Muysken, P. C. (2001). "The origin of creole languages: The perspective of second language learning". In Smith & Veenstra (2001), pp. 157–176.

Nadkarni, M. V. (1975). "Bilingualism and syntactic change in Konkani." *Language*, 51:672–83.

Nespor, M. & I. Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Neumann-Holzschuh, I. & E. W. Schneider (eds.) (2000). *Degrees of Restructuring in Creole Languages*. Amsterdam: John Benjamins.

Nichols, J. (1986). "Head-Marking and Dependent-Marking Grammar". *Language*, 62(1):56–119.

Nissan, E. & R. L. Stirrat (1990). "The generation of communal identities". In Spencer (1990), pp. 19–44.

Nitz, E. & S. Nordhoff (forthcoming). "Subtractive Plural Morphology in Sinhala". In Wohlgemuth, J. & M. Cysouw (eds.) "Rara & Rarissima. Collecting and interpreting unusual characteristics of human languages", Berlin: Mouton de Gruyter.

Noonan, M. (1985). "Complementation". In Shopen (1985b), pp. 42–140.

Noonan, M. (2006). "Grammar writing for a grammar-reading audience". *Studies in Language*, 30(2):351–365.

Nordhoff, S. (2008). "Electronic reference grammars for typology – challenges and solutions". *Journal for Language Documentation and Conservation*, 2(2):296–324.

Nuhman, M. A. (2007). *Sri Lankan Muslims – Ethnic Identity within Cultural Diver-*

sity. International Centre for Ethnic Studies: Colombo.

ODÉ, C. (1994). "On the perception of prominence in Indonesian". In ODÉ, C. & V. VAN HEUVEN (eds.) "Experimental Studies of Indonesian Prosody", Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Leiden University. pp. 27–107.

OSMAN, M. S. & M. A. SOURJAH (2008). "The Malays – an overview of a unique history & culture". In VIJALAKSHMI, E. (ed.) "Cultural minorities – their growth, achievements & relevance today", Colombo: International Centre for Ethnic Studies. 2nd edn., pp. 48–61.

PAAUW, S. H. (2003). "What is Bazaar Malay?" Paper presented at the Seventh International Symposium on Malay/Indonesian Linguistics, Nijmegen.

PAAUW, S. H. (2004). *A Historical Analysis of the Lexical Sources of Sri Lanka Malay*. Master's thesis, York University.

PAAUW, S. H. (2008). *The Malay contact varieties of Eastern Indonesia*. Ph.D. thesis, State University of New York at Buffalo.

PAOLILLO, J. C. (1992). *Functional Articulation in Diglossia: a Case Study of Grammatical and Social Correspondences in Sinhala*. Ph.D. thesis, Stanford University.

PAOLILLO, J. C. (1997). "Sinhala diglossic variation: continuous or discrete". *Language in Society*, 26(2):269–296.

PERCIVAL, R. (1803). *An account of the island of Ceylon*. London.

PIERIS, P. E. (1918). *Ceylon and the Hollanders, 1658-1796*. Tellippalai: American Ceylon Mission Press.

POEDJOESOEDARMO, S. (1982). *Javanese Influence on Indonesian*. Pacific Linguistics Series D-38, Canberra: The Australian National University.

POPE, G. U. (1867). *A Tamil Hand-Book: full introduction to the common dialect of that language*. Madras: Christian Knowledge Society's press.

POWELL, G. (1973). *The Kandyan Wars - The British Army in Ceylon 1803-1818*. London: Leo Cooper.

PRAY, B. (1970). *Topics in Hindi-Urdu Grammar*. Berkeley: Center for South and Southeast Asia Studies.

PRINCE, E. F. (1981). "Towards a taxonomy of given-new information". In COLE, P. (ed.) "Radical Pragmatics", New York, London: Academic Press. pp. 223–255.

RABEN, R. (2000). "Round about Batavia – Ethnicity and authority in the Ommelanden 1650–1800". In GRIJNS, K. (ed.) "Jakarta-Batavia : socio-cultural essays", Leiden: KITLV Press. pp. 93–109.

RAS, J. (1982). *Inleiding to het modern Javaans*. Den Haag: M. Nijhoff.

RENFREW, C., A. McMAHON & L. TRASK (eds.) (2000). *Time Depth in Historical Linguistics*, vol. 2. Cambridge: McDonald Institute for Archeological Research.

RIJKHOFF, J. N. M. (2002). *The Noun Phrase*. Oxford: OUP.

ROBERTS, S. J. (2004). *The emergence of Hawai'i Creole English in the early 20th century: the sociohistorical context of creole genesis*. Ph.D. thesis, Stanford Uni-

versity.

ROBSON, S. (1992). *Javanese grammar for students*. Clayton, Victoria: Centre of Southeast Asian Studies, Monash University.

ROBUCHON, G. (2003). "Malayo Language in Sri Lanka." Paper presented at the 7th International Symposium on the Malay/Indonesian Linguistics in Berg en Dal.

ROGERS, J. D. (1990). "Historical images in the British period". In Spencer (1990), pp. 87–106.

ROSS, M. (1996). "Contact-induced change and the comparative method: cases from Papua New Guinea". In Durie & Ross (1996), pp. 180–217.

ROSS, M. (2007). "Calquing and metatypy". *Journal of Language Contact*, 1:117–143.

ROSS, M. D. (1997). "Social networks and kinds of speech-community event". In BLENCH, R. & M. SPRIGGS (eds.) "Archaeology and language 1: Theoretical and methodological orientations", London: Routledge. pp. 209–261.

ROSS, M. D. (2001). "Contact-induced change in Oceanic languages in North-West Melanesia". In AIKHENVALD, A. Y. & R. M. W. DIXON (eds.) "Areal diffusion and genetic inheritance", Oxford: Oxford University Press. pp. 134–166.

SAID, H. M. (1926). "Ceylon Malays". *Journal of the Royal Asiatic Society (Malayan Branch)*, 4(2):266–268.

SALDIN, B. (2001). *The Sri Lankan Malays and Their Language*. Kurunegala: Nihon, 2nd edn.

SALDIN, B. (2003). *Portrait of a Sri Lankan Malay*. Kurunegala: Nihon.

SALDIN, B. (2007). *A concise Sri Lankan Malay Dictionary – Kamus ringkas bahasa melayu Sri Lanka*. Kurunegala: Nihon.

SAMSURI (1971). *Ciri-ciri prosodi dalam kalimat Bahasa Indonesia*. Flores: Nusa Indah.

SANNASGALA, P. (1976). *A study of Sinhala vocables of Dutch origin*. Colombo: The Netherlands-Alumni Association of Sri Lanka.

SCHACHTER, P. (1996). *The subject in Tagalog : still none of the above*. Los Angeles: University of Los Angeles.

SCHIFFMAN, H. (1999). *A Reference Grammar of Spoken Tamil*. Cambridge: CUP.

SCHROEDER, C. (1999). *The Turkish Noun Phrase in Spoken Discourse*. Wiesbaden: Harrassowitz.

SCHÜTZE, C. (1996). *The empirical base of linguistics – Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.

SCHWEITZER, C. (1931[1680]). *Reise nach Java und Ceylon, 1675–1682*. Haag: Nijhoff. Originally published by Nijhoff in Haag in 1680.

SELBACH, R. (2007). "Lingua Franca of the Mediterranean (1350-1830): finding Nessie?" In URELAND, S., A. LODGE & S. PUGH (eds.) "Language Contact and Minority Languages on the Littorals of Europe", Berlin: Logos. pp. 149–160.

SELBACH, R. (in preparation). *History and Structure of the Lingua Franca of the*

Mediterranean (1500-1830). Ph.D. thesis, University of Amsterdam.

Selkirk, E. (1984). "On the major class features and syllable theory". In Aranoff, M. & R. T. Oehrle (eds.) "Language sound structure: studies in phonology presented to Morris Halle by his teacher and students", Cambridge, Massachusetts: The MIT Press. pp. 107–136.

Selvarajagopal, K. D. (1979). *tamiZHk kilLaimoZHi oliyaniyal Ayivu maTTakkaLappukkilai moZHi (Research on Tamil dialect phonemics: The Batticaloa Dialect)*. Master's thesis, Vidyalankara Campus of the University of Ceylon. Date and author are not indicated on the copy but had to be inferred. The copy is located at the Institut für Indologie und Tamilistik in Cologne.

Shanmugam Pillai, M. (ca. 1986). "Tamil Dialectology". Unpublished manuscript. Copy available at the Institute for Indian and Tamil Studies, University of Cologne.

Shibasaki, R. (2005). "On the grammaticalization of verbal reduplication in Japanese". In Hurch & Mattes (2005), pp. 283–313.

Shopen, T. (1985a). *Language typology and syntactic description*, vol. I. Cambridge: CUP.

Shopen, T. (1985b). *Language typology and syntactic description*, vol. II. Cambridge: CUP.

Silverstein, M. (1976). "Hierarchy of features and ergativity". In Dixon, R. (ed.) "Grammatical Categories in Australian Languages", New Jersey: Humanities Press. pp. 112–171.

Singh, R. (2005). "Reduplication in Modern Hindi and the theory of reduplication". In Hurch & Mattes (2005), pp. 263–281.

Slomanson, P. (2003). "Sri Lankan Malay Grammars: Lankan or Malay?" Revised Version of a paper presented at The 7th International Symposium on Malay/Indonesian Linguistics, Nijmegen June 2003 and the SPCL Summer meeting 2003 University of Hawaii, August 2003.

Slomanson, P. (2006). "Sri Lankan Malay morphosyntax: Lankan or Malay?" In Deumert & Durleman (2006), pp. 135–158.

Slomanson, P. (2008a). "The perfect construction and complexity drift in Sri Lankan Malay". *Lingua*, 118(10):1640–1655.

Slomanson, P. (2008b). "Progressive Aspect in Sri Lankan Malay and the Syntactic Status of *ambe(l)*". Paper presented at ISMIL 12, 26-27 June 2008 University of Leiden.

Smith, I. (1979). "Convergence in Sri Lanka: A Creole Example". *Lingua*, 48:193–222.

Smith, I. R. (2003). "Running Head: Substrate Influence in Sri Lankan Malay". Paper published at 23rd South Asian Languages Analysis.

Smith, I. R. & S. Paauw (2006). "Sri Lanka Malay: Creole or convert". In Deumert & Durleman (2006), pp. 159–182.

Smith, I. R., S. Paauw & B. A. Hussainmiya (2004). "Sri Lanka Malay: the state of

the art." *Yearbook of South Asian Languages*, 2004:197–215.

Sмith, N. & T. Veenstra (eds.) (2001). *Creolization and Contact*. Amsterdam: John Benjamins.

Southwoth, F. C. & M. L. Apte (eds.) (1974). *Contact and Convergence in South Asian Languages*. Trivandrum: International Journal of Dravidian Linguistics.

Spencer, J. (ed.) (1990). *Sri Lanka – History and Roots of the Conflict*. London: Routledge.

Sridhar, S. N. (1976a). "Dative Subjects". In Mufwene, S. S., C. A. Walker & S. B. Steever (eds.) "Papers from the twelfth regional meeting", Chicago: Chicago Linguistic Society. pp. 582–593.

Sridhar, S. N. (1976b). "Dative subjects, rule government, and relational grammar". *Studies in the Linguistic Sciences*, 6(1):130–151.

Sridhar, S. N. (1979). "Dative Subjects and the Notion of Subjects." *Lingua*, 49:1–28.

Starosta, S. (1985). "Relator nouns as a source of case inflection". In Acson, V. & R. Leed (eds.) "For Gordon Fairbanks", Honolulu: University of Hawai'i Press. pp. 111–133.

Steinhauer, H. (ed.) (1991). *Papers in Austronesian linguistics*, *Pacific Linguistics*, vol. 1. The Australian National University.

Steinhauer, H. (2005). "Colonial History and Language Policy in Insular South East Asia and Madagascar". In Adelaar & Himmelmann (2005), pp. 65–86.

Suseendirarajah, S. (1967). *A descriptive Study of Ceylon Tamil (with Special Reference to Jaffna Tamil)*. Ph.D. thesis, Annamali University.

Suseendirarajah, S. (1973a). "Phonology of Sri Lanka Tamil and Indian Tamil Contrasted". *Indian Linguistics*, 34(3):171–179.

Suseendirarajah, S. (1973b). "A study of pronouns in Batticaloa Tamil". *Anthropological Linguistics*, 15(4):172–182.

Tapovanaye, S. (1986). *Analysis of the segmental phonemes of the Sri Lanka Malay language*. Master's thesis, Univeristy of Kelaniya. Written in Sinhala.

Tapovanaye, S. (1995). *Vowel Length in Sri Lankan Malay*. Master's thesis, University of Hawaii.

Teeuw, A. (1984). *Leerboek Bahasa Indonesia*. Groningen: Wolters-Noordhoff, 4th edn.

Tennent, J. E. (1859). *Ceylon, an Account of the Island: Physical, Historical and Topographical*. London.

Thaliph, M. F. (2005). *Book on Malay Grammar (Nahu) – A guide to write and speak grammatical Malay*. Wattala: Terang.

Thomason, S. G. & T. Kaufman (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: Univeristy of California Press.

Thompson, S. (1985). "Adverbial clauses". In Shopen (1985b), pp. 171–234.

Tomlin, R. S. (1987). *Coherence and grounding in discourse*. Amsterdam: John

Benjamins.

TRUBETZKOY, N. S. (1939). *Grundzüge der Phonologie. (Travaux du Cercle Linguistique de Prague 7)*. Prague.

TSUNODA, T. (2004). "Issues in case marking". In Bhaskararao & Subbarao (2004b), pp. 197–208.

VAN VALIN, R. D. & W. A. FOLEY (1997). *Syntax – Structure, meaning and function*. Cambridge: Cambridge University Press.

VAN ZANTEN, E., R. GOEDEMANS & J. PACILLY (2003). "The status of word stress in Indonesian". In VAN DE WEIJER, J., V. J. VAN HEUVEN & H. VAN DER HULST (eds.) "The Phonological Spectrum. Volume II: Suprasegmental structure", Amsterdam, Philadelphia: John Benjamins. pp. 151–175.

VERMA, M. (ed.) (1976). *The Subject in South Asian Languages*. Madison: University of Wisconsin South Asia Program.

VERMA, M. (ed.) (1993). *Complex Predicates in South Asian Languages*. Delhi: Motilal Banarsidass.

VERMA, M. K. & K. P. MOHANAN (eds.) (1990). *Experiencer subjects in South Asian languages*. Stanford: Center for the Study of Language and Information.

VLEKKE, B. H. M. (1943). *Nusantara : a history of the East Indian Archipelago*. Cambridge: Harvard University Press.

WEINREICH, U. (1953). *Languages in Contact*. London: Mouton.

WIERINGA, E. (1997). "A Serat Ambiya in the Hendrik Kraemer Institute at Oegstgeest". *Caraka*, 27–28. Available at http://www.let.leidenuniv.nl/caraka/27-28/research_report.htm.

WIESE, R. (1986). *Silbische und Lexikalische Phonologie. Studien zum Chinesischen und Deutschen*. Tübingen: Niemeyer Verlag.

WITTGENSTEIN, L. (1953). *Philosophische Untersuchungen*. Oxford: Blackwell.

WURM, S. A., P. MÜHLHÄUSLER & D. T. TRYON (eds.) (1996). *Atlas of Languages of Intercultural Communication*. Walter de Gruyter.

ZUBKOVA, L. G. (1966). *Vokalizm Indonezijskogo jazyka*. Ph.D. thesis, Leningrad University.

ZVELEBIL, K. (1960). "Dialects of Tamil III". *Archiv Orientalni*, 28.

ZVELEBIL, K. V. (1966). "Some Features of Ceylon Tamil". *Indo-Iranian Journal*, 9(2):113–138.

ZWICKY, A. M. (1977). *On clitics*. Bloomington: Indiana University Linguistics Club.

ZWICKY, A. M. & G. K. PULLUM (1983). "Cliticization vs. Inflection: English *n't* ". *Language*, 59:502–513.

Texts

This chapter contains five texts. The first and the last texts were provided in writing, while texts two to four are transcriptions of recordings. The texts are orded according to difficulty.

The first text, *How Andara Ate Sugar* is a short story of a jester, Andare, who the king wants to fool, but in the end it is the jester who fools the king. Andare is a character of Sinhalese folklore. This text was provided by Mr Izvan Salim.

The second text was told by Gnei Binthan Muumin, a senior English teacher, and relates the story of her life as a Malay. This text is given in both phonetic and orthographic transcription. Intonation is indicated as well.

The third text covers the history of the Sri Lankan Malays. It is told by Tony Salim, who belongs to the same household as Gnei Binthan Muumin, and is the father or Izvan Salim. It shows more oral features than the preceding ones. All the speakers mentioned above are from Kandy and are native speakers of Sri Lanka Malay. All have command of Sinhala, English and Tamil.

The fourth text was told by Mr Sherifdeen from Mawanella and is about the 'Sri Lankan Robin Hood', *Saradiyel*, who was captured by a Sri Lankan Malay. This text is very oral in nature, and a bit difficult to follow in the written form. Mr Sherifdeen is a retired soldier and speaks Sri Lanka Malay, Sinhala, Tamil, and English.

The fifth text is a translation of the Brothers Grimm's tale *Snow-White and Rose-Red*. The translation was made by Izvan Salim. It is the most complex text, but might show some features more typical of European writing than the other texts.

Some parts of the transcriptions are not included in the text here. This is the case for questions asked by the researcher, other people intervening, or unrelated content (e.g. offering tea).

This text collection was edited in collaboration with Tony Salim and Izvan Salim. I am very grateful for their help, commitment and patience.

How Andare Ate Sugar

This text is a translation of a Sinhalese tale about the jester Andare. Andare being a jester, he has the habit of fooling people. The king in turn wants to fool Andare this time, but in the end, it is again Andare who prevails. This text is clearly of written modality, with full sentences, clear reference tracking, little dropping of referents, no code-switching etc.

(1) *ANDARE GUULA ANÀ-MAAKANG MOSTHOR.*
 Andare sugar PAST-eat manner
 'The way how Andare ate sugar.'

(2) *Hathu haari Andare=le aanak=le mliiga=nang anà-pii.*
 INDEF day Andare=ADDIT child=ADDIT palace=DAT PAST-go
 'One day, Andare and his son went to the castle.'

(3) *Incayang mliiga=nang kapang-pii, Raaja hathu thiikar=ka guula asà-siibar*
 3s.POLITE palace=DAT when-go king INDEF mat=LOC sugar CP-spread
 mà-kìrring simpang su-aada.[1]
 INF-dry keep PAST-exist
 'When he was going to the palace, the King had sprinkled sugar in a mat and had left it to dry.'

(4) *Andare raaja=ka su-caanya inni mà-kìrring simpang aada aapa=yang*
 Andare king=LOC PAST-ask DEM.PROX INF-dry keep exist what=ACC
 katha.
 QUOT
 'Andare inquired from the King what was it that was left [on the mat] to dry.'

(5) *Andare=yang mà-enco-king=nang raaja su-biilang itthu paasir katha*
 Andare=ACC INF-fooled-CAUS=DAT king PAST-say DEM.DIST sand QUOT
 'To tease Andare the King said it was sand.'

---

[1]The main verb in this string of four verbs is *simpang* 'keep'. *Aada* is used to put *simpang* in the perfect tense. *Simpang* has a preceding purposive clause *màkiiring* 'so that it dries'. The three mentioned verbs form a clause, which is preceded by a conjunctive participle clause which describes the previous action, sprinkling sugar on a mat.

(6) *Giithu   asà-biilang=apa raaja Andare=yang su-enco-king.*
   like.that cp-say=after   king Andare=acc past-fooled-caus
   'The King teased Andare by speaking like that.'

(7) *Thapi=le[2] Andare thàrà-jaadi   enco.*
   But=addit Andare neg.past-become fooled
   'But Andare was not fooled.'

(8) *Andare [[bale-king=apa[3]] raaja=yang mà-enco-king=nang] su-iingath*
   Andare turn-caus=after   king=acc inf-fooled-caus=dat past-think
   'Andare thought that he should fool the king instead.'

(9) *Andare aanak[4]=nang su-biilang:*
   Andare child=dat   past-say
   'Then, Andare told his son .'

(10) *Aanak, [[[lu=ppe   umma su-maathi][5] katha bitharak=apa][6]*
    child   2s.familiar=poss mother past-die   quot scream=after
    asà-naangis][6] mari!
    cp-cry   come.imp
    'Son, cry loudly and come towards me saying that your mother died!'

---

[2]Note the use of *two* adversative markers here, *thapi* and *=le*.

[3]The conjunctive participle of *baleking* 'return(tr.)' is used in the meaning of 'instead' here. One could translate in a more literal way: 'Andare thought that he should take his turn and fool the king.'

[4]The child had not been introduced before, but it is inferred from context that Andara must be the father. Furthermore, one would infer that we are dealing with a boy child, the sex of a girl would probably be indicated, by *pompang*.

[5]The reported string 'luppe umma su maathi' still has the second person pronoun *luu* instead of the first person pronoun *see*, which would be expected if we were dealing with direct speech.

[6]This clause chain is formed by first *=apa* and then *asà-*.

(11) *Suda Andare=pe   aanak=le   baapa anà-biilang=kee=jo asà-naangis*
      thus Andare=POSS child=ADDIT father PAST-say=SIMIL=EMPH cry
      ambel[7] su-dhaathang
      take   PAST-come
      'So, Andare's son also did exactly what he was asked to do and came crying[8]
      (to the king).'

(12) *Itthu=nang   blaakang Andare guula thiikar=ka asà-lompath*
      DEM.DIST=DAT after   Andare sugar mat=LOC   CP-jump
      'And then, Andare jumped on to the mat of sugar and'

(13) *lu=ppe   muuluth=ka=le   paasir, se=ppe muuluth=ka=le*
      2S.FAMILIAR=POSS mouth=LOC=ADDIT sand   1s=POSS mouth=LOC=ADDIT
      paasir[9] katha biilang˜biilang baaye=nang baapa=le   aanak=le   guula
      sand   QUOT say˜RED   good=DAT   father=ADDIT child=ADDIT sugar
      su-maakang
      PAST-eat
      'crying out "your mouth is filled with sand, my mouth is also filled with sand",
      repeating this over and over again they both ate sugar well.'

(14) *Suda kanabisan=ka, raaja Andare=yang mà-enco-king   asà-pii, raaja=jo*
      thus last=LOC   king Andare=ACC   INF-fooled-CAUS CP-go king=EMPH
      su-jaadi   enco
      PAST-become fooled
      'Ultimately what happened, the King set out to fool Andare, but the king was
      fooled himself.'

---

[7]The vector verb *ambel* indicates either the beginning of the crying here, or the beneficial nature of the
crying, which will become clear later.
      [8]A more faithful translation would be 'cried and came', with a perfect participle instead of the present
participle 'crying'.
      [9]Dead people will have sand in their mouth. To express his grief, Andare wants to have sand in his
mouth to. This is of course a trick to take profit of the king's assertion that the substance on the mat was
sand. Andare can now put the substance in his mouth, and the king can say nothing against it since he
himself had asserted that it was sand, rather than sugar.

Binthan Muumin's story

This text is about the family of the speaker and the very recent history of the Malays of Peradeniya Road in Kandy. In former times, there were many Malays in this area, but as it develops commercially, the resident Malays sell their property to buy cheaper places in rural areas. This text is given in phonetic transcription with rough indications of intonation, next to the normal orthographic rendering. Especially interesting are the different realizations of schwa, as well as the reduction of some other vowels. In the consonantal domain, the variation in geminated palatal stops can also be seen. Dots indicate parts that are not transcribed, normally the researcher asking a question in order to keep the narrative going.

(1)   H   L
kiṭaŋ   kɛnḍikaɟo
kithang Kandy=ka=jo
1PL   Kandy=LOC=foc
'We are from Kandy.'

(2)   H   L
sɛpːɛ   ʊmːaːbaːpa
se=ppe umma-baapa
1s=POSS mother-father
'As for my parents,'

(3)   H   L H   L H   L
sɛppe   baːpa nːigʌmbɔ, nigʌːmbo guːmbu
se=ppe baapa Negombo Negombo *Guum̌bu*
1s=POSS father Negombo Negombo Negombo
'my father is from Negombo.'

  [...]

(4) L H   L
m̩ma   kʰanḍi
umma Kandi
mother Kandy
'My mother is from Kandy.'

(5)   H   L   H   L
askaːʊŋ   baːpa pɪkɪɟan anagərɩɟ͡ʐa pəliːska
asà-kaaving baapa pukurjan anà-kijja police=*ka*
CP-marry   father work   PAST-make police=loc
'After having married, father worked in the police.'

(6)   H   L   H   LH   H L

pɔliːska pəkʊjan asgɪɟ͡za sːiːnijoː asaɖuːɖu aːnap paɖa
police=*ka pukurjan asà-kijja siini=jo* *asà-duuduk aanak pada*
police=LOC work CP-make here=EMPH CP-stay child PL

   H L
sɐbɾaːnaʔ
su-braanak
PAST-birth
'Having worked in the police and having come to live right here, the children were born.'

(7)   H L
aːnak paːɖa əsbəsːaɾ.
aanak pada asà-bìssar[1]
child PL CP-big
'The children grew up.'

(8) H   L
sːkulnaŋ anəpiː
skuul=nang anà-pi
school=DAT PAST-go
'They went to school.'

(9) H   L   H
skulnan anəpiːnam bla:kɐŋ
skuul=nang anà-pi=nang blaakang
school=DAT PAST-go=DAT after
'After having gone to school'

(10)   H   L
kaːɾaŋ kit̪anːɐ inːi nːi nigɪɾi suˇ su suuk su-bijaːsa
karang kithang inni inni nigiri ... ... su-byaasa
now 1PL DEM.PROX DEM.PROX country ... ... PAST-accustomed
'we have got accustomed to this country now.'

(11) H   L
luʋaɾ nigɪɾi kit̪anːə məpiː t̪əɾɐsuːka
luvar nigiri kithang=nang mà-pi thàrà-suuka
outside country 1PL=DAT NON.PAST-go NEG-like
'We do not want to go abroad.'

---

[1]In this sentence, the conjunctive participle *asà-* is used in a main clause, as indicated by the intonation. This sentence and the following thus do *not* form a clause chain. While the use of *asà-* is possible in main clauses, it is not very common; the use in cosubordinate clauses is more common.

(12)

                    H

karẽ   inːi   nigiɾikaɟɔ   kit̪ampɛ   aːnak buːʋa paɖajaŋ
karang inni   nigiri=ka=jo   kithang=pe aanak buuva[2] pada=yang
now   DEM.PROX country=LOC=EMPH 1PL=POSS   child   fruit   PL=ACC

    H   L

asːumpaŋ
asà-simpang
CP-keep

'It is in this country that we raised our children'

(13)   H   L H LH

inːi   skuːls paɖɛnaŋ aʃkiːɾiŋ
inni   skuuls pada=nang asà-kiiring
DEM.PROX schools PL=DAT   CP-send

'Having send them to these schools '

(14)   H   LH H   L H H   L   H

samːa asgəɾɪɟ͡za kaɾan sʊblaːɟaɾ   pəkʊjan asgəɾɪɟa ambɛ skaːɾan
samma asà-kijja   karang asà-blaajar pukurjan asà-kijja ambel skaarang
all   CP-make   now   CP-learn   work   CP-make take   now

          L   H   L

siːniɟoː   aɾəduːɖuʔ
siini=jo   arà-duuduk
here=EMPH NON.PAST-stay

'Having done all that now, having learned, having taken up work, we are now
living right here.'[3]

(15)

                        H   L

kaɾaŋ inːi   kit̪ampɛ nigiri sʊɟaːd̪i
karang inni   kithang=pe nigiri su-jaadi
now   DEM.PROX 1PL=POSS   country PAST-become

'Now this has become our country.'

---

[2]*aanak buuva* is an idiom to refer to one's children.

[3]The agents of the different events (working, learning, living) change reference. Sometimes the agent includes the speaker, sometimes not. The change of participants between elements of a clause chain is normal, see ⊡→13.3, p. 529. This is especially clear from the events of 'sending to school' (13) and 'learning' (14), which clearly have different agents. The person who sends is not the one who is learning, it is the person being sent who learns.

(16)   H   H   L
se:lõŋlɛ   kiṭampɛ   mla:junal:ɛ   haṭu pagijan anɐa:ḍa
Seelon=le   [[kithang=pe mlaayu=nang=le hatthu bagiyan anà-aada]
Ceylon=ADDIT 1PL=POSS   Malay=DAT=ADDIT INDEF part exist

          H   L
haṭu nigiri suɟa:ḍiʔ
hatthu nigiri]⁴ su-jaadi
INDEF country here=EMPH NON.PAST-stay
'Ceylon also became a country where our Malays also have a part in.'

  […]

(17)   H   L H   L
sam:a ṭumpaṭ sɛ:lɔŋ, sam:a ṭumpaṭ m̥la:ju a:ḍa
samma thumpath Seelon samma thumpath mlaayu aada
all   place   Ceylon all   place   Malay exist
'All over Ceylon there are Malays.'

(18)   H L   H   LH
kɐṭa:ma kiṭaŋ kic͡cɪl mu:siŋka ini   pe:raḍenije ɟa:laŋka
kàthaama kithang kiccil muusing=ka ini   Peradeniya jaalang=ka
before   1PL   small period=LOC DEM.PROX Peradeniya street=LOC

    L   H   L
sama   na:ḍa   m̥la:ju
samma   anà-aada mlaayu
all   PAST-exist Malay
'In former times, when we were small, all the people who lived on the Per-
adeniya Road were Malays.'

(19)   H   L H L   H   L
ba:ẽ m̥la:ju baṭai   m̥la:ju
baae mlaayu bathai   mlaayu
good Malay   courageous Malay
'Good Malays, brave Malays.'

(20)   H   L
baˤ:e m̥la:ju paḍɐ
bae   mlaayu pada
good Malay PL
'Good Malays.'

---

⁴Note the preposed relative clause, which is separated from the head word by the indefiniteness marker
hatthu.

(21)   H   LHL   H L   H L

n̪ːi   ml̥aːju pad̪ɛ ãɾi məpiː məpiː
inni   *mlaayu pada aari mà-pii mà-pii*
DEM.PROX Malay PL   day INF-go INF-go
'As days passed by the Malays'

   [...]

(22) H   L H   L   H   L H L H L H L H L

t̪umpad̪ ɡaɡasaːʋã -- ɡaɡasaːʋa, ɡaːɡa, land, saːʋa, fields, their
thumpath gaga-saava – gaga-saava, ga:ga, land, sa:va, fields, their
place   land-field – land-field land land field field their

    H   L
property --
property –
property[5] –
'places, fields and lands, their property'

(23)   H   L

inːi   samːajaŋ   aslʊpːas t̪aːŋaŋ
inni   *samma=yang asà-luppas thaangang*[6]
DEM.PROX all=ACC   CP-leave hand/arm
'gave up all that.'

(24)   H   L

ʋud̪ik   subəlanɛ̃   səpiː
uudik   *subàla=nang su-pii*
countryside side=DAT   PAST-go
'and went to the village areas.'

(25)   H   L

inːi   sʋaːra liːʋat̪
inni   *svaara liivath*
DEM.PROX noise more
'The noise here is too much.'

---

[5]This is some metalinguistic commentary on the word used.
[6]*luppas thaangang* is an idiom meaning 'to give up/to let go'.

(26)   H   L H   L   H L H

it̪ːusʊbːat̪   d̪əɽam pada ɟ͡zaːlaŋ aɾəgɪɽːa   but̪ːul ɾuːma padajaŋ

itthusubbath derang pada jaalang arà-kijja   butthul ruuma pada=yang

therefore   3PL   PL   street NON.PAST-make correct house PL=ACC

              L

aɾəpicakã̃ŋ

arà-picca-king

NON.PAST-broken-CAUS

'Therefore, they are repairing the roads, they are demolishing the houses.'

(27) H L   H L   H   L

ɾuːma aɾəkic͡cil   it̪usubːat̪   ini   t̪ũmpat̪ padɐjɐm paːsaɾ

ruuma arà-kiccil;   itthusubbath inni   thumpath pada=yang paasar

house NON.PAST-small therefore   DEM.PROX place   PL=ACC   shop

padə̃naŋ   ...

pada=nang ...[7]

PL=DAT   ...

'The houses were becoming small, and therefore these places were (turned) into shops.'

(28) H   L   H

ɔːbat̪   paːsaɾ padə̃nən dispɛnsɛɾinaŋ

oobath   paasar pada=nang dispensary=nang

medicine shop PL=DAT   dispensaries=DAT

'Dispensaries.'

  [...]

(29)   L   H   L H   H L H L   H   L

askasi   t̪riːma d̪ɛɾɐn samːa d͡ʒaːo uːd̪iʔ   vilɛɟ͡z æɾija padɐɲã̃

asà-kaasi thriima derang samma jaau uudik   village area padanang

CP-give thanks 3PL   all   far   countryside village area PL=DAT

      HL

səpiː

supii

PAST-go

'After handing over (the property), they all went to the far away village areas.'

---

[7]The verb is missing in this sentence because the speaker switches to explaining the kinds of shops established before finishing this sentence.

Tony Salim's story

This text is about the earlier history of the Malays, from the very beginning to the present day, including some demographic history and the family history of the speaker, Tony Salim. There is some information about the different generations of Malays, but it is unclear, how the word 'generation' is used, and whether there are five, eight, or yet a different number of generations. Finally, there is some material on the Malay saints, in Sri Lanka. This text is shows some features of spoken language, like repetitions, sentence breaks, low coherence, lots of discourse markers, etc., but few of the reductions found in other texts (dropping of prefixes and clitics for instance), which distinguishes it from the next text, *Saradiyel*.

(1) *Seelon=dika karang ka-umpath* generation.
Ceylon=LOC now ORD-four generation
'Now it's the fourth generation in Ceylon.'

(2) *Ka-umpath* generation *katha arà-biilang se=dang duppang lai thiiga*
ORD-four generation QUOT PAST-say 1s=DAT front more three
generation *pada ini nigiri=ka anà-duuduk.*
generation PL DEM.PROX country=LOC PAST-exist.ANIM
'The fourth generation is to say that before me there were three generations that existed in this country.'

(3) *Se=ppe oorang thuuva pada anà-biilang kithang pada* Malaysia=*dering*
1s=POSS man old PL PAST-say 1PL PL Malaysia=ABL
anà-dhaathang katha.
PAST-come QUOT
'My elders said that we had come from Malaysia.'

(4) *Itthu anà-dhaathang, kithang=pe oorang thuuva pada, samma*
DEM.DIST PAST-come 1PL=POSS man old PL all
Seelong=ka see[1] *arà-duuduk raaja pada=nang mà-banthu=nang.*
Ceylon=LOC 1s??? NON.PAST-exist.ANIM king PL=DAT INF-help=DAT
'Those who came, our forefathers were all in Ceylon to help the kings.'

---

[1]The use of *see* 'I' here seems to be a performance error. The speaker was born after the last king had already been dead for over a century.

(5) *Derang punnu=le anà-dhaathang mà-pìrrang=nang.*
3PL   full=ADDIT PAST-come   INF-wage.war=DAT
'Most of them came to wage war.'

(6) *Spaaru oorang pada arà-biilang kithang=pe oorang thuuva pada=yang*
some man   PL   NON.PAST-say 1PL=POSS man old   PL=ACC
Dutch *Seelong=nang anà-aaji.baa*[2] *katha.*
Dutch Ceylon=DAT   PAST-bring.ANIM quot
'Some people say our forefathers were brought to Ceylon by the Dutch.'

(7) *Spaaru oorang pada arà-biilang Seelong=nang English anà-aaji.baa*
some man   PL   NON.PAST-say Ceylon=DAT   English PAST-bring.ANIM
katha.
QUOT
'Some people say it is the English who brought them to Sri Lanka.'

(8) *Itthu muusing Seelong=ka Dutch=pe=le English=pe=le*
DEM.DIST time   Ceylon=LOC Dutch=POSS=ADDIT English=POSS=ADDIT
raame su-aada.
trouble PAST-exist
'During that period there was trouble between the Dutch and the English.'

(9) *Itthu raame anà-aada vakthu, Kandy=pe raaja=nang kithang=pe*
DEM.DIST trouble PAST-exist time   Kandy=POSS king=DAT   1PL=POSS
ini banthu-an asà-kamauvan su-aada.
DEM.PROX help-NMLZR CP-need   PAST-exist
'While that trouble took place the king of Kandy needed our help.'[3]

(10) *Ini raaja=yang mà-jaaga,*[4] *itthu mà-jaaga=nang anà-baa*
DEM.PROX king=ACC   INF-protect DEM.DIST INF-protect=DAT PAST-bring
mlaayu=dering,[4] *anà-aaji.baa mlaayu=dering*[4] *satthu oorang=jo see.*
Malay=ABL   PAST-bring.ANIM Malay=ABL   one man=EMPH 1s
'Out of the Malays that were brought to protect this king, I am one.'

---

[2]*Aaji baa* can only be used with animates. *Aaji* does not seem to be used outside of this collocation, but *baa* means 'to bring' when used alone.

[3]The speakers construes the community of helping Malays as including him as well, although several centuries separate the soldiers and the speaker.

(11) *Suda, inni,   see karang, blaakang=nang laiskali kapang-pii, see*
     thus DEM.PROX 1s now   after=DAT   again   when-come 1s
     ka-umpath generation *katha kapang-biilang.*
     ORD-four   generation QUOT when-say
     'So, when I repeat again that I belong to the fourth generation,'

(12) *Itthukapang   se=ppe baapa, se=ppe kaake,   kaake=pe*
     DEM.DIST=when 1s=POSS father 1s=POSS grandfather, grandfather=POSS
     baapa, kithang samma oorang Seelong=pe oorang pada.
     father 1PL   all   man   Ceylon=POSS man   PL
     'then my father, my grandfather, great-grandfather, we all are Ceylonese.'

(13) *Ikang, se=ppe kaake=pe   baapa=nang sithari=jo se=dang*
     then   1s=POSS grandfather=POSS father=DAT   that.side 1s=DAT
     mà-biilang thàrà-thaau.
     INF-say   NEG-know
     'Then, I cannot tell you anything beyond my great grandfather.'

(14) *Incayang=jo   sathu vakthu=nang Malaysia=dering dhaathang aada.*
     3s.POLITE=EMPH one   time=DAT   Malaysia=ABL   come   exist
     'He is the one who would have come from Malaysia at some point in time.'

(15) *Ithu=nang=aapa   incayang[5] kapang-dhaathang cinggala incayang=nang[5]*
     DEM.DIST=DAT=after 3s.POLITE when-come   Sinhala 3s.POLITE=DAT
     thàrà-thaau.
     NEG-know
     'But, when he came he didn't know Sinhala.'

---

[4]Three starts to get the correct phrasing of the complex NP right.

[5]The first *incayang* established the topic, while the second one is used to host the semantic role, experiencer, indicated by =*nang*, which had been omitted at the first occurrence.

(16) *Siini Seelong=nang dhaathang=nang blaakang=jo, incayang cinggala*
     here Ceylon=DAT   come=DAT   after=EMPH   3s.POLITE Sinhala
     asà-blaajar=apa siini=pe   raaja=nang mà-banthu anà-mulain.
     CP-learn=after here=POSS king=DAT   INF-help   PAST-start
     'Only after having come to Ceylon, he learned Sinhala and began to help the
     local king.'[6]

(17) *Itthu   vakthu ini   Malaysia=ka anà-duuduk   milt...* military
     DEM.DIST time   DEM.PROX Malaysia=LOC PAST-exist.ANIM milt... military
     regiment Malay Military Regiment *katha.*
     regiment Malay Military Regiment QUOT
     'At that point in time in Malaysia there was the military regiment called the
     Malay Military Regiment.'[7]

(18) *Itthu   regiment=pe   oorang pada=jo..., =nang=jo see=le*
     DEM.DIST regiment=POSS man   PL=EMPH   =DAT=EMPH 1s=ADDIT
     araa...kumpul-an.
     NON.PAST-gather-NMLZR
     'I also belong to the so called military regiment people.'[8]

(19) *Se=ppe=le   kumpul-an itthu   regiment=nang=jo.*
     1s=POSS=ADDIT gather-NMLZR DEM.DIST regiment=DAT=EMPH
     'My association is also with that regiment.'

(20) *Ikang*[9] *Seelong=nang dhaathang aada mlaayu oorang ikkang.*[9]
     then   Ceylon=DAT   come   exist Malay man   fish
     'Then those who came to Ceylon were the Malay fishermen.'

---

[6]Hussainmiya (1987, 1990) states that the Malays first helped the Dutch. Then, some of them deserted and fought for the Kandyan king. There is no indication of Malays going directly to the Kandyan king.

[7]It is probable that this is a reference to the Ceylon Rifle Regiment established in the 19th century by the British in Ceylon.

[8]More precisely, the descendents of the soldiers.

[9]*ikang (=itthukang)* 'then' and *ikkang* 'fish' only differ in the gemination of /k/.

(21) *Derang pada=jo Hambanthota=ka arà-duuduk,*   *iiya.*
3PL   PL=EMPH Hambantota=LOC NON.PAST=exist.ANIM, yes
'They live in Hambantota.'

(22) *Ikang Seelong=nang lai*   *hathu kavanan anà-dhaathang.*
then Ceylon=DAT other one   group   PAST-come
'Then there came another group to Ceylon.'

(23) Slave Island *subla=ka arà-duuduk.*
Slave Island side=LOC NON.PAST-stay
'They stay in Slave Island.'[10]

(24) *Itthu*   *mlaayu, ithu*   regiment *mlaayu pada=nang mà-banthu*
DEM.DIST Malay   DEM.DIST regiment Malay PL=DAT   INF-help
anà-dhaathang oorang pada, itthu   *hathu kavanan laayeng.*
PAST-come   man   PL   DEM.DIST INDEF group   different
'Those Malays came to serve the regiment Malays, that's another group.'

(25) *Karang Kandi=ka arà-duuduk*   military regiment *mlaayu=le;*
now   Kandy=LOC NON.PAST-exist.ANIM military regiment Malay=ADDIT
hathyang kubbong, kubbong mlaayu.[11]
other   garden   garden Malay
'Now those who live in Kandy are the Malays of the Military Regiment and the others, estate, estate Malays.'

(26) *Punnu oorang pada anà-dhaathang thee thee daavong maa...*
full   man   PL   PAST-come   tea tea leaf   INF-...
thaanàm=nang thee pohong=yang mà-thaanàm=nang.
plant=DAT   tea tree=ACC   INF-plant=DAT
'Many people came to plant tea leaves, to plant the tea bush.'

---

[10]The name 'Slave Island' probably has no relation to the Malay slaves.

[11]*kubbong mlaayu* 'estate Malays' is a nonce compound of two nouns, which is right-headed. Established compound, like *oorang ikkang* 'fisherman' are left-headed.

(27) *Suda see arà-biilang   kaarang Kandi=ka   hathu thiga-pulu riibu=kee*
     thus 1s NON.PAST-say now   Kandy=LOC INDEF three-ty   thousand=SIMIL
     mlaayu pada arà-duuduk.
     Malay PL   NON.PAST-exist.ANIM
     'So I will tell you now that there could be about 30,000[12] Malays living in
     Kandy.'

(28) *Libbi   mlaayu samma Kluum̌bu=ka   arà-duuduk   punnu=le.*
     remain Malay all   Colombo=LOC NON.PAST-exist.ANIM full=ADDIT
     'The balance Malays are all living in Colombo.'

(29) *Seelong=ka arà-duuduk   hathu ìnnam-pulu liima riibu=ke*
     Ceylon=LOC NON.PAST-exist.ANIM INDEF six-ty   five thousand=SIMIL
     mlaayu pada.
     Malay PL
     'The maximum that you will find in Sri Lanka is about sixty-five thousand
     Malays.'

(30) *Suda kaarang pada kithang sajja=jo   asà... libbi   arà-duuduk.*
     thus nom   PL   1PL   only=EMPH CP-... remain NON.PAST-exist.ANIM
     'So now, its only we who remain in balance together.'

(31) *Itthu=sàsaama kithang lai=sangke=le   mlaayu arà-oomong.*
     DEM.DIST=COMIT 1PL   more=until=ADDIT Malay NON.PAST-speak
     'We still happen to speak Malay.'

---

[12]The actual number of Malays in Kandy is probably only one tenth of this figure.

(32) *Inni   sumbaring oomong kithang laile kithang=pe ithu   laama*
     DEM.PROX while   speak 1PL   still 1PL=POSS DEM.DIST old
     traditions pada Malaysia=*ka anà-duuduk   mosthor, anà-boole mosthor*
     traditions PL   Malaysia=LOC PAST-exist.ANIM manner PAST-can manner
     kithang itthu=yang   itthu   mosthor=nang arà-baapi,   sumbaring
     1PL   DEM=DIST=ACC DEM.DIST manner=DAT   NON.PAST-bring while
     baapi.
     bring.
     'While we talk we still stick to those old traditions of ours just as we were in
     Malaysia, whatever we can preserve we still continue with it in the process.'

(33) *Pake-yan=ka=le   punnu bedahan   thraa.*
     dress=NMLZR=LOC=ADDIT full   difference NEG
     'Even at our dress there is not much difference.'

(34) *Ithee   mosthor arà-paake,   anà-kamauvan vakthu=nang,*
     DEM.DIST manner NON.PAST-dress PAST-want   time=DAT
     kithang=nang[13] *itthu   mosthor=nang,* Malaysian *hathu mosthor=nang*
     1PL=DAT   DEM.DIST manner=DAT   Malaysian INDEF manner=DAT
     kithang=nang[13] *bole=duuduk.*
     1PL=DAT   can-exist.ANIM
     'We still dress like that, when it is needed some times we still can be in the
     Malaysian way.'

(35) *Itthu   pada samma ithee   mosthor aada.*
     DEM.DIST PL   all   DEM.DIST manner exist
     'They all prevail according to the same custom.'

(36) *Suda see bìsaran=nang nyaari arà-biilang: itthu   mlaayu mosthor*
     thus 1s pride=DAT   today NON.PAST-say DEM.DIST Malay manner
     lai=sangke kithang=ka aada.
     other=until 1PL=LOC   exist
     'So, today I tell you proudly: we still possess the same old Malay traditions.'

---

[13]*kithangnang* is present twice in this clause.

(37) *Hathu vakthu=nang kithang cinggala pada samma=le astha-kumpul aada,*
      one   time=DAT   1PL   Sinhala PL   all=ADDIT CP-gather   exist
      'Some times we may have got mixed together with the Sinhalese.'

(38) *spaaru* Christian *pada astha-kumpul aada, a... mulbar astha-kumpul aada.*
      some   Christian PL   CP-gather   exist a... Tamil CP-gather   exist.
      'Some mix together with Christians, or with Tamils.'[14]

(39) *Itthukapang, kumpul-an=nang=le, punnu mlaayu pada lai   Seelong=ka*
      then   mix-NMZLR=DAT=ADDIT full   Malay PL   more Ceylon=ADDIT
      aada.
      exist
      'Then, and in spite of all this mixing, there are still a good lot of Malays in Sri
      Lanka.'

(40) *Itthee   mlaayu mosthor=nang arà-duuduk.*
      DEM.DIST Malay manner=DAT   NON.PAST-stay
      'They still continue to live according to their old customs.'

(41) *A... what more do you want?*

(42) *Se=ppe baapa* Royal Ceylon Air Force=*ka   anà-bagijja.*[15]
      1s=POSS father Royal Ceylon Air Force=LOC PAST-work
      'My father worked in the Royal Ceylon Air Force.'

(43) Royal Ceylon Air Force=*ka   asà-bagijja itthu=nang   blaakang asà-*retire
      Royal Ceylon Ari Force=LOC CP-work   DEM.DIST=DAT after   CP-retire
      incayang thuju-pul liima thaaun=sangke incayang anà-iidop.
      3s.POLITE seven-ty five year=until   3s.POLITE PAST-live
      'He worked in the R.C.A.F. and then after that, he retired and lived up to 75
      years.'

---

[14]The speaker does not mention the Moors, even if in his own family, there are several marriages with
Moors.

[15]*bagijja* 'work(v)' seems to be an incorporation of *pukurjan* 'work' into the verb *kijja* 'make'. The more
common order would be *pukuran anà-gijja*, with the object preceding the tense marker. But due to its high
frequency and strong conceptual unity, the collocation might have become fused (cf. Mithun 1984).

(44) *ini*[16]*=dìkkath=jo   anà-mniiňǧgal.*
DEM.PROX=vicinity=EMPH PAST-die
'He passed away just recently.'

(45) *Suda ithu=nang   blaakang see arà-duuduk, se=ppe blaakang*
Thus DEM.DIST=DAT after   1s NON.PAST-stay 1s=POSS after
arà-dhaathang se=ppe fifth generation, se=ppe anak klaaki pada.
NON.PAST-come 1s=POSS fifth generation 1s=POSS child male PL
'So after that, now, I am staying; following behind me is the fifth generation, my sons.'[17]

(46) *Suda se=ppe thuuva anak klaaki asàdhaathang dhlapan-blas thaaun.*
thus 1s=POSS old   child male COPULA   eight-teen   year
'So my eldest son is now eighteen years old.'

(47) *Anà-muuda=jo   anak klaaki lai=le   duuva thaaun.*
SUPERL-young=EMPH child male   more=ADDIT two   year
'The youngest son is still (only) two years.'

(48) *Se=dang liima anak klaaki pada aada.*
1s=DAT   five child male PL   exist
'I have five sons.'

(49) *Incayang=yang anà-braanak Udapussallava=dika.*[18]
3s.POLITE=ACC   PAST-be.born Udapussallawa=LOC
'He was born at Udapussalawa.'

(50) *Incayang Udupussallava=ka*[18] *anà-braanak.*
3s.POLITE Udapussallawa=LOC PAST-be.born
'He was born in Udapussallawa.'

---

[16]Note the temporal use of the proximal deictic *ini* in this sentence.

[17]It is not clear why the sons are treated as fifth generation here.

[18]The locative enclitic in its uncommon full form *dika* and the more common reduced form *ka* in the second sentence.

(51) *Se=ppe kaake   asàdhaathang* estate tea factory officer, estate tea
      1s=POSS grandfather COPULA   estate tea factory officer estate tea
      factory officer.
      factory officer
      'My grandfather was an estate tea factory officer.'

(52) *Itthu=subbath=jo   incayang=yang siithu anà-braanak.*
      DEM.DIST=because=EMPH 3s.POLITE=ACC   there PAST-be.born
      'Because of that only he was born there.'

(53) *Siithu asà-braanak incayang=pe   plajaran*[19]*=nang incayang Kandi=nang*
      there CP-be.born 3s.POLITE=POSS education=DAT   3s.POLITE Kandy=DAT
      anà-dhaathang.
      PAST-come
      'After he was born there, for his studies he came to Kandy.'

(54) *Kandi=nang asà-dhaathang,*[20] *Kandi=ka   asà-kaaving=apa,*[20]
      Kandy=DAT CP-come   Kandy=LOC CP-marry=after
      itthu=nang   blaakang[20]*=jo kithang pada anà-bìssar.*
      DEM.DIST=DAT after=EMPH   1PL   PL   PAST-big
      'He came to Kandy and after marrying in Kandy, after that, we grew up.'

(55) *Baapa=pe kaake=pe   naama asàdha..., baapa=pe baapa=pe*
      father=POSS grandfather=POSS name   ...   father=POSS father=POSS
      naama asàdhaathang T N Salim, Thuan Nahim Salim.
      name   COPULA   T N Salim Thuan Nahim Salim
      'Father's, grandfather's name, father's father's name was T N Salim, Thuan
      Nahim Salim.'

(56) *Incayang=nang sithari   se=dang mà-biilang thàrà-thaau.*
      3s.POLITE=DAT that.side 1s=DAT INF-say   NEG-know
      'I don't know anything to say after him any further.'

---

[19]The voiceless stop in *plajaran* stems from an old nominalizing circumfix *pər-...-an*, which is no longer prodcutive today. The unnominalized verb is *blaajar* 'learn', with a voiced stop.

[20]Note the different strategies for marking of subsequence, *asà-, asà-...=apa, blaakang*.

(57) *Incayang=pe   baapa=pe   naama see thàrà-thaau, iiya a...*
    3s.POLITE=POSS father=POSS name   1s   NEG-know   yes a...
    'I do not know his father's name, yes.'

(58) *Derang pada, kithang=pe oorang thuuva pada bannyak dhaathang aada*
    3PL   PL   1PL=POSS man   old   PL   much   come   exist
    Malaysia=*dering.*
    Malaysia=ABL
    'Many of them, of our forefathers came from Malaysia.'

(59) *Spaaru Indonesia=dering dhaathang aada.*
    some   Indonesia=ABL   come   exist
    'Some came from Indonesia.'

(60) *Siini dhaathang=nang blaakang, derang kaaving=apa, itthu,   ithu*
    here come=DAT   after   3PL   marry=after   DEM.DIST DEM.DIST
    maana˜maana[21] *thumpath katha kithang=nang buthul=nang mà-biilang*
    which˜RED   place   QUOT 1PL=DAT   correct=DAT INF-say
    thàrboole.
    cannot
    'And after they came here, they got married; where all[21] they lived, we cannot tell exactly.'

(61) *Hatthu se=dang bàrà-biilang:* Malaysia *samma oorang=pe naama pada*
    one   1s=DAT can-say   Malaysia all   man=POSS name   PL
    'Maas'.
    Maas
    'One thing I can say: the names of all Malaysians are "Maas".'

(62) *Maas=dering=jo arà-start,   derang=pe naama pada samma*
    Maas=ABL=EMPH NON.PAST-start 3PL=POSS name   PL   all
    arà-gijja   buthul Maas=dering.
    NON.PAST-make correct Maas=ABL
    'They make their names with the beginning "Maas".'

---

[21]The reduplicated interrogative pronoun indicates the need for an exhaustive answer. This is difficult to render in English

(63) *Indonesia=dering arà-dhaathang naama pada samma 'Thuan'.*
Indonesia=ABL NON.PAST-come name PL all Thuan
'The names which come from Indonesia are all "Thuan".'

(64) *Suda itthu=dering=jo kithang=nang ini Indonesia=pe*
thus DEM.DIST=ABL=EMPH 1PL=DAT DEM.PROX Indonesia=POSS
oorang=si giithu kalthraa Malaysian *oorang=si katha bara-thaau*
man=INTERR that.way otherwise Malaysian man=INTERR QUOT can-know
ambe, iiya.
take, yes
'Thus, we will come to understand whether this is an Indonesian or otherwise
a Malaysian, yeah.'[22]

(65) *Lai aapa lai aapa mà-biilang aada,* right*?*
more what more what INF-say exist, right
'What else do I have to say?'

(66) *Seelong=ka arà-duuduk a..., se=dang thaau mosthor=nang,*
Ceylon=LOC NON.PAST-stay a... 1s=DAT know manner=DAT
kaarang=nang kà-dhlaapan generation *arà-pii, kà-dhlaapan.*
now=DAT ORD-eight generation NON.PAST-go ORD-eight
'staying in Ceylon, er, as far as I know up to now we are in the eighth genera-
tion, eighth.'

(67) *Ka-dhlaapan* generation *katha kapang-biilang.*
ORD-eight generation QUOT when-say
'When I say it is the eighth generation, ...'

(68) *Seelong=nang duppang duppang*[23] *anà-dhaathang mlaayu asàdhaathang*
Ceylon=DAT front front PAST-come Malay COPULA
oorang ikkang, oorang ikkang, drifted *Indonesians pada.*
man fish man fish drifted Indonesians PL
'The Malays who came as the very first to Ceylon were fishermen, fishermen,
drifted Indonesians.'

---

[22]At least as far as the paternal ancestor is concerned. There is no indication that the different immigrant
groups stayed separate. It is likely that most Sri Lankan Malays have both Indonesian and Malaysian blood
in them.

(69) *Derang pada=jo mlaayu pada, muula anà-dhaathang mlaayu pada.*
3PL   PL=EMPH Malay PL   before PAST-come   Malay PL.
'They were Malays, the Malays who had come in the beginning.'

(70) *Itthu=nang   blaakang dhaathang aada* as slaves.
DEM.DIST=DAT after   come   exist as slaves
'The ones who came after that were slaves.'

(71) *Kà-duuva anà-dhaathang* slaves *pada, soojor   pada anà-baa   oorang*
ORD-two PAST-come   slaves PL   European PL   PAST-bring man
pada.
PL
'The second ones who came were slaves, people brought by the Europeans.'

(72) *Kà-thiiga=joo   *regiment.
ORD-three=EMPH regiment.
'The third, soldiers.'

(73) *Kà-thiiga anà-dhaathang* regiment.
ORD-three PAST-come   regiment.
'The third ones to come were soldiers.'

(74) *Ithu=kapang,   see arà*-belong   *kà-ùmpath* generation=*nang.*
DEM.DIST=when 1s   NON.PAST-belong ORD-four   generation=DAT
'Then I belong to the fourth generation.'

(75) *Ikang asàdhaathang kithang* delay *aada.*
then COPULA   1PL   late exist
'Then we were late.'[24]

(76) *Kithang=pe* regiment=*yang Seelong=nang mà-dhaathang anà*-delay.
1PL=POSS   regiment=ACC Ceylon-DAT   INF-come   PAST-late
'Our regiment was late to come to Ceylon.'[24]

---

[23]Note the reduplication of the temporal adverb to get an intensive reading.

[24]The speaker does/did not belong to the Regiment, but construes it as comprising him as well.

(77) *Ithu=na=apa   kithang=nang duppang lai   duuva bergaada*
DEM.DIST=DAT=after 1PL=DAT   front   more two   generation
dhaathang aada.
come   exist
'So before us, there are two other generations.'

(78) *Itthu   duuva bergaada=jo   kà-thaama ikkang, oorang ikkang.*
DEM.DIST two   generation=EMPH first   fish   man   fish
'those two generations, the first were fishermen.'

(79) *Kà-duuva=jo   slaves pada.*
ORD-two=EMPH slave   PL
'The second slaves.'

(80) *Itthu=le   kithang=pe muusing=ka=jo anà-baa   ...*
DEM.DIST=ADDIT 1PL=POSS   time=LOC=EMPH PAST-bring ...
'But in our time they brought...'

(81) *Itthu,   kithang, hathu vakthu=nang, kithang=nang duppang dhaathang*
DEM.DIST 1PL   one   time=DAT   1PL=DAT   before   come
athi-aada exiles *pada=jo anà-baa.*
IRR-exist   exiles PL=EMPH PAST-bring
'Then, at that time, the exiles, who would have come before us, were brought.'

(82) *Itthu=nang=jo   see   anà-biilang oorang   soojor   anà-baa*
DEM.DIST=DAT=EMPH PAST-say man   European PAST-bring QUOT
katha, iiya.
yes
'That's why I say that we were brought by Europeans.'

(83) *Suda itthu   oorang pada=le Seelong=ka arà-duuduk.*
thus DEM.DIST man   PL=ADDIT Ceylon=LOC NON.PAST-exist.ANIM
'So these people are staying in Ceylon as well.'

(84) *Derang pada bannyak bisaran oorang pada.*
3PL   PL   much   pride   man   PL
'They are very proud people.'

(85) *Konnyong=ke thaakuth thraa.*
few=SIMIL   fear   NEG
'They are not afraid of the slightest thing.'

(86) *Konnyong katha arà-biilang   sdiikith.*
konnyong QUOT NON.PAST-say few
' *konnyong* is to say "few".'

(87) *Thaakuth thraa.*
fear   NEG
'They had no fear.'

(88) *Punnu buthul braani, kalbu braani.*
full   correct brave, mind brave
'They are very brave.'

(89) *Itthu   anà-dhaathang, derang pada=ka=jo Seelong=ka avuliya pada.*
DEM.DIST PAST-come   3PL   PL=LOC=EMPH Ceylon=LOC saint   PL
'Among the one who came, among them, there are a lot of saints.'

(90) You call them saints iiya...

(91) *Mlaayu=ka=jo   bannyak avuliya Seelong=ka aada.*
Malay=LOC=EMPH much   saint
'There are many saints among the Malays in Sri Lanka.'

(92) *See laskali arà-biilang itthu   saints pada, itthu=ka   bàrnaama*
1s again NON.PAST-say DEM.DIST saints PL   DEM.DIST=LOC famous
anà-pii   saints pada=jo, sudaara thuuju.
NON.PAST-go saint PL=EMPH brother seven
'I say again that they are saints. There are saints who have become famous, seven siblings.'

(93) *Thuan Thungku, Thuan Idriis, Thuan Skiilan.*
Thuan Thungku Thuan Idriis Thuan Skiilan
'Thuan Thungku, Thuan Idriis, Thuan Skiilan'

(94)   *Hathyang Thuan Kuddhuus su-biilang, thraa Thuan Kuddhuus Thuan*
       other   Thuan Kuddhuus su-biilang, NEG   Thuan Kuddhuus, Thuan
       Thungku Thuan Skiilan Thuan Idriis.
       Thungku Thuan Skiilan Thuan Idriis
       'another one is called Thuan Kuddhuus, no, Thuan Kuddhuus, Thuan
       Thungku, Thuan Idriis, Thuan Skiilan.'[25]

(95)   *Thuan ...* uncle *hathyang naama pada saapa* uncle.
       Thuan ... uncle other   name PL   who uncle
       'uncle, who(=what) are the other's names?'

(96)   *Binthan* auntie=*ka caanya, Binthan* auntie=*yang konnyong panggel, iiya,*
       Binthan auntie=LOC ask   Binthan auntie=ACC few   call   yes,
       Faathima Naaciaar, Faathima Naaciaar, iiya.
       Faathima Naaciaar Faathima Naaciaar, yes
       'Ask auntie Binthan, call Binthan auntie yeah, Fatima.'

(97)   *Faathimaa Naaciaar=le   hathyang thiiga oorang thiiga oorang laskali*
       Faathimaa Naaciaar=ADDIT other   three man   three man   again
       Malaysia=*nang su-baapi.*
       Malaysia=DAT   PAST-take.away
       'Naciar, and another three men. Three other men were brought to Malaysia.'

(98)   *Seelong=ka su-mniiňğgal, ithu   thiiga oorang=pe naama kithang*
       Ceylon=LOC PAST-die   DEM.DIST three man=POSS name 1PL
       thàrà-thaau.
       NEG-know
       'They died in Ceylon, but we do not know those three men's names.'

(99)   *Derang samma oorang pada* saints *pada anà-jaadi.*
       3PL   all   man   PL   saint PL   PAST-become
       'Those people all became saints.'

---

[25]The speaker and the people around are not too sure about the names and deeds of the saints, and some confusion arises.

(100) *Aapa derang pada sgiithu   braani, derang pada samma raaja aanak pada,*
what 3PL PL that.much brave 3PL PL all   king child PL
samma avuliya anà-jaadi.
all   saint PAST-become
'They were so brave, they were all princes, they all became saints.'

(101) right*, hatthu avuliya aada kithang=pe ruuma dìkkath.*
right one   saint exist 1PL=POSS house vicinity
'There is one saint close to our house.'

(102) *Itthu   Thuan Skiilan, lai hatthu Avuliya su-aada Kandi*
DEM.DIST Thuan Skiilan more INDEF saint   PAST-exist Kandy
town=ka,   karang Malay Mosque katha arà-biilang.
downtown=LOC now   Malay Mosque QUOT NON.PAST-say
'That's Thuan Skiilan. There is another saint in Kandy town, they call it now Malay mosque.'

(103) *Kandi=ka Malay Mosque=pe blaakang=ka incayang=pe zihaarath*
Kandy=LOC Malay Mosque=POSS behind=LOC   3s.POLITE=POSS shrine
aada.
exist
'In Kandy, behind the Malay Mosque, there is his shrine.'

(104) *Ikang derang=pe sudaara pompang=jo aada Hanthane=ka.*
then 3PL=POSS sibling female=EMPH exist Hanthane=LOC
'Then their sister is in Hanthane.'

(105) *Hanthane, iiya, guunung=ka aada.*
Hanthane yes mountain=LOC exist
'Yeah, in Hanthane, on the mountain.'

(106) *Lai hatthu avuliya=pe aada Kluum̌bu Dematamaram=ka, iiya.*
more INDEF saint=POSS exist Colombo Dematamaram=LOC yes
'Ehm, there is another saint in Colombo, in Dematamaram yes.'

(107)   *Derang ùmpath oorang=pe=jo   karang Seelong=ka aada.*
        3PL   four   man=POSS=EMPH now   Ceylon=LOC exist
        'These four saints' are now in Ceylon.'

(108)   *Hathyang thiiga oorang=yang a...* Malaysia=*nang su-ambel baapi   iiya.*
        other   three man=ACC   a... Malaysia=DAT   PAST-take take.away yes
        'Three others were brought to Malaysia.'

Saradiyel

This story is about the 'Sri Lankan Robin Hood', Saradiyel, who robbed the rich to give to the poor and was arrested by a Malay policeman (cf. Osman & Sourjah 2008:52). Generally, Robin-Hood-types carry the sympathy of the speakers, but in this case, his adversary is Malay, which yields split sympathies. On the one hand, Saradiyel is portrayed as a hero, on the other hand as a terrorist. Especially interesting in this text is the tracking of reference between the policeman and Saradiyel. The policeman is referred to by the polite third person singular pronoun *incayang*, while Saradiyel is normally indicated with the impolite form *dee*. This suffices to track reference, and other means of disambiguation are normally not employed.

This text shows more features of oral communication than the preceding ones. There are frequent changes of topic, and new starts, as well as a good deal of code switching between SLM and English and some lexical gaps which are filled by Sinhala or Tamil. Also, the story is told in a somewhat cyclical fashion. In a first run-through, the major events are related, while the details are worked out in subsequent iterations.

(1) *Robin* Hood *katha arà-biilang laama ...*
Robin Hood QUOT NON.PAST-say old   ...
'The person we call Robin Hood is an old ...'

(2) *cinggala=dering arà-biilang naama Saradiyel*
Sinhala=ABL   NON.PAST-say name Saradiyel
'The name he is called in Sinhala is "Saradiyel".'

(3) *de=ppe*[1]   *naama Saradiyel*
3s.IMPOLITE=POSS name   Saradiyel
'His name was Saradiyel.'

(4) *diya*[2] *ini   Seelong=ka anà-duuduk   18, 18 odd=ka, 18*
3s   DEM.PROX Ceylon=LOC PAST-exist.ANIM 18 18 odd=LOC 18
odd=*ka=jo*[3]
odd=LOC=EMPH
'He lived here in Ceylon in 18-odd.'

---

[1]The first occurrence of the impolite pronoun *dee* to refer to Saradiyel indicates that he does not have the sympathy of the speaker.

[2]*Diya* is a neutral third person pronoun, which is used very rarely.

[3]This sentence shows many oral features, like repetition and code switching.

(5) 18 odd=*ka=jo*   *anà-duuduk*   *siini Seelong=ka*
    18 odd=LOC=EMPH PAST-exist.ANIM here Ceylon=LOC
    'He lived here in 18-odd, in Ceylon.'

(6) *de=ppe*   *pukurjan:*
    3S.IMPOLITE=POSS work
    'His work:'

(7) *dee*   *maaling hatthu*
    3S.IMPOLITE thief   INDEF
    'He was a thief.'

(8) *dee*   *arà*[4]*-cuuri*   *baaye kaaya oorang pada=dering*
    3S.IMPOLITE NON.PAST-steal good rich   man   PL=ABL
    'He steals from very rich people.'

(9) *kaaya oorang – monied people – =dering=jo arà-cuuri*
    rich   man   =ABL=EMPH NON.PAST-steal
    'It is from moneyed people that he steals.'

(10) *mà*[5]*-cuuri abbis dee*   *arà-kaasi*   *miskin pada=nang*
     INF-steal   finish 3S.IMPOLITE NON.PAST-give poor   PL=DAT
     'After stealing, he gives the money to the poor.'

(11) *itthunangapa diya asàdhaathang inni*   government, government=*pe*
     however   3S   COPULA   DEM.PROX government government=POSS
     forces *pada=nang=susaama baaye thraa*[6]
     forces PL=DAT=COMIT   good NEG
     'However, he was not good with the government forces.'

---

[4]The tense in the narrative switches from past to the present *arà* to render the story more lively.

[5]Not all speakers would agree with the infinitive marker here.

[6]The structure of this sentence seems to be influenced by English syntax, namely the use of the copula and the comitative.

(12) *dee mana vakthu=le* forces=*nang arà-bukalai*
3s which time=ADDIT forces=DAT NON.PAST-fight
'He fights with the forces all the time.'

(13) forces *pada=samma arà-bukalai*
forces PL=COMIT NON.PAST-fight
'He fights with the army.'

(14) *inni=yang mà-peegang=nang, subla nigiri=dering suda anà-dhaathang.*
DEM.PROX INF-catch=DAT side country=ABL thus PAST-come
'to catch this man, forces[7] from the neighbouring countries came.'

(15) *inni* forces *pada asà-dhaathang, Seelong=ka asà-duuduk, derang*
DEM.PROX forces PL CP-come Ceylon=LOC CP-stay 3PL
anà-peegang dee=yang[8]
PAST-catch 3s.IMPOLITE=ACC
'These forces came and stayed in Ceylon and caught him.'

(16) *itthunang pàrthaama kithang=pe mlaayu=jo anà-peegang*[9]
but before 1PL=POSS Malay=EMPH PAST-catch
'But, before that it was our Malay who caught him.'

(17) *[incayang anà-peegang ∅]*[10] *asàdhaathang Mavanella*
3s.POLITE PAST-catch COPULA Mawanella
'Where he caught him was at Mawanella.'

---

[7]The reference of the helpers is not established but must be inferred from the context. Some unnamed entity came from neighbouring countries. This entity must be in a position to help law enforcement against a thief, so it must also be some kind of force. Also note the switch back to past tense.

[8]Note the SVO word order.

[9]Note the dropped 'object'/patient in this sentence. The agent *mlaayu* is mentioned, but not the patient, Saradiyel, which can be inferred from context.

[10]This is a headless relative clause referring to a location.

(18) *peegang abbis*
      catch   finish
      'After catching him'

(19) *peegang=nang blaakang, incayang*[11] *theeṁbak aada*
      catch=DAT   after   3S.POLITE shoot   exist
      'after catching him, he shot him.'

(20) *incayang theeṁbak abbis,* Ø[12] *salba   laari aada thumpath=nang*
      3S.POLITE shoot   finish   escape run   exist place=DAT
      'Upon shooting, he escaped and ran to the place.'

(21) *itthu=jo,*   first, *kàthaama* police *oorang nya-maathi*[13] *inni*   terrorist
      DEM.DIST=EMPH first first   police man   PAST-die   DEM.PROX terrorist
      hatthu=dering
      INDEF=ABL
      'And that is how the first policeman came to die by the hands of a terrorist.'

(22) 'Terrorist'*=nang apa=yang arà-biilang   mlaayu=dering?*
      terrorist=DAT   what=ACC NON.PAST-say Malay=ABL
      'How do you say "terrorist" in Malay?'

(23) terrorist *hatthu=dering anà-maathi kàthaama oorang=jo incayang,* Kandy,
      terrorist INDEF=ABL   PAST-die   first   man=EMPH 3S.POLITE Kandy
      police man, bandit[14]
      police man, bandit
      'The first person who died by a terrorist was him.'

---

[11]It is not clear whether the police man is agent or patient of the shooting here.

[12]From syntax only, it is not clear whether the shooter or the victim ran away.

[13]It is only this sentence that establishes that the police man was shot, and that the thief was the killer, who then ran away.

[14]The words in this string change reference between the police man and the bandit.

(24) bandit, bandit bandit=*dering anà-maathi* first *kàthaama* police man=*jo*
      bandit bandit bandit=ABL   PAST-die   first first   police man=EMPH
      incayang
      3S.POLITE
      'The first policeman killed by the hands of a bandit was him.'

(25) *laile itthu=yang   arà-..inni..-iingath-kang   inni   police=dering*
      still DEM.DIST=ACC NON.PAST-..HESIT..-think-CAUS DEM.PROX police=ABL
      'Still the police bring into memory this story.'

(26) *[police=ka anà-aada ∅]*[15] ∅[16] first police man to die in Sri Lanka
      police=LOC PAST-exist   first police man to die in Sri Lanka
      'Of those who were in the police, he was the first man to die in Sri Lanka.'

(27) by terrorist
      by terrorist
      'from the hands of a terrorist.'

(28) *itthu=yang   arà-..   laile arà*-celebrate-*kang,   oorang pada ini*
      DEM.DIST=ACC NON.PAST-.. still NON.PAST-celebrate-CAUS man   PL   DEM.PROX
      memorial *dìkkath=ka asà-pii derang pada* parade *samma asà-gijja*
      memorial vicinity=LOC CP-goo 3PL   PL   parade all   CP-make
      derang=yang arà-iingath   maana vakthu=le
      3PL=ACC   NON.PAST-think which time=ADDIT
      'Still, this incident is celebrated, people go to the memorial place and hold a
      parade and think of him all the time.'

(29) *Maana maana thaaun=le*
      which which year=ADDIT
      'Every year'

---

[15]A headless relative clause referring to all persons who are in the police.
[16]The participant is dropped here.

(30) *maana thaaun=le Mavanella=ka, itthu   asàdhaathang Uthuvan*
which year=ADDIT Mawanella=LOC DEM.DIST COPULA   Uthuwan
Kandha Paamula=ka.
Kandha Paamula=LOC
'every year in Mawanella, that is at the Uthawan Kandha Paamula (at the foot of the mountain).'

(31) *Uthuvan Kandha=ka=jo   diya anà-duuduk.*
Uthuwan Kandha=LOC=EMPH 3   PAST-stay
'He stayed at Uthuwan Kandha.'

(32) *Sini=ka aada, blaakang=ka, bìssar guunung hatthu.*
here=LOC exist behind=LOC big   mountain indef
'There is a big mountain here, behind us.'

(33) *Dee   itthu=ka=jo   arà[17]-sbuuni duuduk*
3S.IMPOLITE DEM.DIST=LOC=EMPH NON.PAST-hide stay
'He was hiding there.'

(34) *Itthu=ka=jo   anà-sbuuni duuduk.*
DEM.DIST=LOC=EMPH PAST-hide   stay
'He was hiding right there.'

(35) *Itthu=deri sbuuni, dee   maana aari=le   asà-dhaathang,*
DEM.DIST=ABL hide   3S.IMPOLITE which day=ADDIT CP-come
thìnga-ari vakthu=nang kalthraa maalang vakthu=nang
middle-day time=DAT   otherwise night   time=DAT
'He was hiding right there, and from there he came every day, during the afternoon or otherwise during the night and'

(36) *kàthaama nya-aada lorry bus thraa[18]*
before   PAST-exist lorry bus NEG
'before, there were no lorries nor busses.'

---

[17]The use of *arà-* is surprising here. Time reference is clearly not the present, and relative simultaneous tense seems unlikely as well.

[18]The cooccurrence of *aada* and *thraa* in surprising in this sentence. Strictly speaking, *thraa* is already a

(37) lorry bus *thraa, kreetha kuuda kreetha, ya, kuuda kreetha*
　　 lorry bus NEG　 cart　 horse cart　　 yes horse　cart
　　 'There were no lorries or busses, only horse carts, yeah, horse carts.'

(38) *kuuda kreetha*, horse driven cart

(39) *itthu　　 guunung=ka　 naayek=jo, asà-dhaathang arà*[19]*-attack-kang*
　　 DEM.DIST mountain=LOC climb=EMPH CP-come　　 NON.PAST-attack
　　 itthu　　 oorang pada=yang
　　 DEM.DIST man　　 PL=ACC
　　 'As soon as they climb up the hill, he comes and attacks the people.'

(40) *[Sithu=ka　　 aada bìssar oorang pada]=yang asà-attack-kang,* mail=*nya*[20]
　　 DEM.DIST=LOC exist big　　 man　 PL=ACC　　 CP-attack-CAUS　 mail=ACC
　　 asà-cuuri,
　　 CP-steal
　　 'He attacks the big people who are there and stole the mail.'

(41) Mail=*ka duvith arà-baapi,　　　 bukang*
　　 mail=LOC money NON.PAST-take.away tag
　　 'The mail has money in it, isn't it.'

(42) *Suurath arà-baapi=ka　　　　 duvith aada. Laeng Kandi=dang,*
　　 letter　　 NON.PAST-take.away=LOC money exist　Other Kandy=DAT
　　 Kluum̌bu, Kandi=dering Kluum̌bu=dang suurath arà-baapi.[21]
　　 Colombo Kandy=ABL　 Colombo=DAT letter　 NON.PAST-take.away
　　 'The brought letters contained money. Some to Kandy, some from Kandy to Colombo the letters go.'

---

negated existential, so that there is no need to add the existential *aada*. One could interpret *nyaaada lorry bus* as a relative clause, meaning 'the lorries and busses which existed', but the whole meaning 'the lorries and busses which existed did not exist in former times' does not sound very good either.

[19] A return to the historical present in this sentence.

[20] The meaning of =*nya* is unclear here.

[21] The speaker gets a bit lost in the marking of the spatial semantic roles in this sentence.

(43) *Inni   samma dee   ambel abbis.*
      DEM.PROX all   3S.IMPOLITE take   finish
      'He steals all that.'

(44) *Dee   athi-kaasi miskiin oorang pada=nang.*
      3S.IMPOLITE IRR-give   poor   man   PL=DAT
      'He would give it to the poor.'

(45) *Miskiin oorang, miskiin* means poor people
      poor   man   poor

(46) *oorang pada=nang athi-kaasi.*
      man   PL=DAT   IRR-give
      'would give it to the people.'

(47) *Itthu=yang=jo   incayang=le   diyath anà-duuduk.*
      DEM.DIST=ACC=EMPH 3S.POLITE=ADDIT look   PAST-sit
      'He was always eyeing for that mail.'

(48) *[[Dee   asà-duuduk] anà-gijja   samma]=jo itthu*[22]
      3S.IMPOLITE CP-exist.ANIM PAST-make all=EMPH   DEM.DIST
      'That was what he did during his lifetime.'

(49) *Itthu=nang   blaakang dee=yang   su-peegang siini Seelong=ka.*
      DEM.DIST=DAT after   3S.IMPOLITE=ACC PAST-catch   here Ceylon=LOC
      'After that they caught him here in Ceylon.'

(50) *Peegang abbis*
      catch   finish
      'After being caught'

---

[22]An equational sentence where the first term is a quantifier modified by a complex relative clause. The relative clause consists of a higher clause containing the verb *gijja* 'make' with the participants *dee* 'he' and the head noun *samma* 'all'. This clause has in turn a cosubordinate clause containing the verb *duuduk* 'live'.

(51) *diya siini thàrà-duuduk.*
　　3s　here NEG.PAST-stay
　　'he didn't stay here.'

(52) *Diya laaye hathu nigiri=nang anà-baapi, buunung-king=nang, iiya*
　　3　other INDEF country=DAT PAST-bring kill-CAUS=DAT　　yes
　　'He was taken to another country, to have him executed.'

(53) hero[23] not killed in Sri Lanka so

(54) *Diya=yang anà-baapi　London.*[24]
　　3s=ACC　　PAST-take.away London
　　'They took him to London.'

(55) *Thraa,* England=*nang nya-baapi, dee=yang*
　　no　　England=DAT PAST-bring 3s.IMPOLITE=ACC
　　mà-execute-kang[25]*=nang.*
　　INF-execute-CAUS=DAT
　　'No, they took him to England to execute him.'

(56) *Mà-mathi-king, mà-mathi-king=nang siithu=jo　anà-baapi.*
　　INF-kill-CAUS　INF-kill-CAUS=DAT　there=EMPH PAST-take.away
　　'They took him there to execute him.'

(57) *Itthu=nang　kàthaama, incayang punnung dhraapa　oorang theeṁbak=le,*
　　DEM.DIST=DAT before　3s.POLITE full　how.many man　shoot=ADDIT
　　incayang=nang thama[26]*-kìnna.*
　　3s.POLITE=DAT　NEG.IRR-strike
　　'Before that, however many people shot at him, they would never hit him.'

---

[23]Note that the former 'terrorist' is now treated as a 'hero'.

[24]Note the absence of case marking on *London*.

[25]*-kang* is not used as a causative marker here, but as a loanword integrator, as is clear from the following sentence, which paraphrases this one.

[26]Note the use of the irrealis negator *thama-* to negate a habitual context in the past. *Thama-* is normally used for non-past negation, but habitual contexts also allow the use of *thama-* with past reference.

(58) Police *pada dhraapa   theeṁbak=le, siyang siyang thama-kìnna*
      police PL   how.many shoot=ADDIT   earlier earlier NEG.IRR-strike
      de=dang
      3s.IMPOLITE=DAT
      'How much the policemen would shoot, first they would not strike him.'

(59) *Incayang=pe   kaaki=ka nya-aada kiyang*[27], *hathu* talisman.
      3s.IMPOLITE=POSS leg=LOC   PAST-exist EVID   INDEF talisman
      'In his leg there was a talisman, it seems.'

(60) *Surayak*[28],   talisman, talisman, talisman.
      talisman[Sinh.]

(61) So that he didn't get shot.

(62) Talisman *hatthu anà-aada, kiyang.*
      talisman INDEF PAST-exist EVID
      'A talisman was there it seems.'

(63) Talisman *katha arabiilang   hatthu .. inni   ..; aapa*[29] *itthu=nang*
      talisman QUOT NON.PAST-say INDEF .. DEM.PROX .. what   DEM.DIST=DAT
      mlaayu=dering arà-biilang?
      Malay=ABL   NON.PAST-say
      'The thing they call a talisman, what do they call this in Malay again?'

---

[27]In contradistinction to the preceding sentences, which the speaker did not witness either, this sentence contains the evidential marker *kiyang*, probably to underscore the unbelievable nature of a talisman, which the speaker cannot vouch for.

[28]This lexical gap is filled with Sinhala.

[29]Metalinguistic information is requested in this sentences by means of *aapa*, whereas when querying for the word for 'terrorist' above (22), *apayang* was used.

(64) *Inni kithang arà-thaaro, iiya, iiya, kithang arà-thaaro, aapa*
DEM.PROX 1PL NON.PAST-put yes yes 1PL NON.PAST-put what
thaau=si, 'jiimath'!
know=INTERR talisman
'We put, yes, we put, you know what, a talisman.'

(65) jiimath, jiimath, talisman is jiimath

(66) *Jiimath hatthu thaaro kal*[30]*-aada, jiimath=yang thaaro kalu,*[30] *inni*
talisman INDEF put if-exist talisman=ACC put if DEM.PROX
kaaki daalang=ka=jo bìlla abbis thaaro aada.[31]
leg inside=LOC=EMPH split finish put exist
'When you have put a talisman, when you have put the talisman in this leg, he
had cut open (the leg) and put it.'

(67) *Kaaki poothong, daalang=ka thaaro=apa, jaaith aada, a. talisman*
leg cut inside=LOC put=after sew exist a. talisman
'Cut the leg and put it inside and sewn it, the talisman.'

(68) *talisman thaaro jaaith, a. jaaith*[32]
talisman put sew a. sew
'Put the talisman and sew.'

(69) *Suda itthu=dering de=dang thama-kìnna kiyang vatthu.*
thus DEM.DIST=ABL 3S.IMPOLITE=DAT NEG.IRR=strike EVID time
'So because of that he was never hit it seems.'

(70) *Itthu asàdhaathang* Buddhist priest *hatthu.*
DEM.DIST COPULA Buddhist priest INDEF
'That was a Buddhist priest.'

---

[30]Note the preverbal and postverbal occurrence of the conditional marker *kal(u)*.

[31]This string of four verbs (*bìlla abbis thaaro aada*) should probably receive some TAM-marking on one
of the verbs, although it is difficult to determine on which one.

[32]The speaker gets more and more taken away by the fantastic narrative and the syntactic structure of the
sentences becomes less important.

(71) *[Anà-gijja kaasi ∅] kiyang, incayang=nang giithu   hathu hathu inni*
      PAST-make give   EVID   3S.POLITE=DAT that.way INDEF INDEF DEM.PROX
      power *hatthu, kuvath-ahaan hatthu*
      power INDEF   strong-NMLZR INDEF
      'The one who gave home this power it seems, like this a this this power, this
      power'

(72) *Itthu   Buddhist priest=jo   gijja kaasi aada.*
      DEM.DIST Buddhist priest=EMPH make give   exist
      'That Buddhist priest made that for him.'

(73) *Suda lai   hatthu ma... mosthor arà-biilang*
      thus more INDEF   ma... manner NON.PAST-say
      'Then, another version says'

(74) *incayang=nang inni   Seelong=ka anà-duuduk   mlaayu pada, itthu*
      3S.POLITE=DAT DEM.PROX Ceylon=LOC PAST-exist.ANIM Malay PL   DEM.DIST
      muusing=ka anà-duuduk   mlaayu pada, 1876=ka anà-duudu   mlaayu
      period=LOC PAST-exist.ANIM Malay PL   1876=LOC PAST-exist.ANIM Malay
      pada, itthu   itthu   muusing, those days
      PL   DEM.DIST DEM.DIST period   those days
      'that the Malays who lived here in 1876 back then.[33]'

(75) *Mlaayu pada dhaathang vakthu*
      Malay PL   come   time
      'When the Malays came'

(76) *incayang hatthu mlaayu derang=samma kumpulan   baaye.*
      3S.POLITE INDEF Malay 3PL=COMIT   association good
      'he[34] was a Malay, who was a good friend of his.'

---

[33]The complex NP is uttered three times until the speaker is satisfied with its structure.

[34]At this point in time, the anaphoric pronoun *incayang* has no antecedent. It refers to Saradiyel's friend
Mammale, who has, however, not been introduced yet.

(77) *Mammallan itthu   inni=samma,   derang=pe naama Mammale, itthu*
Mammallan DEM.DIST DEM.PROX=COMIT 3PL=POSS   name   Mammale DEM.DIST
inni   katha, derang pada baaye kumpulan.
DEM.PROX QUOT 3PL   PL   good association
'Mammale was his good friend, his name was Mammale, who was always together with him.'

(78) *Suda kapang-duuduk derang pada*[35] *kaasi kiyang hathu kirris   hatthu.*
thus when-exist.ANIM 3PL   PL   give EVID   INDEF dagger INDEF
'So when they were there, it seems he gave him a dagger.'

(79) *kirris   kirris, a. kirris*   kinissa[36]
dagger dagger a. dagger dagger[Sinh.]

(80) what do you call that, 'dagger'!

(81) *itthu   asàdhaathang [baaye=nang vaasil-kang   aada dagger] hatthu*
DEM.DIST COPULA   good=DAT   blessed-CAUS exist dagger   INDEF
'That was a well blessed dagger.'

(82) waasil means

(83) *aapa dhua, dhua, dhua*
what prayer prayer prayer

(84) *dhua* blessed, blessed

(85) blessed dagger, blessed blessed, blessed dagger

(86) you understand blessed dagger, ya blessed dagger

---

[35]*Derang pada* is technically plural, but is used here with singular reference, which is very polite. The use of plural forms for singular reference is also found in Tamil. Furthermore note that there is no overt indication of semantic role; we do not know who was the giver of the dagger and who the recipient.

[36]The English word *dagger* could not be retrieved at this moment, and the Sinhala word *kinissa* is used instead.

(87) *so itthu=ka*   *mà*[37]*-aada kiyang, thembak-an thama-kìnna hatthu*
SO DEM.PROX=lic INF-exist EVID   shoot-NMLZR NEG.IRR-strike INDEF
oorang=na=le
man=DAT=ADDIT
'So when he has that one, no shot will hit him it seems, by anybody.'

(88) *Itthu*   *nya-aada katha=le   arà-biilang.*
DEM.DIST PAST-exist QUOT=ADDIT NON.PAST-say
'He had that also, they say.'

(89) *Arà-biilang, dee=ka itthu=jo   anà-aada katha=jo   arà-biilang.*
NON.PAST-say 3=LOC DEM.DIST=EMPH PAST-exist QUOT=EMPH NON.PAST-say
'They say that he had this with him.'

(90) *[Dee   arà-sbuuni duuduk   cave] asàraathang sini=ka asàduuduk*
3.IMPOLITE NON.PAST-hide exist.ANIM cave COPULA   here=LOC from
hathu three miles cara jaau=ka.
INDEF three miles way far=LOC
'The cave where he stayed hidden is about three miles away from here.'

(91) Three miles *cara jaau=ka aada dee   anà-sbuuni duuduk cave=yang.*[38]
three miles war far=LOC exist 3.IMPOLITE PAST-hide stay   cave=ACC
'Three miles away from here is the cave where he stayed hidden.'

(92) *Itthu*   *cave=nang kithang=le pii aada, mà-liyath=nang.*
DEM.DIST cave=DAT   1PL=ADDIT go exist INF-look=DAT
'We have also gone to the cave to have a look.'

(93) *Bannyak jaau mà-pii thàràboole.*
much   far INF-go cannot
'You can't go very far.'

---

[37]The use of the infinitive is unexpected here, a conditional would be more readily expected.

[38]Note the use of the accusative marker =*yang* here for a locational predication, but its absence in the preceding sentence.

(94) *Itthu=ka   daalang=ka mà-jaalang naarath,*[39] *mlaarath.*
DEM.DIST=LOC inside=LOC INF-walk   difficult   difficult
'Walking inside there is difficult.'

(95) *Itthu=ka,   daalang=ka giini   aada,* sorungum[40]
DEM.DIST=LOC inside=LOC like.this exist dungeon[Tamil]
'Inside the cave, it is like a dungeon.'

(96) *Giithu=jo   thuuňduk abbis=jo   masà-pii.*
that.way=EMPH bend   finish=EMPH must-go
'You must really walk bowed down.'

(97) *Kapang-pii, gìllap daalang=ka.*
when-go   dark   inside=LOC
'When you go, it's dark inside.'

(98) vavvaal[41] *itthu   ni   aada,* vavvaal *katha binaathan pada*
bat[Tamil] DEM.DIST DEM.PROX exist bat   QUOT animal   PL
'There are vavvaal, the animals called vavvaal.'

(99) different kind of birds what do you call this bats and all this, vavula[Sinh]

(100) vavvaal
bat[Tamil]

(101) *Kiccil vavvaal pada daalang=ka arà-duuduk.*[42]
small bat   PL   inside=LOC NON.PAST-exist.ANIM
'There are small bats inside.'

---

[39]*mnaarath* or *naarath* are occasionally found for *mlaarath* 'difficult'.
[40]Here a Tamil word is used to fill the lexical gap.
[41]The lexical gap is again filled with a Tamil word.
[42]Note the use of *duuduk* as an existential marker for [+animate] bats, which do not have the ability to sit, the original meaning of *duuduk*. This indicates the semantic bleaching of *duuduk* from a full verb to an existential.

(102) *Suda mà-pii thàràboole daalang=ka, light=le   mà-ambel baapi*
so   INF-go cannot   inside=LOC light=ADDIT INF-take take.away
thàràboole, daalang=ka pii=nang blaakang dhraapa   puukul=le
cannot   inside=LOC go=DAT after   how.much hit=ADDIT
thama-kuthumung.
NEG.IRR-see
'So you cannot go inside, you cannot take light inside, after going inside, however much he would hit them,[43] no one would see.'

(103) so dark inside=*ka*
so dark inside=LOC

(104) it is just close by here

(105) it was by this about 3 miles

(106) *mà-jaalang=jo masà-pii itthu=ka,   bukang*
INF-walk=EMPH must-go DEM.DIST=LOC TAG
'You have to walk there, you know.'

(107) *puunu mlaarath guunung atthas=ka ... jaalang kapang-pii athi-kuthumung*
full   difficult mountain top=LOC ... walk   when-go   IRR-see
itthu=yang
DEM.DIST=ACC
'Very difficult on top of the mountain. While walking, you can see all this.'

(108) *Mavanella kapang-liyath athi-kuthumung, Mavanella* bridge=*ka*
Mawanella when-see   IRR-see   Mawanella bridge=LOC
asàduuduk liyath kalu
from   see if
'When you look in Mawanella, you will see, if you look from the Mawanella bridge.'

---

[43]This probably refers to the victims of Saradiyel, although the precise meaning is unclear.

(109) *Saradiyel=pe   baathu arà-kuthumung*
Saradiyel=poss stone   non.past-see
'You see the Saradiyel Rock.'

Snow-White and Rose-Red

This text is a translation of the English version of the Brothers Grimm's tale. It is clearly written modality, with long and complex sentences and explicit reference tracking, some of which might be due to influences from the English text, e.g. the frequent use of the demonstrative *itthu* where English would require the definite article.

(1) *Snow-White=le Rose-Red=le.*
Snow-White=ADDIT Rose-Red=ADDIT
'Snow-White and Rose-Red.'

(2) *Hathu muusing=ka, sendiiri thurung guunung hatthu=ka, kampong=nang*
INDEF time=LOC alone slope mountain INDEF=LOC village
konyong jaau sindari, hathu kiccil ruuma su-aada.
few far PROX.ABL INDEF small house PAST-exist
'Once upon a time, alone by a hill slope, a little far away from this side of the village, there was a small house.'

(3) *Itthu bannyak laama hathu ruuma, itthule ruuma duuva subla=ka*
DEM.DIST very old INDEF house, but house two side=LOC
[panthas rooja kumbang pohong komplok] duuva asà-jaadi su-aada.
beautiful rose flower tree bush two CP-become PAST-exist
'That was a very old house, but still on the two sides of the house, there had grown two beautiful rose bushes.'

(4) *Hatthu komplok bannyak=jo puuthi caaya, hathyeng=yang meera=jo meera*
one bush very=EMPH white colour, other red=EMPH red
caaya.
colour
'One bush was pure white, the other one was pure red.'

(5) *Itthu ruuma=ka [laaki anà-mniiňğgal hathu pompang]=le,*
DEM.DIST house=LOC husband PAST-die INDEF female=ADDIT,
incayang=pe buthul panthas aanak pompang duuva=le su-aada.
3s.POLITE very beautiful child female two=ADDIT PAST-exist
'In that house, there lived a widow with her two very beautiful daughters.'

(6) *Derang duuva oorang=pe naama pada kalu, Snow-White, hattheyang*
3lp   two   man=POSS name PL   cond Snow-White, other
Rose-Red.
Rose-Red
'Their names were: Snow-White and Rose-Red.'

(7) *Kumbang rooja komplok duuva=pe naama=kee=jo.*
flower   red   bush   two=POSS name=SIMIL=EMPH
'Just like the two bushes of roses.'

(8) *Derang=pe panthas=dering=le,   pehel=dering=le,*
3PL=POSS   beautiful=ABL=ADDIT, behaviour=ABL=ADDIT
duduk-an=dering=le   butthul baaye.
live-NMLZR=ABL=ADDIT very   good
'Their beauty, behaviour and the way they lived were very good.'

(9) *Derang derang=pe umma=nang butthul saayang=kee=jo samma ruuma*
3PL   3PL=POSS mother=DAT very   love=SIMIL=EMPH all   house
pukurjan=nang=le anà-banthu.
work=DAT=ADDIT   PAST-help
'They helped her with all the homework, while they loved their mother very much.'

(10) *Ruuma pukurjan abbis=nang blaakang, derang dìkkath=ka   aada*
house   work   finish=DAT after   3PL   vicinity=LOC exist
laapang=nang mà-maayeng=nang su-pii.
ground=DAT   INF-play=DAT   PAST-go
'After finishing their homework, they went out to play at the ground close to their home.'

(11) *Soore=ka,   Snow-White=le   Rose-Red=le   derang=pe*
evening=LOC Snow-White=ADDIT Rose-Red=ADDIT 3PL=POSS
umma=samma appi dìkkath=ka arà-duuduk ambel.
mother=COMIT fire close=LOC SIMULT-sit   take
'In the evening, Snow-White and Rose-Red would retire to sit near the chimney with their mother.'

(12) *Derang=pe umma   derang=nang jaith-an=le,   jaarong pukurjan=le*
     3PL=POSS mother 3PL=DAT   sew-NMLZR=ADDIT needle   work=ADDIT
     su-aajar.
     PAST-teach
     'Their mother taught them sewing and needle work.'

(13) *Derang anà-duuduk samma vakthu=ka=le   derang suuka=ka vakthu*
     3PL   PAST-stay   all   time=LOC=ADDIT 3PL   like=LOC time
     anà-empas.
     PAST-spend
     'All the time when they were together, they spent the time very happily.'

(14) *Hathu haari,*[1] *diinging hathu soore=ka,   kìrras pinthu=nang arà-thatti*
     INDEF day   cold   INDEF evening=LOC strong door=DAT   SIMULT-knock
     hathu svaara su-dìnngar.
     INDEF sound PAST-hear
     'One cold evening, they heard a noise of somebody knocking at the door.'

(15) *Itthu   saapa anthi-aada!*
     DEM.DIST who IRR-exist
     'Who would that be (there)!?'

(16) *Snow-White=le   Rose-Red=le   pinthu su-bukka.*
     Snow-White=ADDIT Rose-Red=ADDIT door   PAST-open
     ' Snow-White and Rose-Red opened the door.'

(17) *Sithu=ka, hathu bìssar beecek caaya Buruan*[2] *su-duuduk.*
     there=LOC INDEF big   mud   colour bear   PAST-exist.ANIM
     'There was a big brown bear.'

(18) *Aanak pompang duuva su-thaakuth.*
     child female   two   PAST-fear
     'The two girls got afraid.'

---

[1] The use of an initial ⟨h⟩ in this word is typical of the written variety.

[2] The bear is always capitalized in this text, which might be taken over from the English original.

(19) *"Thussa mà-thaakuth", Buruan su-biilang.*
    NEG.IMP INF-fear   bear   PAST-say
    'The bear said: "Don't get afraid!" '

(20) *"See lorang=nang thama-sakith-kang, lorang see=yang diinging=dering*
    1s   2PL=DAT   NEG.IRR-hurt-CAUS   2PL   1s=ACC   cold=ABL
    kala-aapith;
    cond-protect
    ' "I will not harm you, if you protect me from the cold;'

(21) *see arà-sumpa,   paanas muusing dhaathang thingka see siini=dering*
    1s   NON.PAST-promise hot   time   come   middle 1s here=ABL
    arà[3]-pii.
    NON.PAST-go
    'I promise, soon when the spring arrives I will leave from here.'

(22) *Aanak pompang duuva=le derang=pe umma=le   Buruan=yang ruuma*
    child female   two=ADDIT 3PL=POSS mother=ADDIT bear=ACC   house
    daalang=nang su-panggel.
    inside=DAT   PAST-call
    'The two girls and their mother invited the bear into their home.'

(23) *Buruan diinging abbis=sangke siithu su-siinga.*
    bear   cold   finish=until   there PAST-stay
    'The bear stayed there until the cold was gone.'

(24) *Demikian, thahun[4] baaru muusing=pe kàthaama haari=ka beecek caaya*
    absolutely, year   new time=POSS first   day=LOC mud   colour
    Buruan, incayang=yang baaye=nang anà-kuvaather aanak pompang
    bear   3s.POLITE=ACC good=DAT   PAST-protect child female
    duuva=nang slaamath katha su-biilang.
    two=DAT   greetings QUOT PAST-say
    'Absolutely, on the first day of the new year, this brown bear bid farewell to
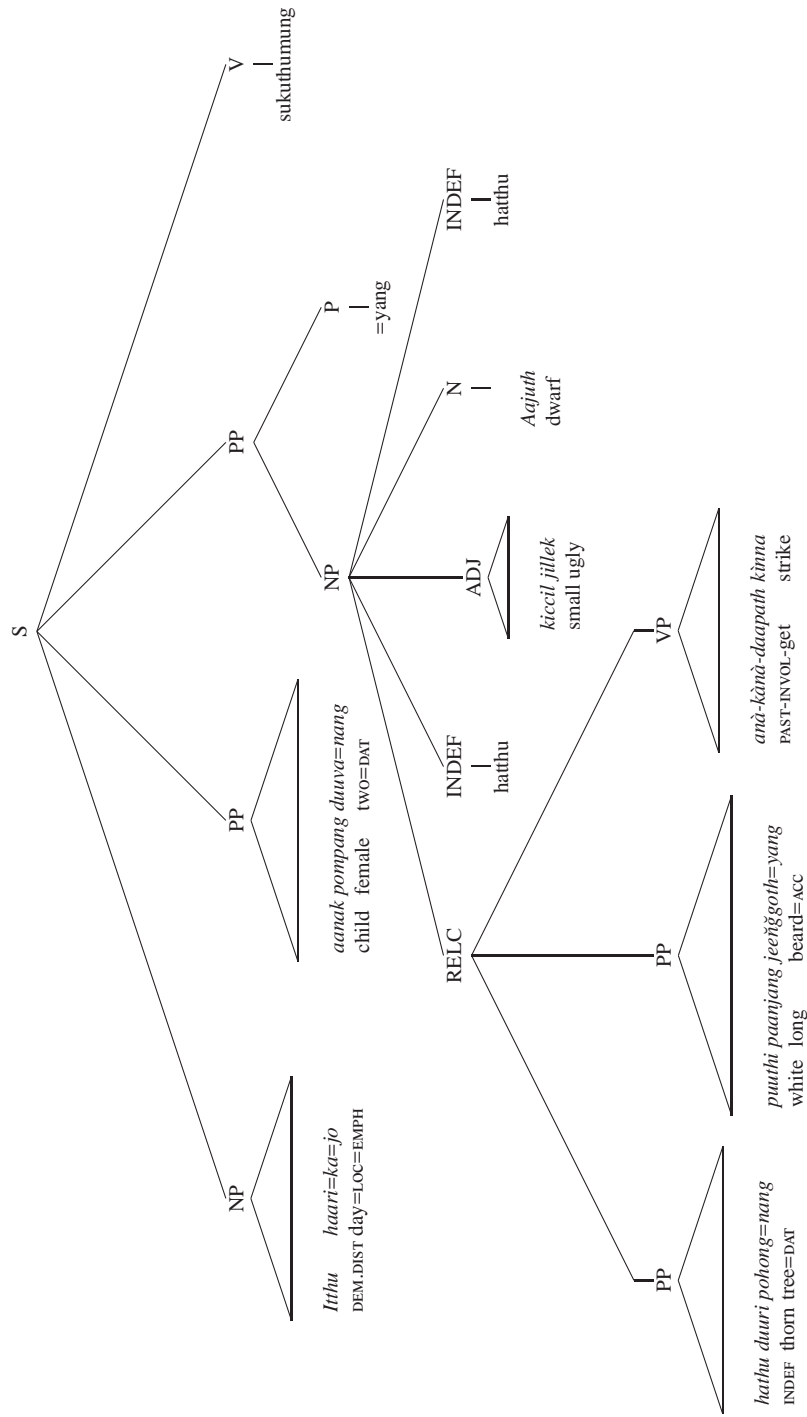   the two girls who had taken good care of him.'

---

[3]The use of the non-past marker *arà-* instead of the irrealis marker *anthi-* confers more commitment to the promise.

(25) *Itthu   haari=ka=jo   aanak pompang duuva=nang [[[hathu duuri*
      DEM.DIST day=LOC=EMPH child female   two=DAT   INDEF   thorn
      pohong=nang] [puuthi paanjang jeeňǧgoth=yang] anà-kànà-daapath
      tree=DAT   white   long   beard=ACC   PAST-INVOL-get
      kìnna] hathu kiccil jillek Aajuth[5] *hatthu=yang su-kuthumung.*
      strike INDEF small ugly dwarf   INDEF][6]=ACC PAST-see
      'On the same day, the two girls saw a small ugly dwarf whose long white
      beard had got stuck in a thorn tree.'

(26) *Derang incayang=yang su-salba-king.*
      3PL   3s.POLITE=ACC   PAST-escape-CAUS
      'They saved him.'

(27) *Incayang, derang=nang thriima jamà-kaasi=nang,   [anà-banthu*
      3s.POLITE 3PL=DAT   thanks NEG.NONFIN-give=DAT PAST-help
      vakthu=ka incayang=pe   jeeňǧgoth=yang asà-thaarek=apa
      time=LOC 3s.POLITE=POSS beard=ACC   CP-pull=after
      incayang=nang su-sakith-kang katha] anà-maaki.
      3s.POLITE=DAT PAST-hurt-CAUS QUOT PAST-scold
      'He did not thank them and instead he scolded them saying that they pulled
      his beard and hurt him while they were helping him.'

(28) *Hathyang aari=ka Snow-White=le   Rose-Red=le   laapang*
      other   day=LOC Snow-White=ADDIT Rose-Red=ADDIT ground
      piiňǧgir=ka, berry *kapang-picca-kang, bìssar hathu buurung derang=pe*
      border=LOC berry when-broken-CAUS big   INDEF bird   3PL=POSS
      atthas=dering su-thìrbang.
      top=ABL   PAST-fly
      'The next day, when Snow-White and Rose-Red were picking berries by the
      side of the ground, there was a big bird flying over them.'

---

[4]The short vowel in *thahun* might be due to it being in a close collocation with *baaru*, thereby being parsed into one phonological word, rather than two. As such, there is no need to build an additional foot in *thahun*, given that one is already present in $b(a_\mu a_\mu)<ru>$.

[5]Like the bear, the dwarf is capitalized throughout the story.

[6]This very complex NP is given as a tree on the next page.

S

NP — *Itthu haari=ka=jo* / DEM.DIST day=LOC=EMPH

PP — *aanak pompang duuva=nang* / child female two=DAT

PP

NP

P — *=yang*

INDEF — *hatthu*

N — *Aqiuth* / dwarf

ADJ — *kiccil jillek* / small ugly

RELC

INDEF — *hatthu*

VP — *anà-kànà-daapath kìnna* / PAST-INVOL-get strike

PP — *puuthi paanjang jeeñggoth=yang* / white long beard=ACC

PP — *hathu duuri pohong=nang* / INDEF thorn tree=DAT

V — *sukuthumung*

(29) *Banthu-an asà-mintha arà-naangis svaara hatthu, derang=nang*
    help-NMLZR CP-ask   SIMULT-cry   sound INDEF 3PL=DAT
    su-dìnngar.
    PAST-hear
    'They heard a noise crying and asking for help.'

(30) *Derang su-kuthumung ithu   buurung=pe kuuku=ka Aajuth asà-sìrrath*
    3PL   PAST-see   DEM.DIST bird=POSS   claw=LOC dwarf CP-stuck
    kìnna arà-duuduk.
    strike SIMULT-exist.ANIM
    'They saw the dwarf stuck in the claws of the bird.'

(31) *Aajuth=yang buurung mà-angkath baapi   su-diyath.*
    dwarf=ACC   bird   INF-lift   take.away PAST-try
    'The bird was trying to take the dwarf with him.'

(32) *[[Kaaki=ka gaaňdas-kang ambel anà-duuduk   Aajuth]=yang*
    leg=LOC   tie-CAUS   take PAST-exist.ANIM dwarf=ACC
    sangke-luppas], [hathu pollu=dering] [Rose-Red] [buurung=nang]
    until-leave   INDEF stick=ABL   Rose-Red bird=DAT
    [su-puukul].
    PAST-hit
    'Rose-Red whacked the bird with a stick until the bird let loose of the dwarf
    who was held under the legs of the bird.'

(33) *Laiskali, inni   Aajuth [incayang=pe jiiva anà-salba-king   itthu*
    again   DEM.PROX dwarf 3S.POLITE=POSS life PAST-escape-CAUS DEM.DIST
    aanak pompang duuva=]nang thriima thàrà-kaasi.
    child female two=DAT   thanks NEG.PAST-give
    'Once again, the dwarf did not thank the two girls who had saved his life.'

(34) *"Lorang=pe guuna thàràsampe! hathyang skali bannyak masà-aapith!"*
    2PL=POSS   respect not.enough other   time much   must-care
    Aajuth su-butharak.
    dwarf PAST-scream
    ' "Your respect is not enough! Next time, you have to be more careful!" the
    dwarf shouted.'

(35) *Blaakang, derang Aajuth=yang laiskali su-kuthumung, incayang=ka*
after 3PL dwarf=ACC again PAST-see 3S.POLITE=LOC
[[bannyak panthas ummas baarang pada=le, bathu inthan[7] pada=le
much beautiful gold item PL=ADDIT stone shine PL=ADDIT
anà-punnu-kang] bìssar beecek caaya hathu bag] su-aada.
PAST-full-CAUS big mud colour INDEF bag PAST-exist
'Later, they saw the dwarf again, he had a big brown sack of very beautiful
golden items and gems with him.'

(36) *Incayang ithu pada=yang kapang-thumpa-king, itthu pada sraathus*
3S.POLITE DEM.DIST PL=ACC when-spill-CAUS DEM.DIST PL 100
binthan pada arà-kiilap=ke su-kiilap.
stars PL SIMULT-shine=SIMIL PAST-shine
'When he spilled them, they shone like a hundred shining stars.'

(37) *Itthu vakthu=ka hathu bìssar beecek caaya Buruan mlaarath=ka*
DEM.DIST time=LOC INDEF big brown colour bear difficulty=LOC
uuthang=dering luvar=nang su-dhaathang.
forest=ABL outside=DAT PAST-come
'At the same time, a big brown bear came out of the woods.'

(38) *[Anà-kaageth kìnna Aajuth] mà-laari=nang su-baalek,*
PAST-startle strike dwarf INF-run-DAT PAST-turn
'The shocked dwarf got back to run.'

(39) *Buruan incayang=yang asà-thoolak, ummas gooni dìkkath=nang su-pii.*
bear 2S.POLITE=ACC CP-push gold bag vicinity=DAT PAST-go
'The bear pushed the dwarf aside and went up to the sack.'

(40) *Aajuth thaakuth=ka su-naangis,*
dwarf fear=LOC PAST-weep
'The dwarf wept in fear:.'

---

[7]Note the short vowel in the first part of the compound *bathu inthan* 'gems'.

(41) *"See=yang luppas, Thuan Buruan.*
      1s=ACC   leave   Sir   bear
      ' "Please release me, Mister Bear.'

(42) *Lorang se=dang mà-hiidop thumpath kala-kaasi, see lorang=nang*
      2PL   1s=DAT INF-stay place   if-give   1s 2PL=DAT
      lorang=pe samma duvith=le,   baarang pada=le anthi-bale-king"
      2PL=POSS all money=ADDIT item   PL=ADDIT IRR-turn-CAUS
      'If you give me a chance to live, I will return back all your money and your
      things." '

(43) *Snow-White=nang=le   Rose-Red=nang=le   ini   hatthu=ke*
      Snow-White=DAT=ADDIT Rose-Red=DAT=ADDIT DEM.PROX INDEF=SIMIL
      thàrà-mirthi.
      NEG.PAST-understand
      'Neither Snow-White nor Rose-Red understood a single thing.'

(44) *Derang anà-kuthumung pada=nang asà-thaakuth, ruuma=nang mà-laari*
      3PL   PAST-see   PL=DAT   CP-fear   house=DAT INF-run
      kapang-pii, derang=nang byaasa svaara hatthu su-dìnngar.
      when-run 3PL=DAT   habit sound INDEF PAST-hear
      'They only got afraid at what they saw, and when they were about to run back
      home, they heard a familiar voice.'

(45) *"Snow-White, Rose-Red, thussa mà-thaakuth!"*
      Snow-White   Rose-Red NEG.IMP INF-fear
      'Snow-White, Rose-Red do not get afraid!'

(46) *Duuduk! See=le   lorang=samma arà-dhaathang".*
      stay   1s=ADDIT 2PL=COMIT   NON.PAST-come
      'Wait! I am coming with you, too." '

(47) *Derang anà-baalek sajja=jo,   sithu=ka panthas hathu Aanak raaja*
      3PL   PAST-turn   only=EMPH there=LOC beautiful INDEF child   king
      su-aada!
      PAST-exist
      'When they turned around, there was a beautiful prince!.'

(48) *Buruan=pe kuulith incayang=pe   kaaki baava=ka   jaatho su-aada!*
bear=POSS skin   3s.POLITE=POSS leg   bottom=LOC fall   PAST-exist
'The skin of the bear had fallen at his feet!.'

(49) *Aanak raaja su-biilang,*
child   king   PAST-say
'The prince said.'

(50) *"Lorang=nang see=yang ingath-an[8]=si?*
2PL=DAT   1s=ACC   think-NMLZR=INTERR
' "Do you remember me?"'

(51) *Diinging vakthu=ka ruuma daalang=nang anà-ambel bìssar beecek caaya*
cold   time=LOC house inside=DAT   PAST-take big   mud   colour
buruan.
bear
'The big brown bear whom you gave shelter at your home during the cold
season (is me).'

(52) *See hathu bìnnar Aanak raaja.*
1s   INDEF real   child   king
'I am a real prince.'

(53) *Sdiikith thaaun=nang duppang, see ini   Aajuth=nang*
few   year=DAT   front   1s   DEM.PROX dwarf=DAT
su-kìnna-daapath.
PAST-INVOL-get
'A few years before, I was captured by this dwarf.'

(54) *Ini   Aajuth se=ppe baapa=nang su-simpa,   hathu ummas gooni*
DEM.PROX dwarf 1s=POSS father=DAT PAST-promise INDEF gold   bag
kala-kaasi, see=yang athi=kasi-bìrrath katha.
if-give,   1s=ACC   IRR-give-heavy[9]   QUOT
'This dwarf promised to my father that he would return me to my father, if my
father gave him a sack of gold.'

---

[8]Note the use of a nominalized form here, instead of the verb *arà-iingath* 'remember'.

(55) *Se=ppe baapa incayang=nang ummas su-kaasi.*
      1s=POSS father 3s.POLITE=DAT gold PAST-give
      'My father gave him gold.'

(56) *Itthule see=yang mà-kiiring=nang duppang incayang see=yang hathu*
      but   1s=ACC INF-send=DAT front 3s.POLITE 1s=ACC INDEF
      Buruan mà-jaadi   su-bale-king.
      bear   INF-become PAST-turn-CAUS
      'Yet, before sending me back, he changed me into a bear.'

(57) *[Se=ppe oorang pada [see saapa katha] thàrà-thaau]=subbath   see=yang*
      1s=POSS man   PL   1s who QUOT NEG.PAST-know=because 1s=ACC
      su-uubar."
      PAST-chase
      'Because my people did not know who I was and they chased me." '

(58) *"Lorang=le, lorang pada=pe umma=le   see=yang baaye=nang*
      2PL=ADDIT   2PL   PL=POSS mother=ADDIT 1s=ACC   good=DAT
      anà-kuaather,
      PAST-care
      'You two and your mother treated me very well.'

(59) *see=yang lorang=susamma diinging muusing sangke-habbis anà-simpang*
      1s=ACC   2PL=COMIT   cold   time   until-finish PAST-keep
      ambel.
      take
      'You all kept me until the cold season ended.'

(60) *See lorang=nang arà-simpa   kaapang=ke see lorang=nang ithu*
      1s 2PL=DAT   NON.PAST-promise when=SIMIL 1s 2PL=DAT   DEM.DIST
      uuthang arà-baayar   katha.
      debt   NON.PAST-pay QUOT
      'I promise you someday I will pay back that debt.'

---

[9]*Bìrrath kaasi* is an idiom meaning 'to admit, to accept'.

(61) *Lorang see=subbath ithu   Aajuth=yang su-salba-king*
       2PL   1s=because DEM.DIST dwarf=ACC   PAST-escape-CAUS
       'You all saved the dwarf because of me.'

(62) *Itthu=nang   blaakang=jo se=dang laiskali se=ppe bìnnar mosthor*
       DEM.DIST=DAT after=EMPH 1s=DAT again   1s=POSS real   manner
       anà-jaadi,
       PAST-become
       'After that only, I returned to my true self.'

(63) *suda lorang=yang ruuma=nang anthi-aaji.baapi*[10]
       thus 2PL=ACC   house=DAT   IRR-take.away.ANIM
       'So, I will take you back home.'

(64) *Thapi, boole lìkkas=ka*[11] *see lorang=yang mliige=nang anthi=panggel"*
       but   can quick=LOC 1s 2PL=ACC   palace=DAT IRR=call
       'But, as soon as possible I will invite you to the palace.'

(65) *[Aanak raaja=pe   perkathahan=yang] [Snow-White=nang=le*
       child   king=POSS word=ACC   Snow-White=DAT=ADDIT
       Rose-Red=nang=le   suka-han=dering su-punnu hathu hidopan]
       Rose-Red=DAT=ADDIT like-NMLZR=ABL PAST-fill INDEF prospect
       su-thunjiking.
       PAST-show
       'The words uttered by the prince showed prospects which filled Snow-White
       and Rose-Red with delight.'

(66) *Sdiikith haari=nang blaakang ini   sudaari duuva=nang*
       few   day=DAT   after   DEM.PROX sister two=DAT
       mliige[12]*=dering su-panggel.*
       palace=ABL   PAST-call
       'After a few days, these two sisters were invited by the palace.'

---

[10]*Aaji* occurs only in collocation with *baapi* 'take (away)', where it indicates that the theme is animate.
[11]*Boole lìkkas=ka* is an idiom meaning 'As soon as possible'.
[12]The word for 'palace' is sometimes written *mliiga*, sometimes *mliige*.

(67) *Aanak raaja laiskali anà-dhaathang=subbath, samma oorang bannyak*
child king again PAST-come=because all man much
suuka=ka anà-duudu.[13]
like=LOC PAST-exist.ANIM
'Because the prince had returned, all his subjects were very happy.'

(68) *Sdiikith haari=nang blaakang Snow-White Aanak raaja=yang=le,*
few day=DAT after Snow-White child king=ACC=ADDIT
Rose-Red incayang=pe sudaara=yang=le su-kaaving.
Rose-Red 3s.POLITE=POSS brother=ACC=ADDIT PAST-marry
'Some days later, the prince was married to Snow-White and his brother was
married to Rose-Red.'

(69) *Derang=pe umma=le derang=samma hatthu=nang*[14] *mà-hiidop=nang*
3PL=POSS mother=ADDIT 3PL=COMIT one=DAT INF-stay=DAT
mliiga=nang su-dhaathang.
palace=DAT PAST-come
'Their mother also came to live together with them in the palace.'

(70) *Ruuma duuva subala=ka su-aada rooja pohong komplok duuva=yang*
house two side=LOC PAST-exist red tree bush two=ACC
asà-baa=apa mliige=pe duuva subla=ka su-thaanàm.
CP-bring=after palace=POSS two side=LOC PAST-plant
'The two rose bushes which were at one time standing at both sides of the
home were brought and planted on both sides of the palace.'

(71) *Maana thaahun=ka=le ini rooja pohong komplok duuva*
which year=LOC=ADDIT DEM.PROX rose tree bush two
kumbang=dering arà-punnu
flower=ABL NON.PAST-full
'Every year, these two bushes fill with beautiful roses.'

---

[13]The form *duuduk* with a *k* would be more common in written texts.
[14]*Hatthunang* has a lexicalized meaning, 'together'.

(72) *Anà-ambel deri*[15] *:-THE BROTHERS GRIMM.*
    PAST-take   from   :-THE BROTHERS GRIMM
    'taken from The Brothers Grimm .'

(73) *Mlaayu=nang anà-bale-king :- SALIM.*
    Malay=DAT   PAST-turn-CAUS :- SALIM
    'translated into Malay: Salim.'

---

[15]*Deri* is used as a preposition here, which cannot be explained as of now.

Voice Onset Time

The following figures show the voice onset times of several consonants in initial and intervocalic positions. The pitch contour can be used as an indicator of voicing: without voice, there is no pitch. The formants indicate the vocalic phases. The offset between the beginning of the pitch contour and the beginning of the vocalic phase is taken here as the determinant of voice onset time.

Generally speaking, SLM voiceless stops have a voice onset time of close to 0ms. Voiced stops have a negative VOT, and tend to be fully voiced, i.e. the voicing does not cease between vowels preceding and following the voiced stop.

An interesting fact is that voiced stops tend to have a negative VOT even if they follow the voiceless fricative /s-/.

The figures were selected to cover a maximum of the parametric variation possible for [±voicing], [±geminate], [±initial], [± intervocalic], and [±following s].
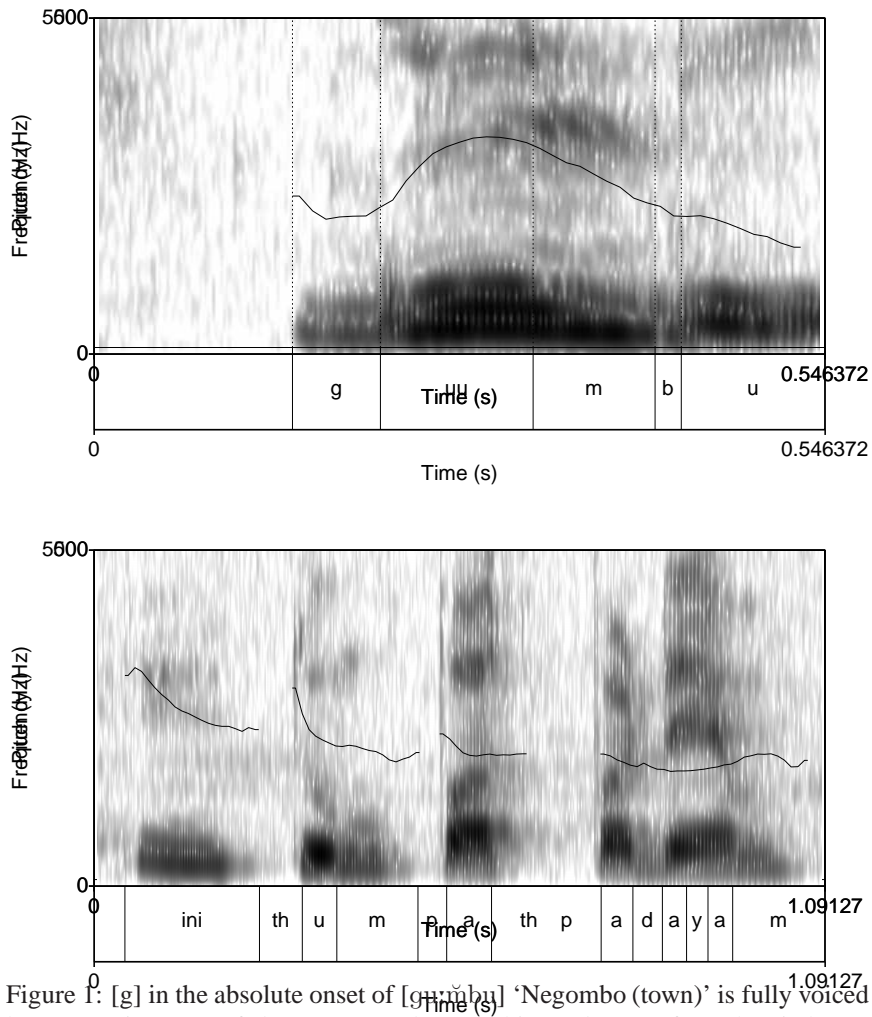
Figure 1: [g] in the absolute onset of [gɐːŋ̆bu] 'Negombo (town)' is fully voiced and has a negative VOT of about 0.068s (above). This can be seen from the pitch contour, which is present during the whole articulation of g. Since pitch can only be measured on voiced parts, the pitch contour can be used to identify voicing. The image below shows that the pitch contour is absent for [t̪] and [p], but voicing sets in as soon as closure of the mouth is released and the vocalic part begins, i.e. we have a voice onset time of close to zero in the onsets [t̪u-, pa-, pa-].
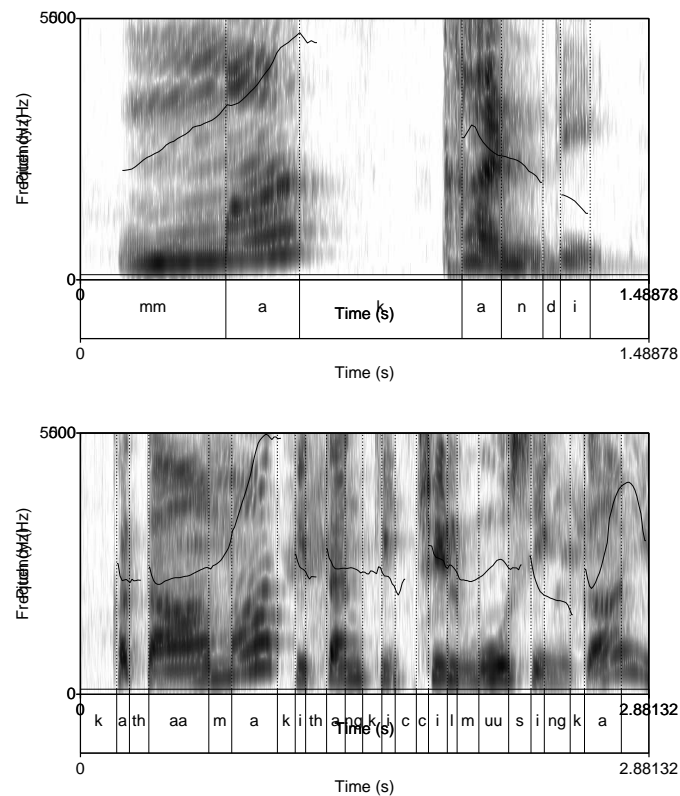
Figure 2: It is rare to find aspirated consonants in SLM, identified by long VOTs, like the [k] in [kan̪d̪i] 'Kandy (town)' the example above, which has a VOT of 0.049s. One can see from the black shading that the closure is already released, but the voicing has not started yet (indicated by the pitch line). It is much more common to have VOTs under 0.020s, as is the case with all the stops in the example below.
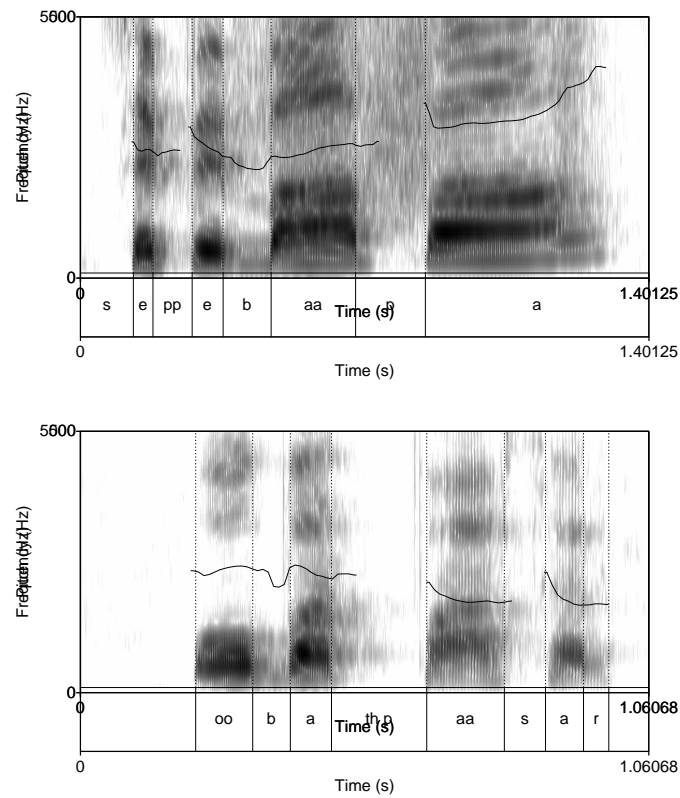
Figure 3: [p] in intervocalic position [a:pa] has a VOT close to zero. [b] in intervocalic position between two morphemes is fully voiced. (above). [b] in morpheme-internal intervocalic position is fully voiced, /p/ in the onset of [pa:sar] has a VOT close to zero (below).
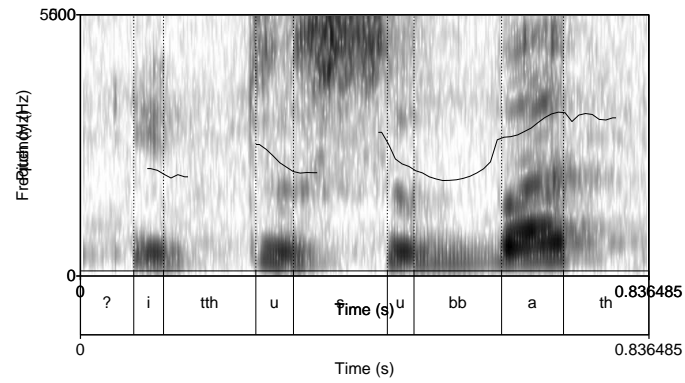
Figure 4: The long dental stop in [it̪:u] has a VOT close to zero, the long labial stop in [sub:at̪] is fully voiced.
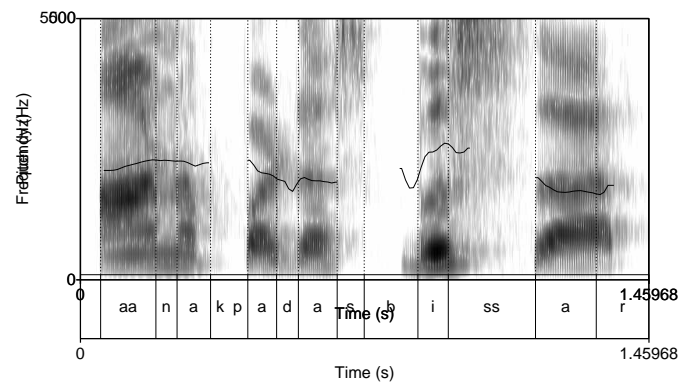

Figure 5: Even after the voiceless fricative [s-], /b/ has a negative VOT of -0.045s. [p] in [pada] has a VOT close to zero.
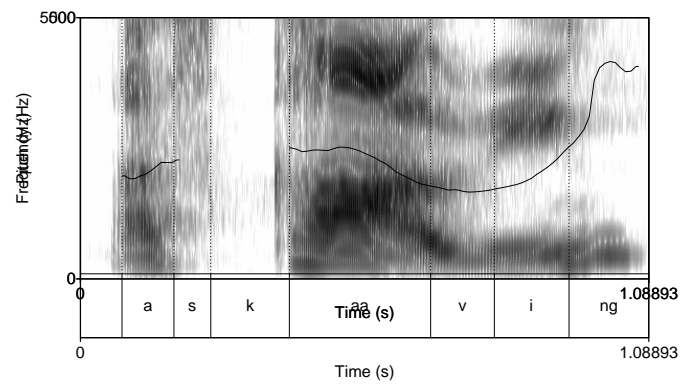
Figure 6: After the voiceless fricative /s/, [k] has a VOT of 0.027s.

Marriage partners

The following table contains all marriages listed in Burah (2006). The first column indicates the descendant of an exile, the second one his or her marriage partner. Some individuals have married more than once. In the first two columns, the names *Gnei, Tuan, Baba, Noona,* and *Mohammed* are abbreviated by their first letter. The remaining text is left untouched, including all idiosyncratic spellings. All abbreviations other than mentioned above are present in the original text. The third column indicates to what ethnic group the partner of the descendent belongs. The following codes are used:

- M: Muslim

- My: Malay

- B: Buddhist

- C: Christian

- H: Hindu

- S: Sinhalese

- T: Tamil

- E: Burgher (European)

- X: Chetty (Indian traders settled in Sri Lanka)

Adjumain Abdul Rahim ∞ G Zamzam Doole	My
Allon Abidin Aleph Sourjah ∞ Amy Grant	CE
Allon Ahamadin Aleph Sourjah ∞ Tamil lady	HT
Allon Akim Sourjan ∞ Sinhalese lady	BS
Allon Halaladin Aleph Sourjah ∞ G Kitchill De Wongso	My
Anggranie Amit ∞ Shiran Wijetunga	BS

Azan Jainudeen ∞ Lorraine	CE
B Abu Cassim Burah ∞ N Araby Muhammed Cassim	My
B Alfi Bahar ∞ Aloma Pietersz	CE
B Angus Perdana Aleph Sourjah ∞ Khairunissa Aniff	M
B Bakki Dole ∞ N Umma	My
B Boorung Ahamath Cassim ∞ G Zuhar Ismail	My
B Boorung Ahamath Cassim ∞ Miss Miskin	My
B Dain Abu Cassim ∞ G Andali Dole	My
B Dain Abu Cassim ∞ unspecified Tamil mother	HT
B Dain Murja Bahar ∞ Wilson Gunaratna	CS
B Dain Sirja Bahar ∞ Ondaatje	X
B Dayon ∞ Puthi Nallawangsa	My
B Dharma Ahamath Cassim ∞ N Ajeema Yusoof	My
B Dowlath Mohammed Aleph Sourjah ∞ Mufeeda Junoos	M
B Fitr Cassim Burah ∞ G Farida Careem	My
B Hadjirin Doole ∞ Packeer Saibo	My
B Hadjirin Doole ∞ Zamzam Miskin	My
B Hakim Ahamath Cassim ∞ Sitti Nihara	My
B Hakim Ahamath Cassim ∞ De Wangsa	My
B Hanif Abu Cassim Burah ∞ G Zama Muhammath Cassim	My
B Hareera Ahamath Cassim ∞ Kitchill	My
B Hureira Cassim Burah ∞ Sellamah (transvestite)	?
B Indra Moela Packeerally ∞ Shirley Packeerally	M
B Ishan Bahar ∞ Dawn Pietersz	CE
B Ismail Bahar ∞ Hilerin Junkeer	?
B Jalal-ud-din Doole ∞ Sarba N Jaimon	My
B Jaleel ABu Cassim Boera X ∞ N Soeri Hadji Bahar	My
B Jamal Jumar ∞ G Dasson Passela	My
B Jayah Abdeen Cassim ∞ G Suriyan Yusoof	My
B Johar Hadji Bahar ∞ Sonia Clausz	CE
B Jumar Hadji Bahar ∞ Zamzam Amath	M
B Junior Burah ∞ Yasmin Amith	M
B Junior Hadji Bahar ∞ Sitti Roekiya Mahath	My
B Junoor Hadji Bahar ∞ N Ardani Cassim (2nd marriage)	My
B Junoor Hadji Bahar ∞ Dane Sirri	?
B Junoor Hadji Bahar ∞ G Jammo Abu Cassim Burah	My
B Kamal-ud-din Doole ∞ N Sufiya	My
B Kereish Cassim Burah ∞ Ummu Jameela Ismail	M

B Kitchil Cassim Burah ∞ N Davy Amoo	My
B Kumar Edward Burah ∞ G Drene Drahin	My
B Lahirin Taliph ∞ G Tasmania Doole	My
B Lambrath Mohammed Aleph Sourjah ∞ Ms Junitha J Amath	M
B Lambrath Mohammed Aleph Sourjah ∞ Junitha Janninoor Ahamath	M
B Mahamoor Dolea Padan ∞ Kitchill Weerabangsa	My
B Mansoor Doole ∞ G Janeena	My
B Mantara Doole ∞ Miss Jamion	M
B Moon Doole ∞ Miss Hadji	M
B Moondu Doole ∞ G Maas	My
B Muhalkin Cassim Burah ∞ G Rakkiath Noordeen	My
B Muhammath Sabreen Aleph Sourjah ∞ G Boodi Taliph	My
B Muhinudeen Pacheerally ∞ Jameela Saheed	M
B Musaddam Abu Cassim Boera ∞ N Araby Muhammed Cassim	My
B Naim Abu Cassim Burah ∞ G Endran Doole	My
B Naim Abu Cassim Burah ∞ G Haniffa Packeerally	My
B Noeri Abu Cassim Burah ∞ Ummu Sahida Singhawangsa	My
B Nuhara Cassim Burah ∞ T Shareef Hassan	My
B Oumar Mohammed ∞ Miss Johar	M
B Pajon Lye ∞ Miss Jayman	?
B Permuda Aleph Sourjah ∞ Makkie Mohammed	M
B Putra Cassim Burah ∞ G Binthan Hassan	My
B Rahim Ahamath Cassim ∞ Sitti Haleema	My
B Ramzan Burah ∞ Rozanna Khan	M
B Ranawijo Doole ∞ G Moeda Abu Cassim Burah	My
B Reendu Doole ∞ Miss Dole	My
B Regent Packeerally ∞ G Elmoon Achon	My
B Rubin Jumar ∞ G Isnain Doole	My
B Saam Jumar ∞ G Kitchill Jaymon	My
B Sakkaf Ahamath Cassim ∞ N Mashoor Yusoof	My
B Shareef Abu Cassim Burah ∞ G Kitchil Junoos Jayan	My
B Soema Ahamath Cassim ∞ N Jammin Doole	My
B Tajul Arfin Dole ∞ Naima Umma	M
B Tariq Cassim Burah ∞ G Endran Hassan	My
B Tasnim Aleph Sourjah ∞ Ayesha Satheen	M
B Thajone Abdeen Ahamath Cassim ∞ G Kitchill Mutaliph	My
B Thaklin Ahamath Cassim ∞ G Boolan Ahamath Cassim	My
B Thyib Mohammed ∞ Zaitoon Samath	M

B Zalfin Mohammed Aleph Sourjah ∞ Razone Barthole	M
B Zoon Taliph ∞ N Jarri Sadikin	My
Binthan (Staria) Aleph Sourjah ∞ B Muthalip Juheer Careem	My
Camaldin Aleph Sourjah ∞ Manja N Doole	My
Camaldin Aleph Sourjah ∞ N Tasminia Doole	My
Canissa Jamdin ∞ Rizvi Jainudeen	My
Corinna Hamid ∞ Razeena Mannan	M
Dain Juwitha Abdeen Cassim ∞ T Yehiya Doole	My
Dane Juwitha Ahamath Cassim ∞ T Yehiya Doole	My
Dani Dole ∞ J. Dawood	?
Dasson N ∞ Mohammed Muthan	M
De Pangeran Soeris (Soerjah) ∞ Din Cujo Ajeone	M
Dilano Burah ∞ Dilani Yusoof	M
Dorain Doella Padan ∞ Cassiere	My
Eileen Haleema Aleph Sourjah ∞ T Mohammed Dane	My
Esmin Sourjah ∞ Gonsalkorale (Australia)	?
F Radin Shiela (G Soeli) ∞ Outschoorn	CE
Fazil Sourjah ∞ Sharmila	M
Ferial Jamdin ∞ Ameer Abdeen	M
G Daleel Cassim Burah ∞ T Bagoes Packeerally	My
G Abinon Abu Cassim Boera ∞ Alim Abdul Rahim Adjumain	M
G Ayn Dole (N Injan Dole) ∞ B Muhammath (B Shroff)	My
G Bagoes Abu Cassim ∞ B Amith Doella Padan	My
G Bagoes Mohammed ∞ Aziz Riaz Ahamath	M
G Bagoos Abu Cassim Boera ∞ B Amith Doella Padan	My
G Bidoora ∞ T Nizam Hanifa	My
G Bidoora Jainudeen ∞ TAM Dain	My
G Binthari Mohammed ∞ T Jaseer Hassan (Tony Hassan)	My
G Bistani Burah ∞ T Harris Dole	My
G Boodi Doole ∞ T Buhary Abu Sallay Dole	My
G Boodi Taliph ∞ T Sabreen Sourjah	My
G Boolan Burah ∞ Ruby Meedin	?
G Buhari Ahamath Cassim ∞ T Noer Doole	My
G Dain Hadji Bahar ∞ T Kitchil Burah	My
G Dason N ∞ Maas Roteh Bangsajayah	My
G Davy N Cassim ∞ T Johore Ismail	My
G Dileema Ahamath Cassim ∞ T Bagoes Yusoof	My
G Edaham Aleph Sourjah ∞ T Mohammed Deen Miskin	My

G Fareeda Buckker ∞ T Victor Jaimon	My
G Ferza Hartini Aleph Sourjah ∞ MIM Nawaz	M
G Freeda Burah ∞ Hafiz Jainudeen	My
G Fuwaizi Jayah ∞ Ibrahim Mohammed	M
G Gunara Jayah ∞ TJ Dole	My
G Jammo Abu Cassim Boera ∞ B Junoor Hadji Bahar	My
G Jarri Abu Cassim ∞ Maas T Kitchil Jayah	My
G Jarri Abu Cassim ∞ T Maas Kitchil	My
G Jarri Burah ∞ TB Jayah	My
G Javi N Doole G Boodi ∞ B Ramblan Oumar Taliph	My
G Javi N Doole G Boodi ∞ T Buhary Abu Sallay Dole	My
G Kartini Burah ∞ Deen Hamid	M
G KAssima Abu Cassim Boera ∞ B Noer Mohamed Doella Pandan	My
G Kassima Mohammed ∞ B Halith Jurampati	My
G Kitchil Lye ∞ Maas T Kitchil Weerabangsa	My
G Kitchill Aleph Sourjah ∞ T Arifin Buxckker	My
G Kitchill Doella Padan ∞ T Arfin Bakkar	My
G Kitchill Doella Padan ∞ Hassan Mohammed Thaha Johar	M
G Moeda Abu Cassim ∞ B Zamzam Doole	My
G Mooda Abu Cassim Boera ∞ B Zain Doole	My
G Mooda Abu Cassim Burah ∞ B Ranawijo Doole	My
G Naima Abu Cassim Burah ∞ Maas Latiff Cassim	My
G Nishana Jayah ∞ Jisthi Kitchilan	My
G Noon Abu Cassim Burah ∞ M Zain Packeerally	M
G Pearl Zareena Burah ∞ T. Faleel Is Ismail	My
G Queenie Burah ∞ TH Jamdin	My
G Queenie Burah ∞ MH Amith	M
G Raleen Burah ∞ B Razick Dole	My
G Razeena Jayah ∞ MH Amit	M
G Rizanz Aleph Sourjah ∞ Sheriffdeen Camaldeen Meedin	My
G Rowena Jayah ∞ Enver Crayon Aliph	My
G Sabooriya Ahamath Cassim ∞ Amir Feisal Thahir	M
G Seenar Ahamath Cassim ∞ MA Sheriffdeen Doray	My
G Seenar Burah ∞ Mohammed Thajudeen Jainudeen	My
G Seenar Burah ∞ T Tariq Rumi Jainudeen	My
G Seenar Doole ∞ MCS Mohammed	M
G Seenar Taliph ∞ TY Amith	My
G Shareefa Cassim Burah ∞ T Johore Ismail	My

G Shirlene Burah ∞ Jurangpathy	My
G Sitara Aleph Sourjah ∞ Che A Wahab	My
G Sulaiha Mohammed ∞ T Moeji Sallay	My
G Surahi Bakkar ∞ Assan Marikkar	M
G Tasnin Doole ∞ T Rumi Jainudeen Samath	My
G Wangsa Aleph Sourjah ∞ T Marhoon De Wongsa	My
G Wasseela Burah ∞ Hamid	M
G Yohan Ahamath Cassim ∞ ATM Johar	My
G Zahuriya Ahamath Cassim ∞ Singhawangsa	My
G Zahuriya Ahamath Cassim ∞ TS Bakkar	My
G Zamzam Jaamiya Doole ∞ Adjumain Abdul Rahim	M
G Zanorra Burah ∞ T Masnavi Jaya	My
G Zarine Doole ∞ TK Moomin	My
G Zaveeni Jainudeen ∞ Dr H Kumarasingha	BS
G Zeenar Burah ∞ Mohammed Thajudeen Jainudeen	My
Hareera Ahamath Cassim ∞ N Sundari Kitchill	My
Hareera Mohammed Thaha Johar ∞ Miss Dole	My
Hareera Mohammed Thaha Johar ∞ Ummu N Johar	My
Hazare Sourjah ∞ Peiris (Australia)	S
HilMy Burah ∞ Charmaine Samsudeen	My
Indrani Meedin ∞ Francis Wickramasuriya	CS
Inthikab Jamdin ∞ Zahira	M
Jean Aleph Sourjah ∞ Lucy BalahaMy	?
Jebeye Doella (B Ibrahim Dole) ∞ N Sabina Abdul Cassim	My
Jebeye Doella (B Ibrahim Dole) ∞ Radin Adjan	My
Kamaldin Aliph Sourjah ∞ Zubaid-ur-Rahman Saleem	M
Kheroon Sohnali Burah ∞ Chanaka Francis de Livera	CS
Kitchil Hamid ∞ Carina Saldin	My
Maas Binthan Kitchill ∞ Dole	My
Maas Chinda Amoo ∞ Wazeera Kanaka	M
Maas Dooray Amoo ∞ B Kitchill Cassim Burah	My
Maas Horoun Aleph Sourjah ∞ Shirani Aher	M
Maas Thymen Amoo ∞ Thilak Maligaspe	?
Maas T Kitchil Jayah ∞ N Rasmani Lye	My
Machan Boera Abu Cassim ∞ K Radin Kaija Adjie Koesoema	My
Manja N Doole ∞ B Thassim Doole	My
Manja N Doole ∞ Zahuriya Jaimon	?
Manouri Kumar Sourjah ∞ Paranagama	S

Mina ∞ Dharmavan Miskin	My
Mohammed Mamudeen Aleph Sourjah ∞ N Jammil Bahoran	My
Mohammed Mamudeen Aleph Sourjah ∞ N Zulaika Bahoram	My
Mohammed Zailanee Ahamath Cassim ∞ Sitti Ruhina Hamid	My
Mohideen Seenar Ahamath Cassim ∞ Kadijah	M
Naim Panangan Kassier Buncoes ∞ Radin Adjan	M
Nazeer Aleph Sourjah ∞ Marins Satheen	?
Nazrij Jainudeen ∞ Ranjith Gunawardane	BS
N Amsiya Doella Padan ∞ Mudliyar B Hakim Mutaliph	My
N Araby Muhammath Cassim ∞ B Musaddam Cassim Burah	My
N B. Aleph Sourjah ∞ ULM Haniffa	M
N Dharmani Jumar ∞ Abdul Cader Mohammed	My
N Gyone Hadji Bahar ∞ Hassan Mohammed Thaha Johar	M
N Gyone Hadji Bahar ∞ T Arfin Bajjar	My
N Injan Doella Padan ∞ B Zain Doole	My
N Injan Dole ∞ B Ahamath (Wirawila B) My	
N Injan Muhammath Cassim ∞ Zainul Abdeen Ahamath Cassim	My
N Jammin Doole ∞ B Sooma Ahamath Cassim	My
N Juwari Cassim Burah ∞ Noor Hassan	My
N Kitchill Jumar ∞ Cassim	My
N Margaret Vermeer ∞ Theidore Adrianne Johannes	CE
N Mera Doole ∞ M Hussain Mandore Jaimon	M
N Rameena Haniffa ∞ T Faliq Meedin	My
N Rohani Ahamath Cassim ∞ TM Thajudeen Hamid	My
N Sabrina Buncoes ∞ Siddiq Naina Muhammath Cassim	My
N Soeri Hadji Bahar ∞ B Jaleel Cassim Burah	My
N Sundari Muhamath Cassim ∞ B Oumar Jumar	My
N Sundari Muhammath Cassim ∞ B Oumar Jumar	My
N Tasmania Doole ∞ Kamaldin Aliph Sourjah	My
N Zarona ∞ Maaz Thumban Amoo	My
N Zavina Aleph Sourjah ∞ Maas Thuban Amoo	My
Macham Boera Abu Casum ∞ Radin Kaidja Ajie Koesoema	My
Perdana Wahab Aleph Sourjah ∞ Ayn Aiffin	M
Raden Sitti Dharmavan Sourjah ∞ Mina	?
Raja Oesman ∞ Sitti Habiba	My
Ranizie Januideen ∞ Rohisa Burah	My
Rehana Jainudeen ∞ Frank David	CS
Riza Aleph Sourjah ∞ Jean Jurampathy	My

Rizwi Jainudeen ∞ Canissa Jamdin	M
Rohan Hanmid ∞ Farhana Zail	M
Roshan Hamid ∞ Faizal Amit	M
Rosie Jainudeen ∞ Imtiaz Mazahry	M
Saiyone Ahlip ∞ Amin Deane	M
Sandhya Kumar Sourjah ∞ Iriyapola	?
Shahnaz Ahlip ∞ Fauzul Khalid	M
Shamil Burah ∞ Zulfi	M
Shamsheer Jamdin ∞ Farzana Hashim	M
Shanaez Hamid ∞ O. Fernando	S
Shihaan Dole ∞ BF Ahlip	My
Shiraz Dole ∞ Imara Ahlip	My
Shirley Rizana Meedin ∞ T Sumari Amit	My
Sitti Devi N Fallil ∞ B Lamboong Doole	My
Sitti Havang ∞ Pangaran Adipati Amangkoerat	My
Sitti Zarina Ahamath Cassim ∞ T Mohammed Hakim Thajudeen	My
Sooriadharma Tasnin ∞ Farleen Jaldin	My
Soraya Jamdin ∞ ... Alvares	?
T Hafiz Jainudeen ∞ G Freeda Burah	My
T Haroun Burah ∞ Binthara Samahon	M
T Hussain Jayah ∞ Freeda Yussoof	M
T Merza Aleph Sourjah ∞ Fatima	M
T Taliph Jayah ∞ Indreswary Arulambalam	HT
T Thajudeen Ahamath Cassim ∞ Sharma Luthfiya Jaimon	M
T. Hisham ∞ Rumi Hadji	M
T. Kitchil Jayah ∞ Shakeela Dissanayaka	BS
T. Manzil Burah ∞ Samath	M
T. Manzil Burah ∞ Karima Amit	M
TC Masnavi Jayah ∞ G Zanorra Burah	My
Tirta Kanchana ∞ TS Jassmin	My
T Abu Cassim Burah ∞ N Marliya Mahamooth	My
T Arfin Machan Burah ∞ Gene Elsie Manel Pereira	CS
T Aziz Burah ∞ G Zareena Suhood	My
T Bagoes Burah ∞ Pearly Meedin	M
T Bishrul Burah ∞ Rumi Hadji	M
T Deep Burah ∞ Marina Asmone	?
T Faazi Packeerally ∞ Kumbang Samahon	My
T Hadji Burah ∞ Dalreena Suhood	M

T Hadjirin Burah ∞ Raazeena Wahid	M
T Haroun Burah ∞ Binthara Samahon	My
T Kathir Burah ∞ Wahida Bahaman	M
T Kathir Burah ∞ Inthan Sheriff	My
T Kitchil Abu Cassim Burah ∞ G Kitchil Dole	My
T Kitchil Burah ∞ G Dain Hadji Bahar	My
T Kitchil Burah ∞ G Surahi Bahaman	My
T Kitchill Doella Padan ∞ Weerabangsa	My
T Mohammed Khalid Cassim Burah ∞ Aynoon Mutaliph	My
T Mubarak Packeerally ∞ G Zareen Jamalon	My
T Nizam Haniffa ∞ G Bidoora Doole	My
T Rafeek Burah ∞ Packeer	My
T Raji Doole ∞ Miss Yusoof	M
T Raschid Burah ∞ Seedin	My
T Riza Dinesh Burah ∞ Coralie Nadine Caspersz	CE
T Roshan Haroun Burah ∞ Shareefa Hanaffi	M
T Shihaan Fareed Burah ∞ Fazweena Akbar	M
T Shums Packeerally ∞ Dottie Rawlin	CE
T Sundara Burah ∞ Fowzia Sikkander	M
T Tahirin Cassim Burah ∞ G Balkis Jaimon	My
T Takilin Ahamath Cassim ∞ G Boelan Ahamath Cassim	My
T Thajudeen Buckker ∞ Suja Dole	My
T Thariq Rumi Jainudeen ∞ G Zeenar Burah	My
T Yehiya Cassim Burah ∞ G Maas Roteh Bangsajayah	My
Ummu Kadija ∞ M Balankaya	M
Ummu N Cassim Burah ∞ T Azeez Hassan	My
Wazeera Hadjirin Burah ∞ Milton Manamperi	CS
Yolande Jainudeen ∞ Fazil Jamaldin	My
Zain Doole ∞ G Moeda Abu Cassim	My
Zaitoon Sourjah ∞ K Gunasekera	S

Summary

Sri Lanka Malay (SLM) is the language of the ethnic groups of the Malays in Sri Lanka. The description of this language in this book is based on a corpus of audio and video files which were collected between 2005 and 2008 in and around Kandy and transcribed in the same period.

In the first part "Background" I sketch the social history of the Sri Lankan Malays. The first sizable groups of Malays were brought by the Dutch in the late 17th century. These included exiles, soldiers, convicts, and slaves. From the very early days onwards, these groups interacted with the local populations, who spoke Sinhala and Tamil. These two languages had considerable influence on the structure of SLM and radically changed its grammar. Based on a critical reassessment of the literature as well as new historical data, I show that earlier claims of intermarriage with the Tamil-speaking ethnic group of the "Moors" have been overstated. I evaluate three competing theories of genesis (Rapid Convergence, Tamil Substrate, and Metatypy) and show that the historical data do not support the first theory. As for the Tamil Substrate Theory, I show that the sociohistorical evidence presented in its favour is insufficient and not supported by a detailed analysis the linguistic data. The linguistic data are inconclusive, but appear to point away from strong Tamil influence. What remains is the Metatypy Hypothesis, which for the moment seems to be the most plausible option.

The second and third parts of the description contains the analysis of the forms of the language. In the domain of phonology, we find 6 vowels and 24 consonants. Particularly interesting are the realization of /ə/ as one of [ə,ɪ,ʊ] and the retroflex and prenasalized stops. SLM is a language which parses words into feet, but all of these feet are headless, i.e. SLM lacks stress.

In the domain of morphology, we can note the existence of vector verbs and relator nouns, which are a clear consequence of adstrate influence. We also find a copula, which is surprising given the areal tendency to lack such a form. Very interesting is the lack of subordinating conjunctions, whose role is taken over by postpositions.

SLM has 13 verbal prefixes, which can be divided into two subsets (genuine and quasi-prefixes) on morphophonological grounds. We find past, non-past, irrealis, conjunctive participle, and infinitive prefixes, among others. There are furthermore two imperative suffixes and seven derivational affixes, of which the involitive prefix *kànà-* deserves special mention. SLM has a large number of enclitics, which are used to

indicate case, but also coordination, information structure, and reported speech.

SLM distinguishes Verbs, Nouns, and Adjectives, but Adjectives can take all of verbal and nominal morphology, i.e. they can convert to become Nouns or Verbs. SLM compounds are normally right headed, but some vestigial left headed compounds still survive.

Phrasal syntax is generally right headed as well, with all modifiers normally preceding the head. An exception to this rule are some modals as well as some adjectives. An intriguing feature is the position of the indefiniteness clitic, which can intervene between a noun and its modifier. This has lead to the analysis of the SLM NP as appositional rather than hierarchical.

Another interesting area of SLM is case. SLM intransitive verbs can assign nominative, accusative, dative, or instrumental to their argument. The instrumental is only used for institutional actors like governments, while the dative is used for involitional actors like experiencers. Certain involitional arguments of intransitive verbs are also marked with accusative case. The normal case is the unmarked nominative, though. Actors of transitive verbs can bear nominative, dative or instrumental case, while undergoers bear nominative, accusative or dative case. This distribution of cases does not point to a clear nominative-accusative or ergative-absolutive system. Instead, it appears that case marking in SLM is driven by semantic considerations. Next to case distribution, a whole range of morphosyntactic phenomena suggests that SLM grammar does not make use of the notion of "subject". It is possible to describe SLM syntax without recurring to that notion.

The SLM sentence is generally predicate-final, but occasionally, we find right-extraction of arguments, i.e. they occur after the predicate. Interrogative and imperative clauses are marked morphologically, while relative clauses are preposed and are not signaled by special morphology. Positional information suffices to identify relative clauses.

In the fourth part I investigate functional domains and discuss how SLM encodes individuals, properties, actions, semantic roles, space, time, possession, kin, and a number of other domains relevant for language typology. Furthermore, information structure is discussed here, which mainly relies on the emphatic clitic =*jo* for a whole array of functions in the domains of topic and focus. A section on the pragmatics of speech acts like requesting confirmation, requesting permission, promising and expressing annoyance closes the descriptive part. The appendix contains five fully glossed and translated texts of both oral and written modality next to some additional historical and phonetic material.

Samenvatting

Sri Lanka Malay (SLM) is de taal van de Maleise gemeenschap in Sri Lanka. Deze taal wordt in dit boek beschreven, gebaseerd op een corpus van audio- en videobestanden en transcripties daarvan die tussen 2005 en 2008 verzameld zijn in Kandy en omgeving.

In het eerste gedeelte "Background" schets ik de sociale geschiedenis van de Sri Lankaanse Maleiers. De eerste belangrijke groepen Maleiers kwamen tijdens het Nederlandse koloniale bestuur naar Ceylon, als bannelingen, huursoldaten, gevangenen en slaven. Vanaf het begin vondt er uitwisseling plaats tussen deze groepen en de inheemse bevolking, die Singalees en/of Tamilisch sprak. Deze twee talen oefenden veel invloed uit op de structuur van het SLM en veranderden zijn grammatica radicaal. Ik bekijk de bestaande literatuur over deze periode en ook nieuwe historische documenten en trek de conclusie dat er voor vroegere claims met betrekking tot gemengde huwelijken met de andere Islamitische groep in Sri Lanka, de "Moors" geen voldoende evidentie is. Ik bespreek vervolgens drie theoriën over de oorsprong van het SLM. Dit zijn Rapid Convergence, Tamil Substrate en Metatypy. Ik laat zien dat de historische data geen aanleiding geven voor de eerstgenoemde theorie. Met betrekking tot de Tamil Substrate Theory laat ik zien dat de sociohistorische evidentie die door haar voorstanders wordt gepresenteerd onvoldoende is. Verder wordt deze theorie ook niet gesteund door een gedetailleerde analyse van de taalwetenschappelijke gegevens. Deze gegevens zijn onduidelijk maar blijken eerder een Tamil invloed tegentespreken. Wat overblijft is de Metatypy Hypothese, die voor het moment de meest waarschijnlijke optie lijkt.

Het tweede en derde gedeelte van de beschrijving bevat de analyse van de vormen van de taal. In de fonologie hebben wij te maken met 6 klinkers en 24 medeklinkers. Bijzonder interessant zijn de realisering van /ə/ als [ə,ɪ,ʊ] en de retroflexe en geprenasaliseerde plofklanken. SLM is een taal die woorden in voeten parst, maar in tegenstelling tot vele andere talen zijn deze voeten geheel 'headless', dwz dat SLM geen beklemtoning kent.

Wat de morfologie betreft zijn 'vector verbs' en 'relator nouns' belangrijke kenmerken. Deze zijn allebei duidelijk het gevolg van invloed van de adstraten. Er bestaat ook een copula, wat verrassend is angezien de verdere afwezigheid van zo een vorm in dit geografische areaal. Interessant is verder de afwezigheid van onderschikkende

807

voegwoorden, wwarvan de functie wordt overgenomen door achterzetsels.

SLM heeft 13 verbale voorvoegsels. Deze kunnen om morfologische redenen worden opgedeeld in twee deelverzamelingen (echte en quasi-voorvoegsels). Er zijn voorvoegsels voor onder andere verleden tijd, niet-verleden tijd, irrealis, conjunctief deelwoord en onbepaalde wijs. Verder zijn er twee achtervoegsels, die de gebiedende wijs uitdrukken, en zeven afleidende affixen, waaronder de involitieve prefix *kànà-*. SLM heeft veel enclitica die worden gebruikt om naamvallen aan te geven, maar ook voor nevenschikking, informatiestructuur en indirecte rede.

SLM onderscheid werkwoorden, zelfstandige naamworden en bijvoegelijke naamwoorden, maar bijvoegelijke naamwoorden kunnen worden gecombineerd met alle verbale en nominale morfologie, da wil zeggen dat ze kunnen worden geconverteerd naar werkwoorden of zelfstandige naamwoorden. Samenstellingen hebben het hoofd gewoonlijk aan de rechterkant, maar sommige oude samenstellingen hebben hun hoofd nog steeds aan de linkerkant.

Zinsdelen hebben hun hoofd gewoonlijk ook aan de rechterkant, vorafgegaan door alle bepalenden woorden of woordgroepen. Een uitzondering op deze regel vormen sommige modale partikelen en sommige bijvoegelijke naamwoorden, die ook op het hoofd kunnen volgen. Een ondoorzichtige kwestie is de positie van het cliticum dat onbepaaldheid uitdrukt. Dit kan tussen de dependens en de hoofd komen te staan. Als gevolg hiervan wordt de SLM NP in dit boek als bijstellend (en niet als hierarchisch) geanalyseerd.

Onovergankelijke werkwoorden kunnen in het SLM nominativus, accusativus, dativus of instrumentalis aan hun argument toewijzen. De instrumentalis kan alleen worden gebruikt voor institutionele actors zoals regeringen. De dativus wordt vooral gebruikt voor actors die niet volitioneel zijn, zoals ondervindende voorwerpen (experiencers). Involitionele argumenten van sommige onovergankelijke werkwoorden kunnen tevens met de accusativus worden gemarkeerd. Dit neemt niet weg dat het ongemarkeerde geval de nominativus is. Met betrekking tot overgankelijke werkwoorden kan men zeggen dat dan actors met nominativus, dativus of instrumentalis worden gemarkeerd en undergoers met nominativus, accusativus of dativus. Deze verdeling van naamvalsenclitica wijst niet duidelijk op een nominativus-absolutivus of ergativus-accusativus-systeem. In plaats daarvan blijkt het gebruik van naamvalsenclitica in het SLM semantisch gemotiveerd. Naast de naamvallen laten ook nog meer morfosyntaktische fenomenen zien, dat in het SLM de notie 'onderwerp' ('subject') niet relevant is. Het SLM kan worden beschreven zonder gebruik te maken van dit begrip.

Zinnen in het SLM hebben normaliter hun predicaat in de laatste positie, maar soms vindt men ook nog argumenten achter het predicaat. Vragende en gebiedende zinnen zijn morfologisch gemarkeerd, maar relatieve bijzinnen worden syntaktisch gekenmerkt door hun positie vorafgaand aan het hoofd. Deze positionele informatie is voldoende om relatieve bijzinnen en hoofdzinnen uit elkaar te houden.

In het vierde gedeelte onderzoek ik functionele domeinen en bespreek hoe het SLM indiviuën, eigenschappen, handelingen, semantische functies, ruimte, tijd, bezit,

verwantschap en nog enkele andere taaltypologisch interessante domeinen uitdrukt. Informatiestructuur wordt ook in dit gedeelte besproken. Informatiestructuur maakt vooral gebruik van het encliticum =*jo*, dat voor heel veel verschillende functies in de domeinen van topic en focus wordt gebruikt. Een bespreking van taalhandelingen zoals eisen van bevestiging, eisen van toestemming, beloftes en uitdrukking van verdriet sluit het descriptieve gedeelte af. De bijlage bevat vijf volledig geglosseerde en vertaalde teksten van mondelijke en schriftelijke modaliteit, naast enig additioneel historisch en fonetisch materiaal.

Curriculum vitae

Sebastian Nordhoff was born in Erlangen (Germany) and studied General Linguistics, Computational Linguistics and Spanish at the universities of Cologne and Seville. He did field work in Paraguay for his master's thesis on word classes in Guarani before a first stay in the Netherlands, an internship in the Max-Planck-Institut für Psycholinguistik in Nijmegen. There, he took care of data curation for the Project "Pioneers of Island Melanesia". After completing this doctoral dissertation about Sri Lanka Malay at the University of Amsterdam, he will move to the Leipzig Max-Planck-Institut für Evolutionäre Anthropologie in his native Germany to develop software for electronic publishing of linguistic data such as grammatical descriptions, dictionaries, and corpora.