Jack-of-all-trades: the Sri Lanka Malay flexible adjective
Sebastian Nordhoff

Abstract

This paper presents the parts-of-speech system of Sri Lanka Malay, which consists of verbs, nouns and adjectives, and argues that the latter are a fully flexible parts-of-speech, which can be used for any discourse function without further measures being taken. The presence of the additional lexical categories of noun and verb distinguishes the Sri Lanka Malay system from other languages with maximally flexible parts-of-speech, which are monocategorial.

This paper argues that noun and verb in Sri Lanka Malay are recent innovations and a result of influence from the adstrates Sinhala and Tamil, which have rigid parts-of-speech systems. This influence triggered the closer association of object-denoting lexemes with the act of reference and action-denoting lexemes with the act of predication. Property-denoting lexemes were left unconstrained as to their possible functions. This closer association was assisted by already existing derivational processes to restrict discourse function. Comparing Sri Lanka Malay to other languages, it appears that the forerunners of specialized parts-of-speech are derived lexemes.

1 Introduction

Discussions of languages with very flexible parts-of-speech systems have focused on monocategorial languages in the past, i.e. languages where all lexemes are said to be in the same word class. Examples are Samoan (Mosel and Hovdhaugen 1992), Riau Indonesian (Gil 1994), Tongan (Broschart 1997), Mundari (Croft 2005, Hengeveld and Rijkhoff 2005), Kharia (Peterson 2005), and Late Archaic Chinese (Bisang 2008a, 2008b). According to these authors, it is impossible to establish morphosyntactic differences between the lexemes in these languages. They are all in the same form class, and no other lexemes are in another class.

This focus on monocategorial languages has obscured the possibility that maximally flexible parts of speech can also exist alongside other lexical categories more limited in their potential.[1] For instance, it would be imaginable that a language has two word classes: a rigid category of verbs and a class of maximally flexible lexemes (i.e. lexemes that could immediately be used in verbal, nominal, adjectival and adverbial functions). Such a language would then not be monocategorial, but polycategorial. This paper presents the polycategorial system of Sri Lanka Malay, where a maximally flexible class of adjectives coexists with the more restricted categories of nouns and verbs.

In Sri Lanka Malay, distributional properties clearly indicate the existence of three morphosyntactically different form classes: verbs, nouns and adjectives. Verbs and nouns are limited in their potential use for different discourse functions. Such a limitation does not

---

[1] The existence of lexical categories with limited flexibility alongside more rigid classes has received more attention, e.g. Bhat (1994).

hold for adjectives. The clear distributional distinction between the three word classes sets Sri Lanka Malay apart from other Malay (and Austronesian) varieties. The reduction of flexibility of the noun class and the verb class is analyzed here as a result of extended contact with the adstrate languages Sinhala and Tamil, which both have a rigid parts-of-speech system.

The paper is structured as follows: After a brief introduction to Sri Lanka Malay (Section 2), we provide an overview of its word classes and their morphosyntactic characteristics (Section 3). Semantic aspects of the various word classes are discussed in section 4. Section 5 is concerned with the discourse functions of members of the various classes in Sri Lanka Malay and section 6, finally, deals with issues such as language contact and the diachronic development of the SLM parts-of-speech system.

2 Sri Lanka Malay

Sri Lanka Malay (SLM) is the language of the ethnic group of the Malays in Sri Lanka. They were brought by the colonial powers of the British and the Dutch between the 17th and the 19th century. While the lexicon has remained quite stable over time, the grammar has been heavily influenced by the adstrates Sinhala (Indo-Aryan) and Tamil (Dravidian) (Adelaar 1991, Smith et al. 2004, Bakker 2006, Slomanson 2006, Ansaldo 2008, Nordhoff 2009, Ansaldo and Nordhoff submitted).

The immigrants were mainly soldiers, some exiled princes and very few convicts and slaves. The mercenaries hailed from all over what is today Indonesia and Malaysia, but were recruited in Jakarta and later Penang (Hussainmiya 1990). In the regiment, some variety of Trade Malay local to Jakarta was therefore most probably the dominant linguistic code (Vlekke 1943, Paauw 2004, Smith et al. 2004, Smith and Paauw 2007).

While the community was quite close-knit until the middle of the 19th century, the dissolution of the Malay regiment in 1873 led to a dispersion of the community. This and the Sinhala nationalist policy following the independence of Ceylon from the UK in 1948 led to attrition and language loss, which we are witnessing today.[2]

Today, the Malays make up 0.3% of the Sri Lankan population. They live in several towns in the Central and Southern Sinhala speaking area, were they make up between 1 and 3% of the local population (30% in Hambantota, over 90% in the Southern hamlet of Kirinda) (Bichsel-Stettler 1989, Ansaldo 2005).

All Sri Lankan Malays can understand each other without problems, but there is a lot of idiolectal variation, which does not pattern geographically, but rather according to family lines (which may be present in several places).

3 Structuralist approach

In the approach established by the American structuralist in the beginning of the last century (e.g. Bloomfield 1933), form classes are defined by the distributional properties of

---

[2] Documentation of the language is currently being carried out as part of a Dobes-project consisting of Umberto Ansaldo, Lisa Lim, Walter Bisang and Sebastian Nordhoff.

lexemes (both morphological and syntactic). Lexemes with the same morphosyntactic properties are grouped in the same class. I will discuss the defining properties of SLM verbs, nouns and adjectives in turn. This classical take will be contrasted with a functional discourse-pragmatic analysis in the next section.

3.1 Verbs

SLM lexemes of the verb class share the ability to take verbal prefixes, verbal preclitics, the nominalizer *-an* and a certain negation pattern. They are distinguished from nouns by the inability to co-occur with the indefiniteness marker *atthu*, the deictics *ini* and *itthu* and the plural marker *pada*. I will give an extensive overview of the properties of verbs, which will be important for the definition of adjectives later on.

3.1.1 Prefixes

The following prefixes can be combined with verbs in SLM:
- the past markers *su-*, examples (1) and (2) and *anà-*, examples (3) and (4)
- the conjunctive participle marker *asà-*, examples (5) and (6)
- the infinitive marker *mà-*, examples (7) and (8)
- the past tense negator *thàrà-*, examples (9) and (10)
- the non-past marker *arà-*, examples (11) and (12)
- the irrealis marker *anthi-*, examples (13) and (14).[2]

The following examples show the use of these prefixes on verbs in constructed sentences (first) and natural discourse from the corpus (second), available at the Dobes-archive.[3] [4] Sentences starting with eli... are taken from the corpus, but are elicited. These are taken if no (appropriate) example could be found in the corpus.[5]

(1) *Incayang* **su-maakang**
     he   PAST-go
     'He ate.' (K081112eli01)

(2) *Kitham=pe* *aanak* *pada=le* *karang* *bae=nang* *cinggala* **su-blaajar**
     1PL=POSS   child   PL=ASSOC   now   good=DAT   Sinhala   PAST-learn
     'Our children have learnt Sinhala well now.' (K051222nar05)[6]

---

(3) *Incayang   anà-maakang*
he   PAST-go
'He ate.' (K081112eli01)

(4) *Seelon=nang lai   hathu   kavanan   anà-dhaathang*
Ceylon= DAT other INDEF   group   PAST-come
'Another group came also to Ceylon.' (K060108nar02)

(5) *Incayang   asà-maakang su-miinung*
he   CP-eat   PAST-drink
'He ate and drank.' (K081112eli01)

(6) *Nyaakith   oorang pada asà-pii thaangang   arà-cuuci*
sick   man   PL   CP-go hand   NONPAST-wash
'The patients go and wash their hands.' (K060116nar03)

(7) *Incayang=nang   mà-maakang(=nang) thàràboole*
he=DAT   INF-eat=DAT   cannot
'He cannot eat.' (K081112eli01)

(8) *Derang   pada=nang   atthu=le mà-kijja=nang thàràboole*
3PL   PL=DAT   one=ASSOC INF-make= DAT cannot=because
'They could not do a single thing, ….' (N060113nar01)

(9) *Incayang   thàrà-maakang*
he   NEG.NONPAST-eat
'He did not eat.' (K081112eli01)

(10) *Puaasa muusing   thàrà-duuduk=si?*
fasting   season   NEG.PAST-stay=INTERR
'You were not here during the fasting period, were you?.' (B060115cvs03)

(11) *Incayang   arà-maakang*
he   NONPAST-eat
'He is eating.'

(12) *Uncle, mana=ka   arà-duuduk?*
uncle   where=LOC   NONPAST-live
'Where do you live?' (B060115cvs16)

(13) *Incayang   anthi-maakang*
he   IRR-eat
'He will eat.'

(14) *Kithan=le*   **anthi-banthu**
     1PL=ASSOC   IRR-help
     'We will help, too.' (B060115nar02)

3.1.2 Negation

As we will see later on (Section 3.3.6), negation patterns provide good cues to the class-membership of lexemes. SLM verbs show the following negation pattern: in the past tense, they are negated by *thàrà-* in examples (9) and (10) above. In the non-past tenses, they are negated by *thama-* (15) and (16) below.

(15) *Incayang*   **thama-maakang**
     he   NEG.NONPAST-eat
     'He does/will/would not eat.'

(16) *Kitham=pe*   *aanak*   *pada* **thama-oomong**
     1PL=POSS   child   PL   NEG.NONPAST-speak
     'Our children don't speak (Malay).' (G051222nar01)

3.1.3 Nominalizer -an

The derivational suffix -an can be used to derive nominals from verbs, like *maakang* 'eat' → *makanan* 'food', *thaksir* 'think' → *thaksiran* 'thought'.

3.1.4 Negative features

So far, we have dealt with morphemes whose presence indicates that the lexeme they combine with belongs to the class of verbs. We now turn to morphemes whose co-occurrence with a lexeme precludes membership of that lexeme in the class of verbs. There are four main morphemes which can be used to exclude a lexeme from the verb class: the indefiniteness markers (*h*)*a*(*t*)*thu* , the deictics *in*(*n*)*i* or *it*(*t*)*hu* or the plural marker *pada*. The examples in (17) show the ungrammaticality of these combinations.

(17) a.   **at(t)hu maakang*
           INDEF   eat
     b.   **i(n)ni/*it(t)hu*   *maakang*
           PROX/DIST eat
     c.   **maakang pada*
           eat   PL

To make the sentences in (17) grammatical, the nominalizer –*an* has to be added.

3.2. Nouns

Lexemes of the class of nouns are the mirror image of verbs. They can combine with the morphemes which are ruled out for verbs just discussed (Section 3.1.4), and they cannot combine with the verbal prefixes or verbal preclitics discussed in Section 3.1.1. The morphemes that indicate the membership of a lexeme in the class of nominals are:

- indefinite article *(h)a(t)thu*
- the deictics *in(n)i* and *i(t)thu*
- the plural marker *pada*
- the nominal negator *bukang*

3.2.1. Indefiniteness marker

The indefiniteness marker can combine with nouns, where it can occur preposed (18-19), postposed (20-21) or in both positions (22).

(18) *Ini   atthu   ruuma*
     this   INDEF   house
     'This is a house.' (K081112eli01)

(19) *Sindbad the   Sailor   hatthu   Muslim*
     Sindbad the   sailor   INDEF   muslim
     'Sindbad the Sailor was a Moor.' (K060103nar01)

(20) *Ruuma atthu   su-aada*
     house   INDEF   PAST-exist
     'There was a house.' (K081112eli01)

(21) *See   awuliya atthu   su-jaadi*
     1s   saint   INDEF   PAST-become
     'I have become a saint.' (K051220nar01)

(22) *Sithu=ka hathu   maccan hathu   duuduk aada*
     there= LOC INDEF   tiger   INDEF   stay   EXIST
     'A tiger stayed there.' (B060115nar05)

3.2.2 Deictics

Both the proximal deictic *in(n)i* in examples (23) and (24) and the distal deictic *it(t)hu* in examples (25) and (26) can combine with lexemes of the noun class. They are always preposed.

(23) ***Ini ruuma*** *bìssar*
this house big
'This house is big.' (K081112eli01)

(24) ***Inni oorang***=*nang*   *itthu thàràthaau*
PROX man= DAT   DIST NEG.know
'The man didn't know that.' (K07000wrt01)

(25) ***Itthu ruuma*** *pada bìssar*
those house   PL   big
'Those houses are big.' (K081112eli01)

(26) ***Itthu watthu***=*ka* ***itthu nigiri***   *pada=ka*   *arà-duuduk*
DIST time= LOC DIST land   PL=LOC   NONPAST-stay
'At that time, (they) lived in those countries.' (N060113nar01)

3.2.3 Plural marker

The plural marker *pada* can occur with many nouns (25) and (26), but not all of them, due to the well-known distinction between count nouns and mass nouns. Co-occurrence with a lexeme is then positive evidence for membership of that lexeme in the class of nouns, but impossibility to co-occur does not provide evidence to the contrary.

3.2.4 Negation pattern

Nouns are always negated by *bukang* when used as predicates.[7] The reverse it not true. While nouns are always negated by *bukang*, negation by *bukang* does not entail that we must be dealing with a noun. *Bukang* can also be used for constituent negation and clausal negation (cf. Gil this volume for the use of *bukan(g)* in other varieties of Malay.). Therefore, this feature is only a cue to class-membership, but not sufficient in itself.

(27) *Incayang (hatthu) doktor   bukang*
3S   INDEF doctor NEG.NONV
'He is/was not a doctor.' (K081105eli02)

3.2.5 Negative features
Nouns are distinguished from verbs by their inability to combine with the verbal prefixes, the nominalizer -*an* or verbal negation patterns listed above. Example (28) shows the ungrammaticality of these combinations.

---

[7] Existential negation is done by *thraa*, e.g. *Bìrras thraa* 'There is no rice'.

(28) a. *su/*anà-/*mà-/*arà-/*anthi-   ruuma
        PAST-/PAST-/INF-/NONPAST-/IRR-   house

   b. *ruma -(h)an[8]
       house- NMLZR

   c. *thàrà-/*thama-   ruuma
       NEG.PAST/NEG.NONPAST house

3.3 Adjectives

We now turn to the category crucial to this paper, the adjective. In order to signal its special status, and to make it clear that there are significant differences with adjectives in other languages, I will use the label 'Flexible Adjective' for this class in the following. This term should not be interpreted as entailing the existence of non-flexible adjectives in SLM. All SLM adjectives are flexible. Lexemes of this class can combine with all features characteristic of the verb class and with all features characteristic of the noun class. There is not one morphosyntactic process Flexible Adjectives could not take part in. Adjectives are the Jack-of-all-trades of SLM parts-of-speech. The following examples show that lexemes of the Flexible Adjective class can take all the morphology typical of the classes of verbs and nouns. The most prominent feature of Flexible Adjectives is thus their versatility.

3.3.1 Verbal prefixes

I have discussed above that a number of prefixes provide good cues to the membership of a lexeme in the class of verbs. This has to be amended, because Flexible Adjectives can combine with the very same prefixes. This means that these prefixes do not establish that a lexeme is a verb, because it could also be a Flexible Adjective. What these prefixes do establish is that the lexemes under discussion cannot be a noun. We list the verbal prefixes again for convenience and give examples of their use with Flexible Adjectival lexemes.

- the past marker *su-*, examples (29) and (30), and *anà-*, examples (31) and (32);
- the conjuctive participle marker *asà-*, examples (33) and (34);
- the infinitive marker *mà-*, examples (35);
- the non-past marker *arà-*, examples (36) and (37);
- the irrealis marker *anthi-*, examples (38) and (39);
- he negators will be discussed in a special section, 3.3.6.

(29) *Oorang* **su-thiiňği**
     man   PAST-tall
     'The man grew up.' (K081112eli01)

---

[8] The nominalizer *–an* is realized as *–han* in certain contexts. Both forms are ungrammatical here.

(30) *Hatthu spuulu liimablas* ~~*tawon*~~*=na jaalang blaakang,*
    INDEF   ten   fifteen   year=DAT go   after
    inni kumpulan **sa-mampus**
    PROX association   PAST-dead
    'About ten, fifteen years after that, the association became defunct.'

(31) *Oorang* **anà-thiiňǧgi**
    man   PAST-tall
    'The man grew up.' (K081112eli01)

(32) *Itthu=nam blaakang=jo, kitham   pada* **anà-bìssar**
    DIST   after=EMPH   1PL   PL   PAST-big
    'After that, we grew up.' (K060108nar02)

(33) *Aanak   asà-bìssar* **su-pii**
    man   CP-big   PAST-go
    'After the child had grown up he went away.' (K081112eli01)

(34) *Aanak   pada asà-bìssar,   skul=nang*   **anà-pii**
    child   PL   CP-big   school=DAT PAST-go
    'After the children had grown up, they went to school.' (K051222nar04)

(35) *Oorang mà-bìssar=nang* **arà-maakang**
    man   INF-big=DAT   NONPAST-eat
    'The man eats to become big.' (K081112eli01)

(36) *Oorang* **arà-thiiňǧgi**
    man   NONPAST-big
    'The man is growing.'

(37) *Ruuma* **arà-kiccil**
    house   NONPAST-small
    'The houses are getting small.' (K051222nar04)

(38) *Oorang* **anthi-thiiňǧgi**
    man   RR-big
    'The man will grow.'

(39) *Ithukapang gaathal* **anthi-kuurang**
    then   itching IRR-little
    'Then the itching will become less.' (K060103cvs02)

9

3.3.2 Nominalizer -an

Flexible Adjectives can be nominalized by *-an*, like *maanis* 'sweet' →*manisan* 'sweets', *sìggar* 'healthy' → *sìgaran* 'health'.

3.3.3 Nominal indefiniteness

The preceding sections have shown that Flexible Adjectives can combine with all verbal morphology.[9] We now examine to what extent Flexible Adjectives can combine with nominal morphology. One characteristic of nouns mentioned was the possibility to co-occur with the indefiniteness marker *atthu*. Example (40) shows that this marker can also co-occur with Flexible Adjectives.

(40) *Se=dang   kiccil   hatthu   kaasi-la*
     1s= DAT   small   INDEF   give- IMP
     'Please give me a small one.' ~~(no source)~~ (K081112eli01)

3.3.4 Nominal deixis

Just like the indefiniteness marker, the deictics can also occur with both nominals and Flexible Adjectives, as seen in example (41).

(41) *Se=dang   ini   kiccil=yang   kaasi*
     1S=DAT   PROX small= ACC   give
     'Give me the small ones.' ~~(no source)~~ (K081112eli01)

3.3.5 Nominal number

The last morpheme we had analyzed as indicative of membership in the noun class was the plural marker *pada*, which can also combine with Flexible Adjectives, as in (42).

(42) *Sedang ini   kiccil   pada kaasi*
     1s= DAT PROX small   PL   give
     'Give me those small ones.' ~~(no source)~~ (K081112eli01)

---

[9] An anonymous review asks: 'Does this only take place when there is a lexical gap, or can a converted adjective substitute for an existing verb with comparable meaning? Are there any adjectives at all that do not convert, and if so, is this strictly a function of the semantics of the adjective? Can exceptional adjectives be grouped in well-definable semantic classes? Are borrowed lexical items converted as easily as etymologically Malay items?' In this paper, I cannot survey each and every SLM adjective. As far as I can tell, conversion can take place with any adjective. I have not found any Adjective that could not convert to a verb (see Gil (this volume) for methodological arguments about the burden of proof). As a consequence, classification of the exceptions is not possible: there are no exceptions. Furthermore, there is no indication of conversion being ruled out when a verb with comparable semantics exists. Finally, conversion can also be found with loanwords like 'late'.

We have seen up to now that Flexible Adjectives can take the morphology we introduced as verbal as well as the ~~one~~ introduced as nominal. This entails that we have to reformulate our observations about the cues these features provide: If the morphosyntactic cues exclude a lexeme from the noun class, we are dealing with a verb; if they exclude it from the verb class, we are dealing with a nominal. If no exclusion cues exist, we are dealing with a Flexible Adjective.

3.3.6 Negation pattern

From the discussion so far it appears that there is not a single distributional property that uniquely identifies members of the class of Flexible Adjectives. Flexible Adjectives are characterized by the possibility to behave as nouns or verbs, as the case may be. We have to amend this conclusion slightly; there is one characteristic that irrefutably indicates membership in the Flexible Adjectives ~~class. This is~~ the negation pattern. While verbs are normally negated with a prefix, and nouns by *bukang*, most Flexible Adjectives are negated by postposed *thraa* (46) and (47).

(46) *Se thiiňǧgi thraa*
  I tall NEG
  'I am/was not tall.' (K081112eli01)

(47) *Samma bannyak responsible thraa*
  every much responsible NEG
  'They are all not very responsible.' (B060115cvs01)

When *thraa* negates a Flexible Adjective, it carries no indication of tense, as examples (46) and (47) show. *Thraa* can also be used with nouns, but then indicates absence as in example (48). This example could not possibly mean 'It is not a Malay political party'.

(48) *Malay political party atthu thraa*
  Malay political party INDEF NEG.EXIST
  'There is no Malay political party.' (K051206nar12)

Furthermore, *thraa* can also be used with verbs, but then indicates the perfect tense, as in (49).

(49) *Kithang baaye mlaayu arà-oomong katha incayang biilang thraa*
  1PL good Malay NONPAST-speak QUOT NEG 3S say NEG.PERF
  'He has not said that we speak good Malay.' (B060115prs15)

The functions of *thraa* are thus different depending on whether it is used in combination with a noun, a verb, or a Flexible Adjective. This can then be used as a shortcut to establish class-membership. Some Flexible Adjectives have a divergent negation pattern where *thàrà-* can be used for all tenses (Slomanson 2006). Because of this temporal vagueness,

they can be distinguished from verbs, where *thàrà-* always indicates past tense.

The two different classes of Flexible Adjectives do not show obvious phonological, morphological or semantic characteristics, and different speakers assign different lexemes to them. The different negation patterns thus provide a means to quickly check the category membership of a certain lexeme (Table 1).

	Past	Perfect	Present	Future
Verb	*thàrà-*V	V *thraa*	*thama-*V	
Noun	<mark>*bukang*</mark>			*bukang thama-jaadi*
FlxAdj1	<mark>FlxAdj *thraa*</mark>			FlxAdj *thraa* *thama-*FlxAdj [10]
FlxAdj2	<mark>*Thàrà-*FlxAdj</mark>			FlxAdj *thraa* *thama-*FlxAdj

Table 1. Negation as a cue for word class membership.[11]

This particular negation pattern is changed, however, when a Flexible Adjective is converted to a noun or a verb. In that case, the normal verbal or nominal negators are used, as described above. As an illustration, the unconverted Flexible Adjective *kaaya* 'rich' denotes a state and is normally negated by *thraa* (50).

(50) *Se **kaaya** thraa*
    1s rich   NEG
    'I am/was/will not (be) rich.' (K081104eli06)

When this lexeme converts to a verb, it denotes a process ('becoming rich'). In this case, it is negated with the verbal negator *thàrà-* for past reference (51) or *thama-* for non-past reference. Note that (51) is not temporally vague but has necessarily reference to the past.

(51) *Se **thàrà-kaaya***
    1S NEG.PAST[verbal]-rich
    'I did not become rich' (K081104eli06)

This conversion can occasionally be found with past and present time references. It is close to obligatory with predications with future time reference. In that case, the conveyed message is not 'will not be PROPERTY' but rather 'will not become PROPERTY' (see example (52)).

---

[10] Negated adjectival predicates with future time reference are normally negated as converted verbs with the meaning 'will not become the case' rather than as Flexible Adjectives with the meaning 'will not be the case'.

[11] Negation can also be used to establish subclasses of Flexible Adjectives (FlxAdj1 and FlxAdj2), but these have no further properties differentiating them.

(52) *Inni pukuran=yang   mà-gijja   **thamau-gampang***
     PROX work= ACC   INF-make   NEG.NONPAST-easy
     'To do that kind of work will never be(come) easy.' (K081106eli01)

If this 'change-of-state'-reading is not intended, the normal negation pattern can be used. This is much less common. Example (53) illustrates this for Type 1 Flexible Adjectives.

(53) *Incayang=pe dudukan   hathiyan   thaaun=ka=le   laile **thàrà-baae***
     3S =poss   behaviour   other   year= LOC=ADDIT   still   NEG-good
     'His behaviour will still not be good (=remain bad) even next year.' (K081104eli06)

Similarly to verbal conversion discussed above in (51), *kaaya* 'rich' can be converted to a noun and then be used referentially. In this case, it is negated with the nominal negator *bukang*.

(54) *Se   **kaaya   bukang**   incayang=jo   kaaya*
     1s   rich   NEG.NONV   3S=EMPH   rich
     'I am not the rich person, the rich person is him.' (K081104eli06)

Membership in the class of Flexible Adjectives can thus be established by the negation with postposed *thraa*, but the other negators can also be found following conversion to a verb or a noun, as the case may be.

3.3.7 Negative features

While there is morphology whose presence can serve to exclude a candidate from the class of verbs (e.g. the plural marker) or from the class of nouns (e.g. irrealis prefix), no such features exist for Flexible Adjectives.

3.4 Summary

SLM lexemes can be categorized according to distributional criteria. There is morphology that cannot combine with nouns, and there is other morphology that cannot combine with verbs. The occurrence of one of these morphemes rules out membership in the respective class. Flexible Adjectives can take any morphology. This is summarized in Table 2.

13

	Noun	Verbs	Flexible Adjectives
Verbal prefixes	-	+	+
Nominalizer	-	+	+
Indefininiteness	+	-	+
Deictic	+	-	+
Plural	+	-	+
Thàrà/thama-	-	+	+
Bukang	+	-	+
Predicative present tense negation with *thraa*	-	-	(+)

Table 2. Overview of defining characteristics of lexical categories (notice that predicate negation with *thraa* is only possible for some Flexible Adjectives).

4 Semantic correlates of the word classes

On a pretheoretical level, adjectives encode properties as their core meaning (Sasse 1993, Sassse 1995, Croft 2001). This core meaning is also found in the following two cases for SLM: modification (55) and property assignment (56).

(55) *Incayang [atthu bìssar orang]*
　　 3s　　　INDEF big　　man
　　 'He is a big man.' (K081112eli01)

(56) *Aanak thiiňǧgi*
　　 child　　tall
　　 'The child is tall.' (K081103eli02)

When Flexible Adjectives are converted to nouns (i.e. they co-occur in a positively nominal context), they encode objects/individuals (cf. Wierzbicka 1986).

(57) *Incayang [kitham=pe bìssar]*
　　 3S.POLITE　1PL=POSS big
　　 'He is our boss.'

14

When Flexible Adjectives are converted to verbs, (i.e. they co-occur in a positively verbal context), they denote processes.[12]

(58) *Incayang   arà-bìssar*
       he   NONPAST-big
       'He is growing up.'

We can summarize this by saying that Flexible Adjectives adapt their semantics to the nominal/object, verbal/process or 'adjectival'/property context that the morphosyntactic frame might require. This means that three cases can be distinguished: 1) In a nominal context, they denote objects, 2) in a verbal context, they denote processes, and 3) elsewhere, they denote properties. This regular shift of semantics makes SLM a member of type (ii) flexibility (Haig 2006).

   The distributional behaviour of Flexible Adjectives is then best captured in the structuralist approach by assuming conversion/zero-derivation.

(59) FlxAdj+$\emptyset$ = N

(60) FlxAdj+$\emptyset$ = V

Since the semantic shift that Flexible Adjectives undergo under conversion is completely regular, this can be analyzed as word-level-conversion (Don and van Lier, this volume). There are a handful of lexemes that can function both as nouns and verbs in SLM (not Flexible Adjectives), but the semantic relation between them is arbitrary and not regular. Examples for these are *nyaanyi* 'sing/song' and *jaalang* 'walk(V)/street'. These pairs are then instances of root-level-conversion in the model Don and van Lier propose.

5 Functional approach

The structuralist distributional approach works for individual languages, but is impossible to apply when comparing languages since the criteria to establish the categories are dependent on every individual language (Haspelmath 2007). To overcome the limited usefulness of the morphosyntactic approach, Kees Hengeveld and colleagues have developed a classification of parts-of-speech systems that is based on discourse-pragmatic functions rather than on morphosyntactic criteria (Hengeveld 1992, Hengeveld et al. 2004, Hengeveld and van Lier 2008, Hengeveld and Mackenzie 2008). While distributional criteria cannot be used to identify word class membership across languages, it is assumed that discourse pragmatic criteria can be used for this purpose. Languages do not all employ the same set of formal (morphosyntactic) categories such as 'plural marker' or 'case marker', but the discourse notions PREDICATION, REFERENCE and MODIFICATION are probably relevant for all languages (cf. Croft 2001: 84-5,87). In fact, Hengeveld and

---

[12] Similar semantic changes are attested for Mauritian Creole (Alleyne 2000: 131) and Kharia (Peterson, this volume).

colleagues use the category labels PREDICATE, REFERENTIAL PHRASE, HEAD and MODIFIER to arrive at the following four discourse functions (Table 3):

	head	modifier
predicate phrase	head of Predicate Phrase	modifier of Predicate Phrase
referential phrase	head of Referential Phrase	modifier of Referential Phrase

Table 3. Discourse-pragmatic phrasal functions according to Hengeveld.

Each lexeme can be tested as to whether it can be used in one of these four discourse functions as it is, or whether additional material like a copula is needed if the lexeme is to occur in this function. Hengeveld (this volume) gives the following illustration: The four functions and their lexical expression can be illustrated by means of the English sentence in (1).

(61) *The tall girl sings beautifully*

English can be said to display separate lexeme classes of verbs, nouns, adjectives and (derived) manner adverbs, on the basis of the distribution of these classes across the four functions identified in Figure 1: verbs like sing are used as heads of predicate phrases; nouns like girl as heads of referential phrases; adjectives like 'tall' as modifiers in referential phrases; and manner adverbs like 'beautifully' as modifiers in predicate phrases. Crucially, none of the content lexemes in (1) could be used directly in another function, i.e. without morpho-syntactic adaptation. Thus, in this example there is a one-to-one relation between function and lexeme class. Parts-of-speech systems of this type are called differentiated, and the lexical classes can all be said to be specialized for a certain propositional function.
    There are other parts-of-speech systems in which there is no one-to-one relation between the four propositional functions identified and the lexeme classes available. These systems are of two types. In the first type, a single class of lexemes is used in more than one propositional function. Such lexeme classes, and the parts-of-speech systems in which they appear, are called flexible.
    In order to arrive at word classes, we can now analyze every lexeme in a given language to see which of these functions it can fulfill without further measures (like copula support) being taken. Some lexemes can fulfill only one function, while others can fulfill more than one. Lexemes that can fulfill exactly the same functions are said to be in the same class. Table 4 lists 14 classes and the functions that can be fulfilled by the members.

	HP	HR	MR	MP
Verb	+	-	-	-
Noun	±	+	-	-
Adjective	±	-	+	-
Manner adverb	-	-	-	+
Predicative	+	-	-	+
Nominal	-	+	+	-
Modifier	-	-	+	+
Non-verb	-	+	+	+
contentive	+	+	+	+

Table 4. Taxonomy of parts-of-speech as defined by discourse-pragmatic functions
(adapted from Smit 2006).[13]

Table 4 contains familiar terms from parts-of-speech research, like noun or verb. A verb is defined in this theory as a lexeme that can be used as head of predicate phrase as it is, but must undergo further measures to be used in other functions. A manner adverb is defined as a lexeme which can only be used as modifier of predicate without further measures being taken, but cannot be used in other functions as is. Nouns are specialized for the function head of referent and adjectives are specialized for the function modifier of referent. Both may or may not be used as head of predicate. Parts-of-speech with only one function are called rigid.

These are the basic four well-known parts-of-speech reinterpreted in a discourse framework. Parts-of-speech which have more than one function are called flexible and are given further down the table, like the nominal, which can be used as head or modifier of referent, but nowhere else; the non-verb, which can be used anywhere but as the head of predicate; and finally the contentive, which excels in all functions.

In this introduction, I have outlined the functional definition of parts-of-speech and the resulting taxonomy of parts-of-speech as well as the parts-of-speech constellations that are

---

[13] 'Noun' and 'Adjective' are different from the other classes in that they have ± for the use as head of Predicate Phrase. They may or may not require further measures to be used in this function. Other combinations of discourse-pragmatic functions (e.g. + - + -) are logically possible, but have not been proposed up to now.

said to exist. In the following, I will apply these definitions to the Sri Lanka Malay distributional classes established above and see which parts-of-speech can be found in this language.

5.1 The verb

Members of the SLM verb class established above can be used as the head of predicate without further measures being taken. They cannot be used for any other function as they are (cf. Table 5). This means that this class is indeed also a verb in the Hengeveldian sense. We will illustrate this in the following sections.

	Head	Modifier
Predicate phrase	no further measure *(62) Kithan=le anthi-banthu* 1PL ASSOC= IRR-help 'We will help, too.' (B060115nar02)	further measure: reduplication . *(63) Kancil **lompath˜lompath** arà-laari* rabbit jump˜jump NONPAST-run 'The rabbit runs away jumping.' (K081112eli01)
Referential phrase	further measure: infinitive formation (64) ***Mà**-banthu baae* INF -help good 'To help is good.' (K081112eli01)	further measure: relative clause formation *(65) [Arà-maakang] oorang* NONPAST-eat man 'The eating man.' (K081112eli01)

Table 5. The SLM verb in different discourse functions.

Example (62) shows that lexemes of the class of verbs can be used as Head of Predicate Phrase without further measures being taken. For the other functions, further measures are necessary. In order to be used as modifiers in the predicate phrase, verbs have to be reduplicated as shown in (63). In order for verbs to head a referential phrase, the infinitive prefix *mà-* has to be used (64).[14] For modifier function in the referential phrase, a relative clause construction must be employed (65).

5.2 The noun

Members of the SLM noun class can be used as head or modifier of the referential phrase. They cannot be used as head or modifier of the predicate phrase (cf. Table 6).

---

[14] An alternative is the nominalizer *–an.*

	Head	Modifier
Predicate phrase	no further measure *(66) 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)	further measure: adverbialization *(67) Incayang **swaara=nang** arà-oomong* he sound= DAT NONPAST-speak 'He speaks with a loud voice.' (eli14122005) (K081112eli01)
Referential phrase	no further measure *(68) Se=ppe baapa incayang=nang ummas su-kaasi* 1S=POSS father NONPAST 3s= DAT gold PAST -give 'My father gave him the gold.' (K070000wrt04)	no further measure *(69) Panthas rooja kumbang pohong komplok duuwa asà-jaadi su-aada* beautiful rose flower tree bush two CP-grow PAST-exist 'Two beautiful rose bushes had grown.' (K070000wrt04)

Table 6. The SLM noun in different discourse functions.

Example (68) shows the use of the lexemes baapa 'father' and ummas 'gold' as heads of the referential phrase. No further measures are required for these lexemes to be used in this position. Lexemes of the SLM noun class can be used to modify a head of referential phrase, like *komplok* 'bush', which modifies *pohong* 'tree' in (69), and is further modified by *rooja kumbong* 'rose flower', itself consisting of the head *kumbang* 'flower' and the modifier *rooja* 'rose'. Lexemes of the noun class in SLM can also be used as head of predicate phrase without having to resort to further measures such as the use of a copula, as shown in (66). In order to be used as a modifier of predicate, SLM lexemes of the noun class have to take the dative marker =*nang*. Without it, the sentence in (67) would be ungrammatical; hence it is a further measure.

5.3 The Flexible Adjective

We have established that the SLM verb is a rigid class, and that the SLM noun is moderately flexible. We now turn to SLM Flexible Adjectives, which can be used for any discourse function. Table 7 shows that Flexible Adjectives can be used in all four discourse functions without further measures being taken. This means that SLM Flexible Adjectives are 'contentives' in the Hengeveldian terminology.

	Head	Modifier
Predicate phrase	no further measure *(70) Samma oorang baae* all man good 'All men are good.' (B060115cvs13)	no further measure *(71) Incayang pullang arà-oomong* he slow NONPAST -speak 'He speaks softly.' (eli14122005) (K081112eli01)
Referential phrase	no further measure *(72) Incayang hatthu iitham* 3S INDEF black 'He is a dark person.' (K081103eli02)	no further measure *(73) Se=ppe bìssar lae ruuma aada* 1S=POSS big other house EXIST 'There is another big house of mine.' (B060115cvs09)

Table 7. The SLM Flexible Adjective in different discourse functions.

Members of the SLM Flexible Adjective class can be used as head of term, as (72) shows. Iitham 'black' is a Contentive, which is used as the head of the referential phrase. In example (73), *bìssar* 'big' modifies the head *ruuma* 'house', without further measures being taken. Predicate phrases can have a Flexible Adjective in modifier position without further measures being taken: example (71) shows the use of the ~~Contentive~~ *pullam* 'slow' to modify the head verb *oomong* 'speak'. Note that this contrasts with the obligatory use of =*nang* in (67) for nouns in this position.

A difference between Flexible Adjectives and verbs is that the former can be used in different predication types. The Property Assignment Predication is exemplified by (70) and (74). No additional material is needed for a Flexible Adjective to occur in this predication type.

(74) *Aanak   thiiňğgi*
     child   tall
     'The child is tall.' (K081103eli02)

Another predication type where Flexible Adjectives can occur in is the Dynamic Predication, signaled by TAM-morphology on the predicate. Members of the SLM Flexible Adjective class can occur in this position without further measures being taken, as (75) shows.

(75) *Itthu=nam blaakang=jo, kitham   pada* **anà-bìssar**
     DIST   after= EMPH   1PL   PL   PAST-big
     'After that, we grew up.' (K060108nar02)

Note that aktionsart is static in (70) and (74), typical for the Property Predication in SLM, while (75) is an accomplishment, indicated by the use of the Dynamic Predication. The following example shows one and the same lexeme *dhiinging* 'cold' used in both predication types:

(76) a *Thee dhiinging*
    tea   cold
    N ADJ
    'The tea is cold.' (K081112eli01)

   b *Thee arà-dhiinging*
    tea   NONPAST-cold
    N   V
    'The tea is getting cold.' (K081112eli01)

This is irrelevant for the Hengeveldian approach, which does not distinguish between different types of predications; nevertheless a more fine-grained approach could be warranted (Haig 2008). The third predication type a Contentive can occur in is the Class Membership Predication, which uses the indefiniteness marker *atthu*. This is not very frequent. An example is given in (78).

(77) *Incayang   hatthu   iitham*
    3s   INDEF   black
    'He is a dark person.' (K081103eli02)

5.4 Summary

We have seen examples of the use of members of different distributional classes for different discourse functions. Table 8 sums up the findings of the discourse functions the SLM parts-of-speech can fulfill and gives the names in Hengeveldian taxonomy: The SLM verb can only be used as head of predicate, and is therefore a verb (Table 4). The SLM noun can be used as a head or modifier of the referential phrase and is therefore a nominal (Table 4). The SLM Flexible Adjective can be used for any function and is therefore a contentive (Table 4).

Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
SLM Verb (verb)	SLM Noun (nominal)		
SLM Flexible Adjective (contentive)			

Table 8. Schematic representation of the SLM parts-of-speech and propositional functions

The functional approach to parts-of-speech started out with the premise of cross-linguistic comparability and chose to use discourse functions instead of distribution classes to establish parts of speech. This procedure could be successfully applied for the redefinition of SLM parts-of-speech.

6 Diachrony of PoS systems

6.1 Specialization and generalization

Parts-of-speech systems are very often seen as static. This makes it easy to deal with them, but on the other hand we know that languages are not static in that there is always some part of the grammar that is changing. We know that languages change their word order, their morphological type, their alignment type, their stress system etc. Parts-of-speech systems can also change over time.[15][16]

An adequate theory of Parts-of-speech must then be able to capture the transitions from one type to another. Basically, we see can two historical developments: loss of discourse functions and gain of discourse functions (Table 9).[17]

	FA	FB			FA	FB
Lexeme at t_0	+	+		Lexeme at t_0	+	-
	↓ ↙				↓ ↘	
Lexeme at t_1	+	-		Lexeme at t_1	+	+

Loss of a function, specialization. The lexeme has both discourse functions FA and FB at t_0, but has lost FB at t_1.

Gain of a function, generalization. The lexeme has only function FA at t_0, but has gained FB at t_1.

Table 9. Direction of change of parts-of-speech systems.

Linguistic change is gradual. This means that not all of the lexemes will be able to be used in a new function at once. Rather, there will be pioneers and latecomers. Pioneers are normally derived lexemes, while latecomers are lexemes in closed classes, a kind of residue. Derived classes are then the first sign of incipient specialization, while closed classes indicate the last stages of generalization. We will take a look at specialization and leave the processes leading to generalization as a future research project.

---

[15] To be in constant change is actually seen as the normal case for parts-of-speech systems in Hengeveld (1992: 69).

[16] Cf. the flexibility of Late Archaic Chinese, which gave way to the more rigid modern system (Bisang 2008b).

[17] Vogel (2000) discusses the emergence of lexical categories, which she calls grammaticalization and the disappearance thereof, which she terms degrammaticalization. The former is similar to specialization, and the latter to generalization as used in this paper. Vogel's use of the terms applies to the emergence/disappearance of differentiated parts-of-speech, not to the gain/loss of discourse functions, although the two are obviously related.

6.2 Diachrony of the SLM system

Sri Lanka Malay is the descendant of varieties of Trade Malay spoken in Jakarta.[18] While we have no description of the parts-of-speech system of this historic variety, two factors make it seem very likely that it was a language with an extremely flexible part-of-speech system: genetic affiliation to Malayic and Austronesian, and its origin as a trade language. Malay languages (and Austronesian languages in general) are characterized by the fact that there are no clear boundaries between the traditional lexical word classes as recognized in the better-known European languages. Himmelmann (2005: 128f.) writes:

> [On the lexeme level], it is frequently noted in descriptions of western Austronesian languages that lexical bases (roots) are underdetermined in allowing both nominal and verbal derivations or uses. Alternatively, a basic distinction, between nouns and verbs (and possibly adjectives) is made for lexical bases but then it is stated elsewhere in the grammar that nominal bases can be used as (morphosyntactic) verbs essentially in the same way as verbal bases. [...]
> Multifunctional lexical bases, which occur without further affixation in a variety of syntactic functions occur [...] with some frequency in [...] many Malayic varieties.

Himmelmann's description can be reworded as 'lexemes can be used in a variety of functions without further measures being taken', matching precisely the Hengeveldian definition of flexible parts-of-speech. An example of this flexibility is given by Ewing (2005: 230f.) for Colloquial Indonesian as spoken in Java.

(78) *Kalau   saya   cerita   begini*   COLLOQUIAL INDONESIAN
     if   1s   story   like.this
     'If I tell a story like this.'

(79) *Makan juga nggak boléh   di-taroh di   situ* COLLOQUIAL INDONESIAN
     eat   also   NEG   allow   UV-put   LOC there
     'Food also isn't allowed to be put here.'

In (78) the word *cerita* 'story' is used without a verbalizer when it serves as the main predicate (head of predicate phrase). In (79) the word *makan* 'eat' is used without a nominalizer when it serves as the head of a referential phrase.[19] Lexical categories in another variety of Malay, Riau Indonesian, have been investigated inseveral papers by Gil

---

[18] Not to be confused with Betawi Malay, spoken as a native language by Jakarta Malays at the same period, and also distinct from Jakartan Indonesian as spoken today. Jakarta Malay as described by Grijns (1991) is also an offshoot of these varieties. See also Paauw (2004) for a sociohistorical and linguistic analysis of the origin of the SLM immigrants and Adelaar (1991) and Adelaar and Prentice (1996) for a different view. Both are discussed in Nordhoff (2009).
[19] Similar ideas are expressed by Prentice (1990: 920), who observes that colloquial spoken Indonesian is characterized by extensive elision of affixes, resulting in a greater degree of overlap between word classes.

(for example, Gil 1994), who claims that Riau Indonesian lacks distinctive word classes.[20] The trade language spoken in Jakarta before the mercenaries left for Sri Lanka evolved thus in an ecology characterized by languages with weak or absent distinctions between word classes. If we add to this that the Malay variety the mercenaries communicated in was not their mother tongue, the lack of morphologically distinct word classes becomes even more likely given that trade languages (pidgins, lingue franche) tend to be morphologically more reduced than the related language spoken by natives. It is extremely likely that the cocktail of colloquial Malay varieties without specialized word classes distilled into a morphologically reduced trade language would result in a language without distinctive word classes. The same analysis has been proposed in Paauw (2004):

> In [Trade Malay], a word can function variously as a noun, verb, adjective or function word without any change in form, depending entirely upon the context and word order for its function.

Having established that the ancestor of SLM was most likely a monocategorial language, we must ask ourselves the question why SLM is no longer monocategorial today. The equivalents of examples (78) and (79) above in SLM are given below. We see that *criitha* 'story' cannot be used as a transitive predicate in SLM in (80). The use of the transitive verb *biilang* 'say' is obligatory. Example (81) illustrates that the use of the nominalizer *-an* is obligatory in SLM. Leaving it out would result in ungrammaticality.

(80) *See   giin   hatthu   criitha=ke   kala-**biilang***
     1S   like.this NDEF   story= SIMIL   if-say
     'If I tell a story like this.' (K081112eli01)

(81) *Makan-**an**=le   siithu   mà-thaaro thàràboole*
     eat-NMLZR=ASSOC there   INF-put   cannot
     'Food also isn't allowed to be put here.' (K081112eli01)

The reasons for the creation of distributional restriction on the use of nominals and verbs in certain syntactic slots can be sought on two different grounds: language contact on the one hand, and cognitive factors relating to time-stability on the other hand. These will be explored in the following two sections.

---

[20] An anonymous reviewer remarks that claims about the radically monocategorial status of Malay varieties are controversial and that 'there are informal claims that [the existence of semantic constraints] makes the radical claim untenable.' In this paper, I stick to published work and do not take into account informal claims, but I acknowledge their existence. As for methodological issues in Malay linguistics with special consideration of parts-of-speech, see Gil (this volume).

6.3 Language contact

SLM has been in intimate contact with Sri Lanka's majority languages Sinhala and Tamil for at least 300 years, and its structure has been heavily influenced by the adstrates (Adelaar 1991, Smith et al. 2004, Slomanson 2006, Ansaldo 2008, Nordhoff 2009). Different authors have argued for Tamil (Smith 2003, Smith and Paauw 2006, Smith et al. 2004) or Sinhala (Ansaldo 2008) being the more important contact language.

   This is difficult to evaluate, since Sinhala and Tamil share a very similar typological markup (Smith 2003). Smith (2003) argued that SLM follows Tamil where Sinhala and Tamil diverge. These claims were critically reassessed in Nordhoff (2009), who shows that the data do not warrant this conclusion. Smith's data could all be shown to not point toward clear Tamil influence. They mostly point towards joint influence, with some phenomena only being accountable in terms of Sinhala influence. In the present paper, I do not take a stance as to whether Sinhala or Tamil is the more important contact language. Following Nordhoff (2009), I take this to be an empirical question which is not answered yet.

   As it happens, Sinhala and Tamil are both rigid languages, i.e. languages which make a distinction between nouns, verbs and adjectives.[21] Sinhala has a large open class of adjectives (Gair 1967: 21), while Tamil also has adjectives, but the class is closed and hosts only a handful of underived lexemes (Schiffman 1999: 125). In both languages, these classes are usually designated as 'adjectives' and mainly host lexemes denoting property concepts. The term 'adjective' is used e.g. in Gair (1967) or Lehmann (1989), which corresponds in this case to Hengeveldian terminology, i.e. they can only be used as modifier of a Referential Phrase without further measures being taken. Sinhala and Tamil both have a class of verbs, which can only be used in predicative function. In order to use verbs in referential function, nominalizations have to be used in Sinhala (Garusinghe 1962: 62ff) and Tamil (Lehmann 1989: 300ff). Nouns can be used both referentially and predicatively in both languages.[22]

   An anonymous reviewer remarks that it is common for Sinhala nouns to appear in 'verbal' (= predicative) position without any additional morphosyntactic marking. This is true (for Tamil as well, actually), but does not present a problem in the Hengeveldian approach, since the function HEAD OF PREDICATE PHRASE has a ± for nouns (see Table 4). Nouns may or may not take additional morphology in order to appear in predicative position. What distinguishes nouns from verbs is that verbs must not be able to

----

[21] At least on the word level. On the phrase level, class membership is less clear cut, at least in Sinhala cf. Gair (2003: 796).

[22] Besides the classical use for nominal predicates (class-inclusion), there are some additional event predicates which are coded by nouns in Sinhala (action nominals in Gair and Paolillo 1989). An anonymous reviewer points out that this Sinhala structure is not found in SLM. This is true. The reviewer continues that Sinhala influence is therefore unlikely. This objection does not hold. It is possible that Sinhala had an influence on the general configuration of SLM parts-of-speech without that influence extending into very particular predicate types. A wholesale copy of the Sinhala system is no requirement for the postulation of partial influence. The SLM system has moved towards Sinhala and Tamil, but very particular structures of Sinhala (action nominals) or Tamil (closed class of adjectives) are not found in SLM. In spite of this, the SLM system is still closer to Sinhala and/or Tamil than to the systems of other Malays varieties.

appear in referential position without further measures being taken. This is the case in Sinhala, where the nominalizers *-ma* or *eka* must be used (Garusinghe 1962: 62ff), and in Tamil, where *-al*, *-ttal*, *-gai* or *-adu* are used (Lehmann 1989: 300ff).

The grammar of SLM has been influenced by its adstrates in all the main areas: phonology (retroflex consonants; Bichsel-Stettler 1989, Tapovanaye 1995), morphology (cases, finiteness distinctions; Smith et al. 2004, Ansaldo 2008, Slomanson 2006), syntax (SOV word order; Adelaar 1991) and semantics (e.g. the modal systems; Slomanson 2006).

Apparently, language contact has also affected the SLM parts-of-speech system. SLM has converged to match the overall structure of the island's languages and become part of the Sri Lankan sprachbund (Bakker 2006, Ansaldo 2008). In linguistic areas characterized by an absence of rigid parts-of-speech like large parts of Indonesia (Gil XXYY), other Malay varieties could keep their lexical flexibility, but this was not possible in Sri Lanka. While these external conditions trigger the language change in the mind of the individual speaker, there must be some cognitive disposition in the minds of the speakers, which allows them to assign lexemes to lexical categories. In the following section I will argue that [±TIME-STABILITY] is the determining factor for the SLM development.

6.4 The cognitive factor: time stability

The specialization of formerly underdetermined lexemes in the history of SLM happened along the semantic fault lines of [±TIME-STABILITY] (Givón 1984: 51): words typically used for objects (time-stable entities) lost the ability to be used as (dynamic) predicates and were grouped in the class of nominals, while words normally used in connection with events (entities that are not time-stable) lost their ability to be used for things and were grouped in the class of verbs.

Aayang 'chicken' is an example of a word that usually denotes a concrete object (a time-stable entity) that can only be used referentially in SLM without further measures being taken, while *maakang* 'eat' is an example of a word that most frequently denotes an activity and cannot be used referentially any more in SLM (cf. (81)). In distinction to, say, Riau Indonesian, morphological operations like infinitive prefixes, nominalizers, reduplication etc. are required in SLM in order to use *maakang* in any other function than head of predicate (cf. Table 5 above).

It is well known that semantic features of lexemes and potential pragmatic functions correlate (Croft 1991, 2001; Sasse 1993). Objects ([+time-stable]) correlate with reference, actions ([-time-stable]) correlate with predication and properties (intermediate) correlate with modification. In the case of the diachronic development of SLM parts-of-speech, words primarily used for [+time-stable] entities (objects) were restricted to their prototypical pragmatic function of reference and barred from other discourse functions. Words primarily used for [-time-stable] entities (events) in turn were restricted to their prototypical pragmatic function of predication and barred from other discourse functions without further measures being taken. The interesting cases are concepts which are intermediate on the scale of time-stability, i.e. properties. For these, there was no clear semantic cue on the time-stability scale as to their prototypically associated discourse function, and hence specialization did not take place: these lexemes can still be used for every discourse function and underlie no restrictions. Being on the middle ground of the

time-stability scale helped them retain the lexical flexibility typical of their ancestor variety.

At this point of the analysis, we recapitulate that a certain semantic feature of a certain lexeme evokes a certain semantic class which in turn attracts a prototypical discourse function. Over time, the lexeme becomes intimately connected with the discourse function, and use in non-prototypical functions has to be signaled (relative clause formation, infinitive prefixes and the like).

Put like this, it follows that specialization of word-classes is the morphosyntactic reflex of cognitive association of lexemes with prototypical discourse functions. The emergence of new word classes therefore has its origins not in morphosyntax nor in discourse proper, but in semantics and its association to discourse pragmatics.

This association between semantics and discourse function is where language contact kicks in: the differentiated associations between words exclusively used for time-stable entities found in the adstrates (and other words for [-time-stable] entities) had to be mastered by the arriving Malays, who learned the adstrates and applied the newly acquired association to their own language as well. They started with the most obvious cases, those which were clearly [+time-stable] or clearly [-time-stable].

The road for this specialization might actually have been paved beforehand by the existence of optional derivational morphology.[23] The nominalizer -*an* for instance is an affix which survived the state of trade language and is used productively today. There is no evidence whether or not it was obligatory in the variety of the mercenaries. Its use might have been optional, as is the case today in Colloquial Spoken Indonesian (cf. (79)). Speakers would then have the possibility to mark restriction to certain discourse functions if they deemed necessary.[24] The existence of a derivational target function indicates that the speakers could indicate this function if so desired, or leave the identification of the discourse function to the hearer if they did not judge the disambiguation necessary. As stated above, the lexemes derived with one of the optional affixes would then be the pioneers in the new lexical categories, while other lexemes slowly follow suit as their association with the particular discourse function grows tighter. Due to the lack of historical records, the precise development is difficult to reconstruct, but we can assume that the ancestor variety was monocategorial, and we know that the variety of today is not. What is missing is the 'pioneer' stage of derived lexemes claiming the new discourse function. This stage is not attested in the development of SLM, but we can find comparable transitory systems in other languages, which I will briefly illustrate in the next section.

---

[23] See Smit (2006) for similar ideas in a slightly different approach. Smit assumes that derivation is the source of the systems called intermediate in Hengeveld (1992). None of the intermediate systems Hengeveld (1992) proposes fits the SLM data, and the idea of intermediate systems has been abandoned by him since then, so that Smit's approach cannot be applied to the SLM case.
[24] Cf. Marchand (1969) on the 'categorizing' function of derivation. I would like to thank Geoffrey Haig for pointing out this reference to me.

6.5 Comparison with other flexible languages

In this section, I will take a look at languages with a class of contentives and at least one other lexeme class. In Hengeveld and van Lier (2008), we find Kambera (based on Klamer 1998), Santali (based on Neukom 2001) and Samoan (based on Mosel and Hovdhaugen 1992), which all fulfill this criterion. The parts-of-speech constellation of these languages is given in Table 10.

Samoan			
Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
Contentive$_O$			
Verb$_D$			

Kambera			
Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
Contentive$_O$			
Verb$_O$			Manner Adverb$_C$

Santali			
Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
Contentive$_O$			
Verb$_O$	Noun$_D$		

Sri Lanka Malay			
Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
Contentive$_O$			
Verb$_O$	Nominal$_O$		

Table 10. Parts-of-speech systems in various languages with a class of contentives.[25]

We observe that in Samoan, it is possible to derive verbs, but that these derived lexemes are the only ones restricted in their discourse function to head of predication. This is different from Kambera, where underived lexemes can also be found in the class of verbs.[26] In Santali finally, the class of verbs is well established, and yet other derived lexemes

---

[25] Subscript D signals that a class only contains derived lexemes; subscript C indicates a closed class and subscript O an open class.

[26] We do not have anything to say about manner adverbs for now and disregard this class.

specialize for head of Referential Phrase, i.e. they are members of the noun class. The SLM system is added in a fourth table, as a system where both classes of verbs and nominals have progressed beyond hosting only derived lexemes.

Samoan, Kambera, Santali and SLM then illustrate different stages in the specialization of parts-of-speech systems (cf. Table 9). New discourse functions are explored by derived lexemes (verbs in Samoan, nouns in Santali), and later populated by underived lexemes (verbs in Kambera, nominals in SLM). From these four examples, it appears that the development of new lexical categories in this manner proceeds from left to right (or from verbs to nouns), but investigation of more languages, including the diachrony of their parts-of-speech system will be needed to test this hypothesis (cf. Hengeveld and van Lier (2008), who claim that all constraints on PoS-classification converge towards the specialization of verbs, before any other word class.).

7 Conclusion

This paper has shown that Sri Lanka Malay has a maximally flexible class of adjectives, which would be treated as Contentives in Hengeveld's terminology. This maximally flexible class co-exists with the more rigid classes of nominals and verbs, showing that lexeme flexibility and monocategoriality do not necessarily accompany each other, but are orthogonal dimensions.

In diachronic perspective, nominals and verbs evolved from contentives having lost some discourse functions and thereby their flexibility. This change towards a more rigid system was probably triggered by the adstrates Sinhala and Tamil, in which the discourse function Predication is strongly associated with events and the function Reference, with objects. The fact that SLM already had the possibility to derive forms for more specific discourse functions probably accelerated the contact induced changes in the SLM parts-of-speech system.

References

Adelaar, K. A. (1991). 'Some notes on the origin of Sri Lankan Malay', in H. Steinhauser (ed.), *Papers in Austronesian linguistics, Pacific Linguistics, vol. 1*. Canberra: The Australian National University, 23-37.

Adelaar, K. A. and Prentice, D. J. (1996). 'Malay: its history, role and spread', in S. A. Wurm, P. Mühlhäusler and D. T. Tryon (eds.), *Atlas of Languages of Intercultural Communication* (Volume II.1). Berlin: Walter de Gruyter, 673-693.

Alleyne, M. C. (2000). 'Opposite processes in creolization', in I. Neumann-Holzshuh and E. W. Schneider (eds.), *Degrees of Restructuring in Creole Languages*. Amsterdam: Benjamins, 125-134.

Ansaldo, U. (2005). 'Typological admixture in Sri Lanka Malay. The case of Kirinda Java'. Unpublished manuscript. ~~Available at: http://home.medewerker.uva.nl/u.ansaldo/index.html.~~

---- (2008). 'Sri Lanka Malay revisited: Genesis and classification', in A. Dwyer, D. Harrison and D. Rood (eds.), '*A world of many voices: Lessons from documented endangered languages*. Amsterdam and Philadelphia: Benjamins, 13-42.

~~Ansaldo, U. and Nordhoff, S. (in preparation). 'Typological alignment and contact language formation in a trilingual ecology'.~~

Bakker, P. (2006). 'The Sri Lanka Sprachbund: The newcomers Portuguese and Malay', in Y. Matras, A. McMahon and N. Vincent (eds.), *Linguistic Areas - Convergence in Historical and Typological Perspective*. Houndmills NY: Palgrave Macmillan, 135-159.

Bhat, D. (1994). *The Adjectival Category: Criteria for Differentiation and Identification*. Amsterdam: Benjamins.

Bichsel-Stettler, A. (1989). Aspects of the Sri Lanka Malay community and its language. MA thesis, Universität Bern.

Bisang, W. (2008a). 'Precategoriality and argument structure in Late Archaic Chinese', in J. Leino (ed.), *Constructional Reorganization* (Constructional Approaches to Language 5). Amsterdam and Philadelphia: Benjamins, 55-88.

---- (2008b). 'Precategoriality and syntax-based parts of speech', *Studies in Language* 32-3: 568-589.

Bloomfield, L. (1933). *Language*. London: Allen and Unwin.

Broschart, J. (1997). 'Why Tongan does it diffferently', *Linguistic Typology* 1-2: 123-166.

Croft, W. (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

---- (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.

---- (2005). 'Word classes, parts of speech, and syntactic argumentation', *Linguistic Typology* 9-3: 431-441.

Ewing, M. C. (2005). 'Colloquial Indonesian', in K. A. Adelaar and N. P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge, 227-258.

Gair, J. W. (1967). 'Colloquial Sinhala Inflectional Categories and Parts of Speech', *Indian Linguistics* 27: 32-45. Reprinted in J. W. Gair (1998), *Sinhala and Other South East Asian Languages*. New York and Oxford: Oxford University Press, 13-25.

---- (2003). 'Sinhala', in G. Cardona and D. Jain (eds.), *The Indo-Aryan Languages*,

London: Routledge, 766-818.

Gair, J. W. and Paolillo J. C. (1989). 'Sinhala non-verbal sentences and argument structure', *Cornell Working Papers in Linguistics* 8: 39-78. Reprinted in J. W. Gair (1998), *Sinhala and Other South East Asian Languages*. New York and Oxford: Oxford University Press, 87-110.

Garusinghe, D. (1962). *Sinhala, the Spoken Idiom*. München: Hueber.

Gil, D. (1994). 'The Structure of Riau Indonesian', *Nordic Journal of Linguistics* 17: 179-200.

Givón, T. (1984). *Syntax: A Functional-Typological Introduction*. Amsterdam: Benjamins.

Grijns, C. D. (1991). *Jakarta Malay*. PhD dissertation, Universiteit Leiden.

Haspelmath, M. 2007. 'Pre-established categories don't exist: consequences for language description and typology', *Linguistic Typology* 11-1: 119-132.

Hengeveld, K. (1992). *Non-verbal Predication: Theory, Typology, Diachrony* (Functional Grammar Series 15). Berlin - New York: Mouton de Gruyter.

Hengeveld, K. and Mackenzie, L. (2008). *Functional Discourse Grammar: A Typologically-based Theory of Language Structure*. Oxford: Oxford University Press.

Hengeveld, K. and Rijkhoff, J. (2005). 'Mundari as a flexible language', *Linguistic Typology* 9-3: 406-431.

Hengeveld, K. and Lier, E. H. van (2008). 'Parts of speech and dependent clauses in Functional Discourse Grammar', *Studies in Language* 32-3, 753-785.

Hengeveld, K., Rijkhoff, J. and Siewierska, A. (2004). 'Part-of-speech systems and word order', *Journal of Linguistics* 40-3: 527-570.

Himmelmann, N. P. (2005). 'The Austronesian languages of Asia and Madagascar: typological characteristics', in K. A. Adelaar and N. P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge, 110-181.

Hussainmiya, B. A. (1990). *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Bangi: Penerbit Universiti Kebangsaan Malaysia.

Klamer, M. (1998). *A grammar of Kambera*. Berlin and New York: Mouton de Gruyter.

Lehmann, C. (1989). Language description and general comparative grammar', in G. Graustein and G. Leitner (eds.), *Reference Grammars and Modern Linguistic Theory*. Tübingen: Niemeyer, 133-162.

Marchand, H. (1969). *The Categories and Types of Present-day English Word-formation: a Synchronic-diachronic Approach*. Munich: Beck.

Mosel, U. and Hovdhaugen, E. (1992). *Samoan Reference Grammar* (Instituttet for sammenlignende kulturforskning, Series B: Skrifter, LXXXV). Oslo: Scandinavian University Press.

Neukom, L. (2001). *Santali*. Munich: Lincom.

Nordhoff, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. PhD dissertation, Universiteit van Amsterdam.

Paauw, S. H. (2004). *A historical analysis of the lexical sources of Sri Lanka Malay*. MA thesis, York University.

Peterson, J. (2005). 'There's a grain of truth in every "myth", or, Why the discussion of lexical classes in Mundari isn't quite over yet', *Linguistic Typology* 9-3: 391-441.

Prentice, D. J. (1990). 'Malay (Indonesian and Malaysian)', in B. Comrie (ed.), *The*

World's Major Languages. London: Routledge, 913-935.

Sasse, H.-J. (1993). 'Das Nomen - eine universale Kategorie?' *Sprachtypologie und Universalienforschung* 46: 187-221.

---- (1995). 'Syntactic categories and subcategories', in J. Jacobs et al. (eds.), *Syntax* (Handbücher zur Sprach- und Kommunikationswissenschaft 9.1). Berlin and New York: Mouton de Gruyter, 646-686.

Schiffman, H. (1999). *A Reference Grammar of Spoken Tamil*. Cambridge: Cambridge University Press.

Slomanson, P. (2006). 'Sri Lankan Malay morphosyntax: Lankan or Malay?', in A. Deumert and S. Durleman (eds.), *Structure and Variation in Language Contact*. Amsterdam: Benjamins, 135-158.

Smit, N. (2006). 'Intermediate PoS-systems - Clues from derivational morphology'. Paper presented at the conference *Universality and Particularity of Parts-of-Speech systems*, Amsterdam, June 2006.

Smith, I. R. (2003). 'The provenance and timing of substrate influence in Sri Lankan Malay; Definiteness, animacy and accusative case marking.' Paper presented at the *South Asian Language Analysis Roundtable XXIII*, Austin, Texas.

Smith, I. R. and Paauw, S. (2006). 'Sri Lanka Malay: Creole or convert', in A. Deumert and S. Durleman (eds.), *Structure and Variation in Language Contact*. Amsterdam: Benjamins, 159-182.

Smith, I. R., Paauw, S. and Hussainmiya, B. A. (2004). 'Sri Lanka Malay: the state of the art.' *Yearbook of South Asian Languages* 2004: 197-215.

Tapovanaye, S. (1995). *Vowel Length in Sri Lankan Malay*. MA thesis, University of Hawaii.

Vlekke, B. H.M. (1943). *Nusantara: a history of the East Indian Archipelago*. Cambridge: Harvard University Press.

Vogel, P. M. (2000). 'Grammaticalization and part-of-speech systems', in P. M. Vogel and B. Comrie (eds.), *Approaches to the Typology of Word Classes*, (Empirical Approaches to Language Typology 23). Berlin and New York: Mouton de Gruyter, 259-284.

Wierzbicka, A. (1986). 'What's in a noun (or: How do nouns differ in meaning from adjectives?)', *Studies in Language* 10: 353-389.

Abbreviations

ACC accusative; ASSOC associative; COMPL completive; CP conjunctive participle; DAT dative; DEM demonstrative; DIST distal; EMPH emphatic; INDEF indefinite article; INF infinitive; INTERR interrogative; IRR irrealis; LOC locative; NEG negation; NMLZR nominalizer; NONV nonverbal; PL plural; POSS possessive; PROX proximal; QUOT quotative; UV undergoer voice.