

Transparency in Sri Lanka Malay

Sebastian Nordhoff

January 13, 2010

1 Introduction

...

2 Interpersonal-Representational

2.1 No crossreference

Hengeveld & Mackenzie (2008:350) speak of cross-reference when ‘person marking on the verb is sufficient by itself and may optionally be expanded by a lexically realized argument.’ There is no person marking on the verb in Sri Lanka Malay, so that this criterion does not apply. The following example shows that the form of the verb is invariant no matter the reference of the argument (represented by X). Whatever the person, number or gender of X, the verb will remain the same, *su-dhaathang* in (1).

- (1) *X su-dhaathang*
past-come
‘I/you/he/she/it/we/you/they came.’

2.2 No apposition

In the mapping the interpersonal level to the representational level, apposition represents a one-to-many relation. This is not expected under transparency. Sri Lanka Malay is not transparent in this regard since it allows apposition. The following example shows an instance of apposition. The referents *Mr Sebastian* and *see ‘I’* are introduced and subsequently referred to by the personal pronoun *kitham* ‘we’. *Kitham* alone is already sufficient on the representational level. However, *duuva* ‘two’ again refers to the same two referents in the world as does *kitham*; we are dealing with apposition here.

- (2) *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)

Another doubling of interpersonal information on the representational level, albeit not apposition, is the common double occurrence of the indefinite article *hatthu*. One occurrence would be enough to signal the unidentifiable status of the referent to the

hearer, yet it is often found twice, as in (3). This is an intransparent mapping between the interpersonal and the representational level.

- (3) *Sithu=ka hathu=maccan=hathu duuduk aada.*
 there=LOC INDEF=tiger=INDEF stay exist
 ‘There was a tiger.’

2.3 No limitations on which semantic units can be chosen as predicates

In transparent languages, we do not expect restrictions on what semantic units can be chosen as predicates.

In Sri Lanka Malay, verbs, adjectives, modals and locatives can be used as predicates without further measures being taken. Only verbs can take TAM-prefixes though. TAM for other predicates must be expressed lexically.¹

- (4) *Aajuth thaakuth=ka su-naangis_V*
 dwarf fear=LOC PAST-weep
 ‘The dwarf wept in fear.’ (K070000wrt04)
- (5) *Dee buthul jahhath_{Adj.}*
 3 very wicked
 ‘He was very wicked.’ (K051205nar02)
- (6) *Deran=anng thumpath maau_{Modal}*
 3PL=DAT place want
 ‘They wanted land.’ (N060113nar01)
- (7) *Itthu bannyak laama hathu ruuma_N*
 DIST very old INDEF house
 ‘That one was a very old house.’ (K070000wrt04)
- (8) *Se=ppe kaake hathu estate=ka_{Loc.}*
 1s=POSS grandfather INDEF estate=LOC
 ‘My grandfather was on an estate.’ (K051205nar05)

While further measures are not necessary, SLM has an optional copula, which can be used for nominal predicates.

- (9) *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)

¹Adjectives can convert to verbs and afford all of verbal morphology, but they are no longer adjectives then.

3 Representational-Morphosyntactic

3.1 No grammatical relations (but semantic or pragmatic alignment)

In the mapping between the representational level and the morphosyntactic level, some languages use an intermediate level of grammatical relations, which can change the direct mapping of semantic function on morphosyntactic expression. In Sri Lanka Malay, there are no grammatical relations, instead the semantic function (agent, patient, recipient, etc) is directly mapped onto morphosyntax, where it is expressed by case postpositions. The following example shows the use of the postpositions =*yang* ‘acc’, =*dering* ‘abl’ and =*nang* ‘goal’. The agent *incayang* ‘he’ is not marked by a postposition.

- (10) *Itthu baathu=yang incayang Seelong=dering laayeng nigiri=nang asà-baapi.*
DIST stone=acc 3s.POLITE Ceylon-abl other country=dat cp-bring
PAT AG SRC GOAL V
‘He brought those stones from Ceylon to other countries.’ (nosource)

Crucially, there is no way of changing the case postposition an argument takes. While the English Passive Alternation as in *she beat him/he was beaten by her* changes the marking of the patient from accusative to nominative and the marking of the agent from nominative to oblique *by-agent*, in SLM such operations do not exist. There is no possibility to promote or demote an argument morphosyntactically.

Nordhoff (2009) has applied an array of tests for subjecthood to Sri Lanka Malay; none of the tests yielded evidence for the category “subject” (or “object”). For reasons of space, the tests will not be repeated here. However, the three traditional coding properties of subjects can briefly be discussed. These are agreement, word order and case marking. As discussed in Section 2.1, there is no agreement in SLM. Word order in SLM is generally verb-final, but the arguments of the verb can occur in any order to the left of the verb. This can already be seen from example (10), where we find the order PAT AG SRC GOAL V. If word order were an important parameter, we would either expect the subject in initial position (SOV) or in direct vicinity of the verb (OSV). In (10), neither is the case. The best candidate for subjecthood, the pronoun *incayang* is neither in initial position nor adjacent to the verb. (10) thus already sheds doubt on word order as a criterion for grammatical relations. Nordhoff (2009) discusses variations in word order in more detail and concludes that word order in the preverbal field is completely free and cannot be used to establish grammatical relations. The last criterion is case marking. If we find that the only argument of an intransitive predicate (S) is always case-marked in the same way as either the actor (A) or the undergoer (P) of a transitive sentence, we have good arguments for subjecthood. This test also fails in SLM. The only argument of an intransitive predicate in SLM can be marked with either zero, =*yang* ‘acc’, =*nang*/= *dang* ‘dat’, or =*dering* ‘abl’.

- (11) *Dee=Ø su-thiidor baava=ka.*
3s PAST-sleep down=loc
‘He slept downstairs.’ (K051205nar05)

- (12) Titanic *kappal*=**yang** *su-thìnggalam*.
 Titanic ship=ACC PAST-sink
 ‘The ship “Titanic” sank.’ (K081104eli05)
- (13) *Go*=**dang** *karang bannnyak thàràsìggar*.
 1s.FAMILIAR=DAT NOW very sick
 ‘I am now very sick.’ (B060115nar04)
- (14) Police=**dering** *su-dhaathang*.
 police=INSTR PAST-come
 ‘The police came.’ (K081105eli02)

For transitive sentences, the most typical case combinations include

- \emptyset - \emptyset
- \emptyset -ACC
- \emptyset -DAT
- DAT- \emptyset
- DAT-ACC
- DAT-DAT
- ABL- \emptyset
- ABL-ACC
- ABL-DAT

There is no space here to illustrate all these patterns, but the following three examples should suffice to illustrate the diversity.

- (15) *Rose-red*= \emptyset *buurung*=**nang** *su-puukul*.
 Rose-red bird=DAT PAST-hit
 ‘Rose-red hit the bird.’ (K070000wrt04)
- (16) *aathi*=**yang** *sajja hatthu oorang*=**nang** *bole*=*ambel*.
 liver=ACC only one man=DAT can-take
 ‘Only one person can take the liver.’
- (17) *See*=**yang** *police*=**dering** *nya-preksa*.
 1s=ACC police=ABL PAST-enquire
 ‘I was questioned by the police.’ (K051213nar01)

Given this diversity of case marking, it is obvious that there is no clear mapping between the cases found in intransitive sentences and in transitive sentences. This means that the third coding property for subjects, case marking, fails as well.

3.2 No discontinuity

Tearing apart in morphosyntax something which belongs together in semantics is an intransparent operation, leading to discontinuous constituents. Such constituents are not found in Sri Lanka Malay.

3.3 Function marking and derivational processes not sensitive to nature of input

The marking of a semantic function is the more transparent the less parameters it depends on. The most transparent relation is found in cases where the only parameter is the semantic function itself, and other parameters from the realm of morphosyntax or phonology do not play a role. That is, it should make no difference to the expression of a function whether the referent is encoded as a noun, a pronoun, an adjective, a clause, or anything else. This is what we find in Sri Lanka Malay: function marking is indeed indifferent to morphosyntactic properties. Semantic roles are marked by enclitic postpositions, e.g. the dative marker *=nang*, which can mark recipient, experiencer, and manner, among other roles. These postpositions can attach to any type of argument. The following examples show *=nang* attached to a noun, a pronoun, a deictic, an adjective and a clause. There is thus no morphosyntactic restriction on the combinatorial properties of *=nang* with this set. The same is true for the other postpositions, like *=yang* ‘acc’, but these are more difficult to combine with clauses for semantic reasons.

- (18) *Laayeng nigiri=pe soojor pada=nang baae lakuvan_N=nang*
different country=POSS European PL=DAT good price=DAT
anà-juuval.
PAST-sell
‘(He) sold (it) to the Europeans from the other countries for a good price.’
(K060103nar01)
- (19) *Kithang_{Pron}=nang baaye=nang mulbar bole=baaca.*
1 PL=DAT good=DAT Tamil can=read
‘We can read Tamil well.’ (K051222nar06)
- (20) *Itthu_{Deic}=nang blaakang aapa nya-gijja?*
DIST=DAT after what PAST-make
‘What did (they) do then?’ (K051206nar07)
- (21) *Baae_{Adj}=nang arà-nyaanyi.*
good=DAT NON.PAST-sing.
‘He sings well.’ (K081103eli02)
- (22) *Suda buthul suuka nyaari siini su-dhaathang_{CLS}=nang .*
thus correct like today here PAST-come=DAT
‘So I very much liked that you came here today.’ (G051222nar01)

Maximally transparent derivational markers should also not show selectivity with regard to their host. This is true for the (derivational) plural marker *pada*. The follow-

ing examples show the use of the plural marker on a noun, a pronoun, a deictic, and a relative clause.

- (23) *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)
- (24) *Itthu=nam blaakang=jo, kitham_{Pron} pada anà-bissar.*
DIST after=EMPH 1PL PL PAST-big
‘After that, we grew up.’ (K060108nar02)
- (25) *Incayang itthu_{Deic} pada=yang kapang-thumpa-king, itthu_{Deic} 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.’
- (26) *[[Seelon=nang anà-dhaathang]_{CLS} pada] mlaayu pada.*
Ceylon=DAT PAST-COME PL Malay PL
‘Those who had come to Ceylon were the Malays.’ (N060113nar01)

While the plural marker is not selective to the nature of its host, the same is not true of the nominalizer *-an* and the causativizer *-king*. *-an* can attach to verbs, adjectives and nouns, but not to pronouns, deictics or clauses.

- (27) *jalang_V-an hatthu arà-pii vakthu*
walk-NMLZR INDEF NON.PAST-go time
‘when we go on a trip’ (K051213nar06)
- (28) *Manis_{ADJ}-an maakang=nang go suuka bannyak.*
sweet-NMLZR eat=DAT 1S.FAMILIAR like much
‘I like very much to eat sweets.’ (B060115prs20)

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.

The selectivity of *-an* is quite clear in SLM, nevertheless, there is one instance of a nominalization after inflection, i.e. nominalization of a phrase rather than of a stem. This is *thradahan* ‘deprivation’ which is composed of the negative 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. This makes the selection restrictions of *-an* less narrow, but it is still true that in the great majority of cases, *-an* cannot be used to nominalize phrases; it can only be used to nominalize stems.

The causativizer *-king* is another derivational morpheme. It can attach to verbs and adjectives and marginally to nouns, but not to pronouns, deictics or clauses either.

- (29) *Itthu=ka 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)
- (30) *Baaye meera caaya kapang-jaadi, **thurung**_V-king.*
 good red colour when-become, descend-CAUS
 ‘When [the food] has turned to a nice rose colour, remove (it) [from the fire].’
 (K060103rec02)
- (31) *Spaaman=yang asà-**kafan**_N-king, spaaman=yang sithu=ka*
 3S.POLITE=ACC CP-shroud-CAUS 3S.POLITE=ACC there=LOC
nya-kubuur-king.
 PAST-buried-CAUS
 ‘The body was wrapped in cloth and then the body was finally buried.’
 (B060115nar05)

It can be noted that *-an* is a suffix on prosodic grounds while *pada* is a clitic. This is mirrored by their morphosyntactic behaviour: *pada* does not care for the nature of its host, while *-an* does. *-king* on the other hand has an intermediate position with regard to prosody, where there are reasons to treat it as a suffix, but also as a clitic. This is not mirrored by its morphosyntactic properties.

4 Morphosyntactic

4.1 No expletive elements

Transparent languages should not have elements in morphosyntax that correspond to nothing on the representational level. This is the case in SLM. SLM has no dummy subjects. The non-existence of items is always difficult to demonstrate, here I use a meteorological verb, where no expletive element is present, and none can be present.

- (32) *Arà-uujang*
 NON.PAST-rain
 ‘It is raining.’

4.2 No tense copying

Transparent languages are expected to show always the ‘real’ tense value from the representational level in morphosyntax. Changing tenses (tense copying, consecutio temporum) are not expected. Thus, constructions like English *He said that he had two brothers*, where the past tense is used despite the present tense meaning are not expected to occur in transparent languages. Indeed, SLM does not show tense copying.

- (33) *Incayang su-biilang duuva kaaka arà-duuduk katha*
 3s PAST-say two elder.brother NON.PAST exist.ANIM QUOT
 ‘He said that he has two elder brothers/ that there exist two brothers (to him).’

While no past tense form is ever used for present contexts, it is possible to find the non-past form *arà* in past contexts. As argued for in Nordhoff (2009:289f), this is due to the polysemy of this form. Besides the more common meaning as ‘non-past’, this form can also be used as ‘simultaneous’, and this is what we find in examples like (34).

- (34) *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)

While polysemy is not exactly transparent either, the intransparency we find here is due to the lexical entry of *arà*-, and not to a morphosyntactic rule of tense copying.

4.3 No raising

In a transparent language, we would expect every argument to surface in the clause where it semantically belongs. Raising constructions like *John seems to be intelligent* are not transparent in the sense that *John* semantically is an argument of the lower clause but shows up in the higher clause in morphosyntax. Such structures are not found in Sri Lanka Malay.

4.4 No grammatical gender, declension, conjugation

A transparent language is not expected to have elements in morphosyntax which are not motivated on semantic grounds, i.e. elements whose form depends on arbitrary criteria like membership in a certain declension or conjugation class, or grammatical gender. Sri Lanka Malay has no declension or conjugation classes. There is no arbitrary gender assignment (as in German or French) either, even natural gender (sex-based classification) is very marginal. The only instance I am aware of is the pair *puthra* ‘prince’, *puthri* ‘princess’.

4.5 No agreement (but pronominal arguments)

As stated above in Section 2.1, there is no cross-reference, and thus no agreement.

4.6 Phrase marking through clitics rather than head marking through affixes

This is also what we find in SLM as discussed in Section 3.3.

4.7 No fusional morphology

Transparent languages are not expected to express more than one meaning per morpheme. This is to say that we do not expect any fused portmanteau forms. While in the great lines, SLM does not have fusional morphology, some allomorphs of case markers could be analyzed as portmanteau forms. This is the case for the allomorph *=dang* of the dative marker *=nang*. This is obligatory for the monosyllabic pronouns *see* '1s', *goo* '1s', *luu* '2s', *dee* '3s'. Another instance is the allomorph *=ppe* of the possessive marker *=pe*, which attaches to the same four items. One can then postulate that both *=dang* and *=ppe* carry a meaning of SINGULAR and PRONOUN besides their normal case semantics. This means that more than one meaning is expressed in one form, an instance of fusional morphology.

5 Morphosyntactic/Phonological

5.1 Phonological phrasing and morphosyntactic phrasing run parallel

In the interface between morphosyntax and phonology, a transparent mapping is found if there is a 1:1 correspondence between morphosyntactic phrasing and phonological phrasing. The Major morphosyntactic constituent in Sri Lanka Malay are the clause CLS, the predicate PRED and the noun phrase NP. NPs are not restricted to nominal heads. Indeed, numerals, adjectives, pronouns, deictics, quantifiers, and even sentences and utterances can head NPs without further measures being taken. This means that the normal structure of the clause can be represented as NP* PRED. On the phonological level, Nordhoff (2009) distinguishes Presuppositive Phrases with a LH boundary tone from Assertive Phrases with a L boundary tone. There seems to be a 1:1 mapping of NPs on Presuppositive Phrases, and PRED on Assertive Phrases, so that the phrasing runs parallel on both levels and is therefore transparent.

5.2 Phonological weight does not influence morphosyntactic placement

A dramatically transparent language would completely ignore phonological weight when determining the order of constituents. Constituent order would be solely determined by semantics. This is not the case for SLM. While normally all arguments are preverbal, very heavy constituents can be shifted to postverbal position. This is frequently found for reported utterances (35), but other constituents can also be shifted, for instance the complement of a modal in (36).

- (35) *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.'

- (36) *Kitha=nang maau [kitham=pe mlaayu lorang blaajar lorang=pe mlaayu*
 1PL=DAT want 1PL=POSS Malay 2PL learn 2PL=POSS Malay
*kitham blaajar]*_{NP}.
 1PL learn
 ‘We want that you learn our Malay and that we learn your Malay.’
 (K060116nar02)

Right extraposition of heavy constituents is one possibility to facilitate parsing, the other one is shifting heavy nominal modifiers to the right. This is what can be found with relative clauses as in (37), where the relative clause consisting of eleven morphemes is shifted to the left, and the short indefinite marker *hatthu* is now found between the relative clause and the head noun. Semantically, *hatthu* should have scope over the relative clause as well, but this is not mirrored in morphosyntax; the phonological weight has taken precedence over semantic considerations when placing the constituents in this sentence.

- (37) [*Seelon=le kithang=pe mlaayu=nang=le hatthu bagiyan anà-aada*]_{REL}
 Ceylon=ADDIT 1PL=POSS Malay=DAT=ADDIT INDEF part exist
*[hatthu]*_{INDEF} *[nigiri]*_N *su-jaadi*
 INDEF country PAST-become
 ‘Ceylon became a country where our Malays also have a part in.’
 (K051222nar04)

6 Phonology

6.1 No sandhi rules

In a transparent languages, there should be no word-external sandhi. In SLM, we can distinguish combinations of base+affix, base+clitic, compounds, and combinations of independent words as candidate domains for sandhi. As for affixes, we find that the numeral suffixes *-blas* ‘-teen’ and *-pulu* ‘-ty’ cause a labial articulation of the final nasal in the word *dhlaapan* ‘eight’ (*dhlapamblas*, *dhlapampulu*) as well as the dropping of the final consonant in *ùmpath* ‘four’ at least for some realizations of *ùmpa(th)pulu* ‘forty’. Combinations of base+clitic also often show assimilation of nasals to the following consonant, e.g. *oorang* ‘man’+*=pe* ‘poss’=*oorampe* ‘of the man’. Compounds and strings of independent words generally do not undergo sandhi.

6.2 No degemination

In a transparent mapping of morphology onto phonology, we would not expect the reduction of geminates caused by the collision of a coda with an identical onset of an affix. SLM shows non-transparent features in this regard. The word *baalek* is an intransitive verb meaning ‘to return’. When the causativizer *-king* is attached to it to yield a transitive verb, the form is not *balekking*, but *baleking*, so that we are dealing with degemination. This is true for affixation. With enclitic postpositions, degemination is not found, so the combination of *aanak* ‘child’ with the locative enclitic *=ka* is pronounced [a:nakka] and not [a:naka].

6.3 No diphthongization

In a transparent language, the pronunciation of a phoneme should remain the same (as far as this is phonetically possible at all) in any environment. A vowel should always be pronounced as a vowel, and never as a semivowel. In some languages, chance meetings of two vowels cause one vowel to be pronounced as a semi-vowel, yielding a diphthong. This phenomenon is not found in Sri Lanka Malay.

6.4 No nasalization

In line with what has been said above about diphthongization, an oral vowel should always be pronounced as oral, regardless of whether there are nasal consonants in its environment. This has not been investigated for Sri Lanka Malay. Given the mechanics of the articulatory tract and basic principles of economy, it is unlikely that speakers make efforts to keep all their vowels oral in a nasal environment. While it might be possible to rapidly move the velum back and forth in a word like *maangga=nang* ‘for the mango’ to switch between oral vowels and nasal consonants, it is much more likely that the speakers will avoid the effort involved and pronounce most if not all of the vowels as nasal.

References

- ADELAAR, K. A. (1985). *Proto-Malayic; The reconstruction of its phonology and parts of its morphology and lexicon*. Ph.D. thesis, Leiden University.
- HENGVELD, K. & L. MACKENZIE (2008). *Functional Discourse Grammar*. Oxford: Oxford University Press.
- NORDHOFF, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. Ph.D. thesis, University of Amsterdam.