Sinhala influence in Sri Lanka Malay

Sebastian Nordhoff

1 Introduction

- Sri Lanka Malay is a contact variety of Malay which has undergone heavy grammatical restructuring towards Indian typology (Smith 2003, Paauw 2004, Smith et al. 2004, Smith & Paauw 2006, Slomanson 2006, Bakker 2006, Ansaldo 2008, Nordhoff 2009)

- SOV word order, postpositions, preposed relative clauses, retroflexes

- this restructuring is obviously due to contact with one or more local languages

- candidate languages: Sinhala (Indo-Aryan) and Tamil (Dravidian)

- opinions are divided on whether Tamil (Smith 2003, Smith et al. 2004) or Sinhala (Ansaldo 2008) is the more important contact language

- Sinhala is a good candidate because it is spoken by 4/5 of the population

- Tamil is a good candidate because it is spoken by the ethnic group of the "Moors", who are Muslims just like the Malays.

2 Question of origin

- In order to tell apart Sinhala influence from Tamil influence, one has to find areas where they differ.

2.1 Smith's proposal

- Sinhala and Tamil are typologically very similar

- in many areas, SLM shows structures which are found in both Sinhala and Tamil

- for those cases, the source of a construction cannot be decided

- therefore: let's take a look at domains where Sinhala and Tamil diverge:

    - number marking
    - accusative marking
    - marking of (in)definiteness

- Influence in any of these domains can tip the scales in favour of either Sinhala or Tamil

2.2 Smith's findings

- number marking

    - Sinhala has obligatory number marking (sg/pl)
    - Tamil has optional plural marking
    - SLM has optional plural marking
    - → Tamil influence

- accusative marking

    - Sinhala has only very marginal accusative marking
    - Sinhala accusative can only be found on pronouns and animates
    - Tamil accusative can be found on all pronouns and nouns
    - SLM accusative can be found on all pronouns and nouns
    - → Tamil influence

- marking of (in)definiteness

    - Sinhala has grammaticalized marking of indefiniteness
    - Tamil has no such marking
    - SLM has no such marking

- – → lack of Sinhala influence

- Smith's conclusion: at least three features of SLM point towards Tamil influence rather than Sinhala influence

- explanation: heavy intermarriage between Tamil-speaking Moor women and Malay soldiers.

- many Tamil speaking wives acquired the language of their Malay husbands, but only an imperfect version

- this modified language was passed on to the children and spread through the generations

- there were no Sinhala wives, so the absence of Sinhala influence is to be expected

- I will challenge this scenario on both linguistic and sociohistorical grounds

  - The claimed sociohistorical facts lack empirical evidence
  - A closer look at Smith's domains reveals that there are actually more parallels with Sinhala than with Tamil
  - this is corroborated by investigation of additional domains where Sinhala and Tamil diverge

3 Sociohistorical evidence

- This is discussed at length in Nordhoff (2009:40-47)

- It boils down to the fact that there is no hard and fast evidence for extensive intermarriage between Malays and Moors

- the relevant documents alluded to in Hussainmiya (1987, 1990) are either lacking, or do not provide the necessary ethnographic information (Ansaldo 2008)

- lay genealogy compiled by Burah (2006) points to exclusively Malay weddings in 62.4% of the cases, plus 24.1% which are either Malay or Moor

- also early intermarriage with Hindus, Burghers, and Sinhalese

4 Review of Smith

4.1 Number marking

- Smith observes that number marking is obligatory in Sinhala, while it is optional in Tamil and SLM

- however, there is no need to invoke language contact here

- this feature seems to be a retention of the normal way to deal with number in Malay varieties

- number marking can simply not be used as a criterion to distinguish Tamil influence from Sinhala influence

- score: Sinhala 0; Tamil 0

4.2 Accusative marking

- Smith claims that the SLM accusative marker aligns closely with the Tamil accusative marker

- the distribution of the Sinhala marker is argued to be different

- it is clear that the development of case is due to language contact

- however, Sinhala, Tamil and SLM actually have three different distributions of the accusative marker

    - Sinhala allows the accusative marker only on pronouns and animate nouns
    - the Sinhala accusative is always optional, at least on nouns (Gair 2003:780)
    - Tamil allows the accusative marker on all nouns
    - the accusative marker is obligatory on all definite objects and on all animate objects (Lehmann 1989:27)
    - inanimate indefinite objects can lack the accusative marker under certain conditions (Lehmann 1989:27f)
    - SLM accusative marking is optional, as in Sinhala, but shows a preference for referents which definite, topical, animate, affected, and singular (Nordhoff 2009:329-332)

– The Tamil constraints against accusative marking on indefinite inanimates do not hold as the following example shows:

(1) *Derang* **hathu** *papaaya=**yang** asà-poothong*
    3PL   INDEF papaw=ACC   CP-cut
    'They cut a papaw' (K051220nar01)

– The Tamil constraint to force accusative on animates does not hold either, as the following example shows

(2) *Kumaareng=le thuuju=so dhlaapan=so **oorang**=∅ asà-buunung.*
    yesterday=ADDIT seven=UNDET eight=UNDET man   CP-kill
    'Again yesterday, seven or eight people were killed.' (K051206nar11)

- this distribution seems to be a mix of Tamil and Sinhala constraints.

- score: Sinhala 0.5; Tamil 0.5

4.3 Indefiniteness marking

- Smith claims that SLM lacks indefiniteness marking just like Tamil, but unlike Sinhala

- It is true that Tamil has no grammaticalized indefiniteness marking, while the Sinhala system is fully grammaticalized

- *Pace* Smith, the claim that SLM lacks grammaticalized indefiniteness marking is empirically wrong

- SLM has obligatory indefiniteness marking for

  – new referents

  – predicates of class membership

  – unspecific referents in general (Nordhoff 2009:319-323)

(3) *Se **atthu**=aade,   se **hatthu**=aade.*
    1s INDEF=younger.sibling 1s INDEF=younger.sibling
    'I am a younger sibling, I am a younger sibling.' (K061120nar01)

(4) ***Hathu**=oorang=pe muuluth=dering **hathu**=criitha kal-dhaathang.*
INDEF=man=POSS   mouth=ABL   INDEF=story   when-come
'When a story comes out of a man's mouth.' (B060115prs15)

- this distribution mirrors exactly what we find in Sinhala

- additionally, the indefiniteness marker *hatthu* is also used with loanwords, whether they are indefinite or not

- this is a clear calque of the Sinhala loanword integrator *eka*, orginally meaning 'one'

(5) ***Inni** mock wedding=**hatthu** mas-gijja.*
Mee mock wedding=*eka*   *karanḍa.oonä*
PROX mock wedding=INDEF   must-make
'I have to do this mock wedding.' (K060116nar10)

- in the domain of indefiniteness marking, SLM has clear parallels with Sinhala, and not with Tamil

- score: Sinhala 1.5; Tamil 0.5

4.4   Intermediate conclusion

- Even with the domains chosen by Smith et al, we find more Sinhala influence than Tamil influence

- Let's take a look at more domains:

  - number of stop series
  - gemination
  - distribution of stops and nasals
  - modal particles
  - involitive derivation
  - dative and accusative subjects
  - zero adclausal nominalization
  - adhortative construction

5 More features

- The following features were chosen because Sinhala and Tamil diverge there

5.1 Number of stop series

- Most if not all Malay varieties have 2 series of stops: they contrast a series of voiceless stops (p,t,c,k) with a series of voiced stops (b,d,j,g).

- Tamil has only 1 series (p, ṯ, (r), ṭ, c, k), with voicing conditioned by phonotactics

- Sinhala has 3 series: voiceless (p, ṯ, ṭ, c, k), voiced (b, ḏ, ḍ, j, g), and prenasalized (m̆b, n̆ḏ, n̆ḍ, ŋ̆g)

- SLM also has 3 series: voiceless (p, ṯ, ṭ, c, k), voiced (b, ḏ, ḍ, j, g), and prenasalized (m̆b, n̆ḏ, n̆ḍ, n̆j, ŋ̆g)

- the SLM stops look very much like the Sinhala stops (with the addition of n̆j)

- score: Sinhala 2.5; Tamil 0.5

5.2 gemination

- Most if not all varieties of Malay do not have gemination

- Tamil, Sinhala, and SLM have gemination

- gemination is thus a good candidate for language contact

- Tamil only has gemination of voiceless stops, and sonorants. There are no geminated voiced stops in Tamil

- Sinhala has gemination of voiceless stops, voiced stops, and sonorants

- SLM also has gemination of voiceless stops, voiced stops, and sonorants (Table 1)

- Sinhala and SLM are closely parallel
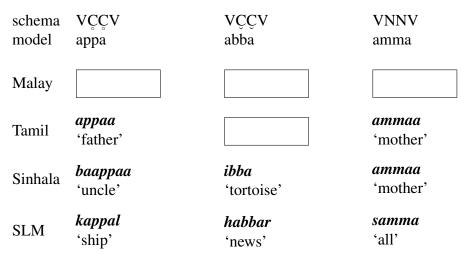
- score: Sinhala 3.5; Tamil 0.5

schema	VC̥C̥V	VC̬C̬V	VNNV
model	appa	abba	amma
Malay			
Tamil	***appaa*** 'father'		***ammaa*** 'mother'
Sinhala	***baappaa*** 'uncle'	***ibba*** 'tortoise'	***ammaa*** 'mother'
SLM	***kappal*** 'ship'	***habbar*** 'news'	***samma*** 'all'

Table 1: Gemination

5.3 Distribution of stops and nasals

- most varieties of Malay have two possible combinations of nasal+stop

  - 'plain' varieties: [mp], [mb] (sampi, ambil)
  - 'prenasalizing' varieties: [mp], [ᵐb] (sampi, aᵐbil)

- Tamil only has [mb]

- Sinhala has [mp], [mb] and [ᵐb]

- SLM also has [mp], [mb] and [ᵐb] (Table 2)

- the distribution of stops and nasals is parallel between Sinhala and SLM

- score: Sinhala 4.5; Tamil 0.5

5.4 modal particles

- modality is expressed in Malay varieties by particles like *boleh*

- Sinhala uses so called quasi-verbs with limited inflectional potential (Gair 1970[1998])

- Tamil uses fossilized third person singular neuter future forms of a number of verbs *veeɳum, poodum (, teriyum, muɖiyum)* etc. (Lehmann 1989:84)

8

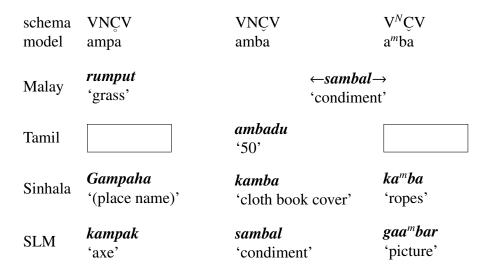| schema	VNC̥V	VNC̥V	V^N C̥V
model	ampa	amba	a^mba
Malay	*rumput* 'grass'	←*sambal*→ 'condiment'	
Tamil		*ambadu* '50'	
Sinhala	*Gampaha* '(place name)'	*kamba* 'cloth book cover'	*ka^mba* 'ropes'
SLM	*kampak* 'axe'	*sambal* 'condiment'	*gaa^mbar* 'picture'

Table 2: Sequences of nasal+stop

- SLM has the particles *boole* 'can', *maau* 'want', and *thussa* 'not.want', *which seem to align more closely with Sinhala* puluvan, oonä, *and* epaa *than with Tamil* muɖiyum, veeɳum, *and* veeɳɖaam.

- *this needs further research and will not be tallied here*

- *score: Sinhala 4.5; Tamil 0.5*

5.5   involitive derivation

- Sinhala has a special process of involitive derivation which allows the use of any verb with a 'dative subject', expressing involuntary action

- *naTənəwa* 'dance'/*naeTenəwa* 'dance without wanting/involuntarily/without intention'

    (6)   a. *ma-mə naʈə-nəwa*
            1s-NOM dance-NONPAST
            'I dance'

          b. *ma-ʈə næʈe-nəwa*
            1s-DAT dance/INVOL/-PRES
            'I get to dancing' (Gair 1976:48)

- this is a productive derivational process

9

(7) *miniha-ʈə diwe-nəwa*
man-dat   run/INVOL/-NONPAST
'The man runs (involuntarily).' (Gair 1991:13)

- this derivation is not found in Tamil (or any variety of Malay)

- it is found in SLM, where we get

(8) a. *se naasi arà-maakang*
1s rice   NON.PAST-eat
'I am eating rice'

   b. *se=dang naasi arà-kànà-maakang*
1s=DAT   rice   NON.PAST-INVOL-eat
'I am eating rice without wanting/unvoluntarily/against my will.'
(Nordhoff 2009:312)

- a very strong case for Sinhala influence can be made for this typologically highly marked feature

- score: Sinhala 5.5; Tamil 0.5

5.6   Dative, instrumental, and accusative subjects

- Tamil has a number of experiencer predicates which assign dative to their argument

- So does Sinhala, but dative subjects are more prevalent there

- additionally, Sinhala has accusative and instrumental subjects, which are not found in Tamil

(9) *miniha-wə liss-una*   SINHALA
man-ACC   slip-PAST.INVOL
'The fellow slipped' (Gair 1976:51)

(10) *ehee polisiy-eŋ   innəwa*   SINHALA
there police-INSTR exist.ANIM
'There are police there.' (Gair 1991:13)

10

- SLM has a good number of dative subjects, as well as the accusative and instrumental subjects found in Sinhala

    (11) Titanic *kappal=yang su-thìnggalam*
          Titanic ship=ACC   PAST-drown
          'The Titanic sank.' (Nordhoff 2009:483)

    (12) Police=*dering su-dhaathang*
          police=INSTR   past-come
          'The police came.' (Nordhoff 2009:483)

- the conditioning factors for the use of accusative and instrumental seem to be the same in Sinhala and SLM

- this is a clear instance of Sinhala influence

- score: Sinhala 6.5; Tamil 0.5

5.7   Zero adclausal nominalization

- In Sinhala, it is possible to put a case marker at the right edge of a zero-nominalized clause

- in Tamil, the nominalizer *adu* is required for argument clauses with case marking

- SLM, like Sinhala, allows the use of case markers on adclausal zero-nominalizations

    (13) *suda butthul suuka **asà**-dhaathang-∅=**nang***
          thus very   like   CP-come(-NMLZR)=DAT
          'So I very much liked that you came.' (Nordhoff 2009:327)

- compare Sinhala *ævillaa-∅=ʈa* 'for the fact that you came' with a zero nominalizer

- Tamil *vand-adu-kku* 'for the fact that you came' with an overt nominalizer *adu*,

(14) *mamə Gunəpaalə ya-nəwa-∅=ʈə*   *kæmətii* Sɪɴʜᴀʟᴀ
　　　1s   Gunapala go-ɴᴏɴᴘᴀsᴛ(-ɴᴍʟᴢʀ)=ᴅᴀᴛ like
　　　'I like for Gunapala to go there' (Gair 2003:796)

(15) *Kumaar [oru tiruʈan̲ va-nt-**at**-ai·p]*   *paar-tt-aan̲*
　　　Kumar a   thief come-ᴘᴀsᴛ-ɴᴍʟᴢʀ-ᴀᴄᴄ see-ᴘᴀsᴛ-3sᴍ
　　　'Kumar saw that a thief had come' (Lehmann 1989:302)

- score: Sinhala 7.5; Tamil 0.5

5.8　adhortative construction

- In both Sinhala and Tamil, the combination of an imperative with a question marker yields a request for command ("Shall I, shall we?", "Is it the case that you order me/us to do something?")

- In Tamil, the imperative form is used in this construction *poo-vaa?*; the infinitive is not used **poo-ha-yaa?*

- in Sinhala, the imperative form is used as well, but it is homophonous to the infinitive: *yanɖa-da?*

- in SLM, the infinitive is used. It seems that the SLM speakers misanalyzed the Sinhala construction. Out of the two options to calque the Sinhala form -*nɖa* they chose the infinitive *mà-* rather than the zero-marked imperative. So we get *mà-pii=si* and not *∅-pii=si*

- the SLM form cannot be explained without recurring to Sinhala influence

- score: Sinhala 8.5; Tamil 0.5

6 Summary

	Sinhala	Tamil	inherited?
number	–	(+)	+
accusative	(+)	(+)	–
indefiniteness	+	–	–
# stop series	+	–	–
gemination	+	–	–
distribution of stops	+	–	–
modal particles	(+)	(–)	(+)
involitive derivation	+	–	–
dative and acccusative subjects	+	–	–
zero adclausal nominalization	+	–	(+)
adhortative construction	+	–	–

- There are 8 domains where there are striking parallels with Sinhala, but not with Tamil

- There are 0 domains where there are striking parallels with Tamil, but not with Sinhala

- How can this asymmetry be explained?

7 Explanations and discussion

- The amount of incontestable Sinhala influence is massive

- On the other hand, there is no clear influence from Tamil grammar

- Sinhala influence must have been more important than assumed earlier

- Two relevant factors

  - Early Sinhala Influence (EaSi): Sinhala as the language of 80 % of the population (90% in the relevant areas) must have had an impact on the formation of SLM (cf. Ansaldo 2008)
  - refer to maps at the end of this handout
  - Late Sinhala Influence (LaSI): The nationalist policies by the Sinhala majority in the second half of the 20th century bolstered Sinhala influence on other languages.

- Given the extent of parallels with Sinhala, it is unlikely that LaSI could have had that much impact in only 50 years (1950-2000)

- Speakers born in the 1930s also show these features.

- Sinhala influence must predate the nationalist policies of the 1950s

- A longer period of contact with Sinhala including EaSI is unavoidable

8 Conclusion

- Sinhala has influenced the phonology, morphology, syntax and pragmatics of SLM

- In all these areas, we find more Sinhala influence than Tamil influence

- The amount and nature of Sinhala influence cannot be a result of the late 20th century alone

- Sinhala influence was present since the beginning of the Malays' stay on Ceylon

- The "intermarriage theory" of Smith et al. is at odds with the sociohistorical data as shown in Ansaldo (2008) and Nordhoff (2009)

- The "intermarriage theory" is furthermore at odds with the linguistic data

- Conversion and metatypy towards both Sinhala and Tamil provide better explanations than creolization and imperfect transmission.

References

ANSALDO, U. (2008). "Sri Lanka Malay revisited: Genesis and classification". In DWYER, A., D. HARRISON & D. ROOD (eds.) "A world of many voices: Lessons from documented endangered languages", Studies in Language, Amsterdam, Philadelphia: John Benjamins. pp. 13–42.

BAKKER, P. (2006). "The Sri Lanka *Sprachbund*: The newcomers Portuguese and Malay". In MATRAS, Y., A. MCMAHON & N. VINCENT (eds.) "Linguistic Areas – Convergence in Historical and Typological Perspective", Houndmills, New York: Palgrave Macmillan. pp. 135–159.

BURAH, T. A. (2006). *Saga of The Exiled Royal Javanese Unearthed*. Dehiwala: Sri Devi.

DEUMERT, A. & S. DURLEMAN (eds.) (2006). *Structure and Variation in Language Contact*. Amsterdam: John Benjamins.

GAIR, J. W. (1970[1998]). "About Colloquial Sinhalese Clause Structures". In Gair (1998).

GAIR, J. W. (1976). "Is Sinhala a Subject Language?" In Verma (1976), pp. 39–63.

GAIR, J. W. (1991). "Subjects, Case and INFL in Sinhala". In VERMA, M. & K. P. MOHANAN (eds.) "Experiencer Subjects in South Asian Languages", Palo Alto: Center for the Study of Language and Information. pp. 13–41. Reprinted in Gair (1998), pp. 65–86.

GAIR, J. W. (1998). *Sinhala and Other South East Asian Languages*. New York, Oxford: OUP.

GAIR, J. W. (2003). "Sinhala". In CARDONA, G. & D. JAIN (eds.) "The Indo-Aryan Languages", London: Routledge. pp. 766–818.

HUSSAINMIYA, B. A. (1987). *Lost Cousins: The Malays of Sri Lanka*, *Occasional Paper*, vol. 2. Kebangsaan: Universiti Kebangsaan.

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

LEHMANN, T. (1989). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

NORDHOFF, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. Ph.D. thesis, University of Amsterdam.

PAAUW, S. H. (2004). *A Historical Analysis of the Lexical Sources of Sri Lanka Malay*. Master's thesis, York University.

SLOMANSON, P. (2006). "Sri Lankan Malay morphosyntax: Lankan or Malay?" In Deumert & Durleman (2006), pp. 135–158.

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. & S. PAAUW (2006). "Sri Lanka Malay: Creole or convert". In Deumert & Durleman (2006), pp. 159–182.

SMITH, I. R., S. PAAUW & B. A. HUSSAINMIYA (2004). "Sri Lanka Malay: the state of the art." *Yearbook of South Asian Languages*, 2004:197–215.

VERMA, M. (ed.) (1976). *The Subject in South Asian Languages*. Madison: University of Wisconsin South Asia Program.