

# **Sri Lanka Malay revisited: genesis and classification.**

*Umberto Ansaldo*

*Universiteit van Amsterdam*

## **Abstract**

This paper presents a fresh take on the origins and nature of Sri Lanka Malay (SLM), based on fieldwork data collected in 2003-4 in Kirinda, in the south-east of Sri Lanka. It departs from previous studies of SLM in that it is based on substantial recordings of spoken data in natural settings as well as coverage of oral history. Work on SLM so far has offered significant insights into the nature of these varieties; due to limited data available, however, some aspects have failed to emerge which are important for our current understanding of SLM. In particular, I aim to show the value of natural linguistic data and first-hand historical information for a reassessment of our current understanding of the field in relation to (a) the nature of adstrate influence in SLM, and (b) the contact dynamics in which varieties such as SLM evolve.

## **0. Foreword**

This paper consists of two parts. Part I critically revisits two fundamental assumptions of historical nature made in previous literature which have not been questioned so far, namely (1) intermarriage and (2) creolization in the evolution of SLM. Based on careful historical analysis, I claim that the views entertained on SLM so far are biased towards Tamil at the expense of Sinhala. Part II presents data from the case system of SLM, showing the interplay of Malay, Sinhala and Tamil in the restructuring process. Finally, I suggest that SLM varieties can best be classified, based on historical as well as structural analysis, as mixed languages with a dual adstrate (Sinhala and Tamil) and a typical Pidgin-Malay-derived (PMD) lexifier (cf. Adelaar & Prentice 1996). Restructured Malay varieties of Sri Lanka are precious for our understanding of contact dynamics as they are among the few contact varieties which have evolved in an environment in which no Standard European acrolect is present. In this case, Malay varieties (cf. below) can be considered as lexifiers, while

the main adstrates are Sinhala and Tamil,<sup>1</sup> which, though genetically belonging to different families, Indo-European and Dravidian respectively, have been in contact over a long period of time and show clear typological convergence (Emeneau 1980). This paper illustrates the role of adstrates and the typological convergence towards Sinhala and Tamil in the development of the SLM nominal domain.

## **PART I Historical foundations**

### **1. Introduction**

Two related assumptions entertained so far in the literature need to be addressed in order to do justice to the history and the nature of SLM varieties: the ‘Tamil bias’ and the ‘creole classification’ idea. Before moving to a critique of these ideas, a sociohistorical background of SLM communities is necessary.

#### **1.1 SLM speech communities**

Sri Lanka Malay (SLM) varieties, so far generally viewed as ‘creoles’ (see e.g. Smith et al. 2004), are currently endangered as they are no longer spoken by the younger generation, with one exception, the community in Kirinda. With the dominant languages Sinhala and Tamil already in conflict, due to political and ethnic struggle, a minority such as the SLM tends to converge towards the dominant linguistic groups in order not to be disadvantaged, which results in the younger generations abandoning their vernacular.

Kirinda is a small fishing village that lies 30 kilometers east of Hambantota in the southeast of Sri Lanka. The community of Kirinda is an ideal setting in which to study SLM as it has remained relatively sheltered from the modernization and globalization that has taken place over the past decade in Sri Lanka and which has led

---

<sup>1</sup> Using the term ‘substrate’ would be misleading in the case of SLM as there was no dominance of Malay over Tamil and Sinhala. All three languages were, if at all, subordinate to the prestige colonial variety until independence (cf. section 1 below).

to looser community ties within the SLM communities and progressive loss of the vernacular.

Table 1 provides an overview of the SLM speech communities today.

Table 1. SLM varieties at present

<b>Community</b>	<b>Characteristics</b>	<b>Vitality</b>
<b>Colombo</b>	Middle-upper class; often bi- or trilingual (Tamil/Sinhala); standardizing in Malay; restricted usage of SLM; English fairly fluent	Endangered: no SLM in younger generation
<b>Slave Island (Colombo)</b>	Lower class; most Tamil influenced; bi- or trilingual; no English	Very endangered; use of SLM discouraged
<b>Kandy (and Hillcountry)</b>	Similar to Colombo community; weak standardization forces	Endangered
<b>Hambantota</b>	Traditionally heavy Sinhalese-speaking area; low-middle class, often trilingual; limited English	Mildly endangered
<b>Kirinda</b>	Lower class; good trilingual competence in middle-younger generations; English limited to a few individuals	Fully vital: mother tongue even in present generation

While the descendants of Malays that inhabit the village of Kirinda are usually classified as part of the ethnic group generally known as Sri Lanka Malays, I refer to their language as ‘Java’ for two reasons: partly in acknowledgment of the fact that this is how the speakers themselves refer to it; and partly because I want to stress the fact that, though from a historical point of view Malay varieties can be regarded as the ancestors of SLM, as can be seen in the lexical domain, there is much structural material in SLM varieties that bears no resemblance to any Malay-based contact

language (Smith et al. 2004). SLM varieties are clearly no dialects of Malay in the traditional sense (Adelaar & Prentice 1996); however, they hardly meet the criteria to be classified as creoles, as I discuss below.

In discussing the origins of SLM speaker, it is important to bear in mind the following points:

- (i) The origins of the SLM speech communities are very heterogeneous, covering an area from Northern Malaysia to the easternmost provinces of Indonesia (Hussainmiya 1987, 1990). Their ethnic and linguistic backgrounds are, likewise, extremely diverse. Under the Dutch, political exiles (as well as convicts) were deported to Sri Lanka from different corners of the Indonesian archipelago and beyond, e.g. Java, Borneo, the Moluccas and Goa, among other places. Typically, the nobility would be deported *together* with their families, and, though contact between these groups was discouraged by the Dutch, intermarriage between the different royal families did indeed take place (cf. Hussainmiya 1987). The largest group of people attributed a Malay origin came as soldiers also from disparate places such as Bali, Java, Riau, Ambon and peninsular Malaysia so that “almost all the major ethnic groups from the region of the Eastern archipelago were represented” (Hussainmiya 1987: 48). The soldiers could also be accompanied by their wives; how common this was during Dutch rule is unclear, however we do know that under the British this practice was encouraged.
  
- (ii) As we can see from the above, the social extraction of the ancestors of the SLM people was of a very varied nature, from exiled princes to slaves and soldiers. It has been suggested that at least two different communities could be distinguished: (a) a rather sophisticated diaspora of noblemen, and (b) a ‘Malay’ garrison, what would become the Ceylon Rifle Regiment under the British (cf. Ricklefs 1974). A third group of convicts may also be identified but little seems to be known about them. It is important to note that, as far as the first two groups are concerned, contacts between the groups were indeed quite frequent, due among other reasons to the practice of employing noblemen as officers of the troops (Hussainmiya 1987, 1990).

(iii) There seems to be a tension between the following two claims, both to be found in Hussainmiya (1987, 1990): (a) intermarriage between individuals of Indonesian/Malay descent and (Tamil) Moors was common; (b) Malay/ Indonesian deportation often consisted of entire family nuclei, not single individuals. Claim (a) is Hussainmiya's own assumption, based on the observation that the two ethnic groups shared the Muslim faith and were therefore naturally in contact (more on this in section 2.1). Claim (b) is actually supported in historical records (cf. Schweitzer 1931) though numerical percentages are not available:

“The wives, which in part are Amboinese [Eastern Indonesian/Malays], in part Singulayans [Sinhalese], and Malabarians [Tamil], say nothing against this [practice of gambling], but when the man games away their little property, they must nourish him and his children as well as they can through the month and await his better fortune at gaming.”

As already noted in (i), the practice of moving whole families, rather than solely male individuals, was quite widespread. Despite this heterogeneous composition, Ricklefs (1974) gives us an image of a rather sophisticated Malay diaspora in Sri Lanka during colonial rule, a close-knit community, where contacts between the different Malay/Indonesian ethnicities, as well as the different social extractions, were maintained through the ranks of the army as well as through common religious practice.

(iv) The claim of intimate relationship between Tamil Moors and SLMs has led to the perception that SLM communities would be linguistically more influenced by Tamil than Sinhala (e.g. Smith et al. 2004). I question this perception in the absence of clear historical evidence; in addition, considering the fact that Sinhalese people having always been in significant numerical majority in the country, Sinhala must have had *at least as much influence* in the evolution of SLM as Tamil. No evidence of social segregation between SLM communities and Sinhalese has been presented so far to my knowledge; I address this issue more thoroughly in section 2 below.

In section 1.2, I discuss the importance of the points outlined above.

## 1.2 The linguistic base of SLM

Following from the above, in terms of linguistic input to SLM, we would expect many different languages such as Javanese, Ambonese, Riau, Malay (colloquial and high), etc. to be involved in the early days of the ‘Malay’ diaspora in Sri Lanka; clearly, a type of Malay-based contact variety, such as Bazaar Malay – the de facto lingua franca of the trade route stretching from Southern China to Northwest India from at least the 15<sup>th</sup> century, based on Low Malay and Low Javanese (typically with borrowed elements from Hokkien, Southern Min) – would have been a good candidate for a language of interethnic communication. Hussainmiya (1990: 47) notes that the Malay must have been some kind of simplified lingua franca or creole, a very likely assumption given that Malay lingua franca had existed for a significant period of time. We do know that varieties of contact Malay were well established and widely used throughout South and Southeast Asia at this point (Adelaar & Prentice 1996: 674).<sup>2</sup> Hussainmiya (1987: 154, 1990) suggests that Batavian Malay, a Malay-based lingua franca, or Jakarta Malay Creole may be good candidates for a common language in the early days of the diaspora, though Low Malay and Bazaar Malay would also have been present. According to Grijns (1991) however, the establishment of Jakarta Malay is unlikely until the 3<sup>rd</sup> quarter of the 19<sup>th</sup> century, while Malay-based lingua franca would have already been in existence. Smith et al. (2004) speak of ‘Vehicular Malay’ as a cover term for a generic colloquial Malay variety of interethnic communication. Adelaar (1991) finds evidence of Moluccan material in SLM and suggests that Eastern Malay dialects (very likely much more widespread across the Indonesian world at that time, Hans den Besten p.c. June 2005) may be involved in the evolution of SLM.

Next to a strong presence of a colloquial Malay variety of the PDM type, there would have been of course Lankan Tamil and colloquial Sinhala. Even if we were to accept the claim that Tamil might have been more closely involved in the evolution of SLM due to the fact that the language of the religious texts and practices would have been

---

<sup>2</sup> However, the issue of whether Bazaar Malay itself may be captured as a single variety or whether it may be better conceived as a cover term for various regional pidgins makes any comparison with a specific list of structural features problematic (cf. Holm 1988).

Tamil (Hussainmiya 1987: 157, but see section 2 below), the sheer numerical and social predominance of Sinhala in Sri Lankan society cannot be ignored, and for this reason we should consider the two languages as, at least, adstrates of similar significance in the evolution of SLM.

## **2. Revisiting basic assumptions**

### **2.1 The ‘Tamil bias’**

Being in the fortunate position of collecting first-hand data and documenting oral history from members of various SLM communities across the island, I here have the opportunity to address some issues that have mistakenly become accepted wisdom when discussing SLM, but are hardly supported by the historical and linguistic data. Due perhaps to the limited data available until recently, there seems to be a general agreement that SLM is a product of creolization; this is a result of the accepted assumption that Tamil Moors would have constituted the dominant group of interaction for the SLM communities. The most specific claim regarding the creolization of SLM is that it developed as a result of intermarriage between Malay men and (Tamil) Moor women (e.g. Smith et al. 2004); this view is based primarily on the historical observations of Hussainmiya (1987, 1990) regarding the records of marriage under the Dutch (*tombos*<sup>3</sup>) which, according to him, show several cases of intermarriage between SLM and Tamil Moors. It is important to verify the validity of this claim for two reasons:

- (i) the specific history of the SLM communities: in recording oral history from the older generations, I found much disagreement with the ‘Tamil bias’ view;
- (ii) implications for our understanding of the evolution of contact varieties: in the current view, a contradictory claim exists, namely that SLM is the product of predominantly Tamil and Malay admixture, which would result in a mixed language, (cf. Bakker 2000), while it is classified as a ‘creole’ (Smith et al. 2004).

---

<sup>3</sup> ‘Tombo’ (from Sinhala *thombuwa* meaning ‘register’ or ‘record’) are the historical records of Sri Lanka. Following the indigenous practice of land records, the Portuguese kept records of their possession until the Dutch conquest of Sri Lanka, during which the Portuguese records were largely destroyed. The Dutch resumed the practice of keeping registers, distinguishing between Land Tombo (registry of land) and Head Tombo (registry of landholders).

The construction of the Tamil bias is summarized below:

- I. The first occurrence of a claim about Malay-Moor intermarriage occurs in Hussainmiya (1987): “a number of” these marriages are reported, next to Malay-Ambonese/ Malabarese/ Sinhalese unions recorded in the Dutch Tombos. The same work however suggests that SLM may be influenced by Sinhala, Tamil or both.
- II. Hussainmiya (1990) suggests religious affinity between Moors and Malays and relates episodes of religious exchange to suggest frequent exchanges between the two communities.
- III. Bakker (2000) develops an account of Sri Lanka Creole Portuguese as originating from the admixture of Portuguese and Tamil. He then extends this genesis scenario to SLM without providing historical data.<sup>4</sup>
- IV. Smith (2003) investigates the influence of Tamil and Sinhala in the Accusative marker of SLM. His conclusions are: “we can only argue for lack of Sinhala evidence, rather than positive Tamil influence”.
- V. Smith et al (2004) appear to be in between claiming influence from Tamil and claiming general ‘Lankan’ influence.
- VI. Heine and Kuteva (2005) treat SLM as Tamil-Malay creole.

From the above, we can see that in the beginning there was only a suggestion of a stronger Tamil influence. By 2005, however, this idea had evolved into a truism. As Hussainmiya (1987) is the only work to directly, though briefly, address this issue, it is important to verify the claim in a precise manner. My investigation of the Dutch Tombos referred to in that work, for the period 1678-1919, in the National Archives

---

<sup>4</sup> Regarding Bakker’s work, which is only tangential to this issue, it is important to note that his arguments actually address Sri Lanka Portuguese (SLP), a ‘creole’ developed from the Eastern port of Batticalhoa which was spoken along the whole Eastern coast of Sri Lanka. This variety, like SLM, displays strong influence from Lankan typology and may well show stronger influence from Tamil than Sinhala. Though Bakker does not present any evidence in favour of this, a historical explanation might be found if we consider that the area of Batticalhoa has typically been Tamil dominated. Be that as it may, Bakker seems to be under the impression that whatever holds for the history of SLP may be extended to the history of SLM; this is completely unjustified, as the two communities evolved under very different ecolinguistic environments. For example, Bakker’s claim of ‘heavy Tamil pressure’ may perhaps have been presented in the SLP community, but there is absolutely no historical evidence that this occurred in the SLM communities.

at The Hague (microfilm copies of the Colombo archives) and the National Archives of Sri Lanka in Colombo yielded the following results:

1. The records for the period up to 1796 are damaged by water, making significant parts of the entries illegible. The most revealing information for identification here are the signatures of the parties. There is however hardly any information of ethnic group, which makes it difficult to identify Malay/Indonesian and Moors given that both groups share the practice of adopting Arabic names. In a particularly interesting section in the *tombos* dedicated to mixed marriages (cf. Hussainmiya 1987), only five of 238 entries clearly refer to individuals of Javanese origin: of these, two records refer to Javanese-Moor marriagee, one to a Javanese-Javanese marriage, and the remaining two are unclear.
2. The following period until 1919, albeit under British rule and therefore less interesting for our claim, shows a more structured archiving system where indication of race is given. Where legible, this reveals still a majority of Western marriages, a growing number of marriages between Eurasians and Burghers (locally born of Dutch/Western heritage), and between Burghers. There are two clear entries involving Malays, one married to a Eurasian (between 1867-1897), and one to a Burgher (1885-1897). From 1897 onwards, race is clearly specified; of 196 entries only one is Malay.

While necessarily brief, the report above of the contents of the Tombos shows that there is hardly any reason to comment on the nature of intermarriage of the Malays in general, and less even to make specific claims about the origins of the parties.<sup>5</sup> While sharing a common religion may have played a role in individual marriages between SLM and Tamil Moors, there is no historical evidence to lend support to a claim of diffuse intermarriage between these two communities, especially of such a magnitude that could conceivably lead to restructuring of the vernacular.

In addition to the historical record, clear evidence against a solid Tamil influence in the development of the SLM community comes from oral history recorded in three

---

<sup>5</sup> Unless the Tombo in Hussaimiya's possession reveal completely different data from the ones of the National Archives, it is safest to discount Hussaimiya's observations.

different SLM communities: in Kirinda, Colombo and Kandy, of approximately 50 families interviewed in total, only two revealed genealogies including Moor-Malay intermarriage. Most families report that marrying outside the SLM community was considered taboo and only allowed in extreme cases. It is only in the present generations that weddings outside the community start being allowed. Moreover, the Moors appear to have had very low status in the eye of the SLM communities and, their low status may be seen as a counteracting force to the hypothetical appeal of religious affinity. Moreover, in at least one community – Kirinda – intermarriage with Sinhalese is well attested in the history of several families.

An example that we need to revisit of implausible linguistic arguments developed precisely on the ‘Tamil bias’ is found in Bakker (2000, 2003), which subsequently led to the Tamil-Malay creole claim in Heine and Kuteva (2005). Bakker argues for rapid typological convergence and admixture with Sinhala and Tamil as a recent development; he assumes that creolization may account for earlier stages of SLM’s history and needs thus to postulate the evolutionary path from a Malay-Tamil jargon to contemporary SLM. However, no evidence of an early jargon exists, nor is there any socio-historical evidence to postulate catastrophic events leading to rapid changes.

Crucially, in the most serious attempt to date to argue for Tamil as a primary ‘substrate’, Smith (2003), in investigating definiteness in the NP, concludes that it is perhaps the absence of Sinhala, not the presence of Tamil that is surprising. This is however not surprising at all. The feature pool from which SLM evolves, can be described as follows: the lexifier would be a Malay *Lingua Franca* with mixed Balinese, Ambonese Malay, Buginese, Javanese and other elements. We do know that such varieties were predominant from around the 15<sup>th</sup> century along trade routes between Southern China and North-West India (Adelaar & Prentice 1996). Adstratal elements according to the feature pool model (cf. Mufwene 2001) would come from Malay, Sinhalese, SL Tamil (as well as some Dutch, English, and perhaps Portuguese elements, mostly in the lexicon). The fact that Sinhala and Tamil show typological convergence accounts for why it is difficult at times to tease apart whether a certain grammatical feature derives from one or the other adstrates; we should expect to find some clear cases as well as a number of opaque ones in tracing the origins of SLM

grammar. That an admixture of these various features is present in the VP of SLM has been suggested in recent work by Slomanson (2004a, b) based on data from informants of the Colombo variety. The data presented in the second part of this paper shows a similar picture for the nominal domain.

## **2.2 SLM is not a creole**

Having established that there is no substantial evidence for constructing Tamil as a prevalent adstrate in the formation of SLM, let us turn to the claim of creolization in the genesis of SLM. As argued for example in Ansaldo and Matthews (2001), it is not always feasible to make a categorical distinction between Creoles and mixed languages, given that hybridization of structure is a matter of degree in high-contact environments. For the sake of argument, however, I will treat the possibility of applying established views of creolization to SLM.

I. Creolization as a sociohistorical process. If we view creolization as a predominantly sociohistorical concept of disenfranchised vernaculars developed in European plantation environments (e.g. Mufwene 2000, 2001), we clearly encounter problems of historical accuracy and social nature. SLM varieties are not products of such conditions, as clearly illustrated in the historical section at the outset: the relationship between Sinhala/ Tamil and Malay cannot be compared to a superstrate-substrate situation. The following points clash with such a view of creolization:

- (a) The Malay variety of predominant use would have been a ‘low’ variety such as Bazaar Malay, devoid of prestige. A contact Malay variety would have not been a target language but the de facto lingua franca of the community. There is no evidence to suggest that a form of High Malay may have been the target; although such a variety was available, it had a clearly defined social rationale (the language of religion and literature) and speakers of Malay varieties would have been accustomed to a form of diglossia of this type.
- (b) Sinhala/Tamil would have been adstrates present from the start, stable but external to the SLM speech community; there is no substantial evidence to suggest intermarriage as the driving force behind the restructuring process and

therefore nativization, if at all viable (cf. Mufwene 1999), is not a valid argument.

Only a very superficial and partial interpretation of this view, namely that of disenfranchised vernacular, would allow one to refer to SLM as a creole variety.

II. Creolization as relexification. SLM grammar is difficult to account for within current views of relexification (cf. Lefebvre 2001): its lexicon is mixed yet derived predominantly from numerous varieties of the Malay/ Indonesian region; syntactic patterns show significant adstrate influence, but we also find Malay features, as for example aspects of its verbal morphology (cf. Adelaar 1991; Slomanson 2004a). In this sense the mixture of SLM grammar appears as ‘modular’, i.e. combining subsets of grammatical systems derived from all of its input languages (cf. Aboh 2004). More importantly, as noted in Bakker (2003), the situation in SLM is the reverse of the typical relexification scenario, in combining ‘new grammar’ with ‘old lexicon’.

III. Creolization as exceptional development. To the extent that this will be shown to be a viable account (for strong reasons against this, see DeGraff 2005; Ansaldo & Matthews in prep.), KJ is surely not a young grammar in the sense of McWhorter’s (2004) as it presents us with complex nominal and verbal morphology that – without further elaborating the still somewhat subjective notion of complexity – does not appear young in McWhorter’s sense and does not look ‘simple’ when compared to its putative lexifier, if we allow for one. If we were to identify any ‘typical’ creole features along the lines of McWhorter’s (1998) prototype, the most likely explanation for them would still be in terms of Pidgin-Malay derived (PMD) features (cf. Adelaar 1991; Adelaar & Prentice 1996).

I limit myself to these three conceptual frameworks as I believe they are, at this stage, the most promising ones within theories of creolization and the only ones that can be positively tested. As pointed out by Don Winford (p.c. 2005), one might still want to view SLM as a ‘special’ case of Second Language Acquisition (SLA) and thus link it to creolization as restricted SLA. I seriously doubt this to be the case: there is no

historical evidence of restricted input in the formation of SLM. Nor is there evidence of a stage of early grammar that could have been material for reanalysis in attempting to acquire a target language. Surely the Malays did not create SLM by trying to acquire Tamil or Sinhala, because if that were the case we would not have a predominantly Malay lexicon. Nor would there have been any plausible reason for Tamils/ Sinhalese to restructure their own varieties in acquiring SLM; they were, after all, speakers of larger, socially more prestigious languages in which the SLM speakers would have been quite competent. Thus, what we do have is language acquisition in an informal context with high degree of bi/multilingualism; there is no evidence nor reason to postulate a break in transmission, an imperfect acquisition process or any other construct typical of creole ideology. Such a scenario, as has been pointed out many times (cf. e.g. Baker 1996; Mufwene 2001; DeGraff 2001, 2004, 2005) is really quite ‘normal’ and common in acquisitional contexts around the globe; as it is notoriously difficult in such situations to decide what may be due to borrowing and what to transfer, I nonetheless leave this possibility open, until the time when the SLA framework of creolization might provide us with a clearer model to be tested. The question of how to classify SLM – assuming we really need to – is addressed in the conclusions.

## **PART II – The Data**

The data presented in this paper were collected during fieldwork in Kirinda, Sri Lanka in December 2003 and January 2004 and amount to roughly 12 hours of recording (Ansaldo and Lim 2005).<sup>6</sup> All the data unless otherwise stated are first-hand data from elicitation sessions and spontaneous speech.<sup>7</sup>

---

<sup>6</sup> This fieldwork was partly supported by the National University of Singapore’s Academic Research Grant (R-103-000-020-112) for the project ‘Contact Languages of Southeast Asia. The role of Malay’.

<sup>7</sup> Until now three sets of data had been available on SLM. The first, on which various papers by Smith et al. are based (for an overview see Smith et al. 2004), consists of a one-hour recording of a speaker from Kirinda from the 1970s (Ian Smith p.c. 2005). The second consists of a dozen very short narratives collected by Bichsel-Stettler in the 1980s, amounting perhaps to half an hour of recording, in the central areas of the country and Colombo (Bichsel-Stettler 1989). The third is a short conversation transcribed by Adelaar on Slave Island (Colombo) in the 1990s (Adelaar 1991, p.c. 2005). A number of grammatical features that I discuss from my data do not appear in any of these earlier sets; this is not surprising as some of these features are highly optional and it is expected that they might not emerge in a short recording. As my data show these features to be present in speakers of the oldest generation, as well as in all other generations, the time lapse in between the recordings cannot be seen as responsible

### 3. Introduction to case in Kirinda Java and its adstrates<sup>8</sup>

Typical accounts of nominal marking in SLM usually recognize at least the following cases: Nominative, Accusative-Dative, Genitive, Locative, Associative and Instrumental. (e.g. Hussainmiya 1987; Smith 2003; Smith et al. 2004 as the most comprehensive). As anticipated above, our corpus (Ansaldo & Lim 2005) shows different grammatical distinctions in some subdomains, as discussed below. In what follows, I rely exclusively on my data for analysis and address previous observations only where relevant. A full discussion of the possible variation in nominal marking across SLM varieties is beyond the scope of this paper.<sup>9</sup>

Table 2 lists the nominal markers of KJ using the semantic roles as a point of departure in order to capture the functional nature of the markers as clearly as possible. Table 3 outlines the semantic functions of case in Sinhala and Tamil.

Table 2. Case in Kirinda Java<sup>10</sup>

Markers	Semantic role	Traditional case terminology
Ø	Agent	Nominative
-naŋ/ -daŋ	Experiencer, also Goal	Dative
-yaŋ	Patient (Pronouns, Animate Nouns Non-volitional, Psych Verbs)	Accusative
-pe	Possessor	Possessive
-ka	Location	Locative
-riŋ	Instrument	Instrumental

for the emergence of them (more on this in section 3.2). Likewise, no catastrophic events marked the last three decades of the SLM communities (until the tsunami of December 26, 2004) and therefore no radical change in the language is to be expected; this would have been easily detectable in generational variation, which it is not.

<sup>8</sup> The Tamil data presented here are limited for two reasons. (1) This paper aims to re-establish a balance between the role of Tamil and Sinhala in the evolution of SLM. As work appeared so far focuses on Tamil, I feel it is importance to highlight the structural features of Sinhala. (2) Tamil data available in the literature are typically derived from Indian Tamil varieties, not Sri Lankan Tamil (SLT). In this sense any Tamil data presented should be taken with the utmost care as variation in Tamil varieties is high and SLT shows clear divergence from Indian Tamil (Wijeratne 2005).

<sup>9</sup> See Ansaldo (2005) for a fuller account.

<sup>10</sup> With the possible exception of the Comitative marker, all the SLM case markers can be lexically related to PMD varieties. Etymological analysis is beyond the scope of this paper, but see Adelaar (1991) and Ansaldo (2005) for more on this aspect.

-le	Association	Comitative
-----	-------------	------------

Table 3. Functions of case in Sinhala/ Tamil (Gair & Paolillo 1997; Schiffman 1999; Asher & Annamalai 2002; Karunatilake 2004)

Sinhala		Tamil	
Nominative (direct case)	Agent (+Animate)	Nominative	Agent
Dative	Experiencer Goal of V <sub>intr</sub> Beneficiary Possessor	Dative	Experiencer Goal of V <sub>intr</sub> Beneficiary Possessor
Genitive-Locative	Temporary possession Location	Locative	Temporary possession Location Path
Instrumental-Ablative	Instrument Source	Ablative	Source
Vocative (+Animate)	Addressee	Associative-Instrumental	Associate Instrument
Accusative (optional)	Human/ animate Definite Goal of V <sub>trans</sub>	Accusative	Definite specific Goal of V <sub>trans</sub>

Comparing Tables 2 and 3, we can make the following observations (with the relevant cells in Table 3 shaded grey):

- (i) Prototypical Agents in KJ (Nominative in SLM) are unmarked, as in the adstrates Sinhala and Tamil.
- (ii) Experiencers and Goals in KJ are marked identically; they correspond to Dative case in the adstrates. Lack of volition or control, just as in Sinhala (and less so in Tamil, Sebastian Nordhoff p.c. August 2005), is the key semantic feature here.
- (iii) Accusative marking parallels the adstrate typology. The KJ definite object marker (ACC, cf. ex. 1-2 below) shares the feature [+definiteness] with Tamil and is optional as in Sinhala.
- (iv) KJ shares the Instrumental-Ablative syncretism with Sinhala.

### 3.1 Features of case in Kirinda Java

The function and interplay of case markers in KJ is illustrated in the examples that follow. In (1), we can see the Agent as zero-marked and the Accusative marker *yan*

in a typical collocation as the direct argument of a transitive verb; the latter is also seen in (2).

(1) ikkaŋ-yaŋ se er-makaŋ  
fish-ACC I DUR-eat  
'I eat the fish'

(2) te-yaŋ er-minum  
tea-ACC DUR-drink  
'I drink the tea'

The marker *naŋ/-daŋ*, though available as a marker of Goal, is most frequently assigned to a range of arguments that appear as Agents involving modal predication as in (3) and (4), and/ or predicates denoting emotions or perceptions as in (5) which can be analyzed as Experiencers. Note that, when the direct argument of the verb is indefinite, no ACC is found:

(3) master-naŋ pena mau  
teacher-DAT pen want  
'Teacher wants a pen'

(4) go-daŋ minum mau  
I-DAT drink want  
'I want/ would like a drink'

(5) go-daŋ Mr. Jalaldeen-yaŋ kutumun  
I-DAT Mr. J-ACC see  
'I see Mr. Jalaldeen'

As these examples illustrate, the marker for Experiencer appears in two alternations (attested in the data collected as well as from elicited information): the form *daŋ* is used for first and second pronouns singular exclusively, while for all the other

pronouns *naŋ* must be used, as illustrated in the paradigm in (6).<sup>11</sup> This suggests a distinction that may be captured along the line of the Animacy Hierarchy where first and second person pronouns are assigned highest animacy and therefore differential marking (Silverstein 1976).

(6)	Go-daŋ	I-DAT
	Lu-daŋ	You(S)-DAT
	Dianaŋ	S/he-DAT
	Kitaŋnaŋ	We-DAT
	Loraŋnaŋ	You(PL)-DAT
	Deraŋnaŋ	They-DAT

When we compare the pronominal paradigm above with the adstrates Sinhala and Tamil, we do not find an identical correspondence to the distinction between 1<sup>st</sup> and 2<sup>nd</sup> pronouns and the rest. Table 4 illustrates the dative pronoun paradigms for Tamil and Sinhala.

Table 4. Dative pronouns in Tamil/ Sinhala<sup>12</sup>

	<b>Tamil</b>		<b>Sinhala</b>
1 <sup>st</sup> sing.	enakku	1 <sup>st</sup> sing.	ma□ə
2 <sup>nd</sup> sing.	onakku	2 <sup>nd</sup> sing.	oyaa□ə
3 <sup>rd</sup> sing. (m.)	avanukku	3 <sup>rd</sup> sing. (anim./inan.)	eyaa□ə/ eekə□ə
1 <sup>st</sup> pl. (incl./excl.)	Nampa□ukku/eŋka□ukku	1 <sup>st</sup> pl.	api□ə
2 <sup>nd</sup> pl.	oŋka□ukku	2 <sup>nd</sup> pl.	oyaalaa□ə
3 <sup>rd</sup> pl. (m/f)	avaŋka□ukku	3 <sup>rd</sup> pl. (anim./inan.)	eyaala□e/ eewa□e

Given that most of the lexical material is derived from PMD varieties (see fn. 10, 11), it is not surprising not to find structural resemblances between the paradigms. As already shown in Table 3 for case distinctions in general, it is the parallel at the

<sup>11</sup> Again the lexical roots of the pronouns reveal PMD origins, most clearly in the Hokkien (Sinitic) forms for 1<sup>st</sup> and 2<sup>nd</sup> person singular, *go* and *lu*. As some speakers, in particular those of higher education/ economic standing, perceive these to be ‘rude’ or ‘vulgar’, variation between *go* and *se* may occur in varieties of SLM. *Se* is the preferred variant in Colombo while in Kirinda *go* is predominant.

<sup>12</sup> There is variation in notation conventions among different authors: here I try to preserve coherence with the choices made to represent SLM sounds as well as IPA symbols. In the glossing of examples from other authors, I respect their original analysis.

semantic level that is significant here (see also section 3.4). Unlike Tamil, the Sinhala system has a natural gender in which the main distinction is between Animate and Inanimate. Inanimate nouns can take four cases: Nominative, Dative, Genitive/Locative, and Instrumental/Ablative, while animate nouns have two additional cases, Accusative and Vocative (Gair & Paolillo 1997). The animacy effect noted in the SLM pronouns could have developed under influence of a more pervasive animacy feature found in Sinhala, the L2 of many SLM speakers. This can be reconciled with a modular view of grammars in contact, in which transfer of semantic features without corresponding syntactic categories constitutes one of the mechanisms of contact-induced transfer (Aboh 2004).

(7) shows how the Dative marker on ‘child’ does not mark Experiencer but Goal; this can be seen as a parallelism to Sinhala, where modal verbs with Nominative subjects will take objects marked for Dative (see Gair & Paolillo 1997: 35), as shown in (8).

(7) go anak-panna sayan  
 I child-PL:DAT love  
 ‘I love (the) children’

(8) mamə ee wagee minissunආ haryəආ kæmətii (Sinhala)  
 I that kind man-PL-DAT really like-ASSN  
 ‘I really like that kind of person’

I am not aware of similar constructions in Tamil; typically a Dative subject is always required in a modal predication, cf. (9).

(9) enakku kaappi veeආum (Tamil, Asher 1982: 169)  
 I-DAT want coffee  
 ‘I want a coffee’

When an active predicate takes two or more arguments, the more directly affected argument is marked as ACC. The secondary argument is marked as Dative to reflect the semantic role of Goal, as shown in (10) and (11).

(10) Sir     aanak-pada-yaŋ     ruma-naŋ     e-luppa  
teacher child-PL-ACC     house-DAT PAST-send  
'The teacher sent the children home'

(11) skola   aanak hattu buk-yaŋ     prinsipal-naŋ   kasi  
school child one book-ACC principal-DAT give  
'The student gives a book to the principal'

In KJ, as in other SLM varieties, *-pe* is a possessive marker, seen in (12) and (13).

(12) goppe   tumman go-yaŋ     e-tolak  
I:POSS friend I-ACC PAST-push  
'My friend pushed me'

(13) goppe     nama (kata) Rihan  
1SG:POSS name (call?) Rihan  
'My name is Rihan'

The marker *-le* marks Comitative case as in (14); it should be noted that, as in South Asian typology, this can also mark emphasis as in (15), among other functions.

(14) ikkaŋ-le     se er-makaŋ  
fish-COM     I DUR-eat  
'I also eat fish (among other things)'

(15) baaru mostor-le     ikaŋ me-pegan   tara<sup>13</sup>  
new method-COM fish catch   NEG-have'  
'We do not have *new* methods to catch fish'

---

<sup>13</sup> The prefix *me-* is originally the agentive marker in Malay; these forms are lexicalized in SLM as the prefix is no longer productive.

Locative *-ka* (16) can be also used in possessive-like function as in (17):<sup>14</sup> such a conflation of existence and possession is not uncommon in languages of South and East Asia.<sup>15</sup> In (18) we see both Possessive and Locative case in the same sentence.

(16) go ikkaṅ pegaṅ er-pi bot-ka  
 I fish catch DUR-go boat-LOC  
 ‘I go fishing in the boat’

(17) nembak oraṅ-ka ada snapan bae  
 hunt man-LOC have rifle good  
 ‘The hunter has a good rifle’

(18) goppe bot-ka ikkaṅ-jo er-pegang  
 I:POSS boat-LOC fish-only DUR-catch  
 ‘I only go fishing in my boat’

The single occurrence of marker *-riṅ* indicates Instrumental case, as in (19). In (20) however, where *-riṅ* marks both ‘market’ and ‘fish’, we observe what appears to be a case of ablative-instrumental syncretism.

(19) dia ikkaṅ-yaṅ er-birsik-ing baru pisu-riṅ  
 he fish-ACC DUR-clean-CAUS new knife-INST  
 ‘He cleans the fish with a new knife’

(20) mahin, market-riṅ ais tra baaru ikkaṅ-riṅ billi bawa  
 Son market-INST ice NEG new fish-INST buy bring  
 ‘Son, buy me some fresh fish from the market’

[Lit. Son, from the market bring me some without ice, new of the fish]

If we also consider the use of the pronominal form *kitaṅriṅ* meaning ‘from us’ and the expression *goriṅ ludaṅ*, ‘from me to you’, we see here, as well as in (20), INST case

<sup>14</sup> As pointed out by an anonymous reviewer, this appears to be at variance with dative-possessor coding strategies in Tamil.

<sup>15</sup> Ian Smith p.c. 2004 and...

encoding ablative or provenance. The same syncretism between ablative and instrumental is found in Sinhala inanimate nouns (cf. Gair & Paolillo 1997: 17; Gair 1998: 66), in (21).

- (21)    *aan*□*uweŋ*            *eekə*□*ə*    *aadaarə*    *denəwa*    (Gair 1998: 66)  
           government-INST    that-DAT    support-PL    give-PRES  
           ‘The government gives support for that’

This feature of syncretism can be found in a number of Indo-European languages (Blake 1994: 175), but is crucially absent from Tamil where two different suffixes mark Instrumental and Ablative case on nouns (Asher 1982), e.g.

Insert Tamil ex.

Another interesting feature of the nominal domain in SLM is its agglutinative morphology, as shown in (22), where ‘Dative’ and ‘Comitative’ markers attach to one another: *naŋ* > *na* + *le* = *nale*. Number and case also combine and appear frequently fused together: the morpheme *panna* [= *pada*+*nang*], seen earlier in (7) and here in (23), provides another example of agglutinative tendencies in the morphology of KJ.

- (22)    *samma*    *aanak-nale*            *larinaŋ*    *suka*  
           All        child-DAT:COM    run-DAT    like  
           ‘All children like running’

- (23)    *siŋhalis-panna*            *ada*    *bissar*    *bot-pada*  
           Singhalese-PL:DAT    have    big    boat-PL  
           ‘The Singhalese have big boats’

These agglutinative tendencies of SLM morphology are interesting from a typological point of view, considering that PMD varieties are typically isolating, while both Tamil and Sinhala are agglutinative languages. In this sense, SLM appears closer to a Lankan morphological type than an Austronesian one.

### 3.2 The Accusative marker

We now focus further on the Accusative and Dative case markers, which have been reported by previous authors to show conflation in SLM (e.g. Hussaimiya 1987, Smith et al 2004). This is clearly not the case in my corpus and could be a case of local variation.

In using different markers for Accusative and Dative, KJ is consistent with Lankan typology, in both adstrates such a distinction is present. As can be seen from examples (24) - (26) below, the marker *-yaŋ* appears clearly as a definite object marker of direct objects:

- (24) pohong-yaŋ potong  
Tree-ACC cut  
'(I) cut the tree'
- (25) inni kendera-yaŋ bapi  
This chair-ACC take:go  
'Take this chair away'
- (26) kaki-yaŋ ambel  
Leg-ACC move  
'Move your legs away!'

The indefinite reading that results from the absence of *-yaŋ* can be seen in examples (3) and (4) earlier. As already noted in Smith (2003), there is a resemblance with the function of the Accusative case in Tamil, illustrated in (27) and (28).

- (27) naan avan-ai. ppaar-tt-een (Annamalai & Steever 1998: 107)  
I-NOM that.man-ACC see-PST-1S  
'I saw him'
- (28) naan anta.p pustakatt-ai vaanj-in-een (Annamalai & Steever 1998: 107)  
I-NOM that book-ACC buy-PST-1S

‘I bought that book’

In Tamil, Accusative case is obligatory for a human, direct object. In non-human direct objects, Accusative case indicates definiteness. In Sinhala, Accusative is only used with animate objects, while inanimate nouns lack Accusative case. Moreover, even with animate nouns, it is highly optional (Gair & Paolillo 1997).

Insert Sinhala ex.

Accusative case marking in SLM, being restricted in terms of animacy and definiteness, reflects influence from both Sinhala and Tamil.

### 3.3 Experiencer/Goal marker

*-Naŋ<sup>16</sup>/-daŋ* is more versatile and complex in its functional domains and represents a typical example of what is usually referred to as the South Asian Dative case, in which the following functions conflate: Experiencer, seen earlier in (5), Goal in (10) and (11) earlier, and in (29) and (30), as well as Beneficiary of an event in (31) and (32).

(29) Rihaŋ maŋke-naŋ eŋ-pi Faizal sama  
R. market-DAT DUR-go F. with  
‘Rihaŋ goes to the market with Faizal’

(30) dia maŋke-le suŋiŋe-naŋ e-pi  
3S night-COM mosque-DAT PST-go  
‘He even went to the mosque at night’

(31) ini buk go-riŋ lu-daŋ  
this book I-INSTR you-DAT  
‘This book is from me to you’

---

<sup>16</sup> Can be shortened to *-na* in fast speech.

- (32) ini foto loraŋ-naŋ  
 This foto 2PL-DAT  
 ‘This photo is for you’

Dative case can even be used to indicate possession, in a syntactic construction with ‘have’, as in (33).

- (33) ni anak-naŋ baek buku-yaŋ attu aada  
 This student-DAT good book-ACC one have  
 ‘This student has a good book’

It should be noted, however, that the conflation of Experiencer, Benefactive, Goal and Possession is in fact not peculiar nor unique to South Asian languages, but can be seen as a universal tendency of Dative case marking, as clearly shown in Blake (1994: 145), who describes Dative as the main non-core case used to mark complements.

For most speakers of KJ, the absence of the Dative marker is not acceptable in most syntactic environments, particularly when in the function of Experiencer as opposed to Goal (or Beneficiary):

- (34) ?go minum mau  
 I drink want

This is particularly relevant since there is variation in the frequency in which other case markers are used. Where context does not require overt specification, speakers freely omit markers and only resort to them for disambiguation purposes, as with the ACC marker in (35) and (36).

- (35) Sudara jariŋ e-buaŋ ikkaŋ e-pegan  
 brother net PAST-throw fish PAST-catch  
 ‘Brother threw the net and caught (some) fish’

- (36) Sudara jariŋ-yaŋ e-buaŋ ikkaŋ e-pegan

brother net-ACC PAST-throw fish PAST-catch  
'Brother threw the net and caught (some) fish'

In terms of frequency and distribution, the most consistent marking found in the NP in KJ relates to what we could call the Agentive-Experiencer/Goal opposition, i.e. zero-marking which appears to be reserved for prototypical Agents, and the DAT marker which covers first arguments in non-agentive roles as well as a range of Patient/Goal roles. It is clear that this marker in KJ shows significant typological convergence with the Sinhala/ Tamil Dative case. Dative Subjects in Tamil appear to be grammatically constrained to occurrence with stative verbs (Schiffman 1999: 64):

(37) adu enakku teriyum (Schiffman 1999: 100)  
that I.DAT know  
'I know that'

(38) enakku kaappi veeṇum (Lindholm 1978)  
I.DAT coffee want  
'I want some coffee'

In Sinhala, Dative marking can be used to distinguish between volitional Agents (marked by Nominative, cf. (39)), and other less volitional entities (marked by Dative, cf. (40)):

(39) miniha duwəṇəwa (Gair & Paolillo 1997: 32)  
man-NOM run-PRES  
'The man runs'

(40) minihaṁə diwəṇəwa (Gair & Paolillo 1997: 33)  
man-DAT run-INVOL-PRES  
'The man runs (involuntarily)'

A distinction between Nominative and Dative as found in KJ can be captured in terms of Control (cf. e.g. Dik 1989) and mirrors the notion of 'Volitional Actor' in

Sinhala (Gair 1976) as well as the concept of ‘Dative subject’ found in the Sinhala and Tamil Dative cases (Schiffman 1999). Dative subjects have high frequency in colloquial Sinhala, but low frequency in Lankan Tamil (Silva 2003; Sebastian Nordhoff, p.c. June 2005). The feature of control as a determining aspect of case assignment, that results in the prominence of Experiencer roles, or ‘Dative Subjects’ (see Bhaskarao & Subarao 2004), is not only a typical Lankan trait but extends to the whole South Indian linguistic area (Masica 1976; Shibatani & Pardeshi 2001).

### 3.4 Summary

Table 5 illustrates the salient features of KJ discussed in the previous sections.

Table 5. Features of nominal domain of KJ and their sources.

<b>Feature</b>	<b>Source</b>
Agglutinative morphology	Lankan/ South Asian
Instr./ Abl. syncretism	Sinhala
Dative subjects	Lankan/ South Asian
Accusative marker	Lankan
Exp./ Ben./ Goal conflation	Universal
Animacy effects	Sinhala <sup>17</sup>

As shown by the right-hand column, the sources for the grammar of KJ’s nominal system show a balanced pool of specific Sinhala and Tamil features, more general Lankan (and even South Asian) traits, as well as universal tendencies. There is no indication of a predominant Tamil linguistic influence in the nominal domain of KJ.

The restructuring process that occurs in SLM in general can thus be captured in a modular view of grammar in which Malay lexical items and Lankan semantics combine to form a new system; this is close to the notion of metatypy (Ross 1996, 2001), and also shares similarities with the idea of Form-Semantics (F-S) mixed languages suggested in Bakker (2003). These two accounts differ, however, on one, crucial parameter: while the former implies gradual development, the latter allows

<sup>17</sup> A less stimulating interpretation of this variation could of course be along the lines of case allomorphy, as pointed out by an anonymous reviewer.

for abrupt changes. In the case of KJ, only the former can apply, as the case system of SLM involves reanalysis of a whole paradigm, something that in any plausible scenario cannot take place in a short time span (cf. Bakker 2003) but must occur over a prolonged period of time as befits the evolution of mature systems such as case (Dahl 2004). The picture that emerges from this study so far is thus one of gradual typological convergence, and finds support in the framework presented in Thomason and Kaufman (1988) as well as in Ross (2001); in this light, SLM would be an advanced case of restructuring in a multilingual setting, a typical product of borrowing under strong structural pressure, involving distinctive phonological features, word-order and morphological material. In the next section I will relate these findings to issues of classification of SLM varieties.

## **4. Conclusions**

### **4.1 Significance for the evolution of SLM**

This study has shown that there is no solid historical evidence for the hypothesis that SLM evolved from the intermarriage of Tamil Moors and people of Malay descent. As we have seen, in the early days of the SLM communities, there were certainly both ‘pure’ Malay/Indonesian families as well as mixed marriages between Malays and locals, be they Sinhalese, Tamil, Moors or others. Though there was a particular bond between the Moors (of Arab descent) and the Malays for religious reasons (cf. Hussainmiya 1990: 42), there is no indication that this led to a dominant Tamil influence in the evolution of SLM.<sup>18</sup> What I hope to have made clear is simply that (1) the historical evidence for the claim of widespread Moor-Malay intermarriage is not there; and consequently (2) the claim of SLM being the linguistic product of predominantly Moor-Malay intermarriage is not supported in the data.

### **4.2 Significance for contact linguistics**

---

<sup>18</sup> In fact, Hussainmiya has been repeatedly drawn into harsh debates in the media and in non-academic papers regarding the picture he paints of the ethnic origins of the SLM, as not everyone agrees with the strong Tamil-Moor connection he implies (cf. Hussainmiya ms).

Having already dealt with the creole notion on theoretical grounds, based on the discussion of case in KJ, we can say that SLM varieties do adhere to the general prototype of mixed languages in combining a dominant lexical influence from one language – the lexifier, in this case a PMD – with a clear grammatical input from one (or more) different language(s) – in this case Sinhala/Tamil (cf. Bakker & Muysken 1995; Matras & Bakker 2003). It is commonly acknowledged that a prototypical definition has to be kept flexible as (a) it is indeed only a small set that is defined by such a definition, and only one example, Michif, may fit as a prototypical example (Matras & Bakker 2003: 2), and (b) the degree of the grammar-lexicon split is sensitive to social and typological variation (Matras 2003). I am not interested here in going into the nitty-gritty of finding the one type of ML to best fit the profile of SLM.<sup>19</sup> This, therefore, is necessarily a simple and likely provisional definition. Considering that mixed languages typically show lexical items predominantly from one source, and grammatical material predominantly from another, SLM fits the bill of a mixed language of trilingual base; its lexicon is primarily of PMD origin while grammatical features are derived from Sinhala and Tamil.

The key to understanding the evolution of SLM lies in the nature of trilingualism typical of these communities. Trilingualism is achieved by members of the SLM community due to contact with the larger group of Tamils and the majority of Sinhalese speakers, not because of social pressure or urge to shift language; Sinhala/Tamil were adstrates and, if at all there were any linguistic pressure, this would have come from the colonial varieties, first Dutch, then English. At the same time, it is not plausible to assume that because of colonial rule SLM speakers would have kept themselves hermetically isolated from the adstrates. Historical sources, as presented in the first part of this paper, are rich in references to the multiethnic, multilingual environment of colonial Sri Lankan society. Inter-marriage was found, though no overwhelming practice involving one particular speech community was found to suggest a major, unidirectional shift. Social circumstances are of the utmost importance in order to assess the outcome and the speed of language change; and

---

<sup>19</sup> Also, the fact that the literature on MLs so far shows almost as many different types as actual attestations of mixed varieties suggests that we are still far from agreeing on prototypical traits of MLs.

there is nothing in the evolution of SLM which points to highly marked circumstances to justify rapid change (cf. Bakker 2000: 616-617). From a structural point of view, the case system of SLM is an instance of a mature system (cf. Dahl 2004: 115) that requires gradual evolution. Trilingualism can explain the source of typological convergence through contact. From a generic Low Malay/Javanese variety, which does not feature nominal marking, the evolution of such a system implies significant realignment of the cognitive-functional principles determining case assignment. This, I suggest, seems possible only due to widespread, long-lasting multilingualism, *in particular trilingualism* which also allows for the type of transfer observed also in situations of second language acquisition (cf. Siegel 1997, 2004). In short, all the signs we have available strongly suggest a gradual type of restructuring.

To sum up, the nature of SLM varieties scores high on a number of points typical of MLs: (a) grammatical admixture, (b) restructuring and (c) conventionalization; it has a clearly identifiable lexifier, but is not particularly isolated nor does it show linguistic separation from the source. In my interpretation of the current state of the mixed language debate (cf. Matras & Bakker 2003), this is more than sufficient evidence for treating SLM as a ML.

### **4.3 Concluding remarks**

Much work remains to be done on the various varieties of SLM and these conclusions need therefore to be taken as provisional generalizations, not state of the art knowledge. The project on the documentation of SLM funded under the DoBeS initiative will avail us with many times the amount of data that has been somewhat available to limited extents so far, and will surely lead us to revise, expand and refine our understanding of these varieties, as well as their contribution to the study of high-contact varieties, in years to come.

### **Acknowledgements**

I would first like to acknowledge the Volkswagen Stiftung's initiative for the documentation of endangered languages (DoBeS) for their financial, technical and intellectual engagement. I am very grateful to members of the Sri Lankan Malay Association (SLMA) and the Conference of Sri Lanka Malays (COSLAM) in Colombo for their support in this research. In particular, I want to thank Mr. T.K. Azoor, Mrs. D. Burah, Mr. T.M.M. Hamin, Mrs. Salma Peiris and Mr. B.D.K. Saldin in Colombo. In Hambantota I want to thank Proctor T.S. Doole and Mr. T.M. Farook. I am very indebted to Principal B. Jalaldeen as well as the whole community of Kirinda for sharing their invaluable knowledge with me. I am also indebted to Hans den Besten, Stephen Matthews, Sebastian Nordhoff, Ian Smith and Don Winford for comments, as well as two anonymous reviewers. Lastly but most importantly, I want to acknowledge Lisa Lim for her substantial collaboration in the collection and glossing of the data and for the many constructive discussions on Kirinda Java.

## References

- Aboh, E. (2004). Toward a modular view of creole genesis. Paper presented at the SPCL summer meeting, Curaçao Creole Conference, Curaçao, August 2004.
- Adelaar, K.A. (1991). Some notes on the origin of Sri Lankan Malay. In H. Steinhauer (Ed.), *Papers in Austronesian linguistics No. 1*. Pacific Linguistics (A-81) (pp. 23-37). Canberra: The Australian National University.
- Adelaar, K.A. and D.J. Prentice (1996). Malay: Its history, role and spread. In S.A. Wurm, P. Mühlhäusler and D.T. Tryon (Eds.), *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas* (pp. 673-693). Berlin: Mouton de Gruyter.
- Annamalai, E. and S.B. Steever (1998). Modern Tamil. In S.B. Steever (Ed.) *The Dravidian languages*. London/ New York: Routledge.
- Ansaldo, U. (2005). Typological convergence and admixture in Sri Lanka Malay. Ms. Universiteit van Amsterdam.
- Ansaldo, U. and S.J. Matthews (2001). Typical Creoles and simple languages: The case of Sinitic. *Linguistic Typology* 5 (2/3), 311-325.
- Ansaldo, U. & S.J. Matthews (Eds.) (in prep.). *Deconstructing "Creole". New perspectives on language creation*. Amsterdam/ Philadelphia: John Benjamins.
- Ansaldo, U. and L. Lim (2005). The Kirinda Corpus. Ms. Universiteit van Amsterdam.
- Asher, R.E. 1982. *Tamil*. Amsterdam: North Holland.
- Asher, R. and E. Annamalai (2002). *Colloquial Tamil*. London/ New York: Routledge.
- Baker, P. (1996). Pidginization, creolization and français approximatif. *Journal of Pidgin and Creole Languages* 11, 95-120.
- Bakker, P. (2000). Rapid language change: Creolization, intertwining, convergence. In C. Renfrew, A. McMahon and L. Trask (Eds.), *Time depth in historical linguistics* (pp. 585-620). Cambridge: McDonald Institute for Archeological Research.
- Bakker, P. (2003). Pidgin inflectional morphology and its implications for creole morphology. In G. Booij and J. van Marle (Eds.), *Yearbook of Morphology 2002* (pp. 3-33). Great Britain: Kluwer Academic Press.

- Bakker, P. and P. Muysken (1995). Mixed languages and language intertwining. In J. Arends, P. Muysken and N. Smith (Eds.), *Pidgins and creoles: An introduction* (pp. 41-52). Amsterdam/ Philadelphia: John Benjamins.
- Bhaskarao, P. and K.V. Subarao (Eds.) (2004). *Non-nominative subjects*. (2 Vols.) Amsterdam/ Philadelphia: John Benjamins.
- Bichsel-Stettler, A. (1989). Aspects of the Sri Lanka Malay community and its language. M.A. dissertation, Bern University, Switzerland.
- Blake, B. (1994). *Case*. Cambridge: Cambridge University Press.
- Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity*. Amsterdam/ Philadelphia: John Benjamins.
- DeGraff, M. (2001). On the origins of creoles: A Cartesian critique of Neo-Darwinian linguistics. *Linguistic Typology* 5, 213-230.
- DeGraff, M. (2004). Against creole exceptionalism (redux). *Language* 80 (4), 834-839.
- DeGraff, M. (2005). Linguists' most dangerous myth: The fallacy of creole exceptionalism. *Language in Society* 34, 533-591.
- Dik, S.C. (1989). *The theory of Functional Grammar. Part I*. Dordrecht/ Providence: Foris Publications.
- Emeneau, M. B. (1980). *Language and linguistic area*. Stanford, California: Stanford University Press.
- Gair, J. (1976). Is Sinhala a subject language, or how restricted is your PNP? In M. Verma (Ed.), *The notion of subject in South Asian languages*. (pp. 39-64). Madison: South Asia Program, University of Wisconsin-Madison.
- Gair, J. and J. Paolillo (1997). *Sinhala*. München: Lincom Europa.
- Grijns, C.D. (1991). *Jakarta Malay: A multidimensional approach to spatial variation*. Leiden: KITLV Press.
- Heine, B. and T. Kuteva (2005). *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Holm, J. (1988). *Pidgins & Creoles* (2 vols). Cambridge: Cambridge University Press.
- Hussainmiya, B.A. (1987). *Lost cousins: The Malays of Sri Lanka*. Kuala Lumpur: Universiti Kebangsaan Malaysia.
- Hussainmiya, B.A. (1990). *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Kuala Lumpur: Universiti Kebangsaan Malaysia.

- Hussainmiya, B.A. (year?) Moor-Malay marriages. Ms. University of Brunei Darussalam.
- Karunatilake, W.S. (2004). *An introduction to spoken Sinhala*. Colombo: Gunasena.
- Lefebvre, C. (2001). Relexification in creole genesis and its effects on the development of the creole. In N. Smith and T. Veenstra (Eds.), *Creolization and Contact* (pp. 9-42). Amsterdam/ Philadelphia: John Benjamins.
- Lindholm, J.M. (1978). Nested case relations and the subject in Tamil, *International Journal of Dravidian Linguistics* 7,1: 75-98.
- Masica, C. (1976). *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.
- Matras, Y. (2003). Re-examining the structural prototype. In Y. Matras and P. Bakker (Eds.), *The mixed language debate: Theoretical and empirical advances* (pp. 151-176). Berlin/ New York: Mouton de Gruyter.
- Matras, Y. and P. Bakker (Eds.) (2003). *The mixed language debate: Theoretical and empirical advances*. Berlin/ New York: Mouton de Gruyter.
- McWhorter, J. (1998). Identifying the creole prototype: Vindicating a typological class. *Language* 74, 788-817.
- McWhorter, J. (2004). Saramaccan and Haitian as young grammars: The pitfall of syntactocentrism in Creole genesis research. *Journal of Pidgin and Creole Languages* 19 (1), 77-138.
- Mufwene, S. (1999). The language bioprogram hypothesis. Hints from Tazie. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony and development* (pp. 95-127). Cambridge, MA: MIT Press.
- Mufwene, S. (2000). Creolization is a social, not a structural process. In E. Schneider and I. Neumann-Holzschuh (Eds.), *Degrees of restructuring in Creole languages* (pp. 65-84). Amsterdam/ Philadelphia: John Benjamins.
- Mufwene, S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.
- Ricklefs, M.C. (1974). *Jogjakarta under Sultan Mangkubumi, 1749-179. A history of the division of Java*. London/ New York: Oxford University Press.
- Ross, M. (1996). Contact-induced change and the comparative method. Cases from Papua New Guinea. In M. Durie and M. Ross (Eds.). *The comparative method reviewed: Regularity and irregularity in language change* (pp. 180-217). Oxford: Oxford University Press.

- Ross, M. (2001). Change in Oceanic languages. In A.Y. Aikhenvald and R.M.W. Dixon (Eds.), *Areal diffusion and genetic inheritance. Problems in comparative linguistics* (pp. 134-166). Oxford: Oxford University Press.
- Schiffman, H. (1999). *A reference grammar of Spoken Tamil*. Cambridge: Cambridge University Press.
- Schweitzer, C. (1931). *Reise nach Java und Ceylon 1675-1682*. Haag: Nijhoff.
- Shibatani, M. and P. Pardeshi (2001). Dative subject constructions in South Asian languages. In P. Bhaskararao and K.V. Subbarao (Eds.), *The yearbook of South Asian languages and linguistics 2001* (pp. 311-347). Delhi: Thousand Oaks; London: Sage.
- Siegel, J. (1997). Mixing, levelling and pidgin/creole development. In A. Spears and D. Winford (Eds.), *The structure and status of pidgins and creoles* (pp. 111-149). Amsterdam/ Philadelphia; John Benjamins.
- Siegel, J. (2004). Morphological elaboration. *Journal of Pidgin and Creole Languages* 19 (2), 333-362.
- Silva, A.W.L. (2003). *Teach yourself Tamil. A complete course for beginners*. Kandana, Sri Lanka; Pubudu Printers.
- Silverstein, M. (1976). Hierarchy of features and ergativity. In R.M.W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112-171). Canberra: Australian Institute of Aboriginal Studies.
- Slomanson, P. (2004a). Sri Lankan Malay Grammars: Lankan or Malay? Revised version of paper presented at the 7th International Symposium on Malay/Indonesian Linguistics, Nijmegen, June 2003, and the SPCL summer meeting 2003, University of Hawai'i, August 2004.
- Slomanson, P. (2004b) The syntax of tense and aspect in Sri Lankan Malay. Paper presented at the SPCL summer meeting, Curaçao Creole Conference, Curaçao, August 2004.
- Smith, I. (2003). The provenance and timing of substrate influence in Sri Lanka Malay: Definiteness, animacy and number in accusative case marking. Paper presented at The South Asian Language Analysis Roundtable XXIII, Austin, Texas, October 2003.
- Smith, I., S. Pauw and B.A. Hussainmiya. (2004). Sri Lanka Malay: The state of the art. In R. Singh (Ed.), *Yearbook of South Asian Languages 2004* (pp. 197-215). Berlin/ New York: Mouton de Gruyter.

Thomason, S. and T. Kaufman (1988). *Language contact, creolization and genetic linguistics*. Berkeley/ London: University of California Press.

Wijeratne, W.M. (2005). Estate Tamil: A new variety of Tamil. Paper presented at the *10<sup>th</sup> International Conference on Sri Lankan Studies*, 16-18 December University of Kelaniya.