Tracing the linguistic crossroads between Malay and Tamil

Tom G. Hoogervorst

1. INTRODUCTION

When Europeans first entered the waters of the Indian Ocean, they encountered a vibrant, interconnected world in which Gujaratis, Persians, Tamils, Swahilis, Arabs, Malays, and a wide range of other peoples traded and settled on shores other than their own. Upon arriving in Malacca in the 1510s, the Portuguese apothecary Tomé Pires noted no less than 61 different nations inhabiting that city, representing much of the Asian continent and the Indian Ocean World. Facilitated by the annual cycle of the monsoon, the Malay-speaking settlements on both sides of the Strait of Malacca formed vital trade entrepôts connecting various parts of Asia and facilitating the dispersal of people, products and ideas. Language contact must have been pervasive in the Malay speech area
since time immemorial. However, while the lexical influence from high-status literary languages such as Sanskrit and Arabic on Malay is relatively well-known (Jones 2007), the impact of spoken vernaculars remains much less so. This is due in part to the fact that many vernacular languages of South and Southeast Asia are themselves understudied, especially in language ecologies characterized by wide-ranging diglossia. Furthermore, language contact between Southeast Asia and other regions of Asia has long been approached as a unidirectional process, reducing Southeast Asia’s populations to mere recipients. There is a modicum of work on the dispersal of pre-modern loanwords from West-Malayo-Polynesian languages to other languages of the Indian Ocean (Hoogervorst 2013), but more could be done in this area. With the exception of Sri Lanka Malay, mixed languages at the interface of Malay and Tamil are almost undocumented.

An historical analysis of language contact between Malay and Tamil, as will be attempted here, provides a better understanding of the past of the Bay of Bengal as an axis of global trade and cultural exchange. This study traces the shared history of two of the largest speech communities of the Indian Ocean World, reconstructing their inter-relationship across several time periods and geographical settings. In the absence of accurate grammatical descriptions of most of the “hybrid” linguistic varieties discussed in this paper, much of my analysis will be of etymological nature. Consequently, this paper cannot be anything but sweeping and remains far from exhaustive. Most of the data and insights presented here are taken from secondary sources, rather than first-hand fieldwork. That being said, the paucity and scattered distribution of scholarship on Malay-Tamil language contact calls for a synthesis and overview of the available data as a first step to determine pathways for further research. In doing so, this study serves to demonstrate what we know, but also what we do not know. It is structured as follows: Section 2 summarizes the long history of contact between Malay and Tamil; Section 3 focuses on relationship between the two languages as reflected in the classical Malay literature; Section 4 introduces the type of Malay spoken by Tamils at present; Section 5 surveys Malay varieties in historical contact with Tamil; Section 6 traces the languages spoken by mixed Malay-Tamil communities; and Section 7 synthesizes our present state of knowledge on the Tamil variety (or varieties) used in Malaysia.

2. History of contact
The archaeological record reveals that contact between South India and Southeast Asia was regular from the first centuries BCE (Ardika and Bellwood 1991; Bellina and Glover 2004). The Old Javanese kakawin literature contains numerous Tamil loanwords, as does classical Malay (Hoogervorst in press a). From at least the ninth century, Tamil inscriptions surface across Southeast Asia (Karashima and Subbarayalu 2009), while different Indian ethnonyms

1 And see Ronkel (1902), Asmah (1966), and Jones (2007) on Tamil loans in modern Malay.
start to feature in the Old Javanese literature around the same time (Christie 1999). For example, early eleventh century Airlangga inscriptions make a distinction between Kling, Áryya, Singhala, and Karṇaṭaka (Krom 1913), while the mid-fourteenth century Nāgarakartāgama adds Goḍā and Kāṇḍīpurī (Pigeaud 1962: 36). South Indian influence is especially strong in North Sumatra. The Dutch orientalist Van Ronkel (1918) was the first to call attention to a number of cultural and lexical peculiarities among the Karo-speaking Sembiring clan, which he connected to the historical presence of Tamil trading guilds in the region. Recent archaeological research supports the settlement of South Indian populations in North Sumatra in medieval times (Guillot and Fadillah 2003; Perret and Surachman 2009). In later times, multi-ethnic Islamic networks between South India, Sri Lanka and the Malay World begin to overshadow earlier Hindu and Buddhist connections (ʿĀlim 1993; Tschacher 2001; Feener and Sevea 2009; Ricci 2011).

The South Indian populations in contact with Maritime Southeast Asia were diverse in terms of religion and caste. By the fourteenth century, Tamil-speaking Muslim communities started to outnumber their Hindu compatriots (McPherson 1990). The first group was then commonly known as Kling or Keling. This ethnonym is probably connected to the Kaliṅga State in present-day Odisha and would later become the generic name for “Indian”, even applied to some Indianized communities in Southeast Asia (compare Damais 1964; Mahdi 2000: 848). At present, the term is regarded as pejorative across the Malay-speaking world. The collective term for South Indian Muslim traders was Chulia or Chuliah.2 The Chulia were seen as distinct from mercantile Muslim groups from Gujarat and other western regions of India, such as the khojā and the Bohrā (compare Hussainmiya 1990; Noor 2012).3 South Indian Muslim communities display a substantial and at times confounding terminological variety (Bayly 1989; ʿĀlim 1993; Tschacher 2001; Hussein 2007; Pearson 2010). One of the terms used for them by non-Muslim Tamils is Jōṉagaṉ (ஜோஙகான்), which is especially applied to Muslims of partly Arabic or Turkish descent. The colonial British censuses typically distinguish the following subgroups of South Indian Muslims:

1. Marakkar or Maricar (Tamil: Marakkāyar; മരക്കയര്‍, Malay: Marikar)
A group claiming ancestry from Arabic merchants, as opposed to less esteemed local converts. They were mostly involved in international shipping trade, inhabited coastal regions, and adhered to the Shāfiʿī school (maddhab) of Islamic jurisprudence (fiqh). The Kāyalār, from the coastal town Kāyalpatṭīnām, are normally considered to be a subgroup of the Marakkāyar.

---

2 Malay Culia, Tamil Cūliyā (சுளியா). The origins of this term are uncertain. See Khoo (2014) for a history of the Chulia community in Penang.
3 Hindu merchants from Gujarat were known as Baniyān.
2. Labbai or Labbay (Tamil: Labbai; లబ్బాయు, Malay: Labai)

Originally an honorary term for an Islamic functionary, but later used to designate a particular Tamil-speaking community of the Hanafi maddhab. They were traditionally involved in trade, pearl-diving and betel-cultivation. The term Labbai is also occasionally applied to non-Marakkāyar Tamil-speaking Muslims as a whole.

3. Mappila or Moplah (Malayalam: Māppila; മാപ്പിള, Malay: Ləbai)

Malayalam-speaking Muslims of partly Arabic ancestry who chiefly resided in the Malabar region (present-day Kerala). The majority follow the Shāfi‘i maddhab.

4. Muslims “from the north”

A container term for predominantly Urdu-speaking Muslims residing in different parts of South India, encompassing the ethnonyms Navaiyat, Sayyid, Shaykh, and Pathān. These groups claim to be descended from non-Dravidian men in service of the Mughal and Deccan sultans. Special mention can be made of the Rowthers (Tamil: Rāvuttar; ராவுட்டர்), a Tamil-speaking group of the Hanafi maddhab claiming descent from Turkish (Tulukkar; Ṭūlūkka) horsemen.

Many Indian merchants who ventured to Southeast Asia married local women. The affluent and influential mixed community that thus emerged became known in Malay as the jawi Parakan ‘local-born Jawi’. These children of merchants were well-connected with the Muslim elites in Southeast Asia and beyond (Fujimoto 1989). In the Straits Settlements, their multilingual background, including in English, qualified them for lucrative employment under the colonial government. They were also involved in the printing press. In 1876, a Singapore-based Malay printing office under the name Jawi Parakan published – at the same time – Southeast Asia’s first Tamil and first Malay newspaper (Birch 1969; Tschacher 2009). In Aceh, mixed people of Tamil ancestry – mentioned by Snouck Hurgronje (1893: 20) as basterd-Klinganeezen – appear to have largely assimilated into the Acehnese mainstream, being only recognizable on a phenotypical level. A still existing hybrid group are the so-called Chitty (Tamil: Cetti; சித்தி), the offspring of Kaling fathers and Malay mothers in Malacca. Their name goes back to Chetty (Tamil: Cetti; செட்டி, Malay: Ceti), a term loosely applied to a number of South Indian mercantile castes and money-lenders in the Malay World. The Chitty people have kept their Hindu

---

4 In Sri Lanka Malay, lebbe still refers to an Islamic scholar (Saldin 1993: 1015). In Indonesia, Labai typically refers to a mosque official.

5 In Penang, the term Jawi Paskan ‘urban Jawi’ is more common. The word Jawi presumably goes back to Arabic Jāwī, an umbrella term for Malays and other Southeast Asian Muslims.

6 The Tamil newspaper was named Taṅgai Sinēhay (தங்கை சின்னைய), the Malay newspaper Jawi Peranakan. Contrary to popular belief, the latter was not the world’s first Malay newspaper. Already in 1869, the Alamat Langkapuri was issued in Colombo, Sri Lanka, by a member of the Malay diaspora (Ricci 2013).
religion to this day, yet can be considered Malay in terms of language and culture (Raghavan 1977). In post-independence Malaysia, however, Chitties have thus far been unsuccessful in claiming *bumiputra*-ship, whereas most Muslim Jawi Peranakan conveniently registered themselves as Malays.

Cultural contact between South and Southeast Asia persisted into modern times. While the British Government had a long tradition of employing Indian personnel in the Straits Settlements, the late nineteenth century saw a substantial increase of labour migration from South India to the Malayan rubber plantations and tea estates. This led to an influx of Indian Tamils, Sri Lankan Tamils, Malayalis, Telugus, and other South Asian communities, then known as ‘cooilies’ (Tamil: *küli*; ்குலி). This system of indentured labour depended on local recruiters and became known as the “kangani system” (compare Guilmoto 1993), from the Tamil word *kangāṇi* (ஏகாணி) ‘supervisor of coolies in plantations’. Around the same time, the Dutch relied on agricultural labourers from South India and other regions to work on the infamous estates of Deli in northern Sumatra (Mani 1993a). The cultural cross-fertilization between South and Southeast Asian populations in Malaysia, Singapore, southern Thailand and parts of Sumatra led, among other things, to the emergence of a mixed Indian-Malay cuisine known as *mamak* food, from the Tamil word *māmā* (மாமா) ‘uncle’. Popular dishes include *roti canai* ‘layered flatbread’, *murtabak* or *martabak* ‘stuffed pancake’, *nasi kandar* ‘steamed rice with various curries’, *mamak rojak* ‘fruit and vegetable salad’ and *teh tarik* ‘pulled tea’.

3. Literary connections

In the light of the trans-regional Islamic networks across the Bay of Bengal, it is not surprising to see South Indian influence reflected in the classical Malay literature. The *Hikayat Sari Rama* – the Malay version of the *Rāmāyaṇa* – for example, displays some uncommon Tamil words. One example is *c-ng-g-l-n-r* ‘a type of water-lily with miraculous powers’, which reflects Tamil *ceṅgalunir* (செங்கலினையிர்) ‘purple Indian water-lily; red Indian water-lily’ (compare Von de Wall 1877-97, appendix: 24; Van Ronkel 1902: 107). Other examples are *parvadam* ‘mountain’ from Tamil *parvadam* (பார்வாடம்) (Juynboll 1899: 66) and *tirisulam* ‘trident’ from *tirisulam* (திரிசலம்) (Van Ronkel 1919: 383). The names of some of the characters, too, suggests that the *hikayat* contains Tamil influence (Table 1).

---

7 The fact that the Chitties have never converted to Islam would either imply that not all Malays were Muslims by the time this hybrid group emerged, or that interreligious marriages were historically seen as less problematic in the Malay World than they are at present (Raghavan 1977: 444-445).

8 Being registered as *Bumiputra* ‘Son of the Soil’ comes with various types of ethnicity-based state benefits (Kessler 1992). In nationalist circles, however, there was no small degree of resentment to the practice among hybrid Indian Muslims (*darah katurunan Kaling*) to claim Malay status (Hussain 2005: 124).
The *Hikayat Hang Tuah* contains more clues to Tamil influence on the classical Malay literature. During his diplomatic journey to the Tamil Land (*Bənna Kaling*), the story’s protagonist surprised his hosts with his fluency in Tamil,

9 This form is evidently rationalized as consisting of the Malay honorific *Sari*, which is Sanskrit *Śrī*.

10 This is the name of a famous Shaiva shrine in the Arcot district. The word was left unidentified by Van Ronkel (1904: 315) and is commonly transliterated as the meaningless compound *sitam barama* in later editions of *Hikayat Hang Tuah*.

11 Van Ronkel (1904: 314) regards this form as a misspelling of *P-l-y-k-t* (Pulicat).

12 Presumably the city of Calicut.

13 This presumably denotes the city of Kandy (Sri Lanka), although Van Ronkel (1904: 314) connects it to the Kannada community in South India.

14 Reflecting Sanskrit *Kṛṣṇarāja*, a common title for South Indian kings.

15 At present, this name is typically transliterated as Nala Sang Guna, which makes little etymological sense.
which he proclaimed to have learned from a Labai from that country. As first pointed out by Van Ronkel (1904), here too we find a number of Tamil-derived onomastics and caste names. Table 2 above lists the Jawi transliterations and their tentative precursors.

A third Malay literary work that casts some light on Malay notions of India is the Hikayat Parintah Nagori Banggala. This text, written in 1811 by the Chulia author Ahmad Rijaluddin, contains a number of uncommon Tamil loanwords, such as bangku ‘dagger’ from vāṅku (வங்கு) and banam ‘rocket’ from vāṇam (வாணம்) (Skinner 1982: 168, 173). In addition, we come across a number of Malay names for South Indian toponyms (Table 3), although it is uncertain what role, if any, these places play in the popular Malay imagination of those days.

<table>
<thead>
<tr>
<th>Malay</th>
<th>Tamil</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harkat</td>
<td>Ārkāṭu (ஆர்காடு)</td>
<td>Arcot</td>
</tr>
<tr>
<td>Macali</td>
<td>Macilippatṭanam (மாசிலிப்பட்டணம்)</td>
<td>Masulipatnam</td>
</tr>
<tr>
<td>Mahu</td>
<td>Māḥē (மாஹெ)</td>
<td>Mahé</td>
</tr>
<tr>
<td>Naga Patan</td>
<td>Nāgappatṭinam (நாகப்பட்டினம்)</td>
<td>Negapatam</td>
</tr>
<tr>
<td>Nagor</td>
<td>Nāgūr (நாகூர்)</td>
<td>Nagore</td>
</tr>
<tr>
<td>Tanjauru</td>
<td>Taṉjāvūr (தஞ்சாவூர்)</td>
<td>Tanjore</td>
</tr>
<tr>
<td>Tipu¹⁶</td>
<td>Srirāṅgappatṭanam (செரிங்கப்பட்டணம்)</td>
<td>Seringapatam</td>
</tr>
</tbody>
</table>

Table 3. Malay names for Tamil toponyms in the Hikayat Parintah Nagori Banggala (Skinner 1982: 26).

Muslims of mixed Malay-Tamil ancestry were often bilingual if not multilingual. By the late nineteenth century, they typically published both Tamil and Malay newspapers in Sri Lanka (Hussainmiya 2008) as well as the Malay World. In the Straits Settlements, Jawi Pāranakan children would have learnt Malay from their mothers and Tamil from private tutors (Fujimoto 1989: 98 fn. 18). This is confirmed in the writings of the famous early nineteenth-century Munshi Abdullah, a Malacca-born author of mixed origins who grew up in a district known as Kampung Pali (Tamil: paḷḷi; பள்ளி) – ‘Kampung of the Mosque’ – a historical part of Malacca known for its mixed population. In his mid-nineteenth century autobiography named Hikayat Abdullah, he wrote that “it had been the custom from the time of our forefathers in Malacca for all the children of good and well-to-do families to learn it [Tamil]. It was useful for

¹⁶ Presumably a reference to Ṭippu Sultān, the late eighteenth-century ruler of Mysore.
doing computations and accounts, and for purposes of conversation because at that time Malacca was crowded with Indian merchants. Many were the men who had become rich by trading in Malacca, so much so that the names of Tamil traders had become famous. All of them made their children learn Tami” (translation from Hill 1955: 48).

The tradition of multilingualism among the literate elite led to a degree of convergence between Islamic Malay and Tamil literature. As previous scholars have pointed out (ʻĀlim 1993: 95-99; Tschacher 2009: 53-54; Ricci 2011: 174), these shared origins become evident upon comparing the ways in which sounds absent in the Arabic alphabet are represented in Arabized Malay (Jawi) and Arabized Tamil (Arwi). This is done in a remarkably similar way, with minor differences on a diacritical level: the voiceless bilabial stop /p/ is written as ǧāʾ <ج> with three upper dots in Jawi and with one lower dot in Arwi, the velar nasal /ŋ/ as ʿayn <ة> with three upper dots in Jawi and with three lower dots in Arwi, and the palatal nasal /ɲ/ as nūn <ن> with three upper dots in Jawi and two lower dots in Arwi. It should be noted here that the Arabicized writing practice dates back to the early fourteenth century in the Malay World, whereas it presumably developed around the late sixteenth century in the Tamil-speaking areas of South India (Tschacher 2001: 27; Ricci 2011: 98), cautioning researchers not to assume an a priori eastward directionality of cultural transmission.

Regrettably, it is not known how many multilingual or otherwise hybrid manuscripts are housed in public and private libraries worldwide. In the late nineteenth-century, there is some circumstantial evidence that the Acehnese kept their administrations records in Klingaleesch, to wit, Tamil (Scherer 1891: 298). One of these manuscripts is mentioned by Voorhoeve (1952: 212) in his inventory on Indonesian manuscripts at the Leiden University Library. That same library also houses the ʿIzām al-fawā'id fi niẓām al-ʿaqā'id, a 1730s manuscript written partly in Tamil and partly in Malay (Van Ronkel 1922; Tschacher 2009: 54). ʻĀlim (1993: 105-106) calls attention to “a book on Muslim Medicine edited in 1807 AD […] written in four languages: Javanese (jawi) [sic!], Persian, Arwi and Arabic” kept in the “Indonesian Manuscript Library at Jakarta”, which is likely to be the same manuscript (B.A. Hussainmiya, personal communication on 12 December 2014). This multilingual Islamic tradition may have persisted into recent times. An undated but modern-looking manuscript discovered in 2008 by Mahyudin Syukri and Siti Aisyah in Balai Jering (Kampar, Riau Province, Indonesia) contains a quadrilingual word-list – Arabic, Malay, Urdu and Tamil – written by a local Muslim scholar (Aisyah 2014). It is hoped that similar works will surface in private collections in Indonesia, Malaysia and Sri Lanka.

4. The Malay used by Tamils
The Malay varieties spoken by communities of Chinese ancestry are relatively well-documented, for example, Lim (1981), Pakir (1986), and Gwee (2006) on “Baba Malay” of the Straits Settlements, Teo (2003) on the variety of
Kelantan, Rafferty (1982) on Malang, Oetomo (1987) on Pasuruan, and Wolff and Poedjosoenardo (1982) on Central Java. The same cannot be said of varieties spoken by Indian communities. Mohamed (2006) describes the lexicophonology of the dialect belonging to the Chitty community, whereas Hassan (1969) and Kader (1971) provide some notes on colloquial Malay as spoken by Tamils. While the majority of local-born Indians speak Tamil at home and some have switched to English, a small part of this group have adopted the Malay language. A systematic study on Malay as spoken by people of South Indian ancestry, however, remains to be conducted to this day.

Research on errors (kasalahan) in Malay offers an otherwise infrequent glimpse into the Malay speech habits of Tamils. I can only give an incomplete set of isolated examples from the sources available to me. On a phonological level, the following tendencies, some of which inconsistent, are observed among Tamil pupils in Malaysian schools (Gan 1982, quoted in Onn 1989: 78-79):

1. Omission of /h/:
   - boleh > bole ‘can’
   - hisap > isap ‘to suck’
   - hujan > ujan ‘rain’
   - puluh > pulu ‘ten’
   - sudah > suda ‘already’
   - tahun > taun ‘year’

2. Substitution of /ə/ by /a/ /e/ or /i/:
   - empat > ampat ‘four’
   - anam > anam ‘six’
   - pəcah > pica ‘broken’
   - salah > silalu ‘always’
   - səndiri > sindri ‘self’
   - təmpat > tempat ‘place’
   - tərus > terus ‘direct’

3. Gemination of word-medial stops:
   - apa > appa ‘what’
   - sadikit > siddikit ‘a little’

17 In 2005, 10.6% of Singapore’s Indian population spoke Malay at home (Aman et al. 2009). The Malaysian statistics are unknown to me.
18 All three variations are attested, for example, in the Tamil renditions of Malay toponyms: Kolantan (Kilāntān; கோலாந்தான்), Malaka (Malākkā; மலாக்கா), Nagari Sembilan (Negiri Sembilān; நாகரி செம்பிளான்), Perlis (Perlis; பர்லிஸ), Tārāngganu (Tiraṅgānu; தைராங்கானு).
(4) Devoicing of /g/
\begin{align*}
gətah & \rightarrow \text{keta} \text{ ‘rubber’} \\
goreng & \rightarrow \text{koring} \text{ ‘to fry’} \\
tiga & \rightarrow \text{tika} \text{ ‘three’}
\end{align*}

(5) Fricativization of /w/
\begin{align*}
wang & \rightarrow \text{van [wang?] ‘money’} \\
wayang & \rightarrow \text{vayang ‘movie’}
\end{align*}

(6) Monophthonization of /ai/
\begin{align*}
kədai & \rightarrow \text{kede ‘shop’}
\end{align*}

Gan (1982) gives no examples on the pronunciation of the diphthong /au/. Mohamed (2006: 88-89) mentions in passing that kalau ‘if’ is pronounced as kalu by Tamil speakers.\(^\text{19}\) She also lists two examples which make it clear that the glottal stop /ʔ/ at the end of a syllable – written in Malay as <k> – is omitted by Tamil mother tongue speakers: anak > ana ‘child’ and kakak > kaka ‘older sister’. In addition, it has been pointed out that Tamil-speakers pronounce orang ‘person’ as worang and barang ‘goods’ as bareng (Hassan 1969: 218). Both examples can be explained through the phonology of colloquial Tamil. The automatic onset of /w/ before close and close-mid back vowels is common across spoken Tamil varieties (Schiffman 1999: 16). In certain dialects, the word-final ending –/aN/ is pronounced as –/ẽ/ (Schiffman 1999: 18), which would explain the transcription of bareng for standard Malay barang. No further examples are given by Hassan (1969) or other scholars to determine the distribution and regularity of this phonological tendency.

On a grammatical level, Kob (1989: 495) provides the following examples of interference from colloquial Tamil on the word order of spoken Malay (here and elsewhere: the spelling, translations and glosses are mine):

(7) \begin{align*}
\text{I} & \text{tu} \ \text{saya} \ \text{pun} \text{ya} \ \text{suka}=\text{lah} \\
\text{DE} & \text{M} \ \text{1SG} \ \text{POS} \ \text{like=}\text{PART} \\
\text{In} & \text{d} \ \text{a} \ \text{u} \ \text{a} \ \text{i} \ \text{ya} \ \text{v} \ \text{i} \ \text{r} \ \text{u} \ \text{p} \ \text{p} \ \text{a} \ \text{m} \ \text{m} \ \text{T} \ \text{a} \ \text{m} \ \text{i} \\
\text{DE} & \text{M} \ \text{1SG,GEN} \ \text{desire} \\
& \text{‘That’s the one I like.’} \\
& \text{(standard Malay: Saya suka itulah.)}
\end{align*}

(8) \begin{align*}
\text{Api} & \ \text{kə} \ \text{ret} \ \text{a} \ \text{j} \ \text{a} \ \text{l} \ \text{a} \ \text{n} \ \text{s} \ \text{u} \ \text{d} \ \text{a} \ \text{h} \\
& \text{fire cart go already}
\end{align*}

\(^{\text{19}}\) This is also the case in Sri Lanka Malay (Paauw 2004).
However, the latter example may also reflect a Hokkien structure (Kob 1989: 495):

(9) Hóe chhia kiâⁿ liáu  
    fire  cart  go  already

Rather than dismissing these phonological and syntactic patterns as erroneous, Hassan (1969) introduces the term ‘Tamil Bazaar Malay’ (Bahasa Malayu Pasar Tamil) and lists a number of characteristics. While the noun phrase-initial position of the demonstratives – as seen in example (7) – is quite common cross-linguistically and could reflect interference from Tamil, Hokkien or other languages, the clause-final position of the verb in Tamil Bazaar Malay specifically points to Indian influence. Hassan (1969: 212) provides the following examples:

(10) Ahmad pokok panjat  
    Ahmad tree  climb  
    ‘Ahmad climbs a tree.’  
    (standard Malay: Ahmad mamanjat pokok.)

(11) Itu budak bola sepak  
    DEM  kid  ball  kick  
    ‘That kid kicks a ball.’  
    (standard Malay: Budak itu manyepak bola.)

In addition to clause-final verbs, Tamil and many other Indian languages also display postpositions instead of prepositions. This impacts on the Malay they use, as Mohamed (2006: 13) demonstrates in an isolated example:

(12) Saya rumah misjid bolakang juga ada  
    1sg  house  mosque  behind  also  is  
    ‘I also have a house behind the mosque.’  
    (standard Malay: Saya juga ada rumah di bolakang masjid.)

The basic SOV word order is common in Indian languages of various families, including Dravidian and Indo-Aryan.
Other characteristics of Tamil Bazaar Malay include a set of distinct personal pronouns and the use of the particle *punya* as a possessive marker. These features are shared with several other Malay contact varieties and will be addressed in more detail in the next section. Hassan (1969: 214) provides the following examples in Tamil Bazaar Malay:

(13) Saya punya rumah boṣar punya anjing ada
1sg pos house big pos dog is
‘There is a big dog in my house.’

(standard Malay: *Di rumah saya ada səekor anjing boṣar.*)

(14) Saya punya rumah puteh (ada)
1sg pos house white (is)
‘My house is white.’

(standard Malay: *Rumah saya bərwarna putih.*)

In a MA thesis on Malay spoken by Tamils, Kader (1971: 446, quoted in Mohamed 2006: 31) provides some additional examples displaying the abovementioned features:

(15) Ittu jam worang tarak banyak
dem hour people neg many
‘Not many people were around at that hour.’

(16) Ittu ujan tara brənti, sampe sattu ari po wora
dem rain neg stop so.that one day part people

*It was raining non-stop, so that nobody showed up for the entire day [and] we had nothing to do.*

The above examples reflect difficulties in terms of transcription. For example, the negative marker *tara* occurs alongside *tara* and *worang* ‘people’ is found alongside *wora*. The forms *<wora>* and *<po>* presumably represent */worã/ (orang) and */põ/ (the particle *pun*). A more systematic phonological analysis of Malay spoken by Tamils, which also addresses the conditions of intervocalic consonant gemination, may help us make sense of these inconsistencies.
5. BAZAAR MALAY AND OTHER CONTACT VARIETIES

As mentioned in the previous section, some of the characteristics of colloquial Malay spoken by Tamils are attested more widely, especially in what is known as “Bazaar Malay” – the container term for Malay contact varieties not spoken as a mother tongue. The Malaysian Tamils presumably learnt this basilectal variety for out-group communication. Studies on West Malaysian Bazaar Malay are not well-distributed beyond a local level and often remain unpublished (Hassan 1969: 210 fn. 3). On the Singaporean variety, Daw (2005) offers the most complete description. Across West Malaysia, Bazaar Malay is spoken by Tamils and Chinese as a contact language. Its personal pronouns (see Table 4), resemble those of other “pidgin-derived Malay varieties”.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>gua</td>
<td>kita (orang)</td>
</tr>
<tr>
<td>2</td>
<td>lu</td>
<td>lu orang</td>
</tr>
<tr>
<td>3</td>
<td>dia</td>
<td>dia orang</td>
</tr>
</tbody>
</table>

Table 4. Bazaar Malay pronominal paradigm (Hassan 1969: 216).

The Bazaar Malay personal pronouns reflect Chinese influence. The 1\textsc{sg} goes back to Hokkien góa (我), whereas the 2\textsc{sg} reflects dialectical Hokkien lú (汝) in the same meaning. The use of orang as a plural marker corresponds to lâng (人 or 傢) in some Hokkien dialects, including in the original meaning of ‘person; people’. This usage is also found in Chitty Malay (Mohamed 2006), Baba Malay (Gwee 2006), Sri Lanka Malay (Nordhoff 2009), Cocos Malay (Adelaar 1996) and several Eastern Indonesian Malay varieties (Paauw 2008). Sri Lanka Malay, along with other Malay varieties, also displays other loanwords from Hokkien, including lō:teng ‘storey, floor’ and kuːwe ‘breakfast’ (Paauw 2004: 45). Other tentative Chinese loans attested as far as Sri Lanka include bangsat ‘bedbug’ and bopeng ‘pock-marked’ (compare Saldin 1993). This implies that the Chinese played some role in the making of a vehicular Malay trade variety. On a grammatical level, this is corroborated by the use of punya as a possessive marker, which is attested in roughly the same pidgin-derived Malay varieties as the above-mentioned Chinese-derived personal pronouns. Pakir (1986: 141-162) demonstrates that the use of mia in Baba Malay – which goes back to punya – largely mirrors the Hokkien syntax:

---

21 In a paper on the language history of Malay, Adelaar and Prentice (1996) distinguish “literary Malay varieties”, “pidgin-derived varieties”, and “Malayic vernaculars”. This distinction roughly corresponds to ‘written Malay’ (Bahasa Melayu Tulisan), ‘regional Malay’ (Bahasa Melayu Daerah) and ‘spoken Malay’ (Bahasa Melayu Lisan), and ‘Bazaar Malay’ (Bahasa Melayu Pasar) proposed by Hassan (1969).

22 Reflecting lâu-téng (樓頂) ‘upper storey, upstairs’ and kóe (粿) ‘cakes’.

23 Presumably from Hokkien bák-sat (木虱) and mō-pang (麻斑).
However, the use of the possessive marker *punya* is by no means limited to Chinese and Chinese-influenced varieties. In a paper on Sri Lanka Malay, Jayasuriya (2002: 49) provides an example in which Sri Lanka Malay, colloquial Sinhala and Sri Lanka Tamil are syntactically identical (the spelling is mine):

(20) Sri: *Laŋka:* =pe te: da:won
    Sri: *Laŋka:* =ve te: ko:lo (Sinhala)
    Sri: *Laŋka:* =qa te:le tu:l (Tamil)
    Sri.Lanka =POS tea leaves
    ‘Sri Lankan tea leaves’

Of the substandard Malay varieties in contact with Tamil, Sri Lanka Malay has received the most academic attention. This variety has an estimated 46000 speakers (Jayasuriya 2002) and its survival is threatened by Sinhala, Sri Lanka’s national language. This Malay variety is spoken by the descendants of soldiers, convicts, slaves and exiles who came to the country from the mid-seventeenth-century onwards, both under Dutch and British rule. These “Malays” came from various places, but the mainstream dialect bears a strong resemblance to East Indonesian Malay, while also displaying some Jakartan influence (Adelaar 1991; Paauw 2004). In-depth analyses of Sri Lanka Malay and its origins are of recent date (Nordhoff 2009, 2014; Slomanson 2011). It has often been argued that Sri Lanka’s Malay population was in close contact with other Muslim communities. Muslims in Sri Lanka constitute a rather diverse demographic
segment, including Malays, groups from Northwest India, and the so-called “Moors”. The term Moor historically referred to Muslims in general and is not regarded as particularly derogatory in a Sri Lankan context. It specifically denotes Tamil-speaking Muslims, who generally do not consider themselves Tamils. Sri Lankan Malays, too, see them as distinct communities; they call the former Keling and the latter Mulbar (Saldin 1993).

Aside from a rather brief description (Hussein 2007: 40-48), systematic and well-distributed linguistic research on the Tamil variety of the Moors – known as Shonam or Sonam and traditionally written in Arwi script – is lacking to this day. There appear to be at least two distinct dialects, one spoken by the inland ‘Sri Lankan Moors’ or Sōnahar and one by the coastal ‘Indian Moors’ or Sammankārar (Nuhman 2007: 25; Hussein 2007: 473). Most scholars of Sri Lanka Malay believe that the language received significant grammatical influence from Shonam. A counterview has been proposed by Ansaldo (2008), who takes issue with what he calls the “Tamil bias” and contends that the rather atypical grammatical features of Sri Lanka Malay may equally well reflect Sinhala influence. While descriptions of Sri Lanka Malay occasionally contain elicited Shonam data (Jayasuriya 2002; Slomanson 2011), this variety deserves a systematic description in its own right if we are to advance the discussion of Sri Lanka Malay origins (Ansaldo 2014: 383-384).

A number of typological features of Sri Lanka Malay can be attributed to Tamil (or Sinhala) influence. These include consonant gemination, the presence of long vowels, SOV word order, postpositions, adjectives preceding nouns, and suffixed conditions (Adelaar 1991; Jayasuriya 2002; Paauw 2004). As we have seen in the previous section, several of these features also occur in Bazaar Malay used by Tamil speakers. Other characteristics of Sri Lanka Malay, such as the position of the demonstratives and the use of a possessive marker, are quite common cross-linguistically – in particular in contact languages – and are therefore poor markers of contact-induced borrowing (compare Ansaldo 2008). Likewise, the omission of /h/ in Sri Lanka Malay (see Saldin 1993: 1001), which could point towards Tamil influence, presumably already took place in the Malay World, where it is quite common across dialects.

As mentioned previously, Sri Lanka Malay bears the greatest typological resemblance to the pidgin-derived varieties of East Indonesia. Paauw (2008) highlights a set of features shared by the varieties of Manado, North Maluku, Ambon, Banda, Kupang, Larantuka and Papua, which he argues go back to an historical “Eastern Indonesian Trade Malay”. In summary, these include: the monophthongization of /au/ to /o/ and /ai/ to /e/, the loss of /h/, the loss of the glottal stop at the end of a syllable, the lowering of /i/ to /e/ and /u/ to /o/ in closed final syllables, the use of punya or a derived form

---

24 These include the so-called Memon from the Sindh region and the aforementioned Khoja and Bohra, who migrated to Sri Lanka in colonial times (Hussainmiya 1990).
25 Shonam corresponds to standard Tamil Cōgam (சீடம்), a term denoting Muslims or (other) foreigners.
26 Corresponding to standard Tamil Cōyakar (சூயகாராள்) and Cammāṅkārar (சம்மங்காராள்) ‘sampan-men’, the latter having been borrowed into Sinhala as Hambankārayā.
as a possessive marker, the shortening of the demonstratives ini ‘this’ and itu ‘that’ for discourse strategies, plural pronouns formed with orang, and tara as a negative particle. Interestingly, several of these features are also found in Bazaar Malay, Chitty Malay, Baba Malay, Sri Lanka Malay, Cocos Malay, and colloquial Indonesian (compare Adelaar 2005). Regardless of whether we prefer the term “pidgin-derived Malay varieties”, “contact varieties”, “trade Malay” or “vehicular Malay”, it is important to keep in mind that varieties displaying this largely shared set of features are geographically attested throughout the Malay speech area – often in a diglossic continuum with standard Malay.

In addition to these shared grammatical features, Malay contact varieties – including colloquial Indonesian – exhibit a set of lexical discrepancies from the acrolectal, inherited varieties of Sumatra and West Malaysia, which stood at the cradle of standardized Malay. In other words, the vocabulary shared by pidgin-derived Malay varieties across a broad geographical range indicates a common origin. Some examples are given in Table 5.

<table>
<thead>
<tr>
<th>Baba Malay (Gwee 2006)</th>
<th>Sri Lanka Malay (Paauw 2004)</th>
<th>Colloquial Indonesian</th>
<th>Inherited Sumatran/West Malaysian Malay</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>bikin</td>
<td>bikin</td>
<td>buat</td>
<td>buat</td>
<td>‘to do, to make’</td>
</tr>
<tr>
<td>bilang</td>
<td>bi:lang</td>
<td>bilang</td>
<td>cakap</td>
<td>‘to say’</td>
</tr>
<tr>
<td>buntot</td>
<td>buntut</td>
<td>ekor, punggung</td>
<td></td>
<td>‘tail, rear part’</td>
</tr>
<tr>
<td>capek (‘lame, limping’)</td>
<td>ca:pe</td>
<td>capek</td>
<td>latih, panat</td>
<td>‘tired’</td>
</tr>
<tr>
<td>dia orang</td>
<td>dia orang</td>
<td>mareka</td>
<td></td>
<td>3sg</td>
</tr>
<tr>
<td>dukun</td>
<td>dukun</td>
<td>bidan, bomoh</td>
<td></td>
<td>‘medical healer, midwife’</td>
</tr>
<tr>
<td>gampang</td>
<td>gampang</td>
<td>mudah, sanang</td>
<td></td>
<td>‘easy’</td>
</tr>
<tr>
<td>gini</td>
<td>gi:ni</td>
<td>gini</td>
<td>bagini</td>
<td>‘like this’</td>
</tr>
<tr>
<td>gitu</td>
<td>gi:tu</td>
<td>gitu</td>
<td>bagitu</td>
<td>‘like that’</td>
</tr>
<tr>
<td>gua</td>
<td>go:</td>
<td>gua</td>
<td>saya</td>
<td>1sg</td>
</tr>
<tr>
<td>kasi</td>
<td>kasi</td>
<td>kasi</td>
<td>bari, bagi</td>
<td>‘to give’</td>
</tr>
<tr>
<td>kamaren (‘two days ago’)</td>
<td>kuma:reng</td>
<td>kamaren</td>
<td>samalam</td>
<td>‘yesterday’</td>
</tr>
<tr>
<td>kuping</td>
<td>kuping</td>
<td>kuping</td>
<td>talinga</td>
<td>‘ear’</td>
</tr>
<tr>
<td>lu</td>
<td>lu:</td>
<td>lu</td>
<td>kau, awak</td>
<td>2sg</td>
</tr>
<tr>
<td>lu orang</td>
<td>lorang</td>
<td>lu orang</td>
<td>kalian</td>
<td>2pl</td>
</tr>
<tr>
<td>pantat</td>
<td>pantat</td>
<td>pantat</td>
<td>punggung</td>
<td>‘buttocks’</td>
</tr>
</tbody>
</table>

27 Borrowed from Tamil cappai (சப்பை) ‘weak, lean’.
Table 5. Some lexical similarities in Malay contact varieties.

Sri Lankan Malay has also adopted some Tamil words at the cost of their Malay equivalents (Table 6).


More Tamil loans in Sri Lanka Malay are given in Scott Paauw’s MA thesis on the lexical origins of Sri Lanka Malay (2004). Some of the words listed in this study appear to display phonological innovations specific to Shonam (compare Hussein 2007), providing some further support to the hypothesis that the latter had influenced the former to a considerable extent. This again underlines the need for an accurate description of this variety if the discussion of Sri Lanka Malay origins is to be taken any further. A modest number of examples are given in Table 7.

Table 7. Shonam loans in Sri Lanka Malay.

28 Borrowed from Tamil tamin (டமன்) ‘male relative or friend’.
Other Tamil loans in Sri Lanka Malay appear to go back to another type of colloquial Tamil, whose precise dialectical origins remain obscure. Some examples taken from Paauw (2004) are listed in Table 8; the envisioned dialectical Tamil etyma are mine.

<table>
<thead>
<tr>
<th>Sri Lanka Malay</th>
<th>dialectical Tamil</th>
<th>literary Tamil</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>e: pong</td>
<td>*e:ppō</td>
<td>ēppam (இலும)</td>
<td>‘burp, hiccup’</td>
</tr>
<tr>
<td>konyong</td>
<td>*koŋŋō</td>
<td>koŋjam (இண்டி)</td>
<td>‘some, little, few’</td>
</tr>
<tr>
<td>ku: re</td>
<td>*ku:re</td>
<td>kūrai (சுறு)</td>
<td>‘roof’</td>
</tr>
<tr>
<td>oː de</td>
<td>*oːde</td>
<td>ōdai (ஆள)</td>
<td>‘canal’</td>
</tr>
<tr>
<td>poː n</td>
<td>*poŋŋ</td>
<td>peŋ (ஆளது)</td>
<td>‘bride’</td>
</tr>
<tr>
<td>reː te</td>
<td>*reːte</td>
<td>iraṭṭai (இறைத்து)</td>
<td>‘twins’</td>
</tr>
</tbody>
</table>


For the sake of completeness, it should be added that Sri Lanka Malay also exhibits some lexical influence from Sinhala, although the role of the latter as a lexifier is more modest than that of Tamil. Some examples from Paauw (2004) are given in Table 9; the tentative Sinhala precursors are mine.

<table>
<thead>
<tr>
<th>Sri Lanka Malay</th>
<th>Sinhala</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>koː ci</td>
<td>kōcci</td>
<td>‘train’</td>
</tr>
<tr>
<td>maheteya</td>
<td>mahattayā</td>
<td>‘Sir’</td>
</tr>
<tr>
<td>nari: ya</td>
<td>nariyā</td>
<td>‘fox’</td>
</tr>
<tr>
<td>poː re</td>
<td>pōra</td>
<td>‘manure’</td>
</tr>
<tr>
<td>pus</td>
<td>pus</td>
<td>‘mould on food’</td>
</tr>
<tr>
<td>rastiyadu (jaː nj)</td>
<td>rastiyādu</td>
<td>‘to roam without a purpose’</td>
</tr>
<tr>
<td>(ruː ma) oː la</td>
<td>ōla</td>
<td>‘cadjan hut’</td>
</tr>
<tr>
<td>siwura</td>
<td>siwura</td>
<td>‘robe’</td>
</tr>
<tr>
<td>tape</td>
<td>tāppe</td>
<td>‘bund, retaining wall’</td>
</tr>
<tr>
<td>teː ro</td>
<td>tera</td>
<td>‘Buddhist monk’</td>
</tr>
</tbody>
</table>

6. Lexical Traces of Mixed Languages

This section highlights a number of poorly described and quickly disappearing Malay varieties influenced by Tamil, whose documentation is typically restricted to small glossaries. The available information is therefore largely lexical. The aforementioned variety used by the Chitty, also known as Malaccan Creole Malay, has around 200 remaining speakers in Malacca, while an even smaller group have migrated to Singapore in the early twentieth century (Dhoraisingam 2006). Some brief notes on their language are given in Moorthy (1997). The most complete analysis of this Malay variety is a lexico-phonological description by Mohamed (2006). Phonologically as well as syntactically, Chitty Malay appears to be quite similar to Baba Malay. In all likelihood, both varieties developed out of a stabilized form of Bazaar Malay.

In terms of phonology, we find several more characteristics of pidgin-derived Malay varieties, including the monophthongization of /au/ to /o/ and /ai/ to /e/ and the deletion of /h/. The glottal stop, however, is retained in Chitty Malay. As in Baba Malay, a word-final glottal stop – written as <k> – is added to a limited number of words (Mohamed 2006: 93):

(21) Addition of word-final glottal stop in Chitty Malay:

\[
\begin{align*}
\text{bawa} & \rightarrow \text{bawak} \quad \text{‘to bring’} \\
\text{cari} & \rightarrow \text{carik} \quad \text{‘to search’} \\
\text{garu} & \rightarrow \text{garok} \quad \text{‘to scrape’} \\
\text{nasi} & \rightarrow \text{nasik} \quad \text{‘boiled rice’}
\end{align*}
\]

Chitty Malay also exhibits the assimilation of consonant cluster /mb/ to /m/ in intervocalic position. This phonological innovation is also attested, for instance, in Baba Malay, Kelantan Malay and subdialects in Kedah and Sarawak (Mohamed 2006: 92), but not among (other) pidgin-derived Malay varieties.

In terms of lexicon, Chitty Malay is predominantly Malay. Some Tamil vocabulary occurs in the domains of family members, religious terms, cultural items, traditional clothes, food, and wedding-related terms (Mohamed 2006: 124-127). A small selection of these loanwords is presented in Table 10.

<table>
<thead>
<tr>
<th>Chitty Malay (Malacca)</th>
<th>Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>aneng</td>
<td>அண்ணை</td>
<td>‘elder brother’</td>
</tr>
<tr>
<td>arjanai</td>
<td>அர்ஜானை</td>
<td>‘a religious ritual’</td>
</tr>
<tr>
<td>besti</td>
<td>வெஸ்டி</td>
<td>‘white man’s cloth’</td>
</tr>
<tr>
<td>kolem</td>
<td>கோலம்</td>
<td>‘ornamental figures drawn on floor with rice flour’</td>
</tr>
<tr>
<td>pandaram</td>
<td>பாண்டரம்</td>
<td>‘assistant temple priest’</td>
</tr>
</tbody>
</table>
Upon comparing the short glossary of Singaporean Chitty Malay given in Dhoraisingam (2006: 94), we see some minor discrepancies in pronunciation (Table 11), possibly due to infrequent usage or dialectical differences.

<table>
<thead>
<tr>
<th>Chitty Malay (Malacca)</th>
<th>Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>parpu</td>
<td>paruppu (පොපු)</td>
<td>‘pigeon peas’</td>
</tr>
<tr>
<td>patriège</td>
<td>pattirigai (பத்திரிக்கை)</td>
<td>‘invitation card for a wedding’</td>
</tr>
<tr>
<td>sudem</td>
<td>sūḍāṇ (சுடான்)</td>
<td>‘camphor’</td>
</tr>
<tr>
<td>talpa</td>
<td>talaippā (தலைப்பா)</td>
<td>‘turban’</td>
</tr>
</tbody>
</table>

Table 10. Some Tamil loans in Malaccan Chitty Malay.

A comparable mixed variety must have existed in Penang. This commercial entrepôt, historically belonging to the Kedah Sultanate, was frequented by South Indian communities from pre-colonial times. Since its acquisition in 1786 by the East India Company, the British encouraged foreign settlement, including from India and China. In a description of this dialect as it was used in the early twentieth century, Hamilton (1922: 57) contends that Penang Malay “is really the Malay of Kedah altered slightly to suit the needs of a cosmopolitan town population with a large element of Southern Indians from the Madras Presidency”. The author lists various lexical items, many of which borrowed from Tamil and Hindustani, that make this dialect stand out among other Peninsular Malay varieties. When I checked these words with Penang Malay speakers in 2014, many of them were only recognized by people older than 40, who associated them with the speech of their parents and grandparents. Penang Malay has recently converged with a more mainstream type of colloquial Malay found, with minor regional differences, across West Malaysia. Table 12 lists the Penang Malay words given in Hamilton (1922) that go back to Tamil.

<table>
<thead>
<tr>
<th>Chitty Malay (Singapore)</th>
<th>Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>aneng</td>
<td>anṇṇa (அண் ணன்)</td>
<td>‘elder brother’</td>
</tr>
<tr>
<td>arshaneh</td>
<td>arccaṇai (அர்சைனை)</td>
<td>‘a religious ritual’</td>
</tr>
<tr>
<td>kaboleh</td>
<td>kuvaḷai (குவாளை)</td>
<td>‘drinking vessel’</td>
</tr>
<tr>
<td>pandarom</td>
<td>paṇḍāram (பண்டாரம்)</td>
<td>‘assistant temple priest’</td>
</tr>
<tr>
<td>prasadom</td>
<td>pirāsādam (பிராசாதம்)</td>
<td>‘food offered in temple’</td>
</tr>
<tr>
<td>talpah</td>
<td>talaippā (தலைப்பா)</td>
<td>‘turban’</td>
</tr>
</tbody>
</table>

Table 11. Tamil loans in Singaporean Chitty Malay.
### Table 12. Tamil loans in (old) Penang Malay.

<table>
<thead>
<tr>
<th>Old Penang Malay</th>
<th>Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>candi</td>
<td>ᵰ싼்தில்</td>
<td>‘stubborn of a horse’</td>
</tr>
<tr>
<td>macan</td>
<td>மாண்டனி</td>
<td>‘the husband of an elder sister’</td>
</tr>
<tr>
<td>maini</td>
<td>மன்னி</td>
<td>‘the wife of an elder brother’</td>
</tr>
<tr>
<td>mambu</td>
<td>வெம்பு</td>
<td>‘the neem tree’</td>
</tr>
<tr>
<td>mami</td>
<td>மாமி</td>
<td>‘aunt’</td>
</tr>
<tr>
<td>mandom</td>
<td>மான்ம்</td>
<td>‘worthless, a broken down horse’</td>
</tr>
<tr>
<td>parli</td>
<td>புராலி</td>
<td>‘to tease, to deceive’</td>
</tr>
<tr>
<td>pili</td>
<td>பிளி</td>
<td>‘a water tap’</td>
</tr>
<tr>
<td>ponen</td>
<td>பெண்ணன்</td>
<td>‘impotent’</td>
</tr>
<tr>
<td>poni</td>
<td>போணி</td>
<td>‘a small, tin vessel’</td>
</tr>
<tr>
<td>ponu</td>
<td>பேணி</td>
<td>‘a bride’</td>
</tr>
<tr>
<td>sule</td>
<td>சுலை</td>
<td>‘a rheumatic swelling in the joints’</td>
</tr>
<tr>
<td>tairu</td>
<td>தாயிர்</td>
<td>‘curds’</td>
</tr>
</tbody>
</table>

In addition, some words in Penang Malay can be attributed to colloquial Malayalam, a language closely related to Tamil. These include *pataras* ‘pride, arrogance’ from *patrās* (பாறத்தாஸ்) and *pokri* ‘a profligate, a blackguard’ from *pōkkiri* (പൊപ്ക்கிரி), as well as the generic Peninsular Malaysian term *tandas* ‘toilet’ from *taṇḍās* (தாண்டாஸ்). \(^{29}\) I have looked at pre-modern Malayalam borrowings into Maritime Southeast Asia in another paper (Hoogervorst in press a), but this topic remains underexplored.

Studies on Malay language contact rarely take into account colloquial speech or dialects, neither of the donor nor the recipient language. We may mention in passing the neighbouring dialect of Kedah, which exhibits the loanwords *kambi* ‘plain metal earring’ (Tamil: *kambi*; கம்பி ‘kind or earring’) and *kawar* ‘thief; trespasser’ (Tamil: *kavar*; கவர் ‘to steal’) (Asmah 1966). A systematic study will almost certainly reveal more examples. On a related note, I would argue that lexicographic practices obscure the actual influence of Tamil on generic Malaysian Malay, as numerous loanwords known to and used by its speech community are omitted in most dictionaries. Table 13 lists some examples.

\(^{29}\) I am indebted to Abdur Rahoof Ottathingal and Mahmood Kooria for bringing this to my attention.
In addition, several Malaysian slang words have escaped the attention of linguists. This is a largely unexplored field requiring knowledge of both Malay and Tamil slang. Elsewhere I call attention to the Malay slang word *rendek* ‘to be together with someone’, which goes back to colloquial Tamil *reṇḍu* (*இரண்டு*) ‘two’ (Hoogervorst in press b). Other examples that I have collected during several brief periods of fieldwork in and around Kuala Lumpur are listed in Table 14.

<table>
<thead>
<tr>
<th>Malaysian Malay</th>
<th>Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>dipawali (~ deepavali)</td>
<td>dipāvali (<em>திப்பாவலி</em>)</td>
<td>‘name of a festival’</td>
</tr>
<tr>
<td>kawadi (~ kavadi)</td>
<td>kāvaḍi (<em>கவாடி</em>)</td>
<td>‘a decorated pole carried on shoulders with offerings’</td>
</tr>
<tr>
<td>taipusam (~ thaipusam)</td>
<td>taippūsam (<em>தீப்புசம்</em>)</td>
<td>‘name of a festival’</td>
</tr>
<tr>
<td>tose</td>
<td>tōsai (<em>தோசை</em>)</td>
<td>‘kind of rice-cake’</td>
</tr>
<tr>
<td>toti</td>
<td>tōṭṭi (<em>தோட்டி</em>)</td>
<td>‘a menial servant’</td>
</tr>
<tr>
<td>umapodi</td>
<td>ōmappodi (<em>ॐம்போடி</em>)</td>
<td>‘kind of confectionary’</td>
</tr>
</tbody>
</table>

Table 13. Tamil loans in Malaysian Malay.

Table 14. Malaysian slang words borrowed from Tamil slang.

<table>
<thead>
<tr>
<th>Malaysian slang</th>
<th>Tamil slang</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>aney</td>
<td>anṇē (<em>அண்ணை</em>)</td>
<td>‘employee of mamak restaurant’</td>
</tr>
<tr>
<td>ayoyo</td>
<td>aiyayō (<em>ஏயோய்</em>)</td>
<td>INTJ</td>
</tr>
<tr>
<td>maca</td>
<td>maccā (<em>மச்சா</em>)</td>
<td>‘buddy’</td>
</tr>
<tr>
<td>manjen</td>
<td>maṇjan (<em>மஞ்சன்</em>)</td>
<td>‘Chinese man (derogatory)’</td>
</tr>
<tr>
<td>nandrek</td>
<td>naṇṛi (<em>நஞ்சியிரி</em>)</td>
<td>‘Thank you!’</td>
</tr>
<tr>
<td>pondan</td>
<td>peṇḍan (<em>பெஞ்சன்</em>)</td>
<td>‘effeminate man’</td>
</tr>
<tr>
<td>porah ~ podah</td>
<td>pōḍa (<em>போடா</em>)</td>
<td>‘Get lost!’</td>
</tr>
<tr>
<td>pundek</td>
<td>puṇḍai (<em>புண்டை</em>)</td>
<td>‘cunt’</td>
</tr>
<tr>
<td>tanggaci</td>
<td>taṅgaicci (<em>தங்கைசி</em>)</td>
<td>‘girl’</td>
</tr>
<tr>
<td>yilek</td>
<td>illai (<em>இலை</em>)</td>
<td>‘absent’</td>
</tr>
</tbody>
</table>

The word final <u> is pronounced as a high central rounded vowel /ʉ/ in spoken Tamil. The colloquial pronunciation of this form, hence, is /reṇḍu/. Interestingly, the same word is attested in the gold traders’ slang of Malang (Table 16).

Original meaning: ‘elder brother’.

Original meaning: ‘no’.

31

32
Kojak form | Tamil | Meaning
--- | --- | ---
aambille | ānbiḷḷai (அந்தில்லை) | ‘husband’
kaliyaanam | kalyāṇam (கல்யாணம்) | ‘marriage’
koobdu | kūppiḍu (கூப்பிடு) | ‘to invite’
maliyu | malivu (மலிவு) | ‘cheap’
nombu | nōbu (நூபு) | ‘to fast’
patche | paccai (பச்சை) | ‘virgin’
pille ~ polle | pillai (பிள்ளை) | ‘children’
pombile | peṇbiḷḷai (பெண்மில்லை) | ‘wife’
salli | salli (சற்றை) | ‘money’
satte | satṭai (சட்டை) | ‘shirt’
soro | sōru (சொறு) | ‘rice’
thatpan | tahappan (தகப்பன்) | ‘father’
vange | vāngu (வாங்கு) | ‘to buy’

Table 15. Tamil loans used by the Kojak community in Semarang.

Indonesia exhibits a slightly different situation. With the exception of North Sumatra (Mani 1993a), Tamil communities were generally too small to remain independent and eventually assimilated into the mainstream. In Indonesia, the most common Malay term to denote Indian Muslims was Koja ~ Kojah ~ Khoja, from the aforementioned Ḵẖojā and ultimately from Persian Ḵẖwāja (خوaja) ‘a man of distinction’. The word is first documented in the late fourteenth century in the Tanjung Tanah manuscript, which was written by a certain Kuja Ali (Mahdi 2015). Javanese exhibits the related word Koja ‘merchant (usually Muslim Indian)’. Historically, several harbours on Java’s north coast had a Pakojan; a quarter where the Koja resided. While the term Ḵẖojā normally implies a northwest Indian origin, the eponymous community in the north Javanese city of Semarang traces their ancestors to Tamil-speaking Marakkāyar, with a small minority of Māppiḷa and Gujarati Ḵẖojā (Mani 1993b: 126). At one point, they may have spoken a mixed language akin to Chitty Malay. Mani (1993b: 126-127) briefly addresses the “secret” language of this mercantile community, which at the time of his research was only used by speakers older than 40. Some examples are given in Table 15 above (spelling of the Kojak forms as in original).

Another Tamil-influenced in-group language has been documented in Malang. In the 1950s and 1960s, the gold trade in this East Javanese city was dominated by Indians. Even when local people took over in the 1970s, their cryptolect still consisted of Tamil words (Pujileksono and Kartono 2007: 23-24; Hoogervorst 2014: 114-115). Some examples are given in Table 16.

33 Original meaning: ‘green’.
The opposite phenomenon has also been documented. In colonial times, the cryptolects of the Tamil-speaking Paṟaiyar community in Jaffna (Lewis 1890) and the merchants of the Coromandel Coast (Pandit 1894) have been identified as partly inspired by Malay or a closely related language (Kern 1894; Hoogervorst 2013: 17-18, 27-28). Unsurprisingly, multilingual traders would have communicated in a language they picked up far away to keep their transactions back home a secret.

7. Malayicized Tamil?
As the previous three sections have explored Malay varieties influenced by Tamil, this section discusses localized Tamil varieties of Maritime Southeast Asia. As has been discussed in Section 2, the historical practice among Indian merchants to marry into local, Malay-speaking families gave rise to a wealthy, influential generation of bilinguals. The linguistic competence of such hybrid communities, however, would typically decrease as they assimilated into the mainstream. In all likelihood, a stabilized, “Malayicized” variety of Tamil could only develop in recent times, with substantial numbers of women to ensure intergenerational transmission. Whether we can indeed speak of such a variety remains open for discussion. It may be pointed out that Tamil has been used as a medium of education and local literary works in Malaysia and Singapore since the independence of both nations (see Willford 2006: 45-52). While the phrase “Malaysian Tamil” occasionally surfaces in the literature (compare Renganathan 2009), it is rarely qualified. In general, scholarship on any aspect of Malaysian Tamil remains limited and poorly distributed (see, for example, Fernandez 2008; Karunakaran and Krishnan 2013). A monograph-length description of colloquial Malaysian or Singaporean Tamil is still needed.
The available information is at times contradictory. For example, Venugopal (1996: 5) contends that Malaysian Tamil is relatively Anglicized. He points towards the English educated background of most Malaysian Tamils and gives some examples of lexical influence from English in Malaysian Tamil literature (Venugopal 1996: 285-286). Conversely, Ramasamy and Moses (2004: 58-59) hold that Indian Tamil contains a comparatively larger proportion of English, implying that Malaysian Tamil has remained closer to its roots. The authors further state that there is little dialectical difference in Malaysia, as opposed to India (and Sri Lanka). In addition, they call attention to a number of common Malay loanwords in Malaysian Tamil, as listed in Table 17. The fact that many Tamil-speakers learned Bazaar Malay would have accounted for this situation.34

<table>
<thead>
<tr>
<th>Malaysian Tamil</th>
<th>Malay</th>
<th>Standard Tamil</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:tta:ppu</td>
<td>atap</td>
<td>kūrai (குறை)</td>
<td>‘(thatched) roof’</td>
</tr>
<tr>
<td>campo:ru</td>
<td>campur</td>
<td>cēr (சேர்)</td>
<td>‘to mix’</td>
</tr>
<tr>
<td>kla:mbare</td>
<td>kalambu</td>
<td>kosuvalai (கோசுவலை)</td>
<td>‘mosquito net’</td>
</tr>
<tr>
<td>pasa:ru</td>
<td>pasar</td>
<td>candai (சாண்டை)</td>
<td>‘market’</td>
</tr>
<tr>
<td>saron</td>
<td>sarung</td>
<td>kaili (காலி)</td>
<td>‘sarong’</td>
</tr>
<tr>
<td>vakulu</td>
<td>bakul</td>
<td>kūḍai (குட்டை)</td>
<td>‘basket’</td>
</tr>
</tbody>
</table>

Table 17. Malay loans in Malaysian Tamil (Ramasamy and Moses 2004: 59).

In his treatise on Malaysian literature in Tamil, Venugopal (1996: 299) lists a number of words he considers to be characteristic of Tamil used in Malaysia. These include some Malay loanwords, such as itik ‘duck’, janji ‘promise’, kapala ‘head’, lalang ‘tall grass’, lampu ‘light’, ringgit ‘dollar’, rokok ‘cigarette’, samsu ‘illicit liquor’, salasai ‘solved’, sēnang ‘easy’, sapatu ‘shoe’ and tukang ‘carpenter’. The Tamil spelling and pronunciation of these words are not given, which makes the examples of limited use for linguistic purposes. The author also calls attention to a number of neologisms specific to Malaysian Tamil, such as āyakkottai (அய்கோட்டை) ‘child centre’, kākkā kadai (காக்கா கைடை) ‘Indian Muslim shop’ and pirattikkalam (பிராட்டிக்களம்) ‘the place where workers give their names for attendance’. More examples of Malaysian Tamil vocabulary are given in a pioneering lexical study of the Tamil dialects in lower Perak (Subbiah 1966). Some instances of Malay loanwords in Lower Perak Tamil are listed in Table 18.

---

34 Venugopal (1996: 301) provides examples of code-mixing and code-switching between Tamil and colloquial Malay in Malaysian Tamil literature, but does not present any first-hand data.
Balasubramaniam (1994) offers another isolated contribution to the study of Malaysian Tamil. In his exploration of the language used by tea pluckers and factory workers in the tea estates of the Cameron Highlands, he calls attention to a number of loanwords and hybrid constructions (Table 19).

<table>
<thead>
<tr>
<th>Cameron Highlands Tamil</th>
<th>Meaning</th>
<th>Malay element</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>boṅkusu kāmarā</td>
<td>‘section where packing is done’</td>
<td>bungkus</td>
<td>‘packet’</td>
</tr>
<tr>
<td>campalac-cūra</td>
<td>‘wage slip, salary slip’</td>
<td>surat</td>
<td>‘letter’</td>
</tr>
<tr>
<td>jāmāṅkoṭṭāy</td>
<td>‘latrine’</td>
<td>jamban</td>
<td>‘toilet’</td>
</tr>
<tr>
<td>kāmarā</td>
<td>‘room type structure; section of factory’</td>
<td>kamāra</td>
<td>‘chamber, cabin’</td>
</tr>
<tr>
<td>kappalā</td>
<td>‘attendant in the estate dispensary’</td>
<td>kāpala</td>
<td>‘head’</td>
</tr>
</tbody>
</table>
Unsurprisingly in the light of the strong connections of Southeast Asian Tamils to their ancestral land, some of these Malay loanwords gained currency back in India. Contemporary examples include *kittā* (கிட்டா) ‘rubber’ (Malay: *gatal*) and *sēvā* (சேசவா) ‘to rent’ (Malay: *sewa*). Elsewhere, I have called attention to other examples of Malay loanwords in Tamil (Hoogervorst 2013). The lexical items highlighted in the latter study were presumably transmitted before interethnic commerce in the Indian Ocean World had become a European-dominated enterprise. A selection is given in Table 20.

Table 19. Loanwords and hybrid constructions in Cameron Highlands Tamil (Balasubramaniam 1994).

<table>
<thead>
<tr>
<th>Tamil</th>
<th>Malay</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lampuk-kāsu</td>
<td>lampu</td>
<td>‘light money (money deducted from the workers’ pay)’</td>
</tr>
<tr>
<td>tīṭṭi</td>
<td>cuti</td>
<td>‘a day off’</td>
</tr>
<tr>
<td>tōmpukoṭṭāy</td>
<td>tong</td>
<td>‘shed for heating water in large drums for bathing’</td>
</tr>
</tbody>
</table>

Unsurprisingly in the light of the strong connections of Southeast Asian Tamils to their ancestral land, some of these Malay loanwords gained currency back in India. Contemporary examples include *kittā* (கிட்டா) ‘rubber’ (Malay: *gatal*) and *sēvā* (சேசவா) ‘to rent’ (Malay: *sewa*). Elsewhere, I have called attention to other examples of Malay loanwords in Tamil (Hoogervorst 2013). The lexical items highlighted in the latter study were presumably transmitted before interethnic commerce in the Indian Ocean World had become a European-dominated enterprise. A selection is given in Table 20.

Table 20. Malay loans in Tamil (Hoogervorst 2013).

<table>
<thead>
<tr>
<th>Tamil</th>
<th>Malay</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>jōṅgu (ஜோங்கு)</td>
<td>jung</td>
<td>‘sea-going ship’</td>
</tr>
<tr>
<td>kaiyāppudai (காயப்புடை)</td>
<td>kayu putih</td>
<td>‘cajuput tree’</td>
</tr>
<tr>
<td>kajaṅgu (கசங்கு)</td>
<td>kajang</td>
<td>‘mat protection against the rain’</td>
</tr>
<tr>
<td>kākkattuvāṅ (காக்கத்வாங்)</td>
<td>kakaktua</td>
<td>‘cockatoo’</td>
</tr>
<tr>
<td>kambīr (கம்பீர்)</td>
<td>gambir</td>
<td>‘a plant used for betel chewing’</td>
</tr>
<tr>
<td>kirisu (கிரிஸு)</td>
<td>karis</td>
<td>‘a kind of dagger’</td>
</tr>
<tr>
<td>sagu (சாகு)</td>
<td>sagu</td>
<td>‘sago’</td>
</tr>
<tr>
<td>sāmbal (சாம்பால்)</td>
<td>sambal</td>
<td>‘chili-based spicy sauce’</td>
</tr>
<tr>
<td>sambāṅ (சம்பாங்)</td>
<td>sampan</td>
<td>‘a type of boat’</td>
</tr>
<tr>
<td>sappaṅgi (சப்பாங்கி)</td>
<td>sapang</td>
<td>‘a type of fragrant wood’</td>
</tr>
<tr>
<td>taṅgāṅ (தங்காங்)</td>
<td>tangahan</td>
<td>‘half’</td>
</tr>
</tbody>
</table>

35 The form *sēvā* (சேசவா) is only in use among the Nāṭṭukkōṭṭai Chetties, a mercantile caste who often migrated to Southeast Asia (Subbiah 1966: 151).
8. Conclusions

In his treatise on Malaysian Tamil literature, Venugopal (1996: 285) contends that “[p]urity of language is not found in the writings of these novelists”. Diglossia and the existence of high-status literary languages have made expressions of the above type almost commonplace in both South and Southeast Asia. Yet notions of language purity obstruct, rather than stimulate, in-depth linguistic research. Academic attention to the colloquial and dialectical registers of Malay and Tamil remains remarkably sparse, although the situation is gradually improving for Malay. The traditional focus on “high language” has left the vernacular dimensions of language contact across the Bay of Bengal largely underexplored, leading to an imbalanced understanding of the cultural history of this part of the world. Indeed, while artists, scholars and scribes feature prominently in South and Southeast Asian historiographies, merchants, middlemen and labourers remain poorly documented. That the latter groups were vital to the introduction of products and ideas across geographical and ethno-linguistic boundaries becomes clear – among many other things – from the lexical and grammatical influence of Tamil and Hokkien dialects on Malay contact varieties, including Sri Lanka Malay. The role of Indian communities in these networks deserves closer academic attention in the future.

If we are to understand the wide range of activities underpinning cultural and linguistic contact in the Bay of Bengal and elsewhere, we need to focus on non-standardized languages. I have mentioned in passing the influence of Malayalam on the Peninsular Malay varieties. This topic merits a more thorough analysis than has been possible here. The same holds true for specific Tamil dialects or sociolects, especially those belonging to mercantile groups (both Hindu and Muslim). Fieldwork-based grammatical descriptions of Shonam, Malaysian Tamil and Malay spoken by Tamils, too, are lacking to this day. Local scholarship in this area is often of great interest, but poor distribution leaves much of it neglected in wider academic circles. Meanwhile, there is no reason to believe that similar mixed languages did not exist before colonial times. Varieties spoken by the hybrid communities highlighted in this paper must have been common historically. The present overview has only been able to scratch the surface of the linguistic processes and phenomena emerging at the crossroads of Malay and Tamil.

Two major problems remain in advancing this area of study. Foremost, much of the scattered scholarship on language contact at the interface of Malay and Tamil is done in a linguistically haphazard way. For phonologically complex languages such as Tamil, either an IPA representation or consistent transliteration with diacritics are prerequisites for up-to-standard data presentation. If the phonology of the languages under research is disregarded, painstakingly collected data will lose much of their value to serious linguistic research. The second problem has to do with post-colonial paradigms in academia. With some notable exceptions, humanities scholars based in South and Southeast Asia tend to treat Indonesian, Malaysian, Indian and Sri Lankan (language) history as teleological narratives, without considering events taking
place outside the present-day borders of the countries under research. Yet at the same time, a series of pivotal, transnational events left a deep impact on all regions involved. One prominent example is the emergence of the Malay printing press in Colombo and later in Singapore, which was spearheaded in both cities by culturally hybrid groups. More comparative work in the fields of historical linguistics, literature, manuscript studies, religion, and other disciplines is required to explore broader perspectives of language contact and etymology in some of the world’s most diverse and exciting language ecologies.

**ABBREVIATIONS USED**

1 : first person  
2 : second person  
3 : third person  
DEM : demonstrative  
GEN : genitive  
NEG : negation  
INTJ : interjection  
PART : particle  
PL : plural  
POS : possessive  
PST : past  
SG : singular  
SOV : subject-object-verb

**REFERENCES**

Aisyah, Siti. 2014. “Religious and commercial networks lexicalized; Interpreting a unique quadrilingual manuscript”. [Paper, Conference “Sailing the waves of convergence; Cultural links and continuities across the Bay of Bengal”, Leiden, 3 December.]

Aman, Norhaida; Viniti Vaish; Wendy Bohkorst-Heng; Aisha Jamaludeen; Durgadevi P.; Feng Ying Yi; Khoo Boon Suan; Mardiana Roslan; Pauline Appleyard; and Tan Teck Kiang. 2009. *The sociolinguistic survey of Singapore 2006*. Singapore: Centre for Research in Pedagogy and Practice, National Institute of Education.


Feener, R. Michael and Terenjit Sevea (eds). 2009. *Islamic connections; Muslim
societies in South and Southeast Asia. Singapore: Institute of Southeast Asian Studies.


Jones, Russell. 2007. Loan-words in Indonesian and Malay; Compiled by the Indonesian Etymological Project. Leiden: KITLV Press.


Ronkel, Ph.S. van. 1918 “Drawidische volksnamen op Sumatra”, *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 74: 263-266.


Ronkel, Ph.S. van. 1922. “A Tamil Malay manuscript”, *Journal of the Straits Branch of the Royal Asiatic Society* 85: 29-35.


Teo, Kok Seong. 2003. *The peranakan Chinese of Kelantan; A study of the culture,*