

## Documenting modern Sri Lanka Portuguese

Hugo C. Cardoso, Mahesh Radhakrishnan, Patrícia Costa & Rui Pereira  
*Universidade de Lisboa, Faculdade de Letras*

### Abstract

Sri Lanka Portuguese (SLP) is a Portuguese-lexified creole formed during Sri Lanka's Portuguese colonial period, which lasted from the early 16<sup>th</sup> century to the mid-17<sup>th</sup> century. The language withstood several political changes and became an important medium of communication for a portion of the island's population, but reached the late 20<sup>th</sup> century much reduced in its distribution and vitality, having essentially contracted to the Portuguese Burgher community of Eastern Sri Lanka. In the 1970s and 1980s, the language was the object of considerable research and documentation efforts, which were, however, curtailed by the Sri Lankan civil war. This chapter reports on the activities, challenges, and results of a recent documentation project developed in the post-war period and designed to create an appropriate and diverse record of modern SLP. The project is characterised by a highly multidisciplinary approach that combines linguistics and ethnomusicology, a strong focus on video recordings and open-access dissemination of materials through an online digital platform (*Endangered Languages Archive*), archival prospection to collect diachronic sources, a sociolinguistic component aimed at determining ethnolinguistic vitality with a view to delineating revitalisation strategies, and a strongly collaborative nature. Here, we describe the principal outputs of the documentation project, which, in addition to a digital corpus of transcribed and annotated materials representing modern manifestations of SLP and the oral/musical traditions of the Burgher community, also include the findings of the sociolinguistic survey, an orthographic proposal for the language, as well as the copies and transcriptions of hard-to-obtain historical sources on SLP (grammars, dictionaries, biblical translations, liturgical texts, and collections of songs).

**Keywords:** Sri Lanka Portuguese, language documentation, ethnomusicology, orthography development, sociolinguistic survey, archival sources

## 1. Introduction<sup>1</sup>

The implantation of the Portuguese language in Asia, which began in earnest in the early 16<sup>th</sup> century, and the many situations of contact that it induced left a long-lasting imprint on Asia's linguistic landscape. The products of this history of interaction were varied (see e.g. Cardoso 2016), one of the most evident being the formation of a string of Portuguese-lexified creoles that, at one point, stretched from India to East Timor (Tomás 1992; Cardoso, Hagemeyer & Alexandre 2015). Many historical varieties have ceased to be spoken over time, and the ones that persist – in India, Sri Lanka, Malaysia, and Macau - are currently endangered to different degrees.

Large parts of the island of Ceylon, modern-day Sri Lanka, were under Portuguese control from the 16<sup>th</sup> to the mid-17<sup>th</sup> century, and this created the conditions for the development and establishment of one such Portuguese-lexified creole, which we identify here as *Sri Lanka Portuguese* [henceforth SLP]. While SLP was arguably the Asian-Portuguese creole with the most robust implantation historically, its use and intergenerational transmission are contracting rapidly (Nordhoff 2013). This fact motivated the development of a documentation project, entitled “Documentation of Sri Lanka Portuguese” [henceforth DSLP], of which this paper provides an overview. DSLP ran from January 2017 to February 2019 and was hosted by the Centro de Linguística da Universidade de Lisboa (University of Lisbon Centre for Linguistics). Since it was funded by the Endangered Languages Documentation Programme (ELDP), it follows ELDP's well-established standards with respect to archiving formats, metadata, ethical requirements, and access. Several fieldwork trips were conducted, mostly in Sri Lanka's Eastern Province, between February and May 2017, in October and November 2017, and between March and May 2018.

The primary objective of the project was to document SLP in its current context, producing an online corpus of SLP speech, transcribed and annotated, to be deposited at the *Endangered Languages Archive* (ELAR).<sup>2</sup> Currently, the online corpus contains 19h of video recordings and 18h of audio materials, fully transcribed and partially annotated, in addition to a few supporting documents such as scans, photographs of written texts, and consent forms. The collection will continue to grow with materials which are currently being collected and processed. While a significant proportion of these materials consist of modern-day SLP speech recorded in very diverse settings, the

---

<sup>1</sup> The authors wish to thank the many members of the Portuguese Burgher community in Sri Lanka who collaborated with the documentation project, as well as several Sri Lankan institutions which also provided essential assistance: the Burgher Union (Batticaloa), the Sri Lanka Portuguese Burgher Foundation, the Burgher Welfare Association (Trincomalee), Burgher Folks, and the International Centre for Ethnic Studies (Colombo). We are also grateful to two anonymous reviewers for their suggestions. The research reported on in this paper was funded by the Endangered Languages Documentation Programme (grant: MDP0357) and the Fundação para a Ciência e a Tecnologia (grant: IF/01009/2012).

<sup>2</sup> *ELAR* is available online at: <http://elar.soas.ac.uk>. The DSLP deposit is accessible at: <https://elar.soas.ac.uk/Collection/MPI1035102>.

project also branches out towards other domains and different chronologies. In fact, this is an intrinsically multidisciplinary project that combines complementary linguistic, sociolinguistic, and ethnomusicological strands, and has an inbuilt component of archival research that aims to unearth early records of the language from the beginning of the 19<sup>th</sup> century onwards. This multifaceted approach, we argue, is advantageous as a way of enhancing the diversity of the collected materials, the societal appeal of the documentation, as well as the academic significance and potential of the resulting corpus. Given SLP's condition as a non-standardized, endangered language associated with a particular minority community within Sri Lanka, we will also highlight the features of the documentation project that have been designed to potentially feed into revitalization initiatives.

This paper describes the ongoing documentation of SLP, since the formal end of the DSLP project has not meant the end of documentation and archiving activities. We begin by providing some information about the language, its external history and current setting, as well as the status of its documentation (Section 2), and then describe the various initiatives developed within DSLP with a view to constituting the corpus and advancing knowledge of the language (Section 3). Finally, Section 4 focuses on steps taken with a view to supporting the revitalization of SLP and the possible impact of the DSLP project in that domain.

## **2. Background**

The presence of the Portuguese in Ceylon lasted for about one and a half centuries, between 1505 and 1658 (Pieris & Naish 1920; Flores 2004). Well into the 16<sup>th</sup> century, they achieved effective control over large portions of the island, in what constituted a rare example of actual territorial domination in their Eastern empire (Serrão 2014: 184; Biedermann 2018), until they were challenged by the Dutch and eventually ousted by the mid-17<sup>th</sup> century. The formation of the local Portuguese-lexified creole is therefore the product of this period, yet it remained a major language of Ceylon well after that, up to the 19<sup>th</sup> century. At the turn of that century, Ceylon Portuguese (i.e. SLP) was described by Sebastião Rodolfo Dalgado, a Goan priest and philologist who was around that time posted to Ceylon, as the most important and well-known overseas “dialect” of Portuguese in Asia, a language used in a wide range of social settings and with a vibrant set of oral traditions, and an important means of interethnic communication on the island (Dalgado 1900: xxii-xxiii). In the 19<sup>th</sup> century, during the British colonial period of Ceylon, there are plenty of references to the widespread use of “Ceylon Portuguese”. An early and significant such reference is that of the Wesleyan missionary Rev. John Callaway, who describes the situation of Ceylon Portuguese in the following terms:

With thousands of inhabitants, this language is the direct medium of intercourse; and the fact of its having for centuries more than maintained its ground, under circumstances which have been fatal to other tongues [...] is no contemptible evidence of its intrinsic worth. (Callaway 1820: ii)

By “circumstances which have been fatal to other tongues”, Callaway is probably referring to the tendency of disposing of language forms associated with a particular dominant group (as is so often the case with creoles lexically-based on European colonial languages) when political circumstances change. In the case of colonial Ceylon, that was not the case, with Ceylon Portuguese preserving its status as a *lingua franca* among certain tracts of the society well into the Dutch and British colonial periods. As an L1, however, it was restricted to particular groups: the *Burghers* and the community known as *Kaffirs*, or *Afro-Sri Lankans*.

In Sri Lanka, the communities that claim Eurasian ancestry are to this day known by the Dutch-derived label of *Burghers*, and traditionally divided into two sections: the “Portuguese Burghers” and the “Dutch Burghers”. The dichotomy is perhaps overstated, as there is plenty of evidence that the two groups share a common history and cultural traits (McGilvray 1982: 236–237) that include cuisine, family names, religion, and indeed also language. As a matter of fact, Ceylon Portuguese is known to have been the language of domestic use of both the Portuguese Burghers and the Dutch Burghers (Coelho 1881: 32; Smith 2013a: 111). In response to the inquiries of a reader about the use of Dutch among the Burghers, in 1952, the editor of the *Journal of the Dutch Burgher Union* states:

In the towns and strongholds which the Dutch took over the people spoke a form of the Portuguese language. The early Dutch settlers who came to live with, and among these people, were more or less compelled to learn this mixed Portuguese language as it was used for all ordinary and domestic purposes. Its liquid sounds and freedom from grammatical restraints helped the new comer to pick it up easily. Dutch children born in Ceylon learnt it from their nurses and used it as a home language with greater freedom than their mother-tongue. The result of all this was a compromise. While Dutch was the spoken and written word of polite society, a hybrid Portuguese in which a large number of Dutch words had found admission, was the language used in the office, the platform, the pulpit and in Dutch households. (JDBU 1952: 94)

In addition to the Burghers, L1 use of Ceylon Portuguese was also associated with the so-called *Kaffir* (Afro-Sri Lankan) community, a group with African heritage settled especially in the Puttalam and Trincomalee districts (Goonetilleke 1983; de Silva Jayasuriya 2007). To this day, those residing in the Puttalam region preserve a unique form of music and dance identified as *manja* or *manha*, for which they have gained wide recognition in Sri Lanka overall, in which songs are sung in SLP.

Historically, then, SLP was a vital and relatively widespread language in Sri Lanka, spoken as an L1 by Burghers and Kaffirs and as an L2 by other communities. More recently, however, the situation has changed dramatically. Over time, the Dutch Burghers abandoned SLP in favor of English, and the latest reports find that there are no longer any fluent speakers among the Kaffir population (Nordhoff 2013), who have shifted to Sinhala. Use of SLP has therefore receded to the Portuguese Burghers, and especially those residing in the Eastern Province, in the Trincomalee, Batticaloa, and Ampara districts. Recently, a preliminary prospection of the Jaffna Portuguese Burgher community done as part of the project has also identified a few SLP speakers there with different degrees of fluency.

Trincomalee and Batticaloa hold the largest concentrations of the Portuguese Burgher population, which is also present in nearby towns such as Eravur, Valachchenai, Kalmunai, and Akkaraipattu. For that reason, the recent documentation of SLP has been taking place primarily in the Eastern Province. However, even there, the Burghers constitute a small minority. Traditionally, the community in these areas has been bilingual in SLP and Tamil, the locally-dominant language and the one used for communication with neighboring communities and most other activities (including education, administration, and religious services). At this point, the community is shifting rapidly to Tamil as a home language, and less so to Sinhala and English. Determining the actual rate of transmission of SLP, as well as the sociolinguistic responses to the pressure of other languages, and the geographical distribution of language preservation and loss, are among the objectives of the documentation project (see §3.4.).

## 2.1. History of Sri Lanka Portuguese documentation

Given the importance of Ceylon Portuguese in colonial Ceylon, the language has caught the attention of observers since at least the 19<sup>th</sup> century, when several books in or about the language were published – for an overview, see Coelho (1881); Goonatilleka (1970: 136–138); Tomás (1992); Smith (2016). Some of these appear to have been secular initiatives – the first grammar (Berrenger 1811) is dedicated “To the English Gentlemen in the Civil and Military Service on Ceylon” –, but most had a religious motivation, connected in particular with the Wesleyan Mission in Ceylon (see Smith 2016). In the 19<sup>th</sup> century, missionaries such as John Callaway, William Buckley Fox, and Robert Newstead were responsible for putting out several dictionaries of Ceylon Portuguese and English and/or Sinhala (Callaway 1818; 1820; 1823; Fox 1819), but also translations of biblical (Newstead 1852) and liturgical texts (Newstead 1827; 1871). The Wesleyan Mission produced several other such works for which the name of the author or translator is not known, and, in the 1880s and the 1890s, even published a magazine in Ceylon Portuguese called *O Bruffador*.

Several of the other Ceylon Portuguese sources produced in the 19<sup>th</sup> century were ethnographic in nature, concerned especially with the documentation of oral traditions in that language – for an overview, see Cardoso, Hagemeyer & Alexandre (2015: 680–682). That is the case of an important manuscript compilation of songs obtained by the British officer Hugh Nevill and first edited by Jackson (1990), or an anonymous article about the traditions of the Kaffirs (Anon. 1895). The first scholarly treatments of SLP also date from the late 19<sup>th</sup> century, and include a study of an SLP reportative particle *diski* by William Goonetilleke (1892), but especially the unpublished draft of a larger article by Hugo Schuchardt, the scholar credited with inaugurating Creole Studies, now edited by de Silva Jayasuriya (1999).

The early 20<sup>th</sup> century saw some more production about SLP, beginning with a book by the Goan priest Sebastião Rodolfo Dalgado (1900) which collates several texts in the so-called “Indo-Portuguese dialect of Ceylon”,<sup>3</sup> as well as his own sermons delivered in that language, complemented with extensive linguistic analysis, a dictionary and a phrasebook. Further collections of oral traditions produced around this time include a series of articles published by Tavares de Mello between 1905 e 1914 – collected in Tavares de Mello (1998) – and an anonymous song collection (Anon. 1914).

After this period, there was a gap of several decades without any significant scholarly production on SLP.<sup>4</sup> In the 1970s, however, several scholars took a fresh interest in SLP and engaged in fieldwork-based documentation, including M.H. Goonatilleka (see e.g. 1983; 1985), Ian Smith (see e.g. Smith 1977; 1979a; 1979b; 1984), and K. David Jackson (see e.g. Jackson 1975; 1985; 1990; 1991). Their work brought the language and its traditions firmly into modern debates of language contact, creolization, anthropology, and colonial history. However, fieldwork research was abruptly interrupted by the start of the Sri Lankan Civil War, in the early 1980s.

## 2.2. Modern Sri Lanka Portuguese

Among the Burghers of Eastern Sri Lanka, the autoglossonym used to refer to SLP is either *baargarsu lingvááy*<sup>5</sup> ‘Burgher language’ or simply *purtugees* ‘Portuguese’, with some sharing the notion that SLP is a form of *vééy purtugees* ‘old Portuguese’ (as

---

<sup>3</sup> Traditionally, the term “Indo-Portuguese” referred to realities encompassing the whole of South Asia. Therefore, SLP is often counted among the wider category of the “Indo-Portuguese creoles”, despite the linguistic heterogeneity that characterizes this group of languages. In fact, the ISO 639-3 code currently used to identify SLP in various publications is identified as “Indo-Portuguese” and applies indistinctly to all Portuguese-lexified creoles of South Asia (India and Sri Lanka).

<sup>4</sup> De Vos (1950) is an exception, as it does record snippets of linguistic data. However, this article was published posthumously.

<sup>5</sup> Throughout this text, we use the SLP orthography developed within the project, except when quoting from other sources. For further information on orthography, see §3.2.

opposed to *noov purtugees* ‘new Portuguese’, reserved for European Portuguese). The term used in Tamil for both the Portuguese Burghers and their language, *parangi*, is considered slightly derogatory by members of the community (see Smith 2014: 111), especially in Jaffna. In our project, we use *Sri Lanka Portuguese*, in agreement with some of the previous researchers of the language such as Ian Smith, and do not feel the need to include the epithet *Creole* in the designation even though the language can very well be included in that category of contact languages.

Assessing the exact number of speakers has been one of the tasks of the documentation project. Prior to it, the available estimates, collated from censuses or the various Burgher associations and reported e.g. in Bandara (2012) and Nordhoff (2013), actually reported on the size of the Portuguese Burgher population (indicating somewhere between 4000 and 5000 Burghers in Batticaloa district, and about 170 families in Trincomalee) and not on the number of speakers among them. For further information on the demographics of SLP speakers and our sociolinguistic survey of the community in Eastern Sri Lanka, see §3.4. Traditionally, the community has been portrayed as relatively homogeneous: overwhelmingly Roman Catholic by religion, with males mostly employed in manual trades such as carpentry, smithing or mechanical work (which earned them the epithet of *Mechanics*, see McGilvray 1982), and females doing housework or tailoring. Currently, while these characteristics are still highly visible, the situation is much more complex. To begin with, occupations have diversified for both men and women, reflecting changes in the education of sections of the Burgher population, and there is increased mobility to other locations in Sri Lanka (including Colombo, Kandy, Polonnaruwa, Kurunegala, etc.) and abroad. In religious terms, there have been a few conversions to non-Roman Catholic Christian denominations and, occasionally, to other religions such as Hinduism or Islam, mainly as a result of a community-reported increase in the practice of exogamy.<sup>6</sup>

Ever since modern-day SLP data has been brought to light, chiefly as a result of the research developed in the 1970s, the language has generated much interest among linguists. As Smith (1979a; 1984) has clearly shown, the removal of sustained direct contact with the Portuguese language from the 17<sup>th</sup> century on has led to the development of an unmistakably South Asian typology in the structure of SLP, all the while retaining the prominently European lexical component (mostly from Portuguese) that reveals the linguistic environment in which it formed. SLP occupies a unique position in the field of Creole Studies, and presents a peculiar theoretical challenge, due to its most “un-creole” morphosyntactic features, such as typical head-final syntax (SOV word order, postpositions), a system of case suffixes, and considerable morphosyntactic complexity (see Smith 1979a; 1979b).

Prior to the start of the DSLP project, the most significant collection of primary oral data for SLP was Ian Smith’s deposit in *PARADISEC* (Smith 1973), which consists

---

<sup>6</sup> Exogamy is, in fact, frequently singled out as the main reason for the break in SLP transmission. The real extent of this practice or the exact effects on language transmission remain to be ascertained.

chiefly of his field recordings made in 1973 and 1974, mostly in Batticaloa. This is an open-access collection containing about 15h of recordings but no transcriptions or annotations. A lot of the linguistic material therein has been published in various of Ian Smith's articles, and his (2013a; 2013b) contribution on SLP for the *Atlas of Pidgin and Creole Language Structures* (Michaelis et al. 2013) has contributed a significant corpus of 234 fully transcribed and annotated sentences which are freely available on the Atlas' webpage.

David Jackson's 1970s research on the oral traditions has also produced a collection of recordings of SLP songs and music from Batticaloa and Trincomalee, now published in the form of 2 CDs contained in a larger collection of ethnomusicological recordings entitled *A Viagem dos Sons* (1998). The lyrics of a few of those songs have also appeared in the author's publications.

In addition to Smith and Jackson's corpora, there are sporadic videos uploaded to online platforms like *YouTube* that contain SLP speech or song performances. These result from very different initiatives, either by community members or community associations, or by outsiders who visit the community.

It is to this body of modern linguistic documentation that the DSLP project set out to contribute. The following section will explain the various components of the project's output and discuss some procedural decisions and activities.

### **3. Documentation and research**

The output of the DSLP project is centered around the constitution of the corpus deposited at ELAR but includes several initiatives and sub-projects which go beyond the mere collection, transcription and annotation of oral data. As an intrinsically multidisciplinary endeavor, it can be broken up in several components, which we will describe here: linguistic documentation (§3.1.), orthographic development (§3.2.), ethnomusicological documentation (§3.3.), sociolinguistic surveying (§3.4.), and archival prospection (§3.5.).

#### **3.1. Linguistic Documentation**

Oral data was collected in planned sessions, in the form of video and/or audio recordings of spontaneous or semi-spontaneous sessions involving single individuals, small groups (e.g. couples, households,...) or large groups (e.g. extended families, gatherings,...). These recording sessions were held in a variety of settings, ranging from informal spaces such as family compounds and the research team's own lodgings, or the interviewees' workplaces, to public and institutional spaces (e.g. associations, churches, festivals,...). The equipment used for audio and video recordings in the field

are listed in Table 1. Occasionally, we also used our personal cameras and mobile phones for meta-documentation – the documentation of documentation activities and related tasks, such as collaborative transcriptions and data-sharing with the community.

**Table 1.** List of equipment used for audio and video recordings

RECORDERS	VIDEO CAMERAS	MICROPHONES
Zoom H6	Panasonic HC-VX980 (2)	Cardioid Rode NT5 (2)
Marantz PMD620		Wireless lapel Sennheiser EW 112-P G3
		Shotgun Rode NTG2
		Vocal Rode NT2-A

The interviews were mostly conducted in SLP but also, occasionally, in English or Tamil. Although individual speakers varied in age and gender, there is a preponderance of middle-aged or elderly interviewees in the corpus, as a result of the current sociolinguistic distribution of SLP fluency (see §3.4.). Informed consent was always sought and obtained for each individual involved in the recordings, making use of a bilingual (English and Tamil) consent form<sup>7</sup> that explained the background of the documentation project, the intended use, dissemination and archiving of the recordings, the interviewees' retention of copyright, and their right to pull down any recordings from the online archive.

Interviews were geared towards an adequate representation of the cultural characteristics of the Burgher community, including elements of music, cuisine, tradition, crafts, occupations, religious practice, and daily life. This produced a diversity of topics and allowed us to document events such as the cooking of certain Burgher food items (e.g. *fófóórchi*, *bruuder*, *dodol*, wine), crafts (e.g. carpentry, blacksmithing), religious rituals, the playing of instruments, and musical and dance performances (see §3.3.). This approach also produced abundant extralinguistic commentary on the composite identity of Burgher people and their history, traditions, ideas and idiosyncrasies, and especially the current and past status of the language, as well as some metalinguistic information about topics such as variation and language-mixing. In addition, complementary elicitation sessions have also been conducted, aimed at correcting the transcription or interpretation of previous recordings, or at collecting lexical information.

As per the team's workflow, once collected, audio and video recordings are edited, which often involves simple crops but sometimes also synching video and audio

<sup>7</sup> Whenever a participant was unable to read the consent form in either language, the contents were explained orally. The forms used in the DSLP project can be accessed at: <https://elar.soas.ac.uk/Record/MPI1278123>.

from different devices. We aim for the production of a rich set of metadata to accompany the raw recordings (see Nathan & Austin 2004), including transcriptions and annotations. Transcription is done using *ELAN* (Wittenburg et al. 2006)<sup>8</sup> and is fully time-aligned. Since SLP is typically unwritten and therefore lacks a standard orthography, we had to develop an orthographic system, which we describe in §3.2. Table 2 explains other transcription conventions that proved necessary.

**Table 2.** List of transcription conventions

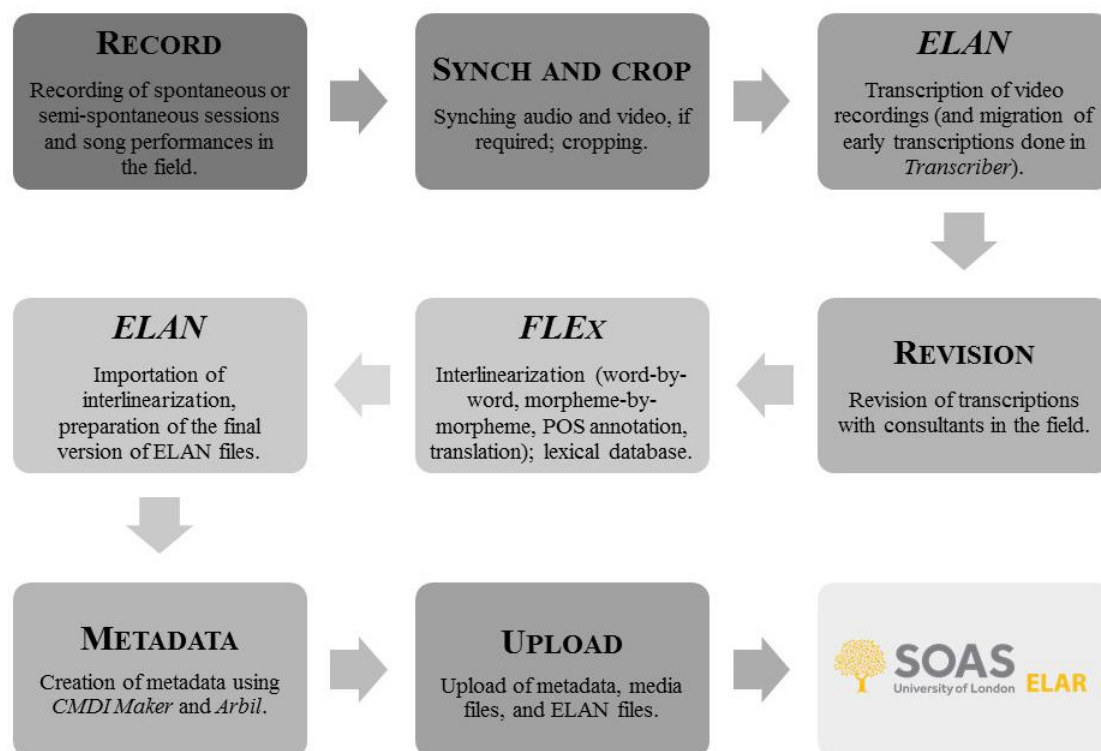
SYMBOLS	DESCRIPTION	EXAMPLES
Quotation mark	Loanwords, foreignisms	<i>isi</i> "wiring", "current" <i>su sirviis</i> (slp005_1)
Apostrophe	Toponyms	'Sri Lanka' <i>ntu</i> (slp006_1)
Square bracket	Filled pauses, hesitations	<i>pikiniin pikiniin sirviispa tandáá</i> [hum] [hum] <i>naanda prenda</i> (slp018_2)
Hash	Abrupt interruptions, truncation	<i>eev taan se# seesda</i> "mile" <i>mee taparáá</i> (slp006_2)
[-]	Non-linguistic content, unclear speech	<i>kii see máál</i> [-] <i>briiya see taam andáá</i> (slp006_1)

Some of the more complex or challenging transcriptions are then revised in the field in collaboration with native speakers, which significantly increases the value and precision of the texts and, concomitantly, expands the team members' knowledge of the language. Subsequently, SIL's *Fieldworks Language Explorer* (FLEX) is used to create morphosyntactic annotations and a lexical database, and then re-imported into *ELAN* for archiving.

The corpus is made available online at the ELAR archive (see footnote 2). The DSLP collection (Cardoso 2017) currently holds over 37h of recordings organised in bundles. Each bundle typically consists of media files (.mp4 for video and/or .wav for audio) and their corresponding annotation files (.eaf and .pfsx) containing a transcription, word-by-word morphosyntactic annotations, translation into English, and, optionally, ethnographic or ethnomusicological notes, following the structure exemplified in Figure 1. Files collected during the DSLP project are given an individual name with the following default structure: slpXXX\_X (or, in the case of files resulting from the 2015 exploratory trip, slp15XX\_X) – in which the digits preceding the underscore identify the recording session (organised chronologically) and the final one distinguishes various clips from the same session, to which is added the file extension: e.g. slp006\_4.eaf. Other bundles include written documents (such as consent forms) or

<sup>8</sup> For some of the earliest audio recordings, done as part of an exploratory visit in 2015, transcription was done using the software *Transcriber*, and later on imported into *ELAN*.





**Figure 2.** Workflow of data collection and processing

### 3.2. Orthographic development

Despite the 19<sup>th</sup>-century and early-20<sup>th</sup> century written production in “Ceylon Portuguese” described above in §2.1. – of which few community members are aware (see §3.5.) – SLP has largely remained an unwritten language. Exceptions include various attempts to record and print ethnographically-relevant texts such as song lyrics (e.g. CBU 1976) or lexicons (e.g. AMI & SLPBF 2011), as well as materials produced for language classes directed at children, which however use widely variable orthographies, not only between resources but often also within single resources. Academic publications about SLP, of course, also contain written words and sentences in the language. Among these, the most consistent orthography – because it is based on detailed phonological research – is that used by Ian Smith, despite some alterations from the earliest version (e.g. Smith 1977) to the latest (e.g. Smith 2013a; 2013b).

In that it is largely an oral language, SLP is not unlike so many other creole languages – or, in fact, so many minority languages – around the globe. All of the Portuguese-based creoles of South Asia, and most other creole languages, lack an accepted set of conventions for written use. Some, such as e.g. Haitian, Cape Verdean Creole, or Papiamentu, have or are in the process of defining standard orthographies; others, such as the Portuguese-based creoles of Macau and Malacca, have some regular written production which can potentially lead to a more or less accepted orthography.

SLP does not fit any of these cases and, as such, the research team was faced with the need to define a writing system for the transcription of the recorded speech.

It is true that it is not always appropriate or possible for linguists to take it upon themselves to define an orthography for the languages they work with (see Lüpke 2011; Oko 2018), and that the graphization of creole languages may be especially contentious (see e.g. Joseph 1987; Romaine 2005; Sebba 2007), but, very often, as Seifart (2006: 275) notes, “the development and implementation of a practical orthography in the speech community is an absolutely necessary task in the early phase of a documentation process”. In the case of SLP, none of the potential sociolinguistic complications to developing an orthography (such as e.g. wide dialectal variation, community-internal disagreement over scripts, or divided ethnic identities) seemed to hold.

While it should be pointed out that the purpose of this task was to assist the production of transcriptions and annotations for the DSLP corpus, and not necessarily produce a writing system to be implemented by the community, our concern with revitalization meant that we aimed for a potentially viable orthography for the Burgher community. One immediate corollary of this option was to avoid the use of IPA symbols, which are featured for some phonological segments in Ian Smith’s notation. In addition, we asked for the input of the Burgher community in the preliminary process of orthographic development. In order to do so, a workshop was held in April 2017, in Batticaloa, with 9 representative community members and the principal investigator. The first point in the agenda was to select a script. Considering the nature and context of the modern SLP speakers, two options were especially obvious: either the Tamil script, since this is the dominant language of the region and one that community members need in order to function in Eastern Sri Lanka, or the Latin script, used to write not only Portuguese but also English, which has an important role in modern-day Sri Lanka and is therefore not an external system. Interestingly, previous attempts to record SLP, starting in the 19<sup>th</sup> century, consistently selected the Latin script. Even though workshop participants demonstrated great ease in applying the Tamil script to the writing of SLP, they did favor the Latin script, a fact which should not be dissociated from identity politics that stress the local distinctiveness and the historical links with Europe of the Burgher community.

Afterwards, workshop participants were asked to write down, however they wished, a set of sample words and sentences, which were then discussed in the group. Several of these stimuli were especially selected to elicit the solutions to some phonological features previously detected and described by Smith (1977), such as vowel length or retroflexion, and also to issues of morphological (non-)integration, in the cases of e.g. grammatical case markers, compounds, or TMA markers. The results of this exercise were taken into account when designing an orthography for SLP, later adopted in the transcriptions of the speech recordings. For instance, it was observed that several participants did select different notations to distinguish short and long

vowels, and therefore it was decided that vowel length should be encoded in the orthography.

Since we have at our disposal Ian Smith's (1977) detailed phonological description of SLP, and therefore the phonemic oppositions of the language are well-established, we opted for a shallow orthography. The correspondence between sounds and graphemes is given in Table 3:

**Table 3.** Sound-grapheme correspondences in SLP orthography

SOUND	GRAPHEME	EXAMPLE	GLOSS
<b>a</b>	á [short]	álagriiya	'happiness'
	áá [long]	lingvááy	'language'
<b>ə, ɐ</b>	a [short]	kuma	'to eat'
	aa [long]	sumaanu	'week'
<b>ɛ, æ</b>	é [short]	éla	3SG.F
	éé [long]	uséenta	'hundred'
<b>e</b>	e [short]	beva	'to drink'
	ee [long]	ree	'king'
<b>ɔ</b>	ó [short]	bósa	2SG.GEN
	óó [long]	nóóyva	'bride'
<b>o</b>	o [short]	oy	'today'
	oo [long]	nooyvu	'bridegroom'
<b>u</b>	u [short]	kumera	'food'
	uu [long]	fuula	'flower'
<b>i</b>	i [short]	minha	1SG.GEN
	ii [long]	siinku	'five'
<b>p</b>	p	pikiniin	'small'
<b>t</b>	t	tudu	'all'
<b>k</b>	k	kera	'to want'
<b>b</b>	b	bóóm	'good'
<b>d</b>	d	deevs	'God'
<b>ɖ</b>	dh	podhiyáás	'child'
<b>g</b>	g	grááandi	'big'
<b>m</b>	m	máám	'hand/arm'
<b>n</b>	n	nistáá	'to need'
<b>ɲ</b>	nh	minha	1SG.GEN
<b>ŋ</b>	ng	ung	'one'

**Table 3.** Sound-grapheme correspondences in SLP orthography (cont.)

SOUND	GRAPHEME	EXAMPLE	GLOSS
<b>f</b>	f	faátus	'things'
<b>v</b>	v	véey	'old'
<b>s</b>	s	disa	'to go down'
<b>z</b>	z	kááza	'house'
<b>dʒ</b>	j	jeentis	'people'
<b>tʃ</b>	ch	chééli	'money'
<b>j</b>	y	papiyáa	'to speak'
<b>w, v</b>	v	eev	1SG
<b>l</b>	l	léés	'to read'
<b>r, r</b>	r	veerdi	'green'

We are aware that, in this system, 2 digraphs could be perceived as problematic, since they are absent from English: <dh> for the retroflex plosive /dʒ/, and <nh> (as in Portuguese) for the palatal nasal /ɲ/. The other two digraphs are retrievable from English spelling and should, in principle, not be so foreign to SLP speakers: <ng> for the nasal velar /ŋ/, and <ch> for the affricate /tʃ/. Another innovative convention in the Sri Lanka context is the use of a diacritic, the acute accent, to distinguish open mid from close mid vowels, or open from mid vowels. While we recognize the limitations of this particular solution in the context of Sri Lanka, the production of an acute accent is easy in most keyboards using the Latin script.

Finally, it is worth noting that the DSLP orthography differs considerably from the previous written records of the language, which could not be taken as valid orthographical models because of their inconsistencies. Despite the unquestionable value of these materials for diachronic studies, as it turns out the orthography they adopted relies excessively either on the European Portuguese norm or on the Dutch conventions, not reflecting the effective use of the language. Callaway (1820: i) himself, one of the first authors to publish books in SLP, revealed some frustration with previous practice: “Many individuals, it may be observed, from inattention to the true orthography, are accustomed to write Portuguese in the Dutch alphabet. The language in so unsuitable a garb, is sufficient to awaken the suspicion and contempt of a stranger, and has probably given birth to the appellation of ‘Corrupt Portuguese’ [...]” – and yet, as Smith (2016: 248) points out, his orthography is also inconsistent. Table 4 exemplifies some of the differences between two of the earliest sources (Berrenger 1811 and Callaway 1820) and our own orthographic notation for modern SLP:

**Table 4.** Some orthographic correspondences in Early and Modern Sri Lanka Portuguese

	<b>BERRENGER 1811</b>	<b>CALLAWAY 1820</b>	<b>CARDOSO 2017</b>
CHILDREN	<i>Criances</i>	<i>Creanças</i>	<i>Kriyáánsas</i>
TO SAY	<i>Falla</i>	<i>Falla</i>	<i>Faláá</i>
EAR	<i>Orelha</i>	<i>Orelha</i>	<i>Oreeya</i>
HAND/ARM	<i>Man/Braço</i>	<i>Maõ/Braso</i>	<i>Máám/(Bráásu)<sup>9</sup></i>
GOD	<i>Deus</i>	<i>Deoes/Deos</i>	<i>Deevs</i>
TEA	-	<i>Tey</i>	<i>Teááv</i>

### 3.3. Ethnomusicological Documentation

Music and dance play an important role as cultural expressions which are iconic and richly laden with specific cultural knowledge and language use that can endure despite language shift. Recognizing this, language documentation work has often included a strong emphasis on the documentation of these temporal arts (Barwick 2012), and its importance continues to be emphasized (Fitzgerald 2017). SLP scholarship since the late 19<sup>th</sup> century has acknowledged the significance of its oral traditions and has included records of verses and descriptions of music and dance along with linguistic description (Cardoso 2011; Cardoso, Hagemeyer & Alexandre 2015: 680ff). The enduring connection between SLP and music is evident in the use of the language in song not only by fluent speakers but also by both younger Portuguese Burghers who may only know a few words of SLP, and Afro-Sri Lankans (Kaffirs), who preserve songs in SLP as a central part of their repertoire of performance despite having shifted to Sinhala in all other domains of communication (de Silva Jayasuriya 1999: 251–253).

The inclusion of ethnomusicological documentation of Portuguese Burgher and, to a small extent, Afro-Sri Lankan performance is justified for this project, not only because of these reasons, but also because Portuguese Burgher and Afro-Sri Lankan musical and dance genres have had an impact on Sri Lankan culture as a whole, influencing and shaping the popular genre of Sri Lankan dance music known as “baila” (Ariyaratne 2001 [1985]; Jackson 1991: 618; Sheeran 1998; 2002; de Silva Jayasuriya 1999; 2018; Chitty 2005). A connection between Sri Lankan baila and Portuguese culture is commonly acknowledged but the significant historical role of these music cultures in the development of baila is still being understood. This is something which this documentation project helps to shed some light on through capturing a range of performances of song, music and dance in a range of contexts, and performers’

<sup>9</sup> In modern SLP, the form *máám* is regularly used for both ‘arm’ and ‘hand’. However, in Smith’s corpus, while that is also the case, the compound form *máám bráásu* is also recorded for ‘arm above the hand’ (see Smith 2013b: ex. 41-174).

knowledge and perspectives on their tradition through ethnographic interviews with them, many of which are in the SLP language. Hence, the ethnomusicological aspect of this project is of relevance for the purposes of language documentation and research, but also to a broader understanding of Sri Lankan music and culture.

Work on SLP performance traditions prior to the DSLP project includes research by Kenneth David Jackson (1990), who conducted some of his fieldwork with Ian Smith in the 1970s, and more recent work by Shihan de Silva Jayasuriya (1996; 2000). Kenneth David Jackson's recordings are available in two CDs (titled *Baila Ceilão Cafrinha* and *Cantigas do Ceilão*) of the aforementioned series *A Viagem dos Sons* (1998). Ian Smith's documentation of the language also includes several music recordings (see Smith 1973: SLP13-SLP17). Comparing the Portuguese Burgher music scenes observed and recorded by Jackson and Smith with the current project, one can see the vitality and fragility of music and dance within the communities. The vitality is evident in the maintenance of a number of old songs and some probable innovations, the continued practice of dancing at weddings (especially the *káfriinha* quadrille), the identification of traditional music as an important part of Burgher identity, and the presence of at least one group involving young singers and musicians in Batticaloa, another one in Kalmunai and a few other young musicians in Batticaloa, Panichiyadi, Jaffna, Trincomalee, and possibly other places. At the same time, observations suggest that there is much less variety in the verses sung today than previously, and there is certainly less knowledge about dancing, with some previous forms, such as the *léensas* (lancers), becoming a form of specialised knowledge. In the case of Trincomalee, it emerges that a previously vibrant Portuguese Burgher music scene is now greatly reduced to the point that either a CD backing is used or one of the two music groups from Batticaloa travel to play for dances at weddings in Trincomalee.<sup>10</sup> There have also been changes to the instrumentation and to the music over time, as reported by participants and also observable when comparing current performances with previous recordings. These developments raise interesting questions in relation to musical vitality and change, another area which warrants further inquiry.

Music and dance were frequent topics in interviews aimed at collecting unconstrained speech, being a subject of conversation in dozens of interview sessions. Discussion about music was typically initiated by the interviewer, who would either ask about wedding customs or music itself, particularly where it was known beforehand that the consultant was a musician or had particular knowledge of the subject. In these interviews, a few consultants describe how weddings traditionally involved specific

---

<sup>10</sup> This observation is confirmed in interviews with participants from Trincomalee, including those previously involved in the Trincomalee Portuguese Burgher music scene. However, as there are still some musicians there, this situation could change again, and will be an area to continue observing. The recent commencement of dance classes for children, discussion about music training, and engagement in traditional music by younger members of the musical families from the past are signs that there is community support behind revitalizing Portuguese Burgher music in Trincomalee.

pieces of music which needed to be played at particular times, as well as specific dance pieces and songs. Consultants also shared knowledge about the life histories of musicians and music groups, descriptions and demonstrations of instruments, or explications of songs and performance sequences. In most cases, these conversations were conducted in SLP but, in other cases, e.g. where the interviewer was not fluent in SLP, either English or Tamil were used.

The main focus of ethnomusicological work within the project has centered on recording and archiving the range of song, music and dance materials, conducting musicolinguistic analysis of songs, ethnographic participant observation and historic work on the whole tradition and its performers. While the focus is decidedly on the musical and dance traditions of the Portuguese Burghers, with 14 full performances already archived and most of the time spent with that community, the project has also made 2 recordings of Afro-Sri Lankan music performance.

Currently, the DSLP collection in ELAR contains over 20 bundles including music and/or dance performance. These bundles represent items from several full performances (either dedicated or public), including performances done during the course of an interview and others during an event within the community (a wedding silver jubilee) which featured music and sets of dancing, as well as a rehearsal for a performance. Overall, this project has observed and recorded 26 formal performances of music - 5 staged performances at public cultural events, 8 performances as part of a Burgher community event, and 13 performances which were given for the purpose of recording. All of them include singing, and 11 also include dance. In addition to these full performances, many informal ones took place during the course of interviews or as part of observation. A total of 196 individual items of performance involving song, music and/or dance have been observed and recorded as part of the project.

Just as in the case of recordings of conversations, music and dance performances are also uploaded to ELAR with an accompanying set of annotation files including transcriptions and translations (in the case of songs) and/or ethnomusicological notes. The lyrics for songs and their meaning are also revised with participants, typically the singers themselves. However, because sung language reflects a different register to spoken language, with occasional archaic language and peculiar syntax – which is, of course, highly relevant for sociolinguistic and diachronic research –, the transcriptions of singing have not been morphosyntactically annotated, unlike speech.

Sung performance forms an important proportion of the overall performances, with 116 items observed and recorded thus far. One of the most important song formats, which comprised 41 of the 116 song items within the Portuguese Burgher performance repertoire, is what is most commonly known in Batticaloa as *straight báyla*, a rhyming strophic song format defined by its tune (though there is much scope for variation between and even within performances) sung over a 6/8 musical rhythm (when accompanied) which is defined by its adherence to a specific primary chord pattern (I-

IV-I, I-V-I) followed for each “verse”,<sup>11</sup> and a set of typical vocal and instrumental melodies based around that structure. The significance of *straight báyla* is evident in its being an iconic musical motif, variants of which are frequently encountered within the broader genre of Sri Lankan *baila* music, in particular the subgenre of Sri Lankan music known as *vāda baila* (‘debate *baila*’). *Straight báyla* is a format with significant history and context in the Portuguese-influenced Indian Ocean which calls for comprehensive research.

Texts sung within a *straight báyla* structure are generally couplets in heptameter, as in the verses in Example 1 (English translations are provided below each verse, in square brackets):

A: *bayláá nóóna, bayláá nóóna, buniitu bayláá,*  
 B: *peyáá bósa sáaya, nóóna, sakudii bayláá,*  
 [Dance, lady, dance, lady, dance beautifully, / Hold up your skirt, lady,  
 shake it (and) dance]  
 \_(slp025\_1, 15:20-15:33)<sup>12</sup>

A: *ááltu murungeera, nóóna, inchiidu verduura,*  
 B: *avóórasu niinas tudu inchiidu gurduura.*  
 [The tall drumsticks, lady, (are) full of greenery, / Today’s girls (are) full of fat]  
 (slp038\_1, 26:52-27:42, slp039\_1, 0:43-0:56)<sup>13</sup>

A: *nóóna fikáá niina, kii tafaya naa kintáál?*  
 B: *eev talaváá léénsu paara jáneeru natáál.*  
 [Girl becoming a lady, what are you doing in the garden? / I’m washing  
 handkerchiefs for New Year and Christmas]  
 (slp055\_8, 1:02-1:17, slp058\_1, 3:19-3:38)<sup>14</sup>

A: *anééla teem ooru, nóóna, sééti péédra yuntu,*  
 B: *keem kera anééla, nóóna, kásáá minha yuntu.*  
 [Ring of gold, lady, with seven stones, / Whoever wants the ring, lady, marry  
 me]  
 (slp059\_4, 4:37-4:57)<sup>15</sup>

### Example 1. Some verses sung to *straight báyla*

<sup>11</sup> Each verse is sometimes referred to by singers using the term *kántiyya*, the SLP word meaning ‘song’ and also sometimes used for whole song items.

<sup>12</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104256>.

<sup>13</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104036>.

<sup>14</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104114>.

<sup>15</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104630>.

The couplets are typically performed in order of the lines, and then the order of the lines is reversed to make for a four-line song structure. The first line is also often repeated. Hence, the lines of text can be said to typically follow an (A)ABBA structure, where the parenthesis indicates that the repetition of the first line is optional. While the text follows the above structure, the melody is framed by an (A)ABAB chord pattern, where A is a I-IV-I chord pattern and B represents a I-V-I chord pattern. Sometimes, (A)ABBA textual structure is not followed and verses are performed as (A)ABAB or, as in two rare cases, structured ABCD, where the verse is comprised of four lines.<sup>16</sup>

In addition to *straight báyla*, Burgher music includes a range of other song types, one category of which comprises song types which draw from the same corpus of verses and which are typically described by the chorus used - for example, *jingali nóóna*,<sup>17</sup> *áyoo mamáá oondi yafooy*,<sup>18</sup> *áádi aavndu kapáátu*<sup>19</sup> or *kuráánjaniita* (also known as *chikóóti*). With the exception of *kuráánjaniita* all the other song formats can also be described as *káfriinha* songs because they can be used to accompany for the *káfriinha* dance. However, there does not appear to be a specific categorization within the community for these other song formats; consultants typically use the term *kantiiya* to describe one song or another or the verses sung to any of these songs. The verses sung between the choruses for these various song formats are typically the same verses as those sung for *straight báyla*. Our observations (backed up by remarks made by consultants) are that just about any of the available verses can be sung for any of these songs, although, in a few cases, there are one or two verses which are canonically linked to a song format and are always sung for that song format.

The other category of song is set compositions, including songs entitled *téera isti téera*,<sup>20</sup> *vii minha ámoor pabayláá* (also known as *ámoor nóóna*),<sup>21</sup> *montááyantu loteem ooru*<sup>22</sup> and *liimpu bistiidu*. These songs are also in 6/8 *káfriinha* rhythm. In fact, in the case of *vii minha ámoor pabayláá*, this song has been sung and performed as an accompaniment to the *káfriinha* dance. The lyrics for these songs, including choruses and verses, are fully composed with minimal variations in the text and music between different singers. In the case of the first three, these were composed within the last 50 years by specific people whom consultants identified (although, in some cases, more than one composer was identified). Interestingly, in the case of *téera isti téera* and *vii minha ámoor pabayláá*, both the songs share the same melodies as Sri Lankan *baila* songs in Sinhala, respectively *mē vadayaki jīvitē*, and *māla giravī* (both also mentioned in de Silva Jayasuriya 2018).

<sup>16</sup> This verse which was, in fact, originally part of another pre-composed song with four-line verses, hence, not normally part of the *straight báyla* repertoire.

<sup>17</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104024>.

<sup>18</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1267782>.

<sup>19</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1266258>.

<sup>20</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104028>.

<sup>21</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1278116>.

<sup>22</sup> Available at: <https://elar.soas.ac.uk/Record/MPI1104016>.

The verses given in Example 1 demonstrate some of the thematic material present in all of the Portuguese Burgher sung repertoire. The themes include love and marriage, music, dancing, merriment, domestic life, social commentary (e.g. about changing times) and verses sung for purely comic effect (Smith 2010). These topics and verses, some of which were also recorded in 19<sup>th</sup> century sources such as Fernando (1894) and beyond Sri Lanka, are historically significant. In fact, Jackson has established links between some of the themes amongst the sung repertoire in Sri Lanka, Malacca, India, and elsewhere (1990: 53ff) and links some of these to thematic material within the Portuguese medieval ballad *Bela Infanta* (1990: 95-104). Cardoso (2011: 156) further states that “the poetic material that circulated among the Luso-Asian communities was subject to creative manipulation, recombination, and adaptation to different realities”.

An important part of the documentation has focused on dance, which is closely linked with the music and occurs in the same performance contexts. As is the case with music, weddings are one of the central contexts for dancing, and the wedding event sequence includes dancing as a significant part of proceedings. The two forms of dancing performed by the community are quadrilles, namely the *káfriinha* and *léensas* (lancers),<sup>23</sup> the latter of which appears to be based on a variant of the lancers quadrille which originated in France and was popular in Britain, Ireland, mainland Europe, Australia, America and other parts of the world in the 19<sup>th</sup> century. While the word “lancers” has been used in relation to the *káfriinha* in S.V.S. Rodrigo’s 1920s score, *Ceylonese Lancers on Caffringha Airs* (cited in Ariyaratne 2001: 88-96), and was used to refer to a range of quadrille dances and music in Trincomalee performed by J.E. Paul and group and recorded by Jackson (see the CD *Baila Ceilão Cafrinha*), the provenance of the “lancers” in Batticaloa is somewhat mysterious. They are present in Smith’s recordings (see Smith 1973: ISLP13-A, 0:00-13:40), but the tunes are not heard or notated in other work, suggesting that they may have been a more recent introduction. While both of these forms of dancing were reportedly performed at weddings, it is only the *káfriinha* that is performed today, with *léensas* only performed during cultural showcases and considered more esoteric.<sup>24</sup> As part of the project, the team observed and recorded both of these dances in cultural showcases and observed and recorded performances of *káfriinha* in the context of a wedding, a wedding silver jubilee and the celebration of a first communion. When performed for the first set of dancers at a wedding or like event or for cultural showcases, the first four *káfriinhas* (in other words, dances of the *káfriinha* suite) are observed to be set to instrumental musical pieces,

---

<sup>23</sup> The full *káfriinha* and one of the *léensas* are available at: <https://elar.soas.ac.uk/Record/MPI1104036>.

<sup>24</sup> Linda and Emmanuel Paul from Trincomalee talked about several different kinds of *léensas* which were performed there when they were young, but not now. Based on the recordings made by Jackson these do not appear to be related to the *léensas* in Batticaloa but appear to be related to Norbert Rodrigo’s ‘Ceylonese Dances’ and S.V.S. Rodrigo’s ‘Ceylonese Lancers on Caffringha Airs’ (included in Ariyaratne 2001: 80–96).

while the fifth *káfriinha* is commonly performed to an accompanied version of the *straight báyla* melody, which often features a violin interlude preceding and interspersing each sung verse.<sup>25</sup> This aspect gives *straight báyla* additional importance in the contemporary performance context as a song which features prominently as both a song and dance. For subsequent sets of dances, observations and recordings show that other songs can be used but that the fifth figure for each set remains *straight báyla*.

### 3.4. Sociolinguistic surveying

Based on opinions expressed in various interviews, the Portuguese Burgher community of Eastern Sri Lanka emphasize that their language is a significant and inseparable part of their identity, although they recognize the challenges that modern SLP faces and the efforts that must be made to maintain and preserve their language and, consequently, their culture. There is a clear sense of the contraction of the language and of a break in transmission which has accelerated in recent times. However, as explained in Section 2., prior to the DSLP project there was no concrete quantification of the extent of the problem, nor of the size of the speech community. Official census publications conflate the Burghers and a few other communities into the category “Others”, which makes it difficult to get a clear picture of even the size of the community. Despite that, as we have already mentioned, recent published figures (Bandara 2012; Nordhoff 2013) indicate the presence of about 170 Burgher families in Trincomalee and 4,250 individuals in Batticaloa (Nordhoff 2013), although they do not specify what percentage of this population speaks SLP. The various Burgher associations of Eastern Sri Lanka (the Sri Lanka Portuguese Burgher Foundation and the Burgher Union, in Batticaloa; the Burgher Welfare Association, in Trincomalee) have their own figures and estimates about the community, which they shared with the team members, but these do not match one another entirely, nor the published data.

To respond to this gap in our knowledge of the SLP-speaking population, the DSLP project included a sociolinguistic survey of the Burgher community, aimed at establishing the current status of SLP regarding its distribution (in terms of geography, demography, and social group), degree of transmission, and domains of use. The survey significantly enriched the DSLP project, providing important statistics for current and future investigations, and for revitalization efforts (see Section 4).

The survey took the form of a sociolinguistic questionnaire<sup>26</sup> which was to be filled by Burgher households (including families in which only some of the members identify as Burghers) in the three districts under study: Trincomalee, Batticaloa, and

---

<sup>25</sup> Discussions with consultants suggest that the performance of *straight báyla* for the fifth *káfriinha* music was a fairly recent innovation. Still, its use as the primary sung piece to accompany dance is significant.

<sup>26</sup> The questionnaire is available at: <https://elar.soas.ac.uk/Record/MPI1278522>.

Ampara. Since the size of the community is relatively small, the survey was designed to be filled in by the entire population, and not a sample of members, although some occasional gaps and oversights are to be expected. To achieve this, we enlisted the assistance of members of the community who were responsible for conducting the survey of their own area of residence, and were paid for this task. This effort involved 17 surveyors and the following areas in which Burgher families reside:

- In Trincomalee District: Trincomalee Town, Orr's Hill, Palaiyoothu, Selvanayagapuram, Jinna Nagar, Barathypuram, Uppuveli, Sampalthivu, Nithyapuri, Alles Garden, 6th Mile Post, Nilaveli, Andamkulam, China Bay;
- In Batticaloa District: Batticaloa Town, Pulliyadikudah, Thandanvelli, Sinna Uppodai, Periya Uppodai, Dutch Bar, Mammangam, Koolavady, Iruthayapuram, Thannamunai, Karuvepankerny/Jeyanthipuram, Kalkudah, Panichiyadi, Eravur, Valachchenai, Thiyawattawan;
- In Ampara District: Kalmunai (Kurunchiadi, Panndiruppu, Manalchenai), Akkaraiyappu, Sharanalaiyapuram, Vinayagapuram, Thirukkivil, Komari, Theevukala, Karaitheevu, Sammanthurai, Veeramunai, Alayadivembu, Ithiyadi.

Each surveyor was given an appropriate number of blank surveys and received a tutorial on how to apply the questionnaire. The collection process extended over nearly 8 months: the logistical structure was established in October and November 2017, and the last completed surveys were collected in April and May 2018. This resulted in 920 completed surveys, each corresponding to a single household. The assistance of the community in the process was vital: the coverage of all the areas would not have been possible without their own knowledge of the community, and this also brought to every Burgher house the discussion of language conservation, language vitality, and identity, contributing to an awareness of the issue.

The questionnaires were completely anonymous, so the respondents could give unconstrained answers. They were also short and simple, so as not to take too much of the respondents' time, and the questions and instructions were given in both Tamil and English. Each member of the household was asked to fill in 4 sections:

- The first section elicited to general socioeconomic information, including age, gender, marital status, education level, profession, and religion;
- The second section focused on language proficiency: the respondents were asked to self-assess their proficiency in various languages (SLP, Tamil,

Sinhala, English, or others), grading their proficiency as one of 5 levels, ranging from “I cannot understand or speak” to “I can understand and speak fluently”;

- The third section aimed at collecting information on language usage patterns: the respondents should state their use of language in a diverse set of communicative contexts (e.g., home, school, religious place), with various interlocutors (e.g., parents, neighbors, classmates), and for different skills (reading, writing), ranking the languages in their multilingual repertoire in decreasing order of preference;
- The fourth section asked for the respondents’ opinion about SLP and Burgher identity, through questions such as “Is SLP important for you?” or “What can be done to preserve SLP?”. This kind of qualitative data helps us understand the nature of the relationship Burghers have with SLP and their identity, and provides important feedback from the community regarding language preservation.

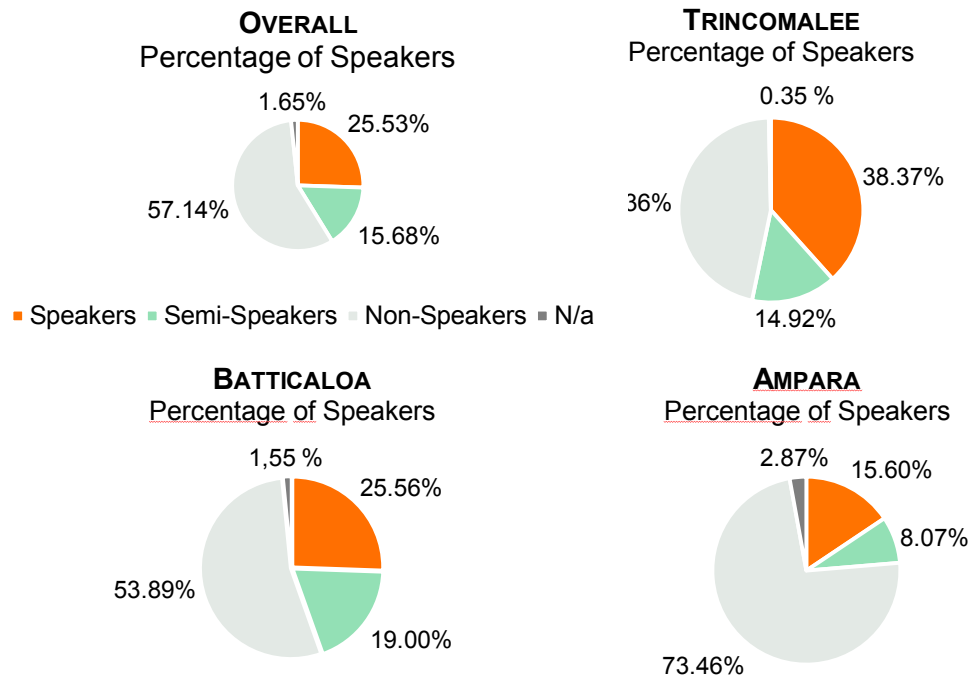
As mentioned above, the survey resulted in 920 filled questionnaires (one per household), totaling 3094 respondents in the districts of Trincomalee, Batticaloa, and Ampara. Here, we present only some initial results of the survey, viz. with regard to the geographical distribution of fluency in SLP. A complete exploration of the data – including distribution of linguistic proficiency across genders, age or socioeconomic groups, as well as domains of use – will be the object of future publications.

Table 5 shows the respondents’ reported fluency in SLP. Here, the category “Speakers” conflates those who responded ‘I understand and speak fluently’ and ‘I understand and speak with little difficulty’; “Semi-speakers” includes those who responded ‘I understand and speak with some difficulty’; “Non-speakers” are those who selected ‘I can’t understand or speak’; and “N/A” (Not Available) refers to respondents who did not answer or did not do so properly.

**Table 5.** Survey results: Figures (number and percentage) according to fluency in SLP (Total + Districts of Trincomalee, Batticaloa, and Ampara)

	OVERALL		TRINCOMALEE		BATTICALOA		AMPARA	
	NO.	%	NO.	%	NO.	%	NO.	%
SPEAKERS	790	25,53%	216	38,37%	460	25,56%	114	15,60%
SEMI-SPEAKERS	485	15,68%	84	14,92%	342	19%	59	8,07%
NON-SPEAKERS	1768	57,14%	261	46,36%	970	53,89%	537	73,46%
N/A	51	1,65%	2	0,35%	28	1,55%	21	2,87%
<b>TOTAL</b>	<b>3094</b>	<b>100%</b>	<b>563</b>	<b>100%</b>	<b>1800</b>	<b>100%</b>	<b>731</b>	<b>100%</b>

In the interest of readability, the percentages in Table 5 are presented as charts in Figure 3 below:



**Figure 3.** Survey results: Fluency in SLP (in percentage)

These results reveal a starkly different scenario in the three districts surveyed. If we look at the proportion of the Burgher community which retains fluency in SLP, it becomes evident that the process of language shift is much more advanced in the Ampara district. In the Batticaloa district, but especially in the Trincomalee district, the percentage of the Burgher community that preserves an active or passive knowledge of SLP is considerably higher, therefore these regions hold the best prospects for language preservation.

### 3.5. Archival prospection

As part of the DSLP project, the team has made it a point to visit key archives and libraries (in Sri Lanka, Portugal, and the UK) to search for sources about the objects of research (language, music and dance traditions), but especially to collect primary linguistic sources. The intention is to complement the ELAR corpus with scans and/or transcriptions of the least accessible literature on SLP published in the 19<sup>th</sup> and early 20<sup>th</sup> centuries, including dictionaries, grammars, vocabularies/phrasebooks, poetry volumes and religious texts. This will benefit, first and foremost, diachronic studies of SLP, thus widening the appeal and applicability of the modern SLP corpus.

These early sources, many of which have never been edited and made available to modern-day academics or to the Portuguese Burghers themselves, are important documents to understand the extent of use and linguistic characteristics of earlier forms of SLP. A look at the language contained in these works reveals wide discrepancies with modern-day SLP in many domains – as do some of the better known early sources, which motivated the hypothesis of a rapid and significant process of language change (through convergence with Tamil) over the past century and a half (see Bakker 2006). However, these works also reveal several inconsistencies that call into question their value as accurate representations of the language spoken at the time. Despite that, Smith (2016) admits that a close critical analysis can go a long way in informing diachronic studies of the development of SLP.

The following primary sources have been collected:

- Anon. n.d. *Cantigas e louvors per missionario servicos*.
- Anon. 1826. *O Livro de oração commum e administração de os sacramentos, e outros ritos e ceremonias de a Igreja, conforme de o uso de a Igreja de Inglaterra: Juntamente com o Psalterio, ou Psalmo de David*. London: G. Ellerton & J. Henderson.
- Anon. 1851. *O evangelho conforme de Santo Mattheos. Traduzido ne Indo-Portugueza*. London: R. Clay.
- Anon. 1863. *O livro de orações usado ne greyas de Wesleyanos ne Ilha de Ceylon*. Colombo: Officio Wesleyano.
- Anon. 1912. *O evangelho conforme de Santo Joaõ*. Colombo: British and Foreign Bible Society.
- Anon. 1914. *Cantigas ne o lingua de Portuguez, Impressado ne Matre, 23 de Juni 1914*. Colombo: Sridhara Press.
- Berrenger. 1811. *A grammatical arrangement on the method of learning the corrupted Portuguese as spoken in India*. Colombo: Frans de Bruin.
- Callaway, John. 1820. *A vocabulary, in the Ceylon Portuguese, and English languages* (2<sup>nd</sup> ed.). Colombo: Wesleyan Mission Press.
- Callaway, John. 1823. *A Ceylon-Portuguese and English dictionary*. Colombo: Wesleyan Mission Press.
- Callaway, John. 1823. *Doze sermans, ne lingoa de Portuguesa de Ceylon*. Colombo: Officia Wesleyano.
- Fox, William Buckley. 1819. *A Dictionary in the Ceylon-Portuguese, Singhalese, and English Languages*. Colombo: Wesleyan Mission Press.
- Nevill, Hugh. 19<sup>th</sup> century. [Manuscript of SLP verse; 40 folios; preserved at the British Library]<sup>27</sup>

---

<sup>27</sup> Photographs of the manuscript ceded to the DSLP project by Prof. K. David Jackson, who collected them from the British Library (c. 1985) and published an edition of the manuscript (Jackson 1990).

- Newstead, Robert (transl.). 1827. *Hum curto catichismo da Biblia, disposto em quarento divisaõs; Todas as repostas per as perguntas sendo ne as palavras de as Santas Escrituras*. London: J.S. Hugues.
- Newstead, Robert. 1852. *O Novo Testamento de Nossa Senhor e Salvador Jesus Christo. Traduzido ne Indo-Portugueza*. Colombo: Officina de Missaõ Wesleyano.
- Newstead, Robert. 1871. *Cantigas per adoraçaõ publico, em lingua portugueza de Ceylon*. Colombo: Officio Wesleyano.

#### 4. Revitalization

While language revitalization is beyond the strict scope of the project, it is never far from the team's preoccupations. One manifestation of this is the inclusion in the survey of questions concerning awareness of language loss and thoughts on revitalization – and indeed the development of the survey itself. In fact, the assessment of the sociolinguistic situation of SLP is an essential step in defining priority areas of intervention towards improving the use and domains of the language, as it reveals which areas and which groups are more susceptible to language loss and also which factors favor or prevent it. Building on the results of the survey, we intend to devise a set of suggestions aiming at language preservation, which will be shared with the Burgher community.

Another domain in which DSLP may contribute towards language preservation relates to the orthographic system devised. The absence of a language from the written domain has been recognized as a factor of linguistic endangerment, to the extent that certain metrics of language vitality (such as the *Expanded Graded Intergenerational Disruption Scale*, see Lewis & Simons 2010) require robust written use in order for a language to be deemed stable. Among the benefits of the development of an orthography, Decker (2014: 1) mentions the documentation of ideas, policies, and facts (Goody 2000), the possibility of communication of ideas over both distance and time, the access to literacy and, ultimately, the reinforcement of oral languages. Because of their complex relationship with their (often socially valued) lexifier languages, the standardization of creoles and other contact languages, including in writing, is especially pressing and sensitive. A writing system can be very beneficial in boosting the speakers' self-assessment of the value of their language and signaling its autonomy from the lexifier, and it is a necessary condition for any literacy-related activities.

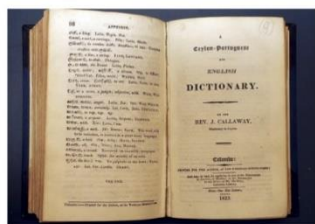
For all these reasons, we found it beneficial to share with the speech community the orthography developed within the DSLP project. To that effect, we designed and printed a pamphlet explaining those orthographic conventions and clarifying the language's long written tradition (Figure 4), which was distributed among the Burgher communities in Batticaloa, Trincomalee, and Jaffna.

orthography:

a set of conventions for writing a language, often based on a standard variety.

#### HISTORY OF THE WRITTEN SRI LANKA PORTUGUESE

Sri Lanka Portuguese has been written down since the early 19<sup>th</sup> century. Between 1819 and 1900, authors such as Berrenger, John Callaway (see below), William Fox or Sebastião Dalgado wrote and published vocabularies/ phrasebooks, religious texts and grammars focusing on the language.



Later, in 1914, an anonymous publication called *CANTIGAS NE O LINGUA DE PORTUGUEZ* presents the transcription of about 100 portuguese *cantigas* - and some of them are still reproduced and remembered today.

#### BASIC PHRASEBOOK:

HOW ARE YOU?	kilááy savóódi?
GOOD MORNING	bóóm pálmíáám
GOOD AFTERNOON	bóóm átáárdi
GOOD EVENING	bóóm ánóóti
GOODEBYE	andáántu vii
YES	seem
NO	naa
PLEASE	merse
THANK YOU	muytu gráándi merse
YOU'RE WELCOME	náada káástiv nuteem
SORRY	perduvää taam
EXCUSE ME	faya merse
WHAT'S YOUR NAME?	botussu nóomi kii?
MY NAME IS...	minha nóomi...



#### VII MINHA AMÓOR PABAYLÁÁ

vii minha ámoor pabayláá  
dóosi minha kórsaam tadiziyáá  
teem isti díiya nósa álagritya  
vii minha juuntu pabayláá.



#### MÁARA NUTEEM FUUNDU

máara nuteem fuundu minha viida pateráá  
ruuva nuteem láárgu minha móórti palevää.

## SRI LANKA PORTUGUESE ORTHOGRAPHY

A guide to knowing, understanding and using the writing conventions used to represent the spoken Sri Lanka Portuguese (SLP).



**Figure 4:** Front page of the pamphlet on Sri Lanka Portuguese orthography<sup>28</sup>

Orthography development is a never-ending activity, always subject to adjustments and improvements, and we are aware that our orthographic conventions may not be seen as adequate by the community. This pamphlet is intended as a proposal, which the community will be free to adapt and transform into a suitable writing system. Our hope is that it may motivate more speakers to write in SLP and get the discussion started on expanding the language into the written domain.

In terms of music and dance, the team has been engaged in sharing resources with the community, in particular showing old archival recordings of performances made by David Jackson and Ian Smith in the 1970s, and making the community aware of old transcribed verses, such as those in the Nevill manuscript compiled in the 1890s.

Finally, the research team, in collaboration with various community members, has also been active in awareness-raising. So far, we have shared information about the Burghers' language and traditions with the Sri Lankan press and conducted workshops in various community-level and academic institutions in Sri Lanka drawing attention to the need to inflect the trajectory of loss. This has been supplemented by the active sharing of information and raising of awareness in academic, press and community circles elsewhere (including Portugal, India, Macau and Australia), giving the issue global relevance.

<sup>28</sup> Pamphlet available at: <https://elar.soas.ac.uk/Record/MPI1279660>.

There is considerable awareness and activism among the community that is being channeled towards cultural – and linguistic – preservation. By providing accessible digital content concerning the Portuguese Burghers’ linguistic and ethnomusicological heritage, and by engaging with so many community members, it is hoped that the DSLP project may have a positive impact in ensuring that the language finds a lasting and stable place within a multilingual and multicultural Sri Lanka.

## References

- AMI & SLPBF. 2011. *Dictionary: English – Portuguese – Sri Lanka Portuguese Creole*. Batticaloa: Assistência Médica Internacional & Sri Lanka Portuguese Burgher Foundation.
- Anon. 1895. The Caffirs of Ceylon and their music. *Monthly Literary Register and Notes and Queries for Ceylon* III(6). 133–134.
- Anon. 1914. *Cantigas ne o lingua de Portuguez, Impressado ne Matre, 23 de Juni 1914*. Colombo: Sridhara Press.
- A Viagem dos Sons*. 1998. Tradisom.
- Ariyaratne, Sunil & Fonseka, E.A. Gamini. 2001. *Baila kaffirinna: an investigation*. Colombo: S. Godage & Bros.
- Bakker, Peter. 2006. The Sri Lanka Sprachbund: The newcomers Portuguese and Malay. In Matras, Yaron & McMahon, April & Vincent, Nigel (eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective*, 135–159. New York: Palgrave.
- Barwick, Linda. 2012. Including music and the temporal arts in language documentation. In Thieberger, Nicholas (ed.), *The Oxford Handbook of Linguistic Fieldwork*, 166–179. Oxford: Oxford University Press.
- Barwick, Linda & Birch, Bruce & Evans, Nicholas. 2007. Iwaidja Jurtbirrk songs: Bringing language and music together. *Australian Aboriginal Studies* 2007(2). 6–34.
- Berrenger. 1811. *A grammatical arrangement on the method of learning the corrupted Portuguese as spoken in India*. Colombo: Frans de Bruin.
- Biedermann, Zoltán. 2018. *(Dis)connected Empires: Imperial Portugal, Sri Lankan Diplomacy, and the Making of a Habsburg Conquest in Asia*. Oxford: Oxford University Press.
- Callaway, John. 1818. *A vocabulary; with useful phrases, and familiar dialogues; in the English, Portuguese, and Cingalese Languages*. Colombo: Wesleyan Mission Press.
- Callaway, John. 1820. *A vocabulary, in the Ceylon Portuguese, and English languages, with a series of familiar phrases*. Colombo: Wesleyan Mission Press.

- Callaway, John. 1823. *A Ceylon-Portuguese and English Dictionary*. Colombo: Wesleyan Mission Press.
- Cardoso, Hugo C. 2016. O português em contacto na Ásia e no Pacífico. In Martins, Ana Maria & Carreira, Ernestina (eds.), *Manual de Linguística Portuguesa*, 68–97. Berlin: Mouton de Gruyter.
- Cardoso, Hugo C. 2017. *Documentation of Sri Lanka Portuguese*. London: SOAS, Endangered Languages Archive. (<http://wurin.lis.soas.ac.uk/Collection/MPI1035102>) (Accessed 2018-05-05.)
- Cardoso, Hugo C. 2011. Oral traditions of the Luso-Asian communities: Local, regional, and continental. In Jarnagin, Laura (ed.), *Portuguese and Luso-Asian Legacies, 1511-2011*, vol. 2 (*Culture and identity in the Luso-Asian world: Tenacities & plasticities*), 143–166. Singapore: Institute of Southeast Asian Studies.
- Cardoso, Hugo C. & Hagemeyer, Tjerk & Alexandre, Nélia. 2015. Crioulos de base lexical portuguesa. In Iliescu, Maria & Roegiest, Eugeen (eds.), *Manuel des anthologies, corpus et textes romans*, 670–692. Berlin: Mouton de Gruyter.
- CBU [Catholic Burgher Union]. 1976. *Portuguese Cantignas*. Batticaloa: Catolico Burgher Uniao.
- Chitty, Gina. 2005. *Public postures, private positions: Baila and Sydney Sri Lankans*. Sydney: Macquarie University. (Doctoral dissertation.)
- Coelho, Adolpho, 1881. *Os dialectos romanicos ou neo-latinos na África, Ásia e América*. Lisbon: Casa da Sociedade de Geographia.
- Dalgado, Sebastião Rodolfo. 1900. *Dialecto Indo-Português de Ceylão*. Lisbon: Imprensa Nacional.
- de Silva Jayasuriya, Shihan. 1996. Indo-Portuguese Songs of Sri Lanka: The Nevill Manuscript. *Bulletin of the School of Oriental and African Studies* 59(2). 253–267.
- de Silva Jayasuriya, Shihan. 1999. Portuguese in Sri Lanka: influence of substratum languages. *Journal of the Royal Asiatic Society* 9(2). 251–270.
- de Silva Jayasuriya, Shihan. 2000. The Portuguese Cultural Imprint on Sri Lanka. *Lusotopie* 7. 253–259.
- de Silva Jayasuriya, Shihan. 2007. A forgotten minority: The Afro-Sri Lankans. *African and Asian Studies* 6. 227–242.
- de Silva Jayasuriya, Shihan. “Sri Lankan Baila.” *The Island*. May 1, 2018.
- de Vos, F.H. 1950. Portuguese Patois. *Journal of the Dutch Burgher Union of Ceylon* 40(4). 134–139.
- Decker, Ken. 2014. *Orthography development for Creole languages*. Groningen: University of Groningen. (Doctoral dissertation.)
- Fernando, C. M. 1894. The music of Ceylon. *Journal of the Royal Asiatic Society (Ceylon)* 13(45). 183–189.

- Flores, Jorge. 2004. A ilha de Ceilão e o império asiático português. In Flores, Jorge, *A Taprobana e a Ponte de Rama. Estudos sobre os Portugueses em Ceilão e na Índia do Sul*, 21–46. Macau: Instituto Português do Oriente.
- Fox, William. 1819. *A Dictionary in the Ceylon-Portuguese, Singhalese, and English languages; to which is prefixed a compendium of the Ceylon-Portuguese language*. Colombo: Wesleyan Mission Press.
- Goody, Jack. 2000. *The power of written tradition*. Washington, D.C.: Smithsonian.
- Himmelman, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In Gippert, Jost, Himmelman, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.
- Goonatilleka, M.H. 1970. Ceilão e Portugal – Relações culturais. *Studia* 30/31. 113–166.
- Goonatilleka, M.H. 1985. A Portuguese Creole in Sri Lanka: A brief socio-linguistic survey. In de Souza, Teotónio (ed.), *Indo-Portuguese History: Old Issues, New Questions*, 147–180. New Delhi: Concept.
- Goonetilleke, M.H. 1983. *Report of an interview with the Portuguese speaking community in Puttalam*. Colombo: Department of National Archives.
- Goonetilleke, William. 1892. The Ceylon Portuguese 'diski'. *Orientalist: A Journal of Oriental Literature, Arts and Sciences, Folklore etc.* 4. 89–90.
- Jackson, K. David. 1975. 'Bela infanta' of Sri Lanka: Ballad fragments in Portuguese Creole communities. *Journal of the Royal Asiatic Society (Colombo)* (new series) 19. 17–25.
- Jackson, K. David. 1985. O folclore do Crioulo Português da Índia e de Sri-Lanka (Ceilão). In Lindley Cintra, Luís Filipe (ed.), *Actas do Congresso sobre a Situação da Língua Portuguesa no Mundo*, 339–346. Lisbon: Imprensa Nacional.
- Jackson, K. David. 1990. *Sing Without Shame. Oral Traditions in Indo-Portuguese Creole Verse*. Amsterdam: John Benjamins.
- Jackson, K. David. 1991. Indo-Portuguese *Cantigas*: Oral traditions in Ceylon Portuguese verse. *Hispania* 74(3). 618–626.
- JDBU. 1952. An interesting letter and the answer. *Journal of the Dutch Burgher Union of Ceylon* 42(2). 90–94.
- Joseph, John. 1987. *Eloquence and Power: The rise of language standards and standard languages*. London: Frances Pinter.
- Lewis, Paul & Simons, Gary. 2010. Assessing Endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2). 103–120.
- Lewis, Paul & Simons, Gary & Fennig, Charles (eds.). 2013. *Ethnologue: Languages of the World*, 17th edn. Dallas: SIL International.
- Lüpke, Friederike. 2011. Orthography development. In Austin, Peter & Sallabank, Julia (eds.), *The Cambridge Handbook of Endangered Languages*, 312–336. Cambridge: Cambridge University Press.

- McGilvray, Dennis. 1982. Dutch Burghers and Portuguese mechanics: Eurasian ethnicity in Sri Lanka. *Comparative Studies in Society and History* 24(2). 235–63.
- McGilvray, Dennis. 2007. The Portuguese Burghers of Eastern Sri Lanka in the wake of civil war and tsunami. In Flores, Jorge (ed.), *Re-exploring the links: History and constructed histories between Portugal and Sri Lanka*, 325–347. Wiesbaden: Harrassowitz Verlag.
- Nathan, David & Austin, Peter K. 2004. Reconceiving metadata: Language documentation through thick and thin. In Austin, Peter K. (ed.), *Language Documentation and Description*, vol. 2, 179–187. London: SOAS.
- Newstead, Robert. 1827. *Um curto catichismo da Biblia, disposto em quarento divisaõs, todas as repostas per as perguntas sendo ne as palavras de a Santas Escrituras* (Transl. of W.F. Lloyd). London: J.S. Hughes.
- Newstead, Robert. 1852. *O Novo Testamento de Nossa Senhor e Salvador Jesus Christo* (Transl. of W.F. Lloyd). Colombo: Wesleyan Mission.
- Newstead, Robert. 1871. *Cantigas per adoraçaõ publico, em lingua portugueza de Ceylon*. Colombo: Wesleyan Mission.
- Nordhoff, Sebastian. 2013. The current state of Sri Lanka Portuguese. *Journal of Pidgin and Creole Languages* 28(2). 425–434.
- Oko, Christina. 2018. Orthography development for Darma (The case that wasn't). *Language Documentation and Conservation* 12. 15–46.
- Pieris, Paulus Edward & Naish, Richard Bryant. 1920. *Ceylon and the Portuguese. 1505-1658*. Tellippalai: American Ceylon Mission Press.
- Romaine, Suzanne. 2005. Orthographic practices in the standardization of pidgins and creoles: Pidgin in Hawai'i as anti-language and anti-standard. *Journal of Pidgin and Creole Languages* 20(1). 101–140.
- Schuchardt, Hugo. 1889 Beiträge zur Kenntnis des creolischen Romanisich. V. Allgemeineres über das Indoportugiesische (Asiportugiesische). *Zeitschrift für Romanische Philologie* 13. 476–516.
- Sebba, Mark. 2007. *Spelling and society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.
- Seifart, Frank. 2006. Orthographic development. In Gippert, Jost & Himmelmann, Nikolaus P. & Mosel, Ulrike (eds.), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.
- Serrão, José Vicente. 2014. The Portuguese land policies of Ceylon: on the possibilities and limits of a process of territorial occupation. In Serrão, José Vicente & Direito, Bárbara & Rodrigues, Eugénia & Miranda, Susana Münch (eds.), *Property Rights, Land and Territory in the European Overseas Empires*. Lisbon: CEHC-IUL.

- Sheeran, Anne E. 1998. *White noise: European modernity, Sinhala musical nationalism, and the practice of a creole popular music in modern Sri Lanka*. Seattle: University of Washington. (Doctoral dissertation.)
- Sheeran, Anne E. 2004. Baila music: European modernity and Afro-Iberian popular music in Sri Lanka. In Silva, Neluka (ed.), *The Hybrid Island: Culture crossings and the invention of identity in Sri Lanka*, 146–170. Colombo: Social Scientists Association.
- Smith, Ian (coll.). 1973. *Indo-Portuguese (Sri Lanka) (IRS01)*. PARADISEC. (<http://catalog.paradisec.org.au/collections/IRS01>) (Accessed 2019-01-28.)
- Smith, Ian. 1977. *Sri Lanka Creole Portuguese phonology*. Ann Arbor: University of Michigan. (Doctoral dissertation.)
- Smith, Ian. 1979a. Convergence in South Asia: a creole example. *Lingua* 48. 193–222.
- Smith, Ian. 1979b. Substrata vs. universals in the formation of Sri Lanka Portuguese. *Papers in Pidgin and Creole Linguistics* 2. 183–200.
- Smith, Ian. 1984. The development of morphosyntax in Sri Lanka Portuguese. In Sebba, Mark & Todd, Loreto (eds.), *York papers in Linguistics*. York: University of York.
- Smith, Ian. 2013a. Sri Lanka Portuguese. In Michaelis, Susanne & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *The Survey of Pidgin and Creole Languages*, vol. 2, 111–121. Oxford: Oxford University Press.
- Smith, Ian. 2013b. Sri Lanka Portuguese structure dataset. In Michaelis, Susanne & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.), *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://apics-online.info/contributions/41>) (Accessed 2019-03-20.)
- Smith, Ian. 2016. The earliest grammars of Sri Lanka Portuguese. *Papia* 26(2). 237–281.
- Tavares de Mello. 1998. *Folclores ceilonenses; Colectânea de textos do crioulo português do Ceilão*. Lisbon: Comissão Nacional para as Comemorações dos Descobrimentos Portugueses.
- Tomás, Maria Isabel. 1992. *Os Crioulos Portugueses do Oriente; Uma Bibliografia*. Macau: Instituto Cultural de Macau.
- Wittenburg, Peter & Brugman, Hennie & Russel, Albert & Klassmann, Alex & Sloetjes, Han. 2006. ELAN: A Professional Framework for Multimodality Research. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, 1556–1559. (<http://www.lrec-conf.org/proceedings/lrec2006/>) (Accessed 2019-01-24.)
- Woodbury, Anthony. 2014. Archives and audiences: Toward making endangered language documentations people can read, use, understand, and admire. In Nathan, David & Austin, Peter K. (eds.), *Language Documentation and Description*, vol. 12, 19–36. London: SOAS.