Factoring sociolinguistic variation into the history of Indo-Portuguese[1]

Hugo C. Cardoso

Universidade de Lisboa

Given that the Portuguese-lexified creoles of South Asia are currently spoken by relatively small and close-knit communities, it is often assumed that little sociolinguistic variation will be found within them and/or that this must have been the case throughout their history. This article, however, explores how modern and earlier (18th-/early-20th-century) descriptions and commentaries on the Indo-Portuguese communities reveal a somewhat stratified linguistic repertoire not unlike that of various other creoles around the world. Based on a survey of sources of different periods and character, I argue that there is good reason to assume a considerable degree of sociolinguistic variation in the Indo-Portuguese communities even when the available sources do not record it, and that failure to do so may compromise the validity of our linguistic studies.

Key words: Indo-Portuguese creoles; Sociolinguistic stratification; Variation; Diachrony; Synchrony.

1. Introduction

The Indo-Portuguese creoles, which were once widespread across coastal Asia, are currently spoken in only a few locations in India and Sri Lanka (for an overview, see e.g. Cardoso 2006). Fortunately, a history of documentation of over two centuries means that we have a considerable corpus of written samples covering even some of the varieties which are now extinct. The geographical distribution of these records, however, is by no means homogenous: some varieties (e.g., those of Mahé or Nagapattinam) are recorded in only one or very few documents, while others, with Sri Lanka having a clear advantage, have a large body of written texts.

Differences between the various Indo-Portuguese creoles are well-established and have been the focus of some recent research (see e.g. the contributions in Cardoso, Baxter & Nunes 2012). But, since these creoles are now spoken in non-contiguous locations and, typically, by relatively small and close-knit communities, one might assume that each one constitutes a homogeneous speech community with little internal variation – and/or that this has been the case throughout their history.

However, closer inspection of recent and early accounts of these languages reveals interesting patterns of linguistic variation whose distribution is governed by specific social variables. This article therefore addresses questions relating to the extent, roots and diachrony of sociolinguistic variation in the Indo-Portuguese creoles. Yet, in the interest of brevity, the study is restricted to the Indian varieties.[2] In order to do so, I survey and discuss evidence of variation in several written sources – whether published or archival. I will begin by discussing the points made by some of the earliest descriptions with respect to community-internal variation (section 2), before moving on to an exploration of more recent descriptive studies (section 3). My argument is that there is good reason to assume a considerable degree of sociolinguistic variation in the Indo-Portuguese communities in the past as well as in the present (even when the available sources do not record it), and that failure to do so may compromise the validity of our research on the Indo-Portuguese creoles (section 4).

2. Early accounts of variation

The first available attestations of the Indo-Portuguese creoles owe much to the work of civil administrators and missionaries. As we will see in this section, some the earliest references to the use of Portuguese-lexified creoles in India (in the 18th century) were made by protestant missionaries, and the first known grammar of one of these creoles (that of Sri Lanka) was composed for the benefit of "the English Gentlemen in the Civil and Military Service on Ceylon" (Berrenger 1811).

The late 19th-century and early 20th-century was a particularly important period for the documentation of the Indo-Portuguese creoles, when

---

[2] The Sri Lanka creoles, in fact, have already motivated important discussions and hypotheses about sociolinguistic variation (see e.g. Smith 1979; Flores 2004) and should also benefit from this exercise. However, addressing these varieties would require the examination of a rather large body of early written evidence, which cannot be done here.

philologists such as Hugo Schuchardt, Adolfo Coelho, Sebastião Rodolfo Dalgado and José Leite de Vasconcelos showed an interest, collected data and wrote about these languages. Their datasets and linguistic descriptions, some of which were never published but can be retrieved from archives, afford us a look at the state of the Indo-Portuguese varieties in various locations: Diu, Daman, the Bombay region, Goa, Mangalore, Cannanore, Mahé, Cochin, Nagappatinam, and Sri Lanka.[3] In terms of approach and scope, this first generation of creolists represents a substantial break with the work of previous authors. Therefore, this section will be sub-divided into two periods: before and after Schuchardt.

2.1. Pre-Schuchardt sources

Early references to internal linguistic variation in Indo-Portuguese are provided by Lutheran missionaries operating in the southeastern Indian city of Tranquebar. In the early 18th century, the Danish monarch Frederik IV (who controlled Tranquebar) invited missionaries from Halle (Germany) to set up a Lutheran mission in that city – the so-called Danish-Halle Mission, established in 1706. Upon their arrival in South India, the Halle missionaries encountered substantial Portuguese-speaking Christian communities, in Tranquebar and elsewhere, which became an important target of their activities – this is evidenced by their extraordinary production of Portuguese-language liturgical and pedagogical literature in the 18th century (see e.g. Iken 2000).

The letters and documents sent back to Halle by the Tranquebar missionaries were later collected and published in a series of volumes wherein frequent mention is made to the local 'Portuguese' (for an overview, see Pfänder & Costa 2006).[4] One of the earliest mentions of linguistic variation among this community can be found in a letter written by Christoph Theodosius Walther dated September 1726:

> In Portuguese, the difference between the high language (as it is spoken in Portugal and is used in books) and the low language (as it has been corrupted here in India) is like the Italian used in Turkey, only much greater. (quoted and translated in Pfänder & Costa 2006: 1157-1158)

---

[3] For further details on the corpus of Indo-Portuguese texts from this period, see e.g. Tomás (1992a); Ladhams (2009); Cardoso, Hagemeijer & Alexandre (forthcoming).

[4] These sources provided a great deal of the information that Schuchardt published in his general description of Indo-Portuguese (Schuchardt 1889a).

This brief description is significant on several counts. On the one hand, it clearly establishes that the Portuguese spoken in India was a "corrupted" form very distant from the one spoken in Portugal; presumably, the reference to the Italian language spoken in Turkey could reveal a consciousness that this process of differentiation owed much to language contact resulting from out-migration. The reference to written language is not as clear as one might wish, but it probably indicates that the "high" (and not the "low") variety was used for writing in India as well.

Another important document in the Danish-Halle archives, included in the 1733 volume, reinforces these ideas but provides much more detail, including the fact that this "high" variety of Portuguese was not alien to India. This is a report called 'Nachricht von den Portugiesen in Indien' [Information about the Portuguese in India] written by Nikolaus Dal, a general description of the Portuguese-speaking Christian population of South Asia covering several aspects: their historical origin, clothing, colour (i.e. race), geographical distribution, religion, occupation and language. For Dal, in India, the term "Portuguese" was applied to anyone who spoke Portuguese and dressed in a Portuguese fashion – these being their most salient identity markers – but covered a considerable racial diversity: whites, blacks (in his words), and several intermediate categories reflecting interracial ancestry ("Mestissen", "Castissen", "Postissen"). Significantly, the section devoted to language (full transcription in Appendix A; see Lopes 1936: 53-57 for a partial Portuguese translation) describes the Portuguese spoken in India in a similarly stratified fashion:

> It should thus only be mentioned that the language is not the same for everyone. For as you have the Portuguese divided into three classes, you can also distinguish three types of the language, namely, 1) the proper, 2) the semi-corrupt, and 3) the entirely corrupt. (Dal 1733: 919, my translation)

Linguistic variation is described here in terms of a continuum of linguistic strata defined in relation to a model of "purity" (the "proper Portuguese" as spoken by Europeans and their direct descendants). Dal also provides a general characterization of the "semi-corrupt" and "corrupt" types of Portuguese, though brief and limited to the domain of the verb phrase. Concerning the intermediate type, the "semi-corrupt", he says:

> This type of language is characterised mainly by the fact that they are unable to use verb conjugations correctly. (Dal 1733: 919, my translation)

Incorrect verbal inflection (from the perspective of European Portuguese or, to be precise, of an L2 speaker of European Portuguese) therefore seemed to be the most salient characteristic of this variety, though there would have been other differences. "Incorrect", however, does not necessarily mean "inconsistent" or "irregular", simply different from an established standard. As a matter of fact, Dal appears to indicate that the "semi-corrupt" Portuguese did in fact have verbal inflection of some sort; this becomes clear when contrasted with the third type, the "entirely corrupt" Portuguese, described in the following terms:

> This kind of language is different from the previous one mainly in that people do not use conjugation at all, instead expressing future tense with the particle *lo*, and past tense with *ja*[d], and construct the infinitive by omitting R.

> (d) *Lo* is derived from logo (as some pronounce it), i.e. immediately, soon. An Indian must have heard a Portuguese say e.g. *logo virei*, I will come soon; and gathered for instance that a future thing must be expressed with *logo*. *Ja* means already. A Portuguese must have said e.g.: *ja ouvi*, I have already heard it; and the Indian thought that past things must take *ja*. The many endings must have become too difficult for him, as one can easily imagine. (Dal 1733: 919-920, my translation)

The verbal system of this variety is therefore described as consisting of invariant verb forms modified by preverbal particles *ja* for past and *lo/logo* for future. These particles are well-established in most Luso-Asian creoles (see Ferraz 1987) with these functions, so Dal seems to be describing an Indo-Portuguese creole. The absence of -*r* in infinitives is also consistent with what we know of the Indo-Portuguese creoles; its mention in this particular context could also indicate that the base form of verbs in this "entirely corrupt" variety derived from Portuguese infinitives, but this is not made entirely clear.

Given their characteristics, one can perhaps consider the "entirely corrupt" variety to correspond to the basilectal form(s) of Indo-Portuguese creole, with "semi-corrupt" describing the acrolectal form(s) of the creole. Dal further indicates that these varieties were distributed in society essentially along racial/ethnic lines; but he also clarifies that certain speakers had command of more than one of these and that there was diglossia, with the basilect being disfavoured in particular settings and media:

> 1. The proper Portuguese is mainly spoken by the European Portuguese and also by those who descend from them. 2. The semi-corrupt speech is generally spoken by those of mixed parentage. But the blacks also make use of it in their letters. […] 3. The entirely corrupt is generally heard from the Portuguese of mixed descent, but especially from those who are completely black. […] This language is considered silly and

ridiculous in letters. One should think that it is not very appropriate for speaking; which is especially true of coherent speech.

This is why the black Portuguese use the half-corrupted language when they pray to God from their hearts. But in incoherent speech this language can still be used: it is generally spoken when the English and the Danes, the Dutch and the French, Europeans and Armenians meet in India and it is not possible for them to communicate in their native languages.[e]

[e] These nations often exchange letters with each other in the proper or semi-corrupt Portuguese language, depending on whether they can get a good writer or not. (Dal 1733: 919-920, my translation)

Since this is a general description of the "Portuguese" in South Asia, and not of a particular location, one might wonder whether or not various strata coexisted in particular communities. It is especially doubtful whether, at this time, European Portuguese had any currency in all of the locations in which a creole existed. But, at least with respect to the interaction between acrolectal and basilectal varieties of Indo-Portuguese creole, Dal is rather explicit when he describes the option for the acrolect in prayers. Another hypothesis which this description seems to dispel is that the acrolect which one finds in writing was not used in unconstrained oral discourse. In fact, the group described as being of "mixed parentage" is said here to speak the acrolect in general, even if some also had more basilectal varieties in their repertoire.

Another important description of the use of "Portuguese" throughout South India in the late 18th century is found in Abraham Hyacinthe Anquetil-Duperron's 1786 'Des recherches historiques et géographiques sur l'Inde' [Historical and geographical researches on India]. Anquetil-Duperron was one of the leading French Orientalists of the 18th century, with a particular interest in Zoroastrianism. He resided in India between 1755 and 1761 and was especially familiar with South India. This particular work is a collection of studies about Indian history and archaeology but opens with a summary of the languages and types of government of the subcontinent. At a given point, he makes reference to the currency of Portuguese in South India (transcribed in Appendix B; for a partial Portuguese translation, see Lopes 1936: 84-85) in the following terms:

We know that the Portuguese were the first Europeans to establish settlements in India. Since, earlier on, this nation did not take the trouble to learn their [i.e., the local] languages, they have therefore, somehow, forced the Natives to learn their own. The descendants of the first Indians dominated by the Portuguese, Christian or otherwise, in general know it. It has expanded with them along both coasts; and, since most of the servants in the service of the French, English, Dutch, and Danes descend from these

first Indians, or from the Abyssinian Slaves of the Portuguese, all foreigners, when they arrive, find themselves in need of learning *Portuguese.*

As a result, the Indian, Moor, Arab, Persian, Parsi, Jewish, and Armenian merchants who deal in the European trading posts, as well as the Blacks who wish to serve as Interpreters, are forced to speak this language; it also serves as means of communication between the European nations established in India.
[…]

In our Establishments, affairs are generally carried out with the Natives of the country, and even with the other European nations, by means of the *Portuguese Jargon* that I have just mentioned; otherwise, we would be limited to the expedient, which is at least dangerous, of having to employ renegades or other highly suspicious people. (Anquetil-Duperron 1787: XI-XII, my translation)

One lesson to glean from this description is that, in the late 18th century, the status of Portuguese – and, specifically, the variety described as *Portuguese Jargon* – as a *lingua franca* of trade and domestic communication in South India would have made for a much larger speech community than one is used to imagining, and one that encompassed large numbers of L1 and L2 speakers alike. As we will discuss in a moment, the size of a speech community is likely to correlate with the degree of internal variation – especially if the community is spread out over such a wide territory as this description indicates. This fact alone would be enough for one to hypothesize a very high degree of variation. But Anquetil-Duperron makes other comments about this *Jargon* which are rather explicit in that respect, and consistent with the Danish-Halle missionary reports we discussed earlier:

But this language is far from being the pure Portuguese, called in India *Reinol Portuguese*. The one that is written down comes closest to it, especially in the Malabar Coast, where that nation had numerous settlements: spoken Portuguese is in effect no more than a jargon of 150 or 200 words, with almost no structure. (Anquetil-Duperron 1787: XI)

While Anquetil-Duperron is less explicit than Nikolaus Dal in his classification of different registers of "Portuguese", a similar type of variation can be deduced from this text. *Reinol* was the term applied in Portuguese India to a person from the *reino* 'kingdom', i.e., from Portugal. By *Reinol Portuguese*, Anquetil-Duperron is therefore referring to L1 European Portuguese or a close local variant. His description of the *jargon* effectively used as a means of interethnic communication ("150 or 200 words, with almost no structure") and spoken by communities of Indians descended from those "dominated by the Portuguese" is very much in tune with the prejudices then associated with pidgins and creoles.

Like Dal, Anquetil-Duperron is also clear about the existence of a register that came particularly close to European Portuguese and which was the only one committed to writing. Unlike Dal, however, he appears to believe that this acrolectal form was exclusive to the written medium, and did not occur in oral speech. Another important detail in Anquetil-Duperron's report is that the written variety was especially acrolectal in the Malabar, the southwestern coast of India where the Portuguese established several of their earliest and most important settlements, such as Cochin, Quilon and Cannanore.

2.2. The first generation of creolists

The last point made about Anquetil-Duperron's description is especially relevant as we approach the significant body of work that Hugo Schuchardt dedicated to Indo-Portuguese, given that a large part of it concerned data collected in southwestern India. This is the case of his articles on the creoles of Cochin (1882), Mangalore (1883b) and Cannanore/Mahé (1889b). Schuchardt had little control over the quality of his samples, since his method of data collection involved the intervention of different correspondents in the field. As a result, his data is full of contradictory information and inconsistencies, which Schuchardt himself detected and tried to make sense of. Let us take his Cochin data as an example. Schuchardt divides the corpus he received from the Anglican bishop of Travancore, who was at the time John Martindale Speechly, into two sections: A (consisting of several dialogues and short texts) and B (several texts of a religious nature). Despite the scarcity of information, Schuchardt nonetheless advances a generalization which appears supported by Nikolaus Dal and Anquetil-Duperron's observations:

> Die Sprachfärbung ist in beiden eine wesentlich verschiedene, was ich mir nur so zu erklären vermag, dass *A* die kreolische Mundart in ihrer natürlichen, charakteristischen Ausprägung darstellt, *B* jedoch in einer der Schriftsprache angenäherten Gestalt, wozu sich auch die grösstentheils religiöse Materie eignet. (Schuchardt 1882: 3)

> [The shade of the language is substantially different in the two of them, such that I can only surmise that A is the creole in its natural and characteristic form, while B approaches the written form, which is generally more suited to religious matters.]

To illustrate the differences between the two, consider the following sentences:

Corpus A

(1)   *Por   quem   ja   vi,   tama   jada;   nos   tama*

      DAT   who   PST   come   also   PST-give   1p   also

      ja   cume (Schuchardt 1882: 804)

      PST   eat

      'We gave [some food] to whoever came; [and] we also ate'.

Corpus B

(2)   *Abrahão   recebeo   este   homem   com   todo   bondade,*

      Abraham   host.PST   DEM   man   with   all   kindness

      labou   seu pe   e   aperelhou   cea   e   deu

      wash.PST   his   foot   and   prepare.PST   supper   and   give.PST

      ele   lugar   por   sentar (Schuchardt 1882: 808)

      him   place   to   sit

      'Abraham welcomed him with great kindness, washed his feet,
      prepared supper and gave him a place to sit'.

The distinction between the two is somewhat (though not entirely) predictable. In general, it is the case that, when the two corpora differ, B tends to resemble Portuguese more closely whereas A has more traits that we recognize from modern-day Cochin creole. If we focus on the verbal system, for the moment, we will notice that verb forms look rather different in the two examples. Whereas, in (1), past tense is expressed with a pre-verbal marker *ja* (as in modern-day Cochin creole) – e.g. *ja vi, ja cume* –, in (2) verbs inflect for past tense producing forms which we recognize as Portuguese 3s past tense forms – e.g. *recebeo*, *lavou*. In (2), word order is strictly SVO, but the first clause in (1) has the indirect object preceding the verb. As a matter of fact, modern-day Cochin creole is strictly verb-final, while Schuchardt's corpus B is strictly verb-medial (and corpus A has both, see below).

On the other hand, there are certain differences in which this type of affiliation is not as straightforward. Notice, for instance, how the recipient in (1) is introduced with a dative preposition *por*, whereas in (2) the recipient is bare. An object marker *pə* (derived from *por* or *para*) is used in modern Cochin creole, though it is postposed rather than preposed to the noun; on the other hand, Portuguese cannot easily account for the construction in (2), given that a pronoun such as *elle* (which is a non-clitic nominative form) could not occur as an indirect object without a preposition. English double-object constructions may have motivated this type of construction. As a matter of fact, close inspection of corpus B will reveal several cases of what looks like word-for-word translations of English constructions, adding another significant layer to the roots of the variation encountered in Schuchardt's materials. Two further examples of this are given in (3) and (4):

(3)   *Bom   palavras   custa   nada*

    good   word.PL   cost   nothing

    'Good words cost nothing.'

(4)   *Eu   suffrio   elle   este   centa   annos*

    1s   suffer.PST   3s   DEM   hundred   year.PL

    'I have suffered him for a hundred years.'

In a sentence such as (3), both Portuguese and modern-day Cochin creole would require negative concord, but not English. English influence is also visible in (4) in the use of the verb 'to suffer' with a human object resulting in a reading close to 'to endure'; this use of 'to suffer' is not attested in Portuguese.

The realization that there is a good deal of variation in both these corpora raises another issue concerning the language used therein. In corpus A, for instance, which can generally be considered to reflect oral basilectal speech more closely, one still finds certain inconsistencies and competing strategies. Consider, for instance, the shape of genitive constructions. In modern-day Cochin creole, as in most Luso-Asian creoles, genitives have the structure [Possessee-*sə* Possessed], while in Portuguese the structure is [Possessed *de* Possessee]. Corpus B, as expected, features Portuguese-like genitives (e.g. *misericordia de Deos* 'God's mercy'), but corpus A contains

both: *cabeça de bagri* 'fish's head' vs. *Manchu su luguer* 'the boat's rental'. Yet another case of variation in this corpus concerns the coordination of noun phrases, which in Portuguese is operated by the sentential coordinator *e* and in modern Cochin creole (as in many other Luso-Asian creoles) by *ku*, derived from the Portuguese comitative/instrumental *com*. Here, again, corpus A contains both options: *Senhor Padrinho e Madrinho* 'Mister Godfather and Godmother' vs. *prezidente com hum servidor* 'the president and a servant'. Finally, another domain in which variation can be identified in corpus A concerns word order. Whereas, as we have already mentioned, OV clauses do occur (as in the modern-day creole), in reality VO is the norm here as well.

There are several possibilities to account for these corpus-internal inconsistencies. One, of course, is that they reflect grammatical subtleties which the limited amount of data does not fully clarify. Another is that the various texts were produced by different informants, in which case they would reflect individual lects; one complicating factor in this scenario, however, is that contrasting options often occur within single texts. Finally, it is possible that normative corrections surface even in texts which are more sensitive to the depiction of basilectal speech (such as corpus B).

If indeed the identified inconsistencies are the result of normative correction, then, a follow-up question is whether these were the speaker's doing or were added by the data collector (in case they are not one and the same person). A different set of data published by Schuchardt, from Mahé, helps to clarify this somewhat. To collect data in the French colony of Mahé, close to Cannanore, Schuchardt enlisted the help of pastor W. Schmolck. The pastor, in turn, resorted to two informants to complete Schuchardt's elicitation questionnaire: Mr. Rozario and Mr. d'Cruz. As a result, Schuchardt received two different versions of the same set of sentences. This is especially significant in that the two sources often contain starkly different strategies. In his publication on the creole of Cannanore and Mahé, Schuchardt condenses the contradictory data in a single line, with the following format:

> *Quilai* [*Como*] *tem vos?*
> *Hoje te faze muito callor* [*calor*].
> […]
> *Vos jà olha* [*Ja olha vos*] *palacio de Ré* [*do Rey*]*?*
> *Tinteira tem riva* [*riba*] *de meza* } = [R. giebt nur das Erstere]
> *Mezasse riva tinteira tem*
> […]   (Schuchardt 1889b: 516-517)

97

In hindsight, it is not difficult to determine whether a particular option is more basilectal (for instance, the interrogative form *quilai* 'how', attested in many Luso-Asian creoles) or more acrolectal (the corresponding Portuguese interrogative *como* 'how']. However, Schuchardt transcribes the data in such a way that one is unable to gather which of the two informants was responsible for which of the variants, or which variants go together. The only time that any mention is made to the source concerns the last two sentences transcribed above, which constitute two very different responses to the same question. In front of them, Schuchardt added "R. giebt nur das Erstere", meaning that R. [i.e. Rozario] gave only the first of the two constructions – which implies that d'Cruz offered both. Interestingly, the two sentences are highly divergent and it is clear that the second option (*Mezasse riva tinteira tem* 'the inkpot is on the table') is much more like the modern-day creoles of the region: the existential verb (*tem*) is clause-final and the genitive (*-se*) is postposed to the noun; the first sentence, on the contrary, is much more Portuguese-like, which can be seen in the medial position of the verb and the prepositional genitive construction with *de*. The fact that only d'Cruz produced both sentences indicates not only that a single speaker could command the variation internal to the speech community, but also that he may have been responsible for all the more basilectal options. Close scrutiny of the two manuscripts (preserved in the Hugo Schuchardt archive at the University of Graz, doc. 11.23.14.5), which allows us to tease apart the production of both informants, confirms this. As illustration, I transcribe in Table 1 the translations that Mr. d'Cruz and Mr. Rozario provided for a few of the sentences in the questionnaire:

Table 1: A sample of Mr. d'Cruz and Mr. Rozario's translations

Sentence	Mr. d'Cruz	Mr. Rozario
a. How are you?	*Quilai tem vos?*	*Como tem vos?*
b. Why hasn't the master come see us?	*Porque mestro nunca vi olha por nos?*	*Porque he que mestre nunca vi olha por nos?*
c. You must not blame anyone.	*Não deve culpar ninguém.*	*Não deve vetuperar a ninguém.*
d. The inkpot is on the table.	*Tintera tem riva de meza. / Mezasse riva tintera tem.*	*Tinteira tem riba de meza.*
e. We are eating fruits.	*Nos fruitos te cume.*	*Nos te come frutos.*
f. Roof of the house.	*Cazás cumé.*	*Cume de caza.*
g. The doctor has gone to Luis' house twice.	*Dois vez doutor jafoi Luisse caza.*	*Dois vez doutor jafoi na caza de Lui.*
h. Can you dance?	*Vos te sabe dança?*	*Sabe vos dançar?*

In order to interpret the variation in Table 1, it is important to consider what we know of the modern-day creoles of the region (especially that of nearby Cannanore) but also Portuguese and French, which was used as a language of administration in 19th-c. Mahé. All in all, it is clear that, as suggested earlier, d'Cruz's data is much more like the modern creole – we will use the term 'basilectal' for short – and Rozario's shows much more influence from Portuguese and French. Let us look at some evidence of this:

- Though inconsistently, d'Cruz is the only one to provide verb-final sentences: this is true of the sentence we have discussed earlier (sentence d.) and also of sentence e., in which the transitive verb *cume* 'eat' occurs after the direct object *fruitos* 'fruits';

- d'Cruz is also the only one to produce the post-nominal genitive: we have already seen this for sentence d., but here it occurs also in f. (where *–s* in *cazás* 'of the house' surely stands for *-sə*) and in g. (*Luisse = Luis-sə* 'of Luis'); for all of these, Rozario uses the Portuguese-style construction with the preposition *de*;

- In modern-day Cannanore creole, locative marking of goals is not necessary with verbs of motion: this is reflected in d'Cruz's translation of sentence g., in which *Luisse caza* has no locative marker; Rozario, on the other hand, uses the locative preposition *na*;

- Rozario shows French influence in the syntax of polar questions: in sentence h., we see an inversion of the verb and the subject characteristic of French (cf. Fr. *savez-vous dancer?*);

- The choice of forms is also relevant: in sentence a., d'Cruz is responsible for the typically creole interrogative *quilai* 'how' and Rozario for the Portuguese form *como* 'how'; in c., Rozario uses a form of the learned and infrequent Portuguese verb *vituperar* 'to revile', whereas d'Cruz chooses the more colloquial verb *culpar* 'to blame'.

In the case of Mahé, then, acrolectal and basilectal speech coexisted, even though some speakers (like Rozario) may have admitted only one of them in writing. In fact, Schmolck's letter makes a further observation which is quite revealing. He explains that Mr. d'Cruz was the first one to fill out the

questionnaire, and that Mr. Rozario had access to it while producing his own version; Rozario's version was, therefore, in effect a correction of d'Cruz's, one that eliminated forms and constructions we now find in the modern creole in favour of Portuguese-inspired (and French-inspired) ones. This is perhaps the clearest example we have of how the written data collected for Schuchardt actively concealed established basilectal speech forms.

In other articles, Schuchardt's corpora also open the possibility of a particular type of linguistic variation based on ethnicity. This concerns the African and African-descended section of the Indo-Portuguese-speaking communities (see also Tomás 1992b; Cardoso 2010). His article on the creole of Diu includes three songs classified as *Negerlieder*[5] 'Negro songs', which formed part of the particular oral traditions of this population. One of these is a Christmas song, recorded in this article (Schuchardt 1883a: 13) as:

Sam Paulo, já bate cino,	Saint Paul's, the bell has struck
Meia noite, já nacê minino,	Midnight, the child is born
Meia noite, já nacê minino.	Midnight, the child is born

Interestingly, the final verse resurfaces in Schuchardt's study of Mangalore (1883b: 889):

Pai Jose ja mata cavallo,	Father Joseph broke fast [Lit. "killed the horse"]
Secco secco manda bata sal,	Very dry, has salt added [to it]
Cafrinha, ja repica sino	The black [person] has struck the bell
Meia noite, ja nasce menino.	Midnight, the child is born.

In this particular text, the song is not specifically attributed to the African section of the population. However, the reference to a *cafrinha* (a diminutive of *cafre*, a common word in Portuguese India to designate a person of African descent) is surely not a coincidence. Besides, there is even further evidence that this particular song was associated with the African population in the fact that another scholar of Indo-Portuguese, Sebastião Rodolfo Dalgado, later identified the same verse in a Goan song which he attributed to the presence of African(-descended) slaves there (Dalgado 1921):

Já tocá xinu Xanta Dominga; mea noiti já naxê Minino

'The bell struck at Saint Dominic's; midnight, the child was born'.

---

[5] *Cantigas dos prêtos de Diu* in the original document (Hugo Schuchardt archive at the University of Graz, doc. 11.23.14.8).

Later on, António Moniz also published a collection of Damanese creole songs associated with the African population. An excerpt is transcribed below (Moniz 1925: 570):

Minha senhora Maria	My lady Mary
Já traze para vos tangê, hêhêhê,	Brought for you to play, *hê hê hê*
Tange sua bobra	To play your gourd
Buli seu pê Ai lê lê lê	And move your feet *Ai lê lê lê*
Tangê sua bobra buli seu pê	Play your gourd and move your feet
Muzungo pequeninha	The child [Lit. "small person"]
Qui já nascê, Ai lê lê lê lê lê	Which is born, *Ai lê lê lê lê lê*
[…]	

All of this confirms that, with respect to oral literature at least, the African(-descended) population of Portuguese India had a specific repertoire which set it apart from other ethnic groups. Whether or not this corresponded to a separate linguistic lect is another matter. To be clear, there is nothing in the grammar of these songs that strikes us as peculiar in the context of the Indo-Portuguese creoles. Only some lexemes do, but even these are not very numerous. In Moniz's song, for instance, we have the occurrence of *muzungo* 'person', a word with cognates in several East African languages, in which it refers to a 'white man/European' (Cardoso 2010). Despite the weak evidence, however, Dalgado was convinced that there had once been a peculiar version of Indo-Portuguese creole spoken by the African population, which he called *cafreal* creole (see Dalgado 1921). The possibility therefore remains open that, in earlier times, Indo-Portuguese comprised different ethnolects.

In Dalgado's descriptions of the Indo-Portuguese creoles, which (in addition to the study of this Goan song) cover the varieties of Daman, the Bombay area, Nagapattinam and Ceylon (Sri Lanka), there are other references to community-internal variation. His study of the Bombay region is especially interesting in this respect. For this study, Dalgado relied on data collected by a local correspondent, Mr. Pascoal Rebello. The region analyzed, the Bombay area, was in fact large enough and the speech community numerous enough to allow for considerable variation, and this is reflected in the data provided by Mr. Rebello. He resided in the parish of Tecelaria (Taná district) but also sent samples reflecting the varieties of Chevai and of the city of Bombay proper. Dalgado treats these as "subdialects" and offers the following observations:

> Observa o meu correspondente que o dialecto fallado, com exclusão de outra língua, pelos trezentos habitantes da sua freguesia, differe notavelmente dos das outras partes e se aproxima muito do de Damão. […]
>
> Pelo exame dos espécimes dos subdialectos se deprehende que o de Bombaim e suburbios conserva algumas flexões nominaes e verbaes, tem menos alterações phoneticas e se resente mais da influencia do inglês, não só na lexicologia, mas até na morphologia, devendo, por isso, considerar-se como misto e hybrido. O de Tecelaria e os de outras localidades distantes, deixados á livre evolução, sem grandes pressões estranhas, desenvolveram-se espontaneamente em crioulo propriamente dito, com algumas peculiaridades individuaes, de pouca monta, que se não generalizaram, ou por motivo de isolação, ou por antagonismo de fórmas preexistentes. (Dalgado 1906: 144)

> [My correspondent observes that the dialect spoken, in detriment of another language, by the three hundred inhabitants of his parish differs considerably from those of the other areas and comes very close to that of Daman. […]
>
> Through the examination of the specimens of the subdialects, we notice that the one of Bombay and its suburbs retains some nominal and verbal inflection, has less phonetic changes and suffers greater influence from English, not only in the lexicon but even in morphology, and therefore it must be considered mixed and hybrid. The one of Tecelaria and other remote places, left free to evolve without much external pressure, developed spontaneously into a creole proper, with some minor individual peculiarities which did not become generalised, whether because of their isolation or the competition of prior forms.]

In this passage, the distribution of acrolectal and basilectal speech in the Bombay area is described in geographical terms. This begs the question of whether the different subpopulations did interact with one another, i.e., whether one can really speak of a single speech community in this territory. Dalgado does hint at a certain isolation of e.g. Tecelaria, but the fact that Mr. Rebello, a resident of Tecelaria, was aware of the speech of other areas probably says otherwise. If indeed there was regular interaction between the various groups, then we are probably dealing with a kind of community-internal variation along the lines of what Clements has described in recent times for Daman (see below), in which geographical variation in fact reflects differences in the social profiles of the inhabitants. That could very well have been the case in 19th-century Bombay and surroundings. It is perhaps significant that the acrolectal variety, in this description, is associated with the city of Bombay, an urban area in which one may expect a white-collar, more highly-educated and affluent population to have resided. But speech was not homogeneous even within the city. In fact, Dalgado's informant also notes that

> […] a fala das mulheres em Bombaim é mais crioulizada que a dos homens, mas não indica a differença; deve naturalmente ser ao mesmo tempo lexica, phonetica e morphologica.

> [the speech of the women in Bombay is more creolised than that of the men, but he does not indicate the differences; naturally, they must be at once lexical, phonetic and morphological.]

This is the only explicit reference in our study to a gender-based pattern of variation. Even though we do not know the details, this short comment clarifies that women's speech was more basilectal than men's. This is consistent with our proposal that variation in the Bombay region depended on social profiles more than just geography; in 19th-century Bombay, women can be expected to have had a lower degree of formal education and less public occupations than men – and, as consequence, less exposure to European languages such as English or Portuguese.

3. Recent accounts of variation

The outstanding production of the first generation of creolists was followed by a hiatus, in the mid-20th century, in which very little attention was paid to Indo-Portuguese. Scholarly attention resumed from the 1970s onwards, producing linguistic descriptions of the extant varieties, including those of Sri Lanka, Korlai and Diu. Descriptive work is still progressing, not only for these varieties but also for those of Daman, Cochin and Cannanore.

Despite this, however, few recent studies of the Indo-Portuguese creoles of India have specifically aimed at correlating linguistic variation with social variables. The ones that make reference to this matter, more or less explicitly, refer to the modern-day creoles of the northern area (Korlai, Daman, Diu). Clements, whose work has focused primarily on the creoles of that region, is particularly active in this respect, and one of his earliest studies of Korlai (Clements 1990) is especially relevant here. In it, the author explores what is interpreted as an ongoing typological change transforming Korlai Portuguese from a verb-medial (SVO) into a verb-final (SOV) language. In order to capture the trend and identify its roots, Clements focuses on a few instances of linguistic variation seen as indicators of this process: a) the realization or deletion of redundant preverbal tense-aspect markers *ja* (past) and *tE* (present perfect/continuous); b) the realization or deletion of prepositions of goal/location with predicates of motion; c) the order of nouns and adjectives; d) the location of adpositions relative to the noun; e) the order of verbs and objects/complements. To determine the direction of linguistic change, the study has a diachronic component which compares the modern

state of the language with early 20th-century written records from the Bombay region but, interestingly for us, it also analyses the current social distribution of the competing options for a), b), d), and e). In order to do that, Clements divides his informants into three groups which essentially diverge with respect to their place of residence (in Korlai or outside) and the degree of contact with Marathi (the regionally dominant language) and Portuguese (the major lexifier):

- Group A: speakers who lived outside of Korlai for a large part of their lives;
- Group B: residents of Korlai who had significant contact with speakers of Marathi and of Portuguese;
- Group C: residents of Korlai who had no significant contact with speakers of Marathi or of Portuguese.

The additional variables considered include: age and sex of the informants, narrative type, process of data collection, fluency in English, or competence (native/nonnative) of the interlocutor in Korlai Portuguese. The author concludes that the sociolinguistic variables determining the distribution of linguistic variants are not the same for all of the instances of variation analyzed:

- The retention of the past marker *ja* is especially characteristic of group A and, in the other groups, occurs only (in much smaller proportions) in storytelling. In Clements' interpretation, both long-term absence from Korlai (representing a certain cut-off point from community-internal linguistic changes) and storytelling favor an archaic register of the language;
- The retention of the present perfect/continuous marker *tE* does not seem to respond to sociolinguistic distinctions but rather to a set of linguistic constraints which hold across the various groups analyzed;
- Group B is the only one that disfavours prepositions of goal/location with predicates of motion; Clements' interpretation hinges on the fact that Marathi, to which this group has a particularly strong exposure, does not have any preposed markers and, therefore, may have influenced the speech of these informants in the direction of deleting all such markers in Korlai Portuguese;
- Adpositions may be preposed or postposed to the noun, either optionally or obligatorily; and all groups of speakers admit the various combinations of these factors; however, congruently with the previous results, it is also

group B that has the least proportion of obligatorily preposed adpositions, which, once again, is ascribed to the influence of Marathi;

- VO word order is preferred by group A, who have been away from Korlai the longest and have had sustained contact with English, whereas group B has the clearest preference for OV word order; the interpretation, once again, goes back to the influence of Marathi, a strongly OV language. In a follow-up article, Clements (1991) adds another group to the study of word order, viz. children (based on the speech of 8- to 10-year-olds); at 61%, the preference of this group for OV over VO is much more robust than in the older informants.

With respect to the linguistic features studied, Clements' study indicates that synchronic variation reveals a diachronic process of linguistic change, and can be interpreted in terms of apparent time. These processes of change are mostly (but not only, cf. the issue with the deletion of redundant preverbal markers) motivated by the increasing influence of Marathi within the community, and the distribution of the competing variants essentially depends on the informants' exposure to Marathi and on their participation in the process of negotiating innovative speech forms in Korlai. Archaizing speech is recognized not only in the production of older informants less exposed to Marathi but also, more specifically, associated with oral traditions (traditional stories), which are more likely to be memorized by rote. The study also shows that long-term absence from Korlai, where the language is used on a daily basis and in a wider number of social contexts, implies a certain crystallization of the language, further favoring the retention of archaic features.

In another important study, Clements (2009a) compares the creoles of Korlai, Daman and Diu in terms of the proportion of non-Portuguese-derived lexemes in the core lexicon, pronominal systems, bimorphemic question words, number of non-Portuguese phonemes, relative order of V and O, and of N and adposition. If we disregard Daman for the moment, what emerges is that the differences between Korlai and Diu are such that the Korlai features are either more archaic (e.g., a familiar second person pronoun derived from Ptg. *vós*) or more Indic in character (e.g., postpositions), while the Diu features reflect the influence of modern Portuguese. The systems could, in general, be classified as more 'acrolectal', in the case of Diu, and more 'basilectal' in Korlai. To explain this pattern, Clements appeals to the fact that the Portuguese language has been influential in Diu until much more recently than in Korlai, but also that, currently, the speech community is essentially white-collar in Diu and blue-collar in Korlai.

The differences between Korlai and Diu therefore constitute an instance of dialectal variation with social and historical roots, but this is less relevant for us in the sense that, at present, the two speech communities have little (if any) contact. Daman is more interesting because, for some of the linguistic domains investigated (bimorphemic question words and the pronominal system), both Diu-like (acrolectal) and Korlai-like (basilectal) variants co-exist. The history of contact with Portuguese is virtually the same as for Diu, which explains the development of the acrolectal system, but a section of the population retains more archaic features. For Clements, the size of the community impacts on the degree of community-internal variation. As a matter of fact, the Daman speech community is considerably larger than any of the other two, which makes for a higher level of social and linguistic variation. Revealingly, the acrolectal variety is by and large associated with an area (Small Daman) characterized by a largely white-collar population, whereas basilectal forms are used chiefly in another part of the territory (Big Daman and Badrapur) where the population has blue-collar jobs.

In a related study, Clements (2009b) analyzes two recent developments in Daman creole which also result in sociolinguistic variation: the transformation of *sənãw* 'if not, otherwise' into a conjunction 'and' linking sentences (a function normally performed by *i* 'and'), and the extension of the functional range of the preposition *də* 'of, from' to encompass the readings 'in, on, to'. As causes for these particular developments, the author proposes the high perceptual salience of *sənãw* and, in the case of *də*, the influence of English (via *in the*, potentially pronounced [ndə]). The English-influence scenario would be consistent with the fact that the extended locative use of *də* is the prerogative of young speakers, who are increasingly fluent in this language as a result of their English-medium education. In addition, it appears that it is young speakers from the basilectal-speaking area of Badrapur who are spearheading these innovations, "in part, because of the lack of influence on the part of normative Portuguese" (Clements 2009b: 71) among this particular group of speakers.

As Clements pointed out, the relatively small size of the Diu speech community makes for a certain linguistic homogeneity. But, while sociolinguistic variation may not be as clear-cut and regular as in Daman, the fact is that Diuese speakers as a whole also have different registers at their disposal. One important reason for this is that, as a result of the recent decolonization of Diu in 1961, standard Portuguese has played a significant role in the community and still exerts some normative pressure. This issue is taken up in Cardoso's (2007) study of Diu, which tries to ascertain who, in the

community, has Portuguese in their repertoire and who is capable and/or willing to introduce modern Portuguese features (e.g. the palatalization of syllable-final /s/,[6] or person inflection on the verb) into their production of Diu creole. Age is a factor, in that fluency in Portuguese has been diminishing since 1961, but adherence to Portuguese models is also favored by high levels of education and affluence (as well as a history of permanence in a Portuguese-speaking region).

Cardoso's (2009) grammar of Diu creole describes further instances of variation in the corpus collected in Diu that correlate with social variables. One case is that of the occurrence of Portuguese-like subjunctive verb forms, which are infrequent and restricted to the speech of "certain adult speakers with a more solid knowledge of SP [i.e., standard Portuguese]" (Cardoso 2009: 241). Another interesting case has to do with the use of nominal reduplication with a pluralizing function; from the available 19th-c. sources (especially Schuchardt 1883a), we know this to have been the standard pluralizing strategy in Diu creole but, in the modern corpus, only elderly speakers and a middle-aged man with a low level of education make use of it (Cardoso 2009: 175-176; see also Cardoso 2013). In addition, some Damanese (families) have been residing in Diu, often for long periods, and this has introduced certain Damanese characteristics into the speech community. One significant difference between Diu and Daman creoles is that Daman creole has a verbal form derived from the Portuguese gerund (e.g. *kantan*) but not Diu creole (except very marginally); the distribution of such forms in Cardoso's corpus collected in Diu, however, shows strong family effects in language transmission in that they only occur in the speech of native Damanese or their direct descendants (Cardoso 2009: 24).

4. Discussion

The sources and studies discussed in this article are not very numerous, and the scarcity of available records means that sociolinguistic variation is not equally well documented (if at all) for all of the varieties of Indo-Portuguese; yet, I believe they clarify that sociolinguistic variation is and has been part and parcel of the Indian creoles as a whole. In reality, I would venture, the fact that the evidence of sociolinguistic variation is not more robust is probably a consequence of our incomplete knowledge of these languages, which have

---

[6] See also Cardoso (2009: 282).

been under-researched so far. It is highly likely that, if more scholars were to focus on these communities and their speech production, more instances of sociolinguistic variation would be identified.

Some of the records analyzed here are actually quite explicit as to the connection between social stratification and linguistic differentiation, and they also show how different registers and diglossic situations developed within the Indo-Portuguese communities. A number of social variables have been shown to preside over the distribution of linguistic variation across Indo-Portuguese creole speakers. Age, gender and ethnicity were all invoked, as were differences in educational and social status.

One important factor, which can derive from any of these social variables, is the relationship that a particular individual establishes with other languages in the ecology – be it the lexifier (Portuguese), a locally dominant language (e.g. Gujarati, Marathi) or other languages of wide use or prestige (e.g. English, French). Some of the studies surveyed, in particular the more recent ones, clarify that individual histories of exposure to one of these languages can have a significant impact on a speaker's idiolect – regardless of whether or not the innovations introduced will become established at a community level. Linguistic attitudes are also essential, as different speakers have different degrees of willingness to adopt normative practices. We have seen this in action in the case of Schuchardt's Mahé corpus, for instance, but also in the various 18th-century accounts of how some registers were frowned upon for certain settings or media.

Exactly when the different patterns of variation formed is, of course, not entirely clear from the data available; what we can say is that linguistic stratification was established as early as the first half of the 18th century. It is highly likely, however, that this is an older trend in Indo-Portuguese, perhaps as old as the process of contact that produced it. Referring to the Portuguese-lexified creoles of East and Southeast Asia, Baxter (1996: 301) voices a similar idea:

> […] it seems highly likely that continua of L2 Portuguese arose, ranging from rudimentary L2 to L2 approximating to L1 European Portuguese, according to the degree and nature of contact. It also seems possible that the lower levels of the continua may have gelled (leading to pidgin formation). Varieties of L2 Portuguese would have provided significant input to L1 acquisition for certain groups of the local population. Where the L2 (and L1) input was significantly watered down in comparison with the original Portuguese L1 model of the colonizers, L1 acquisition by locals led to creolization. Continua of local L1 varieties of 'Portuguese' would have been formed.

But, just as it is legitimate to ask when and how linguistic stratification formed, it is equally important to consider when and how stratification waned. The truth is that, for some of these communities, the degree of variation evidenced in earlier accounts (with well-defined acrolectal and basilectal registers and a host of competing strategies) simply cannot be found at present. Among the variants which have been recently recorded, this is especially true of those of southwestern India (Cochin and Cannanore): the data collected in these locations is much more homogeneous than what Schuchardt's corpora show. However, the speech communities are also severely reduced in these locations: the last fluent speaker of Cochin has recently passed away, and only a handful of speakers remain in Cannanore. It seems, therefore, that the reduction of variation, at least in these cases, is a consequence of obsolescence and of the stark reduction in the size of the speech community – which lends support to Clements' proposed correlation between the size of the community and the degree of internal variation. What is also interesting in these cases is that the register that was lost in this period was the acrolect, which surely reflects a decline in the exposure to Portuguese and its normative pull.

The recognition of the extent to which sociolinguistic variation has impacted the Indo-Portuguese creoles has a number of methodological consequences. The first of these relates to the interpretation of older written documents. All too often, when dealing with little-documented languages, whatever few samples exist of past stages are assumed to be adequately representative, and whatever discrepancy is identified with modern-day data is taken to prove simple diachronic progression. In the case of the Indo-Portuguese creoles, I believe there is now ample evidence that this should not be taken for granted, as it is clear that the early written data typically ignored certain registers. Disparaging though this may sound, the case of the twin Mahé questionnaires shows that some progress can be made nonetheless. It is, therefore, worth investing in the study and contextualization of the primary sources.

The other domain for which it is important to consider sociolinguistic variation is the documentation and description of the modern Indo-Portuguese creoles. What the sociolinguistically-oriented studies (especially Clements') show us is that subtle kinds of variation may exist even in relatively small communities such as Korlai, and that this can be important evidence to reconstruct the past and project the future of these languages. Therefore, documentation projects should a) be sensitive to social categories and aim at

building as representative as possible a sample of the speech community; and b) rely on a wide range of different speech settings and linguistic media.

References:

Anquetil-Duperron, Abraham Hyacinthe. 1786. Des recherches historiques et géographiques sur l'Inde. In *Description historique et géographique de l'Inde*, vol. II, Jean Bernoulli (ed). Berlin: Imprimerie de Pierre Bourdeaux.

Baxter, Alan N. 1996. Portuguese and Creole Portuguese in the Pacific and Western Pacific Rim. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*, vol. II.1, Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tryon (eds), 299-338. Berlin/New York: Mouton de Gruyter.

Berrenger. 1811. *A grammatical arrangement on the method of learning the corrupted Portuguese as spoken in India,* 2nd edition. Colombo: Government Press.

Cardoso, Hugo C. 2006. Challenges to Indo-Portuguese across India. In *Proceedings of the FEL X*, R. Elangayan, R. McKenna Brown, N. Ostler & M.K. Verma (eds), 23-30. Mysore: Central Institute of Indian Languages.

Cardoso, Hugo C. 2007. Linguistic traces of colonial structure. In *Linguistic identity in postcolonial multilingual spaces*, Eric Anchimbe (ed), 164-181. Newcastle: Cambridge Scholars Publishing.

Cardoso, Hugo C. 2009. *The Indo-Portuguese language of Diu*. Utrecht: Landelijke Onderzoekschool Taalwetenschap.

Cardoso, Hugo C. 2010. The African slave population of Portuguese India: Demographics and impact on Indo-Portuguese. *Journal of Pidgin and Creole Languages* 25(1): 95-119.

Cardoso, Hugo C. 2013. Diu Indo-Portuguese. In *The Survey of Pidgin and Creole Language*, vol. 2: *Portuguese-based, Spanish-based, and French-based Languages*, Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds), 90-101. Oxford: Oxford University Press.

Cardoso, Hugo C., Alan Baxter & Mário Pinharanda Nunes. 2012. *Ibero-Asian Creoles: Comparative perspectives*. Amsterdam/Philadelphia: John Benjamins.

Cardoso, Hugo C., Tjerk Hagemeijer & Nélia Alexandre. Forthcoming. Crioulos de base lexical portuguesa. In *Manuel des anthologies, corpus et textes romans*, Maria Iliescu & Eugeen Roegiest (eds). Berlin: de Gruyter.

Callaway, John. 1820. *Vocabulary in the Ceylon Portuguese, and English languages, with a series of familiar phrases*. Colombo: Wesleyan Mission Press.

Clements, J. Clancy. 1990. Deletion as an indicator of SVO → SOV shift. *Language Variation and Change* 2: 103-103.

Clements, J. Clancy. 1991. The Indo-Portuguese Creoles: Languages in transition. *Hispania* 74(3): 637-646.

Clements, J. Clancy. 2009a. Accounting for some similarities and differences among the Indo-Portuguese creoles. *Journal of Portuguese Linguistics* 8(2): 23-47.

Clements, J. Clancy. 2009b. Gradual vs. abrupt creolization and recent changes in Daman Creole Portuguese. In *Gradual creolization: Studies Celebrating Jacques Arends*, R. Selbach, H.C. Cardoso & M. van den Berg (eds), 55-75. Amsterdam: John Benjamins.

Dal, Nikolaus. 1733. Nachricht von den Portugiesen in Indien. In Gotthilf August Francken (publ.) *Drey und dreyßigste CONTINUATION Des Berichts Der Königlichen Dänischen Mißionarien in Ost-Indien / Worin eine Ausfuehrliche Nachricht Von der gantzen Verfassung der Mißion, Wie auch Die Fortsetzung des Tage-Registers vom Jahr 1731. und einige Briefe der Herren Mißionarien enthalten, Nebst einem Anhang Von der Evangelischen Mission zu Madras*, 917-924. Halle.

Dalgado, Sebastião Rodolfo. 1906. Dialecto indo-português do Norte. *Revista Lusitana* 9: 142-166; 193-228.

Dalgado, Sebastião Rodolfo. 1921. Berço de uma cantiga em Indo-Português (à memória de Ismael Gracias). *Revista Lusitana* 22: separata.

Ferraz, Luís Ivens. 1987. Portuguese creoles of West Africa and Asia. In *Pidgin and Creole languages*, Glenn Gilbert (ed.). Honolulu: University of Hawaii Press.

Flores, Jorge. 2004. *A Taprobana e a Ponte de Rama; Estudos sobre os Portugueses em Ceilão e na Índia do Sul*. Macau: Instituto Português do Oriente.

Iken, Sebastião. 2000. O português "por uso das malabaricas e portuguesas escolas" da missão protestante alemã de Tranquebar do século XVIII: Contribuição para o estudo diacrónico da sintaxe do português na Ásia. In *Portugal, Indien und Deutschland = Portugal, Índia e Alemanha : actas / der V deutsch-portugiesischen arbeitsgespräche = do V encontro luso-alemão*, Helmut Siepmann (ed), 303-334. Cologne: Universität Köln / Lisbon: Universidade Nova de Lisboa.

Ladhams, John. 2009. The formation of the Portuguese-based Creoles: Gradual or abrupt?.In *Gradual creolization: Studies Celebrating Jacques Arends*, R. Selbach, H.C. Cardoso & M. van den Berg (eds), 279-303. Amsterdam: John Benjamins.

Lopes, David. 1936. *A Expansão da Língua Portuguesa no Oriente nos Séculos XVI, XVII e XVIII*. Barcelos: Portucalense Editora.

Moniz, António Francisco. 1925. The Negroes and St. Benedict's feast. In *The mission field. The diocese of Damaun*, 570-572. Bombay: published by S. R. Santos.

Pfänder, Stefan & Alessandra Castilho Ferreira da Costa. 2006. Linguistic Variations in Everyday Life: The Portuguese Language in the Protestant Mission of Eighteenth Century South East India. In Andreas Gross, Y. Vincent Kumaradoss & Heike Liebau (eds.), *Halle and the Beginning of Protestant Christianity in India, Vol. III: Communication between India and Europe*, 1155-1162. Halle: Verlag der Franckeschen Stiftungen.

Schuchardt, Hugo 1882. Kreolische Studien II. Über das Indoportugiesische von Cochim. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (Philosophisch-historische Klasse)* 102: 799-816.

Schuchardt, Hugo 1883a. Kreolische Studien III. Über das Indoportugiesische von Diu. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (Philosophisch-historische Klasse)* 103: 3-18.

Schuchardt, Hugo 1883b. Kreolische Studien VI. Über das Indoportugiesische von Mangalore. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (Philosophisch-historische Klasse)* 105(III): 882-904.

Schuchardt, Hugo 1889a. Beiträge zur Kenntnis des creolischen Romanisch. V. Allgemeineres über das Indoportugiesische (Asioportugiesische). *Zeitschrift für Romanische Philologie* 13: 476-516.

Schuchardt, Hugo 1889b. Beiträge zur Kenntnis des kreolischen Romanisch: VI. Zum Indoportugiesischen von Mahé und Cannanore. *Zeitschrift für Romanische Philologie* 13: 516-524.

Smith, Ian R. 1979. Convergence in South Asia: A creole example. *Lingua* 48: 193-222.

Tomás, Maria Isabel. 1992a. *Os Crioulos Portugueses do Oriente; Uma Bibliografia*. Macau: Instituto Cultural de Macau.

Tomás, Maria Isabel. 1992b. A presença africana nos crioulos portugueses do Oriente: o crioulo de Damão. In Ernesto d'Andrade & Alain Kihm (eds.), *Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa*, 97-108. Lisbon: Colibri.

APPENDIX A

- Dal, Nikolaus. 1733. Nachricht von den Portugiesen in Indien. In Gotthilf August Francken (publ.) *Drey und dreyßigste CONTINUATION Des Berichts Der Königlichen Dänischen Mißionarien in Ost-Indien / Worin eine Ausfuehrliche Nachricht Von der gantzen Verfassung der Mißion, Wie auch Die Fortsetzung des Tage-Registers vom Jahr 1731. und einige Briefe der Herren Mißionarien enthalten, Nebst einem Anhang Von der Evangelischen Mission zu Madras*, 919-920. Halle.

4. Sprache

Aus obigem ist denn auch klar, daß diejenigen, so den Namen Portugiesen führen, sich auch der Portugiesischen Sprache bedienen, als welche, nebst der Kleidung, die vornehmste Ursache ist, daß sie so genennet werden. Hier ist also nur zu erwehnen, daß die Sprache nicht bey allen einerley ist. Denn gleichwie man die Portugiesen in drey Classen getheilet hat, so kann man auch drey Arten von der Sprache setzen, nemlich 1) die rechte, 2) die halbverdorbene, und 3) die gantz verdorbene. 1. Die rechte Portugiesische Sprache reden hauptsächlich die Europäischen Portugiesen, und dann auch, die von ihnen herstammen. 2. Die halb-verdorbene reden insgemein, die von vermischter Abkunft sind. Schwartze aber bedienen sich auch derselben in ihren Briefen. Diese Art von Sprache bestehet hauptsächlich darin, daß sie die <u>verba</u> nicht recht nach ihren <u>Coniugationibus</u> zu gebrauchen wissen. 3. Die gantz verdorbene höret man, im gemeinen reden, zwar auch von Portugiesen vermischter Abkunft, doch vornemlich von denen, die gantz schwartz sind. Diese Art von Sprache unterscheidet sich von der vorhergehenden hauptsächlich darin, daß die Leute gar keine <u>Coniugation</u> gebrauchen, sondern nur die künftige Zeit durch die <u>particulam</u> *lo*, und die vergangene durch *ja*[d] aussprechen, und sich dabey des <u>Infinitivi</u> mit Auslassung des <u>R</u> bedienen. Zu Briefen wird diese Sprache für läppisch und <u>ridicul</u> gehalten. Man sollte auch dencken, daß sie zum Sprechen auch nicht sonderlich geschickt sey; welches von einer an einander hängenden Rede auch allerdings wahr ist. Wie denn aus solcher Ursache auch die schwartzen Portugiesen sich der halb-verdorbenen Sprache bedienen, wenn sie aus ihrem Hertzen ein Gebet zu Gott thun. In abgebrochenen Reden aber kann man sich deren noch ziemlich bedienen: wie sie denn auch insgemein gesprochen wird, wenn Engelländer und Dänen, Holländer und Frantzosen, Europäer und Armenier, u.d.g. hier in Indien zusammen kommen, und einer den andern in seiner Muttersprache zu verstehen nicht vermögend ist.[e]

[d]*Lo* ist gemacht aus <u>logo</u>, (wie einige auch sprechen,) d. i. gleich, alsobald.. Es hat nemlich ein Indianer von einem Portugiesen gehöret, z.B. <u>logo virei</u>, ich will gleich kommen; da hat er gemeynt, eine künftige Sache müsse man mit <u>logo</u> ausdrücken. *Ja* heißt schon, bereits. Es hat z.B. ein Portugies gesagt: <u>ja ouvi</u>, ich habe es schon gehöret; da hat der Indianer gedacht, etwas vergangenes müsse man mit <u>ja</u> anzeigen. Die vielen <u>terminationes</u> sind ihm, wie leicht zu erachten, zu schwer geworden.

[e]Also wechseln gedachte <u>Nationes</u> oftmals Briefe mit einander, und zwar in der rechten oder halb-verdorbenen Portugiesischen Sprache, je nach dem sie einen guten Schreiber überkommen können.

APPENDIX B

- Anquetil-Duperron, Abraham Hyacinthe. 1786. Des recherches historiques et géographiques sur l'Inde. In *Description historique et géographique de l'Inde*, vol. II, Jean Bernoulli (ed), XI-XII. Berlin: Imprimerie de Pierre Bourdeaux.

"III.

Moyens de communication employés jusqu'ici par les Européens avec les peuples de l'Inde.

On sait que les portugais sont les premiers Européens qui ayent fait des Etablissements dans l'Inde. Cette nation ne s'étant point d'abord donné la peine d'en apprendre les langues, a par là, en quelque sorte obligé les Naturels de parler la sienne. Les descendants des premiers Indiens soumis par les Portugais, Chrétiens ou autres, la savent généralement. Elle s'est répandue avec eux le long des deux Côtes ; & comme la plupart des domestiques au service des François, Anglois, Hollandois, Danois, sont issus de ces premiers Indiens, ou des Esclaves Abyssins des Portugais, les étrangers, en arrivant, se trouvent dans la nécessité d'apprendre le *Portugais*.

En conséquence les Marchands Indous, Maures, Arabes, Persans, Parses, Juifs, Arméniens, qui trafiquent dans les Comptoirs Européens, ainsi que les Noirs qui veulent faire la fonction d'Interprête, sont obligés de parler cette langue; elle sert encore de communication entre les nations Européenne établies dans l'Inde.

Mais il s'en faut bien que ce soit le Portugais pur, appelé dans l'Inde le *Portugais Reinol*. Celui qui s'écrit en approche d'avantage, surtout à la Côte Malabare, où cette Nation a eu de nombreux Etablissements : le Portugais parlé n'est proprement qu'un jargon, consistant en 150 ou 200 mots, presque sans construction.

Les Anglois & les Hollandois exigent que leurs Interprêtes, & souvent leurs domestiques, entendent l'Anglois, le Hollandois; & le Sécrétaire du Conseil, chez eux, sait toujours le François.

Dans Nos Etablissements les affaires se sont généralement traitées avec les Naturels du pays, & même avec les autres nations Européenes, par le moyen du *Jargon Portugais* dont je viens de parler ; ou bien on étoit réduit à l'expédient, au moins dangereux, d'employer des transfuges & autres gens d'une fois plus que suspecte.

Depuis quelques années, lors de la derniere révolution, en 1778, l'Interprête en Chef de Pondicheri, Chrétien Malabar élevé par les Jésuites, savoit le François ; c'étoit un vice de moins dans l'Administration ; mais ce foible secours suffisoit-il?