BRILL

JOURNAL OF LANGUAGE CONTACT 14 (2021) 524–556

JOURNAL OF LANGUAGE CONTACT

brill.com/jlc

The Gulf of Guinea Proto-Creole and Its Daughter Languages: From Liquid Consonants to Complex Onsets and Vowel Lengthening

Manuele Bandeira
Assistant Professor of Linguistics, Institute of Humanities and Languages, University of International Integration of the Afro-Brazilian Lusophony (Unilab), São Francisco do Conde, Brazil
manuelebandeira@unilab.edu.br

Gabriel Antunes de Araujo
Associate Professor of Linguistics, Faculty of Arts and Humanities, University of Macau, Macau SAR, China; Faculty of Philosophy, Languages and Human Sciences, University of São Paulo, São Paulo, Brazil
gabrielaraujo@um.edu.mo

Thomas Finbow
Assistant Professor of Linguistics, Faculty of Philosophy, Languages and Human Sciences, University of São Paulo, São Paulo, Brazil
thomas.finbow@usp.br

Abstract

Four Portuguese-based Creoles are spoken on the islands in the Gulf of Guinea: Santome, Angolar, Lung'Ie, and Fa d'Ambô. These languages are descendants of the Portuguese-based Gulf of Guinea Proto-Creole, which emerged at the beginning of the sixteenth century on São Tomé Island. Based on Bandeira (2017), we discuss the development of liquid consonants in Santome, Lung'Ie, Angolar and Fa d'Ambô using data from the reconstruction, and we examine the developments in the daughter-languages of the proto-phonemes *r and *l that led to the synchronic systems and the present configurations in the daughter languages, since the liquid consonants evolved differently from the proto-creole. We show that the relation between long vowels and liquid consonants, both in coda and in complex onsets, can be better understood if we consider the modern lexical items in these four languages as continuations of proto-forms, with characteristic modifications in each language governed by regular processes.

DOI:10.1163/19552629-14030002

Keywords

Proto-creole – phonology – Creole languages – Africa – Portuguese – historical linguistics – liquids – rhotics.

1 Introduction

Four Portuguese-based creole languages are spoken on the islands in the Gulf of Guinea: Santome and Angolar on São Tomé, Lung'Ie on Príncipe, and Fa d'Ambô on Annobón (Ferraz, 1979; Hagemeijer, 2009). All four languages (henceforth "daughter languages") are descendants of the Portuguese-based Gulf of Guinea Proto-Creole (GGPC), which emerged at the beginning of the sixteenth century from contact between Portuguese colonists and African populations brought as slaves to São Tomé Island. Isolation, migration of certain groups from São Tomé to the other islands, and linguistic contributions from African languages by way of the constant renewal of the enslaved segment of the population contributed to the proto-creole's speciation. Santome (ST)[1] developed in the colonial centres on São Tomé Island, while Angolar (AN) arose among the descendants of runaway slaves who escaped from the plantations and founded maroon communities in the island's interior, far from the colonial centres. Proto-creole speakers were taken from São Tomé to the islands of Príncipe and Annobón, where local conditions contributed to diversification, giving rise to Lung'Ie (LI) and Fa d'Ambô (FA), respectively.

Initially, the African slaves taken to São Tomé came from the region surrounding the Niger Delta. Later groups were taken from the Congo region and from Angola (Ferraz, 1979: 12–14; Landhams, 2007: 6–7; Hagemeijer, 2009). Ferraz (1979: 15) points to the first stage of Portuguese settlement as being crucial for the formation of the original Gulf of Guinea Creole. São Tomé was a port of call for many ships and some slaves must have remained there, although most passed through to other destinations. Additionally, at that time, there was no large-scale economic activity on the island besides slaving, which resulted in reduced spatial segregation, i.e., there was not the kind of

1 The ISO 639-3 codes (Part 3: Alpha-3 code for comprehensive coverage of languages. See https://www.iso.org/standard/39534.html) for the daughter languages of the GGPC are cri for Santome, pri for Lung'Ie, aoa for Angolar, and fab for Fa d'Ambô. Since these ISO 639-3 codes are not intuitively recognisable, we shall employ our own habitual abbreviations: ST for Santome (also known as Forro), LI for Lung'Ie (Principense), AN for Angolar (Ngola), and FA for Fa d'Ambô (Pagalu).

distancing that typically occurred between plantation slaves and estate owners and employees, and this proximity fostered intense contact between speakers of Portuguese and speakers of African languages from the Congo and Benin areas.

Bandeira (2017) proposed a phonological and lexical reconstruction of GGPC based on a set of 536 cognates from the four daughter-languages (ST: 523 items, FA: 391, LI: 507, and AN: 475) collected during fieldwork and taken from previous work; e.g., Segorbe (2010), Araujo and Hagemeijer (2013) and Agostinho (2016). In this article, we discuss the development of liquid consonants in ST, LI, AN and FA using data from Bandeira's (2017) reconstruction and analyses of the GGPC's phonological processes. To this end, we shall examine the development of the proto-phonemes *r and *l[2] in the daughter languages. We consider the proto-creole's phonological system and the phonological processes that led to the present configurations of the synchronic systems, since the liquid consonants diverged from the proto-creole in each daughter language. Of the four languages, only FA permits laterals in syllable codas. LI admits rhotics in onsets, while ST, AN and FA simply do not possess any rhotic segments, although AN has signs of an incipient [r] in recent loanwords, especially from ST. Additionally, ST allows complex onsets with /l/ as the second element, while cognates in LI, AN and FA do not exhibit complex onsets, having long vowels instead.

Liquid consonants have been dealt with previously by Ferraz (1987) and by Rougé and Schang (2006). Ferraz's objective was a comparative analysis of the four Gulf of Guinea creole languages that gave special attention to the liquid consonants in clusters. However, Ferraz (1987) compares only one language (ST) with Portuguese. Rougé and Schang (2006) proposed studying the origin of the liquid consonant /l/ in ST in another attempt to explain the differences in behaviour found between ST and LI, and between AN and FA. As indicated by the authors, Rougé and Schang (2006) follow the line of reasoning developed by Ferraz (1987) but differs in presenting a quantitative analysis. Both Ferraz (1987) and Rougé and Schang (2006) assume that Portuguese is the common ancestral language when they compare the Gulf of Guinea creole languages. Ferraz (1987) denies the linguistic kinship of ST, LI, AN and FA, claiming the

2 The representation in the official orthography is printed in boldface, e.g., **dhumbo** 'moustache' in AN. This orthography is in accordance with the *Alfabeto Unificado para as Línguas Nativas de São Tomé e Príncipe* (ALUSTP), approved by the government of São Tomé e Príncipe for representing ST, LI and AN alphabetically. Although these languages are mutually unintelligible, they share a considerable number of lexical and grammatical features, which justifies the unified orthography. ALUSTP proposal was discussed by Araujo and Agostinho (2010).

languages have no common ancestor, since they present too many differences amongst themselves to admit speciation. Ferraz (1987: 348) attributes the four languages' similarities to their development in close geographical proximity, concluding, "Hence different languages developed in the archipelago rather than dialects of one contact language."

Thus, for Ferraz, different languages developed in the archipelago rather than evolving from the dialects of one original contact language. Campbell (2004[1999]: 108) claims that there can only be a genetic relation if we assume that, in the past, there were dialects of a single language:

> We say that languages which belong to the same language family are genetically related to one another: this means that these related languages derive from (that is, 'descend' from) a single original language, called a proto-language. In time, dialects of the proto-language develop through linguistic changes in different regions where the language was spoken - all languages (and varieties of language) are constantly changing - and then later, through further changes, the dialects become distinct languages.
> CAMPBELL, (2004[1999]: 108)

Our analysis, therefore, differs from the treatment of Ferraz (1987) and Rougé and Schang (2006) in two central aspects: (i) their analyses essentially investigate ST, while this study compares all four creole languages; and (ii) the studies that juxtapose ST and Portuguese treat the latter as the former's direct ancestor, while we consider the proto-creole to be not only ST's direct precursor but also the immediate ancestor of LI, AN and FA. These differences in approach produce different results. For example, the two previous studies claim that deletion of the liquid consonant was the primary rule in LI, AN and FA for dealing with complex onsets. However, suppressing the liquid is only one strategy these languages employ. Depending on the phonological inventory and restrictions on syllable type, liquid-deletion can be followed by compensatory lengthening, anaptyxis (inter-consonantal vowel epenthesis), re-syllabification, and lambdacism, amongst others. Moreover, by not comparing the four languages simultaneously, the studies cited above are unable to make systematic and regular generalizations. Ferraz (1987), for example, claims that compensatory vowel lengthening is not characteristic of AN, yet the process occurs repeatedly in the language, e.g., *baanku* < *blanku 'white' (PT *branco*). Rougé and Schang (2006: 24) describe a "massive presence of the consonant /l/ in ST, that is not found to the same degree in the other languages". If we consider Portuguese as the creoles' direct ancestor, it is possible that deletion does indeed appear to be the strategy employed not only by LI, AN and FA, but also by ST (*bega*

(ST) ‘belly’ (PT *barriga*) (Rougé and Schang, 2006: 29). However, when we take the proto-creole as the common ancestor, one can see that what is happening in the daughter languages is maintenance, instead of deletion, as in *kura > [kuˈla] [kuˈla] [kuˈra] [kuˈla] ‘to heal’ (PT *curar*).

Araujo, Bhatt and Hagemeijer (2012) was the first study to identify the coronal consonants as determinative for the historical treatment of liquid consonants in ST, FA, LI and AN. These researchers’ pioneering work was the starting point for the reconstruction of *r and *l in codas and *l in complex onsets. Although Araujo, Bhatt and Hagemeijer (2012) did not reconstruct the proto-creole, their work provided a pivotal clue to this study, indicating which direction to take in our analyses. The present study, therefore, coincides with the focus of the investigations mentioned above in exploring the diachronic treatment of liquid consonants. Beyond that, however, the similarities cease, given the different object of analysis, the use of the historical comparative method, and this study’s assumption that ST, LI, AN and FA descend from the GGPC, a different language to Portuguese, and which, in addition, cannot be described as a variety of that language.

We conjecture that the entry of the rhotic in AN and the formation of complex onsets in LI are recent features in those languages, given that LI exhibits second element deletion in complex onsets in the cognate sets analysed (Bandeira, 2017), such as in *ubaanku* ‘white’ < **blanku* (PT *branco*) and *peetu* ‘black’ < **plɛtu* (PT *preto*), just as AN adapted *r in two ways: deletion, e.g., *pɛtu* ‘close’ < **pɛtu* (PT *perto*) or lambdacism, e.g., *liku* ‘rich’ < **riku* (PT *rico*). More details about the adaptations will be shown in the data analysis.

Furthermore, as mentioned above, the liquid consonants *r and *l, both in the onset and in the coda, are not always preserved or retained *in loco,* owing to the different restrictions on syllable-form that operate in each daughter language. Our description is therefore important, because it enables the investigation of four systems that select sometimes one and at other times other strategies from a set of universal options (loss, insertion, compensatory lengthening, metathesis, etc.) according to the syllable structures that are available in each language, independently of the others. Such outcomes can therefore vary depending on the daughter language, the nature of the proto-phoneme (lateral or rhotic) and its position in the syllable (onset or coda). Moreover, our analysis provides evidence that the trigger for compensatory vowel lengthening may not be limited to the rime and can be provoked by segment loss in the onset. This observation is important, since, until now, only Samothracian Greek has been studied amongst the rare cases in which the trigger for compensatory lengthening is loss in the syllable onset rather than in the coda, as happens in most languages (Gess, 2011: 5).

The rest of the text is organized in the following manner: in section 2, we present the methodology. Then, in section 3, we demonstrate the reconstruction of lateral consonants in the GGPC and discuss some issues regarding the contact situation. In section 4, we discuss the changes occurring in simple onsets, codas, and in complex onsets, comparing the proto-forms to the modern cognate liquids across the four daughter languages. Section 5 presents the article's final considerations.

2 Materials and Methods

The corpus for this study comes from the list of cognates drawn up by Bandeira (2017), starting from a set of lexical items extracted from the basic vocabulary of the Gulf of Guinea Proto-Creole's four daughter languages that have been recorded in theoretical studies, dictionaries and grammars of ST, LI, AN and FA. We also worked on data collected from fieldwork in Príncipe (for LI data collection in 2013) and São Tomé (the Angolar data were collected in 2014 and 2018). There are studies written in the nineteenth century on ST (Negreiros, 1895), LI (Ribeiro, 1888), and (Barrena, 1957) on FA, as well as scattered references in Schuchardt (1882, 1888, 1889) and Adolfo Coelho (1880). However, works published before the second half of the twentieth century generally present descriptive limitations and contain incongruencies in data recording, as well as in the analyses of the languages' phonological systems. For these reasons, Bandeira (2017) used contemporary linguistic data as the basis for her reconstruction, which were collected during fieldwork and gathered from the following sources: Ferraz (1979, 1987), Maurer (1995, 2009), Segorbe (2010), Hagemeijer (2009, 2011), Araujo and Hagemeijer (2013), Araujo *et al.*, (2013), Araujo and Agostinho (2014), Agostinho (2016). In addition to these sources, information available on-line in the *Atlas of Pidgin and Creole Language Structures* was consulted (Michaelis *et al.*, 2013), such as grammatical and lexical features of ST, LI, AN and FA, as well as examples of each language.

Bandeira collected 1896 lexical items with semantic and phonetic equivalence (ST – 523, FA – 391, LI – 507, AN – 475). The items she used in her phonological reconstruction were selected from the four CREOLES' shared basic vocabulary, following Fox (1995). Basic vocabulary is needed in comparative reconstruction to try to avoid introducing lexical items resulting from naturalised loans into the analysis, as basic vocabulary is to be more likely to contain inherited material. Bandeira then followed the Historical-Comparative Method for language reconstruction, guided by the principle that recurring regular correspondences between the phonemes in the daughter languages typically descend from a single proto-phoneme in the mother language, unless other diachronic phonological processes intervene (Kaufman, 1990; Hock, 1991; Crowley, 1997 [1992]; Campbell, 2004 [1998]). In this way, Bandeira (2017) identified 536 cognates across the four daughter languages. The comparison of these cognates with their reconstructed etyma in the GGPC forms the basis of our investigation

Besides the principles of the Historical-Comparative Method, our work's theoretical foundations follow non-linear phonological models, such as Autosegmental Phonology (Goldsmith, 1976 [1979]; Clements and Hume, 1995), Syllable Theory (Selkirk, 1984), Lexical Phonology (Kiparsky, 1982) and Moraic Phonology (Hayes, 1989).

3 Reconstructing Liquids

In this section, we present the data underpinning our reconstruction of the liquid consonants (the lateral *l and the rhotic *r) from the cognate sets, based on Bandeira (2017). First, we describe the GGPC's consonant system and indicate the daughter languages' convergent and divergent features. In 3.1, we present the correspondences that support our reconstruction of the liquids and, in 3.2, the correspondence sets for reconstructing the alveolar vibrant. In 3.3, we show how certain aspects of the contact situation played a role in preserving these features in the proto-creole.

The GGPC's consonantal system comprised eighteen consonants, with six manners of articulation: seven stops, four fricatives, three nasals, one lateral, a vibrant and two approximants. The stops and fricatives formed voiced-voiceless pairs (*p *b; *t *d; *k *g; *f *v; *s *z). There were bilabial, alveolar and palatal nasals (*m, *n, *ɲ), a voiced labial-velar stop (*g͡b),[3] an alveolar lateral (*l), an alveolar vibrant (*r), and bilabial and palatal approximants (*w, *j).

The GGPC vowel system contained seven oral vowel phonemes (*i, *e, *ɛ, *a, *ɔ, *o, *u). There were no nasal vowel phonemes; the nasal consonants conditioned phonetic vowel nasalisation, which was restricted to certain conditioning contexts which are: *a*) when a nasal consonant occupies a word-medial syllable-coda, such as **baN*.ku 'bench' > [ˈbẽ.kʊ] (ST), [ˈbẽːkʊ] ~ [ˈbẽ.kʊ] (FA), [uˈbẽ.kʊ] ~ [uˈbẽːkʊ] (LI), and [ˈbẽ.kʊ] (AN); *b*) when a nasal consonant occupies a stressed syllable's onset, the vowel in the preceding syllable can be nasalised. This second type of nasalisation is optional in the daughter languages:

3 Hagemeijer (2011: 119) proposed the reconstruction of this proto-phoneme. Bandeira and Araujo (2021) presents further supporting arguments.

[bɐ̃ˈna] ~ [bɐˈna] 'banana' (ST, FA, LI, AN) (Araujo and Agostinho, 2010; Balduino, Agostinho, Araujo and Christofoletti, 2015). GGPC syllable structure allowed the complex onsets *Cl, *Cw or *Cj. Syllable codas admitted nasal and fricative archiphonemes (*N *S), the approximants (*w *j), and the liquids *r and *l. Primary stress usually fell on the penultimate syllable in nouns, but moved to the final syllable when that was heavy. Verbs were stressed on the final syllable.

(1) a. Simple onset: **liN*.pu 'clean' (PT *limpo*), *rizu 'hard';

b. Coda: **pɛr*.tu 'near' (PT *perto*), *mal.*da*.di 'evil', 'wrong' (PT *maldade*);
c. Complex onset: **gle*.za 'church' (PT *igreja*).

The GGPC's phonological system contained two liquid consonants: a lateral (*l) and a rhotic (*r). These two consonants have been reconstructed for initial position, as in (1-a). Both *l and *r could occur in syllable-final position, as in (1-b), but only the lateral could appear in the second slot of a complex onset, as in (1-c).

The proto-liquids' reflexes in the daughter languages are not free to occupy any position in the syllable. The daughter languages have also undergone changes regarding the ancestral liquids' form. At times, the changes are shared between some or all the languages and at others, each language's evolution differs. We analyse the liquids in the following subsections according to the type of consonant, either rhotic or lateral, and according to the position it occupies in the syllable, either in the onset or in the coda.

Our analysis of the proto-phonemes will proceed in accordance with the place of articulation.

3.1 *Alveolar Lateral*

The GGPC possessed one lateral consonant segment, *l. The reconstruction of the phonological system for this lateral is based on cognate sets that exhibit systematic regularities between the genetically related languages (see Table 1).

The systematic correspondences between the languages indicate that the proto-phoneme evolved differently depending on the position in which it stood, either in the onset (see Section 4.1) or in the coda (4.2). In initial

TABLE 1 GGPC lateral.

Lateral	Alveolar
	*l

position, *l was retained in all the languages. Syllable-finally, *l was relocated to the onset in a process of metathesis in certain contexts in ST. In LI and AN, *l was lost in syllable codas. In FA, depending on the context, *l was either lost or metathesized. In LI and AN, compensatory vowel lengthening could follow loss, if the consonant was adjacent to a coronal. Below, we describe *l regarding syllable position.

First, in syllable onsets in word-initial and word-medial position, the proto-phoneme *l evolves into phonetically identical reflexes[4] in all four daughter-languages, as shown in Table 2 and 3.

In second position in complex onsets, *l can be preceded by the stops *p, *b, *t, *k and *g, and by the fricatives *f and *v. In this context, ST maintains the liquid *in situ*, LI and AN have elided the proto-phoneme, and FA lost the liquid in stressed and unstressed syllables, exhibiting compensatory vowel lengthening. However, in contexts in which *l was in the initial syllable of a word, LI and AN exhibit long vowels, like FA, as well as elision, e.g., **bla*.su, **blaN*.ku, **gle*.za and **pla*.tu (see Table 4).

TABLE 2 *l in word-initial onsets.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	l	l	L
'to shine' (**PT *luzir***)	*lu.zi	[lu'ʒi]	---	[lu'ʒi]	[lu'zi]
'to wash' (**PT *lavar***)	*la.ba	[la'ba]	[la'ba]	[la'va]	[la'ba]
'leprosy' (PT ***lepra***)	*lɛ.pla	['lɛ.plɐ]	['lɛpa]	['lɛ.pɐ]	['lɛ.pɐ]
'far' (PT ***longe***)	*loN.di	['lõ.d͡ʒi]	---	['lõ.ʒi]	['lõ.d͡ʒi]

*l can occur in onsets before all oral vowels: *i, *e, *ɛ, *a, *ɔ, *o, *u (see Table 3).

TABLE 3 Word-medial *l.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	l	l	l
'skin disease'	**la*.lu	['la.lu]	['la.lu]	[u'la.lu]	['la.lu]
'hall' (PT ***sala***)	**sa*.la	['sa.lɐ]	---	['sa.lɐ]	['sa.lɐ]
'first' (PT ***primeiro***)	*pli.*me*.lu	[pli'me]	[pi'me.lu]	[pi'mew]	[pũ'be.lu]
'honey' (PT ***mel***)	**mɛ*.lɛ	['mɛ.lɛ]	['mɛ.lɛ]	['mɛ.li]	['mɛ.lɛ]

4 According to Campbell (2004 [1998]: 112), 'the descendant in a daughter language of a sound of the protolanguage is said to be a reflex of that original sound'.

TABLE 4 *l in word-initial complex onsets.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	Ø:	Ø:	Ø:
'arm' (PT ***braço***)	**bla*.su	[ˈbla.su]	[ˈbaːsu]	[uˈbaːsu]	[ˈbaːsu]
'white' (PT ***branco***)	**blaN*.ku	[ˈblɐ̃.ku]	[ˈbɐ̃ːŋku]	[ˈbɐ̃ːku]	[ˈbɐ̃ːku]
'church' (PT ***igreja***)	**gle*.za	[ˈgle.zɐ]	[ˈgeːza]	[ˈgeːzɐ]	[ˈŋgeːðɐ]
'plate' (PT ***prato***)	**pla*.tu	[ˈpla.tu]	[ˈpaːtu]	[ˈpaːtu]	[ˈpaːtu]

TABLE 5 *l in left aligned-syllable complex onsets.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	Ø:	Ø:	Ø:
'snake' (PT ***cobra***)	**kɔ*.blɔ	[ˈkɔ.blɔ]	[ˈxo.bo.lo]	[ˈkɔ.bɔ]	---
'to cover' (**PT *cobrir***)	*ku.*bli*	[kuˈbli]	[kuˈbi.li]	[kuˈbi]	[kuˈbi]
'fever' (PT ***febre***)	**fɛ*.blɛ	[ˈfɛ.blɛ]	[ˈfi.bi.li]	[ˈfɛ.bi]	[ˈfɛ.bɛ]
'poor' (PT ***pobre***)	**pɔ*.bli	[ˈpɔ.b.li]	[ˈpo.bi.li]	[ˈpɔ.bi]	[ˈpɔ.bi]

On the other hand, if *l occurs in a complex onset at the right edge of a word, independently of stress, ST retains the liquid *in situ* and AN and LI exhibit a simple vowel. FA presents the lateral consonant [l] preceded by an epenthetic vowel copied from the nucleus of the ancestral syllable to break up the cluster (anaptyxis), as in Table 5.

There are, however, two cognate sets in which the daughter languages, except for ST, failed to render the pattern above. In *glɔ.rja, FA retained the lateral consonant by anaptyxis (inserting a vowel between two consonants), even though the complex onset was in the first syllable. In LI and AN *l is lost without vowel lengthening in both lexical items, despite occurring in the conditioning context of a word-initial syllable cluster (Table 6).

In the GGPC, *l could occupy syllable codas and the onset but ST, LI and AN do not permit /l/ in codas. As for FA, although the language allows /l/ in the coda, according to Segorbe (2010) and Araujo *et al.* (2013), the proto-phoneme was not always preserved. In codas before a coronal consonant, *l is lost in ST, LI and AN, while FA presents vowel lengthening as well as loss, as shown in Table 7.

In codas before a non-coronal consonant, LI and AN lose *l and FA once again exhibit long vowels, as well as loss. In ST, on the other hand, *l evolved differently in this context, leaving the coda, and passing into the onset's second position, as the reconstructed item in Table 8 shows.

TABLE 6 *l in word-initial syllable-onset second position – exceptions.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	l	Ø	Ø
'glory' (PT ***glória***)	**gl*ɔ*.rja	[ˈŋglɔ.jɐ]	[gɔˈlɔ.lja]	[ˈgɔ.rjɐ]	[ˈgɔ.ljɐ]
	*l	l	Ø > V:	Ø	Ø
'flower' (PT ***flor***)	**flɔ*.li	[ˈflɔ.li]	[ˈfɔːl]	[uˈfɔ.li]	[ˈfɔ.li]

TABLE 7 *l in coda before a coronal consonant.

Gloss	GGPC	ST	FA	LI	AN
	*l	Ø	Ø > Vː	Ø	Ø
'badness' (PT ***maldade***)	*mal.*da*.di	[maˈda.dʒi]	[maːˈda.di]	---	[maˈda.di]
'curse' (PT ***maldição***)	*mal.di.*saN*	[ma.diˈsõ]	[maːdʒiˈsɐ̃]	[ma.dʒiˈsɐ̃]	---
'fish stew' (PT ***caldeirada***)	*kal.dɛ.*ra*.da	[ka.dɛˈla.dɐ]	---	[ka.dɛˈra.dɐ]	[ka.dɛˈla.dɐ]
'lack' (PT ***faltar***)	*fal.*ta*	[faˈta]	[faːˈta]	[faˈta]	[faˈta]

TABLE 8 *l in coda before non-coronal consonants.

Gloss	GGPC	ST	FA	LI	AN
	*l	l	Ø > V:	Ø	Ø
'to save' (PT ***salvar***)	*sal.*va*	[ʃlaˈva]	[saːˈva]	[saˈva]	[θaˈva]

3.2 *Alveolar Vibrant*

There is only one voiced alveolar vibrant in the GGPC's phonological system. The systematic correspondences between the daughter languages indicate that this proto phoneme took three different evolutionary paths: total loss (FA, AN), loss through fusion with *l (ST) and survival (LI). Only LI's inventory exhibits the phoneme /r/ and even in LI, the segment can occur only in the onset, never in the coda. Although ST does not possess a rhotic, it is, however, the only one of the four languages that conserves complex onsets *in situ*, forming them with all the stops, with two fricatives and a bilabial nasal (/p b k g t d/, /f v/, /m/)

TABLE 9 *r in word-initial onset.

Gloss	GGPC	ST	FA	LI	AN
	*r	l	l	r	l
'to tear, to **rip out/off'** (PT ***arrancar***)	*raN.*ka*	[lɐ̃ˈka]	[laŋˈxa]	[rɐ̃ˈka]	[lɐ̃ˈka]
'rich' (PT ***rico***)	**ri*.ku	[ˈli.ku]	[ˈli.ku]	[ˈri.ku]	[ˈli.ku]
'to snore' (PT ***roncar***)	*roN.*ka*	[lɔ̃ˈka]	[lõ.xuˈa]	[ˈrɔ̃ˈka]	[lɔ̃ˈka]
'rose' (PT ***rosa***)	**rɔ*.za	[ˈlɔ.zɐ]	[ˈlɔ.zɐ]	[ˈrɔ.zɐ]	[ˈlɔ.ðɐ]

and the lateral liquid. Thus, as ST's phonemic inventory has no /r/, *r must have changed into [l] and merged with /l/ < *l.

Systematic correspondences amongst the reconstructed forms confirm the reconstruction of the voiced alveolar vibrant proto-phoneme *r in word-medial syllable onsets. In word-initial and word-medial onsets, *r corresponds to the reflex /l/ in ST, FA and AN and the identical reflex /r/ in LI, as shown in Table 9.

*r can appear in word-initial and word-medial onsets before all the oral vowels (*i, *e, *ɛ, *a, *o, *ɔ, *u) and before an approximant (*j or *w) (see Table 10).

Unlike its daughters, in the GGPC, *r could occur in syllable codas. As ST, FA and AN's inventories do not include /r/, only LI could possibly present an identical reflex in coda position. However, LI only keeps the vibrant in syllable onsets; the coda cannot contain /r/. In non-final codas, therefore, before a coronal consonant, *r is deleted in ST, LI, and in AN, while FA presents a lengthened vowel, as well as deletion, as in [ˈkuːtu] 'short' (PT *curto*), [ˈpɛːtu] 'close' (PT *perto*) and [ˈxaːni] 'meat' (PT *carne*) (see Table 11). In non-final codas before a non-coronal consonant (see Table 12), LI, FA and AN not only exhibit *r-deletion but also have long vowels.[5] On the other hand, ST – which permits

5 Maurer (2009) and Agostinho (2016) do not present the items 'to purge', 'to undo' and 'pig' with lengthening, i.e., their data show **puga**, **baga**, **poko**, rather than **puuga**, **baaga** and **poko**. For them, compensatory lengthening (*CVr > Vː) simply does not occur, whatever the following consonant (**petu** 'close'; **poko** 'pig', **puga** 'to purge' and **baga 'to undo'**). However, in **petu** 'close', the lengthening would not be expected, because of the proximity of the coronal consonant in the following syllable. As for the other examples mentioned, we have found evidence for **puuga**, **baaga** and **pôôko**, respectively (Bandeira, field notes). We must also not discount the possibility of dialect variation in LI: some varieties with lengthening (*CVr -> Vː) and others without, as pointed out by one reviewer. We do not consider, however, the differences between the transcriptions of data from different authors, e.g., Traill and Ferraz (1981), Maurer (2009), Agostinho (2016), Bandeira (2017, field notes), Agostinho and Araujo (in preparation) to be a strong enough reason to discredit our notion that the coronal consonants are relevant to the application of the rule, as we propose.

TABLE 10 Word-medial *r.

Gloss	GGPC	ST	FA	LI	AN
	*r	l	l	R	l
‘yellow’ (PT *amarelo*)	*a.ma.*rɛ*.lu	[maˈlɛ.lu]	---	[maˈrɛ.lu]	[a.maˈlɛ.lu]
‘to run’ (PT *correr*)	*ko.*re*	[koˈle]	[xoˈle]	[koˈre]	[koˈle]
‘to toast’ (PT *torrar*)	*tɔ.*ra*	[tɔˈla]	[tɔˈla]	[tɔˈra]	[tɔˈla]
‘saw’ (PT *serra*)	**se*.ra	[ˈse.lɐ]	---	[ˈse.rɐ]	[ˈθe.lɐ]
‘dark’ (PT *escuro*)	**ku*ru	[ˈku.lu]	[ˈku.lu]	[uˈku.ru]	---

TABLE 11 *r in coda before coronal consonants.

Gloss	GGPC	ST	FA	LI	AN
	*r	Ø	Ø > Vː	Ø	Ø
‘meat’ (PT ***carne***)	**kar*.ni	[ˈka.ni]	[ˈxaːni]	[uˈka.ni]	[ˈka.ni]
‘short’ (PT ***curto***)	**kur*.tu	[ˈku.tu]	[ˈkuːtu]	[ˈku.tu]	[ˈku.tu]
‘close by, nearby’ (**PT *perto***)	**pɛr*.tu	[ˈpɛ.tu]	[ˈpɛːtu]	[ˈpɛ.tu]	[ˈpɛ.tu]

TABLE 12 *r in coda before non-coronal consonants.

Gloss	GGPC	ST	FA	LI	AN
	*r	l	Ø > Vː	Ø > Vː	Ø > Vː
‘to purge’ (**PT *purgar***)	*pur.*ga*	[pluˈga]	[puːˈga]	[puːˈga]	[puːˈga]
‘undo’ (PT***?(des)embargar***)	*bar.*ga*	[blaˈga]	[baːˈga]	[baːˈga]	[baːˈga]

onsets formed of stops or the phonemes /f/ and /v/, plus a liquid /l/, – presents a different development for *r: it becomes /l/. After lambdacism, metathesis occurs, moving the lateral from the coda into the second position of a complex onset, e.g., *pur.*ga* > *pul.*ga* > /pluˈga/ ‘to purge’ (PT *purgar*) and *bar.*ga* > *bal.*ga* > /blaˈ*ga*/ ‘undo’ (PT?(*des*)*embargar*), as described previously by Araujo, Bhatt and Hagemeijer (2012).

3.3 *The Contact Situation*

From the fifteenth to the nineteenth century, the phonological inventory of Portuguese possessed, in addition to the laterals /l/ e /ʎ/, the distinctive pairing

/ɾ/ and /r/, as in *caro* /ˈka.ɾo/ 'dear/expensive' and *carro* /ˈka.ro/ 'car'. During the period in question, however, in non-intervocalic contexts, the two rhotic phonemes occurred in complementary distribution: /ɾ/ was realized as an alveolar tap everywhere, except for word-initial position and before or after a nasal consonant, where /r/ occurred (Clements, 2014).

It is also necessary to consider the extent of substrate influences on the formation of the proto-creole. Edoide languages from the Niger Delta region, which today include Bini, Degema, Emai, Engenni, Etsako, Ghothuo, Isoko and Urhobo (Dryer and Haspelmath, 2013), were important in the early years of the colonization of São Tomé and Príncipe and possess a rhotic segment in their phonologies (Hagemeijer, 2011). The same cannot be said about the Western Bantu languages (Bentley, 1887). Since LI was potentially one of the first daughter languages to branch off from the GGPC (at a time before the arrival of Bantu-speaking slaves on São Tomé), the language exhibits favourable conditions for preserving a Portuguese and African rhotic segment through linguistic convergence fed by the inputs to the linguistic melting-pot that was São Tomé at that time. LI is notably the only language in the cluster that possesses a vibrant /r/ in its phonemic inventory. ST, AN and FA do not exhibit rhotic elements in their inventories, perhaps due to their having greater contact with the Bantu substrate or maybe because of the lack of convergence, which fostered the emergence of distinct features. Regarding the phonemic opposition /ɾ/ and /r/, Clements (2014: 352) noted that the contrast was not kept in the Portuguese-based creole languages of the Gulf of Guinea, because of their ecology and the dynamics[6] in the process of change in each language. Our analysis of the data corroborates Clements's (2014) claim. According to the reconstructed items, since its emergence, the GGPC did not maintain the phonological opposition /ɾ/ and /r/, retaining only one rhotic segment, reconstructed as *r. Based on the daughter languages, only ST admits /l/ in the second onset position, as in [ˈplɛ.tu] 'black' (PT *preto*). Thus, at present, the data point to *l as the only proto-liquid in onset second position. In syllable-final position, on the other hand, the evidence points to a distinction between the rhotic and the lateral in proto-forms, as illustrated by their reflexes (descendant forms) in FA, which is the only one of the four languages to permit the lateral in codas, the opposite of the situation for the rhotic, as in *pɛr.tu > [ˈpɛːtu] 'close' (FA), i.e., elision of the rhotic followed by compensatory lengthening, versus *kal.ma > [ˈxal.ma] 'calm' (FA). In addition to this, LI, which is the only language in the cluster to exhibit a rhotic phoneme in the corpus data, as well

6 For more details concerning the reasons that prevented the preservation of the phonemic distinction between /ɾ/ and /r/ in the GGPC's daughter languages, see Clements (2014).

as the only language to allow /r/ in simple onset position, did not maintain the phonological distinction into modern times. In current LI, [ɾ] and [r] occur in free variation in syllable onsets, e.g., [ˈsɛ.rɐ] and [ˈse.ɾɐ] 'saw' (PT *serra*) and, in syllable-initial position, before an approximant, e.g., [uˈrja] e [uˈɾja] 'ear' (PT *orelha*). Nevertheless, /r/ is realized only as [ɾ] when it is the second element of an onset, as in [kɾaˈva] 'to nail/stick/embed' (Agostinho, 2015: 40).

The GGPC liquids' reflexes in the daughter languages cannot occupy all positions in the syllable. The ancestral liquids have therefore developed in ways that are sometimes similar, but which may also vary from language to language. These will be analysed in the following subsections, depending on the class of segment, either rhotic or lateral, and the position the consonant occupies in the syllable, either in the onset or in the coda.

4 Liquids

In the following subsections, we analyse the liquids' evolution in onsets (4.1), codas (4.2) and in complex onsets (4.2.2) by comparing proto-forms and their cognate reflexes in the four daughter languages.

4.1 *Liquids in the Onset*

In initial position, *l is fully maintained. The vibrant *r is retained in onset position only in LI, in whose inventory it occurs. On the other hand, /r/ does not occur in the phonemic systems of ST, FA and AN. These languages have undergone lambdacism of *r (substitution by the alveolar lateral). In these daughter languages, therefore, *l and *r merge as /l/. In (2-a), we see *r in word-initial onset position being substituted by /l/. Example (2-b) shows *r being substituted by /l/ in word-medial onsets.

(2) *r → l / __V

a) **ra*.bu 'tail''(PT *rabo*) > **labu** [ˈla.bu] (ST), **labu** [ˈla.bu] (FA), **labu** [ˈla.bu] (AN)
b) *sɛ.*ra* 'to smell' (PT *cheirar*) > **sela** [sɛˈla] (ST), **sela** [sɛˈla] (FA), **thela** [θɛˈla] (AN)

It is noticeable that the only possible consonant in the second position in GGPC complex onsets was the alveolar lateral *l. In second position in complex onsets, *r has not been reconstructed because there is no evidence that the segment could occupy that position, unlike *l. We cannot reconstruct *bra.su 'arm', for example, basing ourselves on the Portuguese cognate or even on the substrate languages, for, beyond knowing that those languages took part in the formation of the GGPC, there is no linguistic evidence for /r/ in complex onset positions: the reflex found in ST is *blasu.* Although all four daughter-languages possess /l/, only ST allows it to form complex onsets. This characteristic of syllable-formation sets ST apart from the other languages regarding liquid consonant development, both in the second slot of complex onsets (reconstructed in the proto-forms as *l), and in the coda (reconstructed as *r and *l) (see Section 4.2). Regarding the complex onset in the GGPC, ST is the only language to maintain it *in situ*[7] as in (3):

(3) a) **klo*.su 'fruit stone', 'pit' (PT *caroço*) > [ˈklo.su] (ST)

b) **dle*.tu 'law', 'right' (PT *direito*) > [ˈdle.tu] (ST)
c)*dle.te 'to melt' (PT *derreter*) > [dleˈte] (ST)
d) **bla*.su 'arm' (PT *braço*) > [ˈbla.su] (ST)
e) *ku.*bli* 'to cover' (PT *cobrir*) > [kuˈbli] (ST)

Conversely, FA, LI and AN exhibit several developments for *l as the second onset element in those cases in which the structure has been reconstructed in the proto-form. Thus, if the cluster is in a non-final syllable in the proto-language, LI, AN and FA undergo syncope of *l and compensatory vowel lengthening, as in (4).

(4) a. **klo*.su 'fruit stone', 'pit' (PT *caroço*) > [ˈxoːsu] (FA), [ˈkoː] (LI), [ˈkoːθu] (AN)

b. **dle*.tu 'law', 'right' (PT *direito*) > [ˈdeːtu] (FA), (LI), (AN)
c. **bla*.su 'arm' (PT *braço*) > [ˈbaːsu] (FA), [uˈbaːsu] (LI), [ˈbaːsu] (AN)

On the other hand, if *l was in a complex onset in a final syllable, only AN and LI will exhibit simple loss of the liquid consonant, without compensatory lengthening, as in (5).

(5) a. *ku.*bli* 'to cover' (PT *cobrir*) > [kuˈbi] (LI, AN)

b. **fɛ*.blɛ 'fever' (PT *febre*) > [ˈfɛ.bi] (LI), [ˈfɛ.bɛ] (AN)

In FA, in complex onsets in pre-final syllables of the proto-form, GGPC *l was kept by an additive strategy – a copy-vowel. The alveolar lateral thus passes into the onset of the new syllable created by this anaptyxis, a context that permits the lateral's survival, as in (6):

7 Although there are also some cases of metathesis, as in [ˈmla.gu] 'thin' (PT *magro*).

(6) a. *ku.*bli* 'to cover' > [ku'bi.li] (FA)

b. **fɛ*.blɛ 'fever' > ['fi.bi.li] (FA)
c. **pɔ*.bli 'poor' > ['po.bi.li] (FA)
d. **kɔ*.blɔ 'snake' > ['xo.bo.lo] (FA)

One might surmise an epenthetic solution for the complex onsets in ST, just as in the FA forms in example (6), since Hayes (1989b), confronted with similar data from Samothracian Greek, claims that the /CrV/ cluster is split up by an epenthetic vowel that results in a [CVrV] sequence (a solution discussed by Smith (2003) for Suriname creoles). Thus, intervocalic /r/-deletion would apply, being followed by [VV] fusion into [V:]. However, our rejection of epenthesis as a solution is also because the GGPC stress was sensitive to syllable weight. Primary stress fell on the penultimate syllable in nominal forms in which that syllable was light (e.g., **bla*.su 'arm'), but shifted to the last syllable when that was heavy (e.g., *bo.*toN* 'button'). Verbs were always stressed on the final syllable (e.g., *be.*be* 'to drink'). Therefore, a word's morphological category must be considered during stress allocation. Furthermore, trisyllabic proto-forms are not common (Bandeira, 2017). If we analyse stress in FA (a daughter language which does contain trisyllabic reflexes) in (6), we notice that the proto-form is stressed on the penultimate syllable, as in (6b) **fɛ*.blɛ. As stress-application is the first rule to operate, according to Lexical Phonology (Kiparsky, 1982), it is only afterwards that FA applies anaptyxis, owing to restrictions on syllable-type. This is why we find the stress on the antepenultimate in ['fi.bi.li] from **fɛ*.blɛ. Proposing a proto-form like *fɛ.*bɛ*.lɛ would mean that stress would be attributed in an irregular and unbalanced way, contravening a principle of linguistic reconstruction that states: "Reconstructions should fill gaps in phonological systems rather than creating unbalanced systems." (Crowley, 1997 [1992]: 95).

Loss plus compensatory vowel lengthening and anaptyxis when *l occurs in complex onsets is not conditioned by lexical stress in the daughter languages. The independence of these processes regarding word-stress is observable in (8), in which a complex onset appears in an unstressed syllable. One may therefore conclude that the complex onset position in non-final syllables is a conditioning factor that triggers the loss and compensatory vowel lengthening attested in the other three daughter languages, as well as being a factor that inhibits copy-vowel anaptyxis in FA. On the other hand, by simply allowing complex onsets, ST will invariably maintain such structures *in loco* in any syllable in the word.

FA, LI and AN may not present the same patterns of compensatory vowel lengthening in two situations: (i) when lexical items have more than two syllables – except in FA; (ii) when a syllable with a complex onset also possesses a coda

– in all three languages. In both contexts, one observes the lateral's elision without compensatory vowel lengthening, as in (7), for context (i), and in (8), for (ii).

(7) a. *si.kle.*ve* 'to write' (PT *escrever*) > [ʃi.keˈve] (LI), [ʃi.keˈve] (AN)
b. *o.bli.*ga* 'to force', 'to compel', (PT *obrigar*) > [o.biˈga] (LI), [o.biˈga] (AN)
c. *i.gli.*gu* 'smoke' > [i.giˈgu] (LI), [i.liˈgu] (AN)

(8) a. *pleN.*de* 'to lose' (PT *perder*) > [pẽˈde] (FA), [peˈde] (LI), [pẽˈde] (AN)

b. **flɛS*.ku 'fresh', 'cool' (PT *fresco*) > [ˈfɛs.ku] (FA), [ˈfɛʃ.ku] (LI), [ˈfɛ.ku] (AN)

The lack of compensatory vowel lengthening when the complex onset syllable exhibits a coda can be accounted for if one assumes that these three languages reject syllables with more than two morae. Thus, one may imagine that, in languages that are sensitive to syllable weight, the mora is attributed to the nucleus and the coda, thereby blocking vowel lengthening.

In addition, in verbs in LI and AN, vowel lengthening is not observed, despite the same context as in (9). In this case, the grammatical category "verb" may have proven opaque to the process in the languages' development.

(9) a. *kle.*se* 'to grow' (PT *crescer*) > [keˈse] (LI, AN)
b. *plɛ.*sa* 'to loan' (PT *emprestar*) > [pɛˈsa] (LI), [pɛˈθa] (AN)
c. *plɛ.*ga* 'nail', 'to preach' (PT *pregar*) > [pɛˈga] (LI, AN)
d. *kla.*ga* 'to carry', 'to load/to charge' (PT *carregar*) > [kaˈga] (LI)

LI and AN do not exhibit compensatory vowel lengthening in items that come from verbs. This absence may be related to the final stress pattern of verbs in the infinitive in Portuguese. Indeed, there was a historical /r/ in the coda in these forms that was regularly erased in the proto-creole. However, the stress was retained on the formerly heavy, i.e., bi-moraic, final syllable, as the GGPC was sensitive to syllabic weight for attributing stress, just as its daughter languages are. This topic, however, requires more attention in future investigations.

4.2 *Liquid Evolution in Syllable Coda*

In our data, the daughter languages differ in whether *r or *l appears in the coda. Lexical items with coda *r will be analysed first, followed by lexical items with coda *l.

4.2.1 Vibrant in Coda

The development of coda *r separates ST from the other languages. While the liquid may either undergo metathesis or suffer deletion in ST, in LI, FA and AN, *r is always deleted, and the loss may or may not be followed by compensatory vowel lengthening.

In cases of *l-deletion followed by compensatory lengthening in FA, LI and AN, the factor that determines deletion, except in FA, is the feature [coronal]. Before a coronal consonant, (in the examples in 10, the coronal consonants are [n] and [t]), the vibrant is deleted in ST, LI and AN, without compensatory vowel lengthening.

(10) a. **kar*.ni 'meat' (PT *carne*) > [ˈka.ni] (ST, AN), [uˈka.ni] (LI)
b. **kur*.tu 'short' (PT *curto*) > [ˈku.tu] (ST, LI, AN)

In FA, on the other hand, deleting the vibrant also produces vowel lengthening, as in (11).

(11) a. **kar*.ni 'meat' (PT *carne*) > [ˈxaːni] (FA)
b. **kur*.tu 'short' (PT *curto*) > [ˈkuːtu] (FA)

The independence of vowel lengthening with respect to the adjacent coronal feature in FA is an important piece of evidence that there was a rhotic element in the coda in the GGPC, for two reasons. The first is the fact that, if lengthening occurred in the language, a deleted element had once existed there, as revealed by the FA data. In addition, confirmation that this lost element was in the coda can be acquired from the cognates in the other daughters, bearing in mind that in these languages, the coronal feature is only relevant as a conditioning factor for compensatory vowel lengthening if the liquid is in the coda, as in (11), i.e., *r directly adjoined the lengthened vowel. If the liquid appears in onset second position, its potential coronality exerts no influence, owing to it not being adjacent. The second factor involves the nature of the reconstructed liquid consonant in contexts like (11). According to the FA data, in conjunction with their syllabic restrictions, one can affirm that the liquid in GGPC was *r, since FA frequently exhibits its capacity to maintain a consonant in the coda if it is *l, as in **kal*.ma > [ˈxal.ma] 'calm' (PT *calma*). Confronted by these facts, *r was reconstructed in coda position in **kar*.ni 'meat' and **kur*.tu 'short', as in (11-a) and (11-b).

When *r is found in the coda before a non-coronal consonant in the GGPC, however, LI and AN, together with FA, delete the liquid, just as they present long vowels, as in (12).

(12) a. *pur.*ga* 'to purge' (PT *purgar*) > [puːˈga] (FA, LI, AN)

8 There are no reflexes in the daughter languages that indicate the presence of a rhotic in the second position of complex onsets. Even LI, whose phonemic system includes a rhotic, does not support *r when the proto-form contains a complex onset. Thus, proposing the presence of *r in the second onset position and then claiming a lambdacism rule operated does not seem to us to be supported by the data.

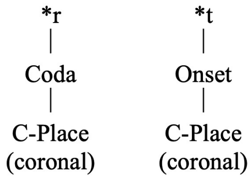
b. *bar.*ga* 'to undo' (PT?(*des*)*embargar*) > [baːˈga] (FA, LI, AN)

Before non-coronals, ST does not exhibit deletion but rather lambdacism, followed by metathesis. We conjecture that *r undergoes lambdacism while still in the coda.[8] Although attested, metathesis in Classical Portuguese was a marginal process and as such it probably did not play a role in the formation of the GGPC. At an intermediate stage, coda *l was metathesized into the second element in a complex onset.

(13) a. *pur.*ga* 'to purge' > *pul.*ga* > plu.*ga* [pluˈga] (ST) (*pur > *pul > plu)
b. *bar.*ga* 'to undo' > *bal.*ga* > bla.*ga* [blaˈga] (ST) (*bar > *bal > bla)

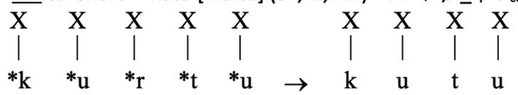
Having noticed the influence of the coronal feature for blocking or allowing total deletion (in the case of LI and AN), and lambdacism and metathesis (in ST), one may infer that such conditioning results from the action of the *Obligatory Contour Principle* (OCP) (Leben 1973, McCarthy, 1986). According to the OCP, identical adjacent segments in a tier must be avoided. Thus, in (14), syllable-final *r is lost when it precedes a coronal consonant, in this case *t:

(14) ***kur.tu** *r > Ø / __ $ C [coronal]



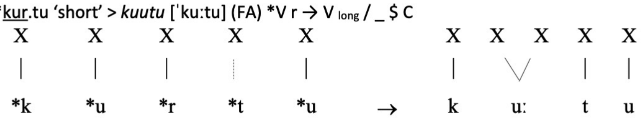
Therefore, the [coronal] feature of *r triggers the OCP when it is adjacent to another consonant with the same feature (in (15), *t). The principle leads to the total loss of the coronal coda in ST, LI and AN, as observed in (15).

(15) *kur.tu 'short' > *kutu* [ˈku.tu] (ST, LI, AN) *V r → Ø _ $ C cor



In FA, however, *r is lost in this context but, nevertheless, the proto-phoneme leaves a time-unit (represented by X) connected to the adjacent vowel root, resulting in vowel lengthening.

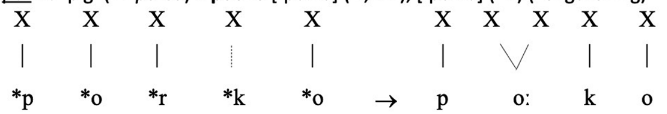
(16) *kur.tu 'short' > *kuutu* [ˈkuːtu] (FA) *V r → V long / _ $ C



In contexts in which coda *r comes before a non-coronal (dorsal or labial) consonant, LI and AN mirror FA, exhibiting loss followed by compensatory vowel lengthening, as in (17) and (18):

(17) **por*.ko 'pig' (PT *porco*) > **pôôko** [ˈpoːko] *Vr → Vlong / _$C[dor, lab] (LI, AN)

(18) *<u>por</u>.ko 'pig' (PT *porco*) > **pôôko** [ˈpoːko] (LI, AN), [ˈpoːxo] (FA) (Lengthening)



The loss of the vibrant in FA, LI and AN happens because codas in these languages do not admit rhotics. The historical coda rhotics have therefore been deleted: totally if the following consonant has an identical feature (*r > Ø / _ $ C [cor.]), as in LI and AN, or partially, should the OCP not be activated (*V r > Ø > Vː / V __ $ C [dor, lab]), leading to compensatory vowel lengthening.

In ST, if the OCP is not activated, *r first undergoes lambdacism (*r > [l] > /l/: [l] < *l), followed by resyllabification. As /l/ cannot occur in the coda, it is moved to the second position in the onset, forming a cluster – a structure that ST allows. These processes, however, arise after the break-up of the GGPC, since the other daughter languages do not exhibit lambdacism followed by metathesis (see 19).

(19) *<u>por</u>.ko 'pig' > *<u>pol</u>.ko (lambdacism) > [ˈplo.ko] (ST) (metathesis)

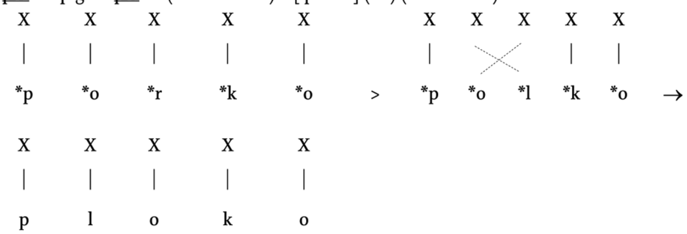


TABLE 13 ST, LI and AN – *r in coda: OCP before [coronal].

Process	OCP	*r deletion
***V r > Ø /_ $ C** $_{\text{[cor.]}}$	Active	√

9 According to Traill and Ferraz (1981), Maurer (2009) and Agostinho (2016), coronal consonants do not seem to be important in the evolutions of liquids consonants. Traill & Ferraz's examples show *CVr > Vː, whatever the following consonant, as in *tooto* 'to skew' and *fuuta* 'to steal' in

TABLE 14 ST – coda *r: no OCP before [dorsal] or [labial].

Process	OCP	*r > l	Metathesis
$*V\ r_{coda} > l_{onset}\ /\ \_\ \$\ C_{[lab,\ dor]}$	Irrelevant	√	√

TABLE 15 LI and AN – *r in coda: OCP before [dorsal] or [labial].

Process	OCP	*r deletion₁ and CL₂
$*V\ r > V{:}\ Ø\ /\ V\ \_\ \$\ C_{[lab,\ dor]}$	Irrelevant	√

We can, therefore, propose some orders for rules involving *r in coda in the daughter languages. The first is the importance of the context of *r in word-medial codas, whether contiguous with the feature [coronal] (/ *r $ C [cor]), in ST, LI and AN.

Contact with the feature mentioned above triggers the OCP, which causes the total deletion of the coronal coda or, in ST, the segment's metathesis. In LI and AN, on the other hand, the process leads either to total deletion or to compensatory vowel lengthening.[9] In ST (Table 14), if the OCP is activated, the coronal coda is eliminated. If the OCP is not triggered, however, the liquid first suffers lambdacism (*r > *l), followed by metathesis (l coda > l onset).

As for LI and AN, the OCP's action entails total deletion when *r is immediately adjacent to a [coronal] feature (/ *r $ C [cor.]). In the absence of the context for applying the rule, both languages first delete *r and afterwards exhibit lengthening of the vowel to the left of the deleted consonant, i.e., *V r > V: Ø (Table 15).

LI. We also recorded these forms, but the former is in variation with *toto* 'skew' which leads us to have reservations about the variant with the long vowel. Additionally, we have found several cases of *CVr before a coronal consonant without long vowels in LI, e.g., *karta > *kata* 'letter', *korta > *kota* 'to cut', *kurtu > *kutu* 'short', *kurtisa > *kutʃisa / kutxisa* 'cork', *lerta > *leta* 'alert' (Agostinho and Araujo, in preparation; Bandeira, field notes). Maurer (2009) and Agostinho's (2016) examples, on the other hand, do not show *CVr > V:, regardless of the following consonant (*petu* 'close'; *poko* 'pig'). In view of the divergent data from LI, two possibilities arise. The first is that coronal consonants are indeed involved in the application of the compensatory lengthening rule. The alternative scenario is that coronal consonants play no part in the application of the compensatory lengthening rule. In the latter case, LI would join the FA in this regard. Whatever the final verdict on LI, it is still the case that ST and AN remain examples of languages that demonstrate the conditioning influence of coronal consonants on compensatory lengthening. One cannot, therefore, disregard such conditioning, if one is to understand the development of the PGG's liquid consonants into some of its daughter languages' systems.

TABLE 16 FA — *r in coda: no OCP (independent of [coronal] in following consonant).

Process	OCP	*r deletion₁ and CL₂
$*V\ r > V{:}\ Ø\ /\ V\_\ \$\ C_{[cor]}$	Irrelevant	√
$*V\ r > V{:}\ Ø\ /\ V\_\ \$\ C_{[lab,\ dor]}$	Irrelevant	√

FA (Table 16) reveals it is not subject to the OCP, since its forms show deletion followed by compensatory vowel lengthening, without regard for context, i.e., *V r > Vː Ø.

4.2.2 Lateral Consonant in Coda

The development of *l in codas also indicates that the coronal feature in the following syllable is a conditioning factor in all the daughter languages, except FA. With regard to *l in coda before coronals in the GGPC, ST, AN and LI delete the liquid, as shown in (20) and (21):

(20) *mal.*da*.di 'badness', 'evil' (PT *maldade*) > [maˈda.dʒi] (ST), [maˈda.di] (AN)

(21) *fal.*ta* 'to lack' (PT *faltar*) > [faˈta] (ST, LI, AN)

In contrast, FA exhibits compensatory vowel lengthening, as well as the deletion of the liquid (see especially 4.2.3), as we can see in (22) and (23)

(22) *mal.*da*.di 'badness', 'evil' (PT *maldade*) > [maːˈda.di] (FA)

(23) *fal.*ta* 'to lack' (PT *faltar*) > [faːˈta] (FA)

In contexts in which *l occurred before non-coronal consonants, ST exhibits metathesis, with the *l from the coda rising to complex onset second position (in 24), while LI and AN simply delete the consonant (25).

(24) *sal.*va* 'to save' (PT *salvar*) > [ʃlaˈva] (ST)

(25) *sal.*va* 'to save' (PT *salvar*) > [saˈva] (LI), [θaˈva] (AN)

FA, however, deletes the liquid and lengthens the preceding vowel, as in (26):

(26) *sal.*va* 'to save' (PT *salvar*) > [saːˈva] (FA)

In Tables 17, 18 and 19, the ordering of the rules for coda *l can be compared across ST, FA, LI and AN.

In the set of cognate reflexes for **al*.ma, even without /l/ appearing before a coronal feature, ST does not exhibit metathesis. One may imagine metathesis as an adaptive strategy that was employed whenever the syllable contained an onset. Lacking an onset in the initial syllable, which the metathesis of *l would

TABLE 17 ST coda *l: OCP active.

Process	OCP	*l deletion	Metathesis
*V l > Ø /_ $ $C_{[cor.]}$	Active	√	Blocked
*l_{coda} > l_{onset} /_ $ $C_{[lab, dor]}$	Irrelevant	--	√

TABLE 18 FA coda *l: OCP irrelevant.

Process	OCP	*l $deletion_1$ and C l_2
*V l > Ø / V _ $ $C_{[cor.]}$	Irrelevant	√
*V l coda > V l onset /_ $ C [lab, dor]	Irrelevant	√

TABLE 19 LI and AN coda *l: OCP active.

Process	OCP	*deletion	* $deletion_1$ and C l_2
*V l > V Ø / V_ $ $C_{[cor]}$	Active	√	Blocked
*V l > VV Ø / V _$ $C_{[lab, dor]}$	Irrelevant	--	√

have turned into a complex onset, ST aligns itself instead with the other languages, undergoing anaptyxis that causes resyllabification, as in (27) Because FA allows alveolar liquids in the coda under certain circumstances, in the case of 'soul', the language maintains *l *in situ*.

(27) **al*.ma 'soul' (PT *alma*) > [ˈa.li.ma] (ST, LI, AN), [ˈal.ma] (FA)

By means of anaptyxis in ST, LI and AN, it was possible to transform a closed syllable into an open one.

4.2.3 Compensatory Vowel Lengthening – Complex Onset and Coda

We argue here that in the daughter languages – with the exception of ST, the only language that does not possess long vowels – compensatory vowel lengthening is triggered by the deletion of *r and *l in the x tier, which leaves a gap that is filled by the lengthened vowel, formerly in the same syllable as the deleted segment. Except for ST, the GGPC's daughter languages exhibit progressive and regressive lengthening. In each case, the vowel is lengthened after the liquid consonant is deleted to preserve the position in the segmental layer. In LI and AN, lengthening may be blocked if the liquid consonant in the coda precedes a consonant with a [coronal] feature. Both types of lengthening will be dealt with below.

Regressive lengthening in FA, LI and AN occurs in a right-to-left direction, as in (28):

(28) *bar.*ga* 'to undo' > [baːˈga] (FA, LI, AN)

However, lengthening in LI and AN can be blocked in this CVC context if the trigger consonant *r or *l is adjacent to a coronal consonant in the following syllable, as in (29):

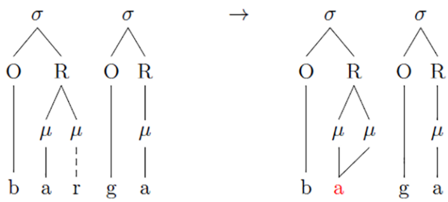
(29) **pɛr*.tu 'close' (PT *perto*) > [ˈpɛ.tu] (LI, AN)

Besides the CVC context, it is possible that vowel lengthening is activated by an *l preceding the vowel in a left-to-right movement in CCV syllables, as in (30).

(30) **bla*.su 'arm' (PT *braço*) > [ˈbaːsu] (FA, AN), [uˈbaːsu] (LI)

Whilst the first kind of CVC > CV:Ø lengthening shown in (29) is extremely common amongst the world's languages (Gess, 2011), the kind presented in (30) is rare and is controversial to linguists, mainly because it provides a counter-example to the structure of Moraic Phonology (Hayes, 1989). According to the traditional exposition proposed by Hayes (1989), compensatory lengthening is based on mora conservation. Thus, lengthening happens to preserve the mora left behind by segment deletion; although a segment is deleted, its time-unit (the mora) is not. The mora left by the deleted segment is re-associated with an adjacent melodic unit. According to this theory, the lexical item *bar.*ga* 'to undo' (< PT?(*des*)*embargar*), for example, first syllabifies and receives morae. Next, the coda *r is deleted but its mora survives and is reconnected to the preceding vowel, producing the vowel lengthening, as in (31):

(31) *bar.*ga* 'to undo' (PT?(des)embargar) > **baaga** (FA, LI, AN)



According to Hayes (1989), compensatory lengthening should only occur when a moraic segment is deleted, i.e., only the deletion of a vowel or consonant in a moraic coda causes lengthening. In Moraic Phonology, this generalisation is based on the alleged absence of records of deletion in the onset followed by lengthening (Campos-Astorkiza, 2011). However, in languages such as Samothracian Greek, compensatory lengthening is triggered by rhotics in word-initial position (r V) or in post-consonantal position (C r V) but does not happen when the segment is intervocalic (V r V) or in word-final position (V r) (Topintz, 2006: 2). Despite Samothracian Greek being generally referred to as one of the rare cases of genuine compensatory lengthening after the loss of an onset segment,[10] FA, LI and AN also join the group of languages that exhibit lengthening in the context mentioned in (31). Unlike the Greek variety, however, the Gulf of Guinea daughter languages also exhibit the more common type of vowel lengthening caused by liquid consonants in CVC codas, as in (30).

There are two main treatments in the theories that analyse the phenomenon: the phonetic conservation model and the phonemic conservation model. According to phonetic conservation (see Timberlake, 1983), compensatory lengthening is thought of as a functional process that aims to preserve part or all the duration of the lost segment's physical material. One of the principal criticisms of this model is the assumption that the morae associated with consonants are equivalent in the duration to those associated with vowels. According to the phonological conservation model (Hayes, 1989), however, compensatory lengthening is classified as a functional process aimed at preserving some aspect of phonological representation (a supra-segmental unit) that is associated with the loss of segmental material. The principal objection to this treatment, advanced by Optimality Theory (OT), is that consonants must have their weight attributed before loss occurs, suggesting a serial analysis (Gess, 2011: 1–15).

Hayes' model attributes compensatory lengthening to mora preservation, postulating additionally that onsets do not bear morae and onsets are, thus, prohibited from activating the process. The data from the GGPC's daughter languages, however, except for ST, constitute counterexamples to this vision. Confronted by data such as those of Samothracian Greek, Hayes (1989) proposes that the cluster /C r V/ is broken up by anaptyxis, resulting in a [C V r V] sequence. Next, the rule deleting intervocalic /r/ is applied, followed by the fusion of [VV] into [Vː]. From this perspective, lengthening is merely an epiphenomenon of vowel fusion, rather than being the direct result of compensatory lengthening after deletion in the onset (Topintzi, 2006: 5). In the case of FA, LI and AN, in which vowel lengthening occurs after the deletion of *l in ClV onsets, Hayes' (1989) analysis is inappropriate, since it would be necessary to imagine the existence of morae in the daughter-languages' input, as well as admitting

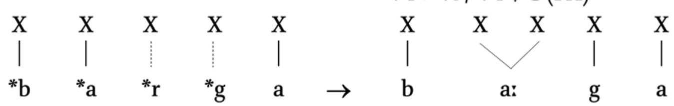
10 Topintzi alleges that the onset may not be the only trigger for lengthening, as in Samothracian Greek but also the target for lengthening, as other languages indicate. For more information about onset behavior and compensatory lengthening, see Topintzi (2006) and Campos-Astorkiza (2011).

unattested intermediate stages. If one considers that the input for FA, LI and AN – the GGPC – can exhibit complex onsets like **bla*.su 'arm', for instance, one cannot accept that, in addition to the vowel, the *l in the onset bears a mora, something Hayes' (1989) structure does not allow. According to Hayes' (1989) model, onsets cannot cause compensatory lengthening because they do not bear morae. Codas, on the other hand, can cause it, but only if they are moraic. The nucleus, in turn, can be both the target and the trigger since it bears morae. Therefore, the loss of the nucleus should, according to this assumption, always cause compensatory lengthening, which does not always occur (Topintzi, 2006: 105). In languages such as Chukchi, Tangale (Kenstowicz, 1994: 96), Klamath (Odden, 2005: 121) and Icelandic (Odden, 2005: 189–190), as well as the loss of /r/ in some contexts not leading to lengthening, deleting the nucleus also does not trigger the process, which raises the question as to the strength of the connection between mora and compensatory lengthening. Topintzi (2006: 32) points out that segment elimination, even of undeniably moraic segments, is no guarantee of compensation by lengthening. Such facts make the explanation for the absence of compensatory lengthening as the result of a lack of mora less convincing. Finally, Topintzi (2006: 32) judges that the absence of the process after the loss of the onset in most languages may be related simply to the scarcity of lengthening in the languages and not to the lack of weight itself.

Moreover, Gess (2011: 15) ponders whether preserving the timing associated with a weight-bearing unit may be favoured in general, since all segments, moraic or otherwise, possess a physical duration associated with them, prejudicing units that do not bear weight. This does not mean, however, that compensatory lengthening cannot occur because of the loss of non-moraic segments. Our analysis, therefore, allies itself with the views proposed by Topintzi (2006) and Gess (2011), since the GGPC's daughter languages, except ST, exhibit compensatory vowel lengthening that result from the loss of consonantal segments, whether moraic or not.

In (32) and (33), the loss of the liquid consonants *r and *l subsequently leads to vowel lengthening. Thus, rather than referring to compensatory lengthening as the result of segment deletion in the onset or coda, it seems preferable to refer to the process as the outcome of *r- and *l-deletion only, transferring the syllabification process to another derivational stage, as envisaged in the Lexical Phonology model.

(32) ***bar.<u>ga</u> 'to undo'** > **[baː'ga] (FA, LI, AN)**
***V r > Vː / V r $ C $_{[dor], [lab]}$ (LI, AN)**
***V r > Vː / V r $ C (FA)**



(33) *p<u>l</u>ɛ.tu 'black' (PT *preto*) > [ˈpɛː.tu] (FA, LI, AN)
*V r > Vː / V r \$ $C_{[dor], [lab]}$ (LI, AN)
*V r > Vː / V r \$ C (FA)

X	X	X	X	X		X	X X	X	X
*p	*l	*ɛ	*t	u	→	p	ɛː	t	u

5 Final Remarks

In this paper, we have investigated how the Portuguese-based Gulf of Guinea proto-creole, a language that allowed lateral and vibrant consonants in syllable onsets, in coda position and /l/ as the second consonant in complex onsets, evolved into four daughter languages with restrictions on the occurrence of these consonants at various levels. From the reconstruction in Bandeira (2017), in which the occurrence of complex onsets in the GGPC was proposed (as still occur in ST), instead of long vowels (as in FA, LI and AN), we sustain that by admitting complex onsets, such as /C l/, ST maintains the same kinds of onsets as the GGPC *in situ*, e.g., [ˈpla.sɐ] (ST) < **pla*.sa 'square', 'plaza' (PT *praça*). LI, AN and FA, on the other hand, exhibit liquid deletion, as they do not admit such syllable structures, and this loss may or may not be followed by compensatory vowel lengthening. Thus, if the complex onset appears in the first syllable of a (disyllabic) proto-form, independent of word-stress, LI and AN exhibit *l-deletion and subsequent vowel lengthening, e.g., [ˈpaː.sɐ] (LI), [ˈpaː.θɐ] (AN) < **pla*.sa 'square', 'plaza'. On the other hand, if the complex onset appears at the right edge of the proto-form, only deletion occurs, e.g., [ˈma.gu] (LI), [ˈmɛ̃.gu] (AN) < **ma*.glu 'thin' (PT *magro*). By contrast, vowel lengthening in FA is not sensitive to the syllable with the complex onset's position in the word and so occurs more consistently than in its two sister languages mentioned above. Moreover, FA also exhibits anaptyxis (generally a copy-vowel) to break up clusters, e.g., [ˈmũ.gu.lu] (FA) < **ma*.glu 'thin'. The outcomes also change when the proto-liquid consonants occurred in syllable codas: LI and AN exhibit loss of *r and *l in contexts in which the proto-phonemes preceded coronal consonants, e.g., [ˈpɛ.tu] (LI, AN) < **pɛr*.tu 'close' (PT *perto*). However, if coda *r occurred before a dorsal or labial consonant, the two languages exhibit compensatory vowel lengthening, as well as loss, e.g., [ˈpoː.ko] (LI, AN) < **por*.ku 'pig' (PT *porco*). On the other hand, in both LI and AN, coda *l was generally lost without vowel lengthening, even when it was not adjacent to a coronal feature, e.g., [ka.dɛˈra.dɐ] (LI), [ka.dɛˈla.dɐ] (AN) < *kal.dɛ.*ra*.da 'fish stew'. Once again, FA reveals itself to be unaffected by conditioning rules, which in this case would be contiguity with a coronal feature, e.g., [xaː.daˈlɐ̃] (FA) < *kal.dɛ.*ra*.da 'fish stew'. As for ST, it exhibits lambdacism (in the case of *r) and metathesis (of /l/ < *r and *l) when the ancestral liquid in the coda occurred immediately before a consonant with a dorsal or labial feature, e.g., [ˈplo.ko] (ST) < **por*.ko 'pig'. However, if the original liquid in the coda appeared before a consonant with a coronal feature, ST does not exhibit lambdacism or metathesis but just deletes the liquid, e.g., [ˈpɛ.tu] (ST) < **pɛr*.tu 'close'. The inhibitory influence of an adjacent coronal feature on the occurrence of /l/-metathesis or compensatory vowel lengthening confirms the original position of *r and *l in the coda, since the coronal feature exerted no influence whenever the liquid consonant was found in the second position in the onset, i.e., coronality does not prevent compensatory vowel lengthening in LI, AN and FA, e.g., [ˈgle.zɐ] (ST), versus [ˈgeːzɐ] (LI, FA), [ˈŋgeːðɐ] (AN) < **gle*.za 'church' (PT *igreja*).

How can one explain the intricate pattern of outcomes for liquids in syllable codas, as in (34) and (35), with recourse only to the Portuguese or ST etymon? Since development patterns change when the liquid consonants are found in coda position, LI and AN exhibit deletion of *r and *l in those contexts in which the proto-phonemes preceded coronal consonants, e.g., **petu** (LI, AN) < **pɛr*.tu 'close'. At the same time, whenever *r in coda preceded a dorsal or labial consonant, both languages exhibit compensatory vowel lengthening, as well as *r-deletion, e.g., **pôôko** (LI, AN) < **por*.ko 'pig'. However, examples such as (35) show that only by postulating *r in the proto-form's coda can one derive the lengthening in FA, bearing in mind that the language is not subject to the action of the OCP.

(34) **por*.ko 'pig' (PT *porco*) (GGPC) >
[ˈplo.ko] (ST), [ˈpoːxo] (FA), [ˈpoːko] (LI, AN)

(35) **pɛr*.tu 'close' (PT *perto*) (GGPC) >
[ˈpɛ.tu] (ST), [ˈpɛːtu] (FA), [ˈpɛ.tu] (LI, AN)

In this paper, we have shown that the relation between long vowels and liquid consonants, both in coda position and in complex onsets, in ST, AN, LI and FA, can be better understood if we consider the modern lexical items in these four languages as continuations of proto-creole forms, with characteristic modifications in each daughter language that are governed by regular processes. Nevertheless, relating the processes described here with the interaction of stress and syllable weight still requires further investigation.

However, the phonological and lexical reconstruction provides evidence of genetic kinship that is the result of descent from a shared proto-system. This empirical verification has consequences for all studies that compare the daughter languages' lexicons with Portuguese etyma. Until now, the studies that undertook comparative analyses of the Portuguese-based creole languages of the Gulf of Guinea did so taking Portuguese as their direct ancestor – see Rougé's dictionary (2004) and Maurer's glossaries (1995, 2009). Yet again, in (36), in which there is a favourable context for vowel lengthening in LI, FA and AN, the three languages do not exhibit it, probably because the coda was occupied by a nasal consonant in the GGPC, as is demonstrated by the proto-form and its reflexes. Thus, how does one explain the blocking of vowel lengthening without considering the nasal consonant in the proto-form's coda, which has reflexes in the items in the cognate set, and instead consider only the ST lexical item that exhibits no nasality?

(36) **pleN*.de 'to lose'

[pleˈde] (ST)
[pẽˈde] (FA)
[peˈde] (LI)
[pẽˈde] (AN)

The present configuration in the daughter languages: the example of the long vowels, in the case of LI, Fa d'Ambô and AN, or of complex onsets, in the case of ST, and the way the liquid consonants have evolved in general, is the result of interactions between the phonemic make-up of the GGPC and a series of regular phonological processes that operated in the various speciation settings on each island.

In addition, the reconstructed items provide evidence that the Gulf of Guinea Proto-Creole was distinct from Portuguese and, therefore, should not be described as a variety of that language, since it exhibits phonological features that are not found in any vernacular variety of Portuguese. Analogously, just as it is not possible to define the GGPC and its daughter languages as reflexes or direct copies of Portuguese, neither can we classify them as reproductions of one of the substrate languages.

References

Agostinho, Ana Lívia. 2015. *Fonologia e Método Pedagógico do Lung'Ie*. Tese (Doutorado em Filologia e Língua Portuguesa) — Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

Agostinho, Ana Lívia. 2016. *Fonologia do Lung'Ie*. Lincom Studies in Pidgin and Creole Linguistics 15. Munich: LINCOM GmbH.

Araujo, Gabriel Antunes de and Ana Lívia Agostinho. 2010. Padronização das línguas nacionais de São Tomé e Príncipe. *Língua e Instrumentos Linguísticos* 26: 49–81.

Araujo, Gabriel and Tjerk Hagemeijer. 2013. *Dicionário Santome-Português/Português-Santome*. São Paulo: Hedra.

Araujo, Gabriel Antunes de, Parth Bhatt and Tjerk Hagemeijer. 2012. Revisiting the liquid consonants in the Gulf of Guinea Creoles. Paper presented at the Formal Approaches to Creole Studies Conference III, 15–16/11/2012.

Araujo, Gabriel Antunes de, Ana Lívia Agostinho, Alfredo Christofoletti Silveira, Shirley Freitas, and Manuele Bandeira. 2013. Fa d'ambô: língua crioula de base portuguesa de Ano Bom. *Cadernos de Estudos Linguísticos* 55(2): 25–44.

Araujo, Gabriel Antunes de & Agostinho, Ana Lívia. 2014. Fa do Vesu, a language game of Fa d'Ambô. *PAPIA*, São Paulo, v. 24, n. 2, p. 265-281.

Balduino, Amanda Macedo; Agostinho Gabriel Antunes de Araujo, Gabriel, and Alfredo Silveira Christofoletti. 2015. A nasalidade vocálica em Santome e lung'ie. *PAPIA*, São Paulo, v. 25, n. 1, p. 7–25.

Bandeira, Manuele. 2017. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. USP, Tese de doutorado.

Bandeira, Manuele & Araujo, Gabriel Antunes de. A estratégia reflexiva no protocrioulo do Golfo da Guiné. *Documentação e Estudos em Linguística Teórica e Aplicada*, v. 38, n. 2. Pre-print. DOI: https://doi.org/10.1590/1678-460×202147848

Barrena, Natalio. 1957. *Gramatica anobonesa.* Madrid: Instituto de Estudios Africanos.

Bentley, William Holman. 1887. *Dictionary and grammar of the Kongo language*. London: Baptist Missionary Society.

Campbell, Lyle. 2004 [1998]. *Historical linguistics*: an introduction. Edinburgh: Edinburgh University Press.

Campos-Astorkiza, Rebeka. 2011. What drives compensatory lengthening? Beyond moraic conservation. In: Socarrás, Gilda and Almitra Medina (eds.), *Philological Research*. Athens: Athens Institute for Education and Research, p. 9–24. Available in: https://bit.ly/2CHDDw2 (accessed July 14, 2014).

Clements, George Nick and Hume, Elizabeth. 1995. The Internal Organization of Speech Sounds. In: Goldsmith, John (Ed.). *Handbook of Phonological Theory*, 245–306. New Jersey: Wiley-Blackwell Publishing.

Clements, J. Clancy. 2014. The status of Portuguese/Spanish /R/ and /r/ in some Iberian-based creole languages. *PAPIA* 24(2): 315–328.

Coelho, Adolpho. 1880. Os dialetos românticos ou neolatinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa*. 2. série, v. 3, 29–96. Lisboa: Academia Internacional da Cultura Portuguesa.

Crowley, Terry 1997[1992]. *An introduction to historical linguistics*. Auckland: Oxford University Press.

Dryer, Matthew S. and Martin Haspelmath (Ed.). 2013. *The World Atlas of Languages Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Available in: <http://wals.info/>. (access October 23, 2016).

Ferraz, Luis Ivens. 1979. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Ferraz, Luis Ivens & Traill, Anthony. 1981. The interpretation of tone in Principense creole. *Studies in African Linguistics*, v. 22, n. 2, p. 205-15.

Ferraz, Luis Ivens. 1987. The Liquid in the Gulf of Guinea Creoles. *African Studies* 46: 287–295.

Fox, Anthony. 1995. *Linguistic reconstruction*: an introduction to theory and method. Oxford: Oxford University Press.

Gess, Randall. 2011. Compensatory Lengthening. In: Van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume, and Keren Rice (Eds.), *The Blackwell Companion to Phonology*. Nova Jersey: Wiley-Blackwell Publishing. Blackwell Reference Online. Available in: http: //www.companiontophonology.com. (access July 13, 2016).

Goldsmith, John. 1976 [1979]. *Autosegmental phonology*. PhD Dissertation, MIT. New York: Garland.

Hagemeijer, Tjerk. 2011. The Gulf of Guinea Creoles: Genetic and typological relations. *Journal of Pidgin and Creole Languages* 26: 111–154.

Hagemeijer, Tjerk. 2009. Initial vowel agglutination in the Gulf of Guinea creoles. In: Aboh, Enoch and Norval Smith (eds.), *Complex processes in new languages*, 29–50. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253–306.

Hock, Hans Henrick. 1991. Initial strengthening. In: Dressler, W. U., H. C Luschützky, O. E. Pfeiffer, and J. R. Rennison (eds.), *Phonologica 1988*: Proceedings of the 6th International Phonology Meeting, 101–10. Cambridge: Cambridge University Press.

Kaufman, Terence. 1990. Language history in South America: what we know and how to know more. In: Payne, Doris (ed.). *Amazonian* Linguistics, 13–31. Austin: University of Texas Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In: Hulst, Harry van der and Norval Smith (Eds.), *The structure of phonological representations*, 93–106. Dordrecht: Foris.

Ladhams, John. 2007. Article Agglutination and the African Contribution to the Portuguese-based Creoles. Available online: shorturl.at/uzIN1 (Access: Oct 20, 2017).

Leben, William. 1973. *Suprasegmental Phonology*. PhD Thesis. Massachusetts: MIT.

Maurer, Philippe. 1995. *L'Angolar: Un créole afroportugais parlé à São Tomé; Notes de grammaire, textes, vocabulaires*. Hamburg: Helmut Buske Verlag.

Maurer, Philippe. 2009. *Principense. Grammar, texts, and vocabulary of the Afro-Portuguese creole of the Island of Príncipe, Gulf of Guinea.* London: Battlebridge Publications.

McCarthy, John J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 19: 451–475.

Michaelis, Susanne Maria et al. 2013. Atlas of Pidgin and Creole Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online: <http://apics-online.info>. (access Sept 6, 2015).

Negreiros, António Lobo de Almada. 1895. *História Etnographica da Ilha de S.Tomé*. Lisboa: Antiga Casa Bertrand.

Odden, David. 2005. *Introducing Phonology*. Cambridge: Cambridge University Press.

Ribeiro, Manuel Ferreiro. 1888. Dialecto da ilha do Principe. MS número 11.23.12, *Schuchardt Archive* da Universidade de Graz, Austria.

Rougé, Jean-Louis. 2004. *Dictionnaire étymologique des créoles portugais d'Afrique*. Paris: Karthala.

Rougé, Jean-Louis and Emmanuel Schang. 2006. The origin of the liquid consonant in Saotomense Creole. In Bhatt, Parth and Ingo Plag (eds.), *The Structure of Creole Words*, 23–37. Tübingen: Niemeyer.

Schuchardt, Hugo. 1882 [2008]. On the creole Portuguese of São Tomé (West Africa). Annotated translation from the German of Ueber des Negerportugiesische von S. Thomé. Sitzungsberichte Wien, v. 101, n. 2, p. 889–917. In: Hagemeijer, Tjerk and John Holm (Eds.). *Contact Languages*: Critical Concepts in Linguistics, v. 1, 31–156. London, New York: Routledge.

Schuchardt, Hugo. 1888. Ueber das Negerportugiesische von Annobom. *Sitzungsberichte Wien*, v. 116, 193–226.

Schuchardt, Hugo. 1889. Ueber das Negerportugiesische der Ilha do Príncipe. *Zeitschrift für romanische Philologie*, v. 13, 463–75.

Segorbe, Armando. 2010. *Gramática descriptiva del fa d'ambô.* Barcelona: CEIBA Ediciones.

Selkirk, Elizabeth. 1984. *Phonology and syntax: the relation between sound and structure.* Cambridge: MIT Press.

Smith, Norval. 2003. New evidence from the past: to epenthesize or not to epenthesize. In Plag, Ingo (Ed.), *Phonology and Morphology of Creole Languages*, 91–107. Tübingen: Niemeyer.

Timberlake, Alan. 1983. Compensatory lengthening in Slavic 2: Phonetic reconstruction. In: Flier, M. S. (Ed.). *American Contributions to the 9th International Congress of Slavists: Linguistics*, p. 293-319. Slavika: Kiev.

Topintzi, Nina. 2006. A (not so) paradoxical instance of compensatory lengthening: Samothraki Greek and theoretical implications. *Journal of Greek Linguistics* 7: 71–119. Available online: https://bit.ly/2CDwEnR. (access in May 14, 2016).