

The common African lexical core of the Upper Guinea Creoles and its historical significance¹

Nicolas Quint & Karina Moreira Tavares

LLACAN – UMR8135 (CNRS/INALCO/USPC) - France

In this paper, we scrutinize and compare the African lexical elements shared by several varieties of Upper Guinea Portuguese Creoles, belonging to the three branches of this group: (i) Continental (Bissau, Cacheu, Casamance, and Geba), (ii) Insular (Fogo and Santiago), and (iii) ABC islands (Curaçao). This is the first study to provide comparative data for a wide range of different creoles of this group. The comparison is based on a standard list of 96 African-derived terms attested in Santiago Capeverdean Creole and builds on three main comparative criteria: (i) presence/absence of a given African-derived term, (ii) phonetic similarity and (iii) semantic similarity. The results show how these items help us understand better (i) which African languages were the main contributors to the formation of the early Upper Guinea Proto-Creole that must have been spoken around the turn of the 15th and 16th centuries, as well as (ii) the historical and phylogenetic relationships existing between the various Upper Guinea Portuguese Creoles at stake. Particularly worthy of mention is the examination of the African element in Fogo Capeverdean and the considerations about the internal classification of the Continental branch of this group of creoles.

Keywords: African substrate, lexicon, comparative linguistics, historical linguistics, reconstruction, Upper Guinea Portuguese Creoles.

1. Introduction

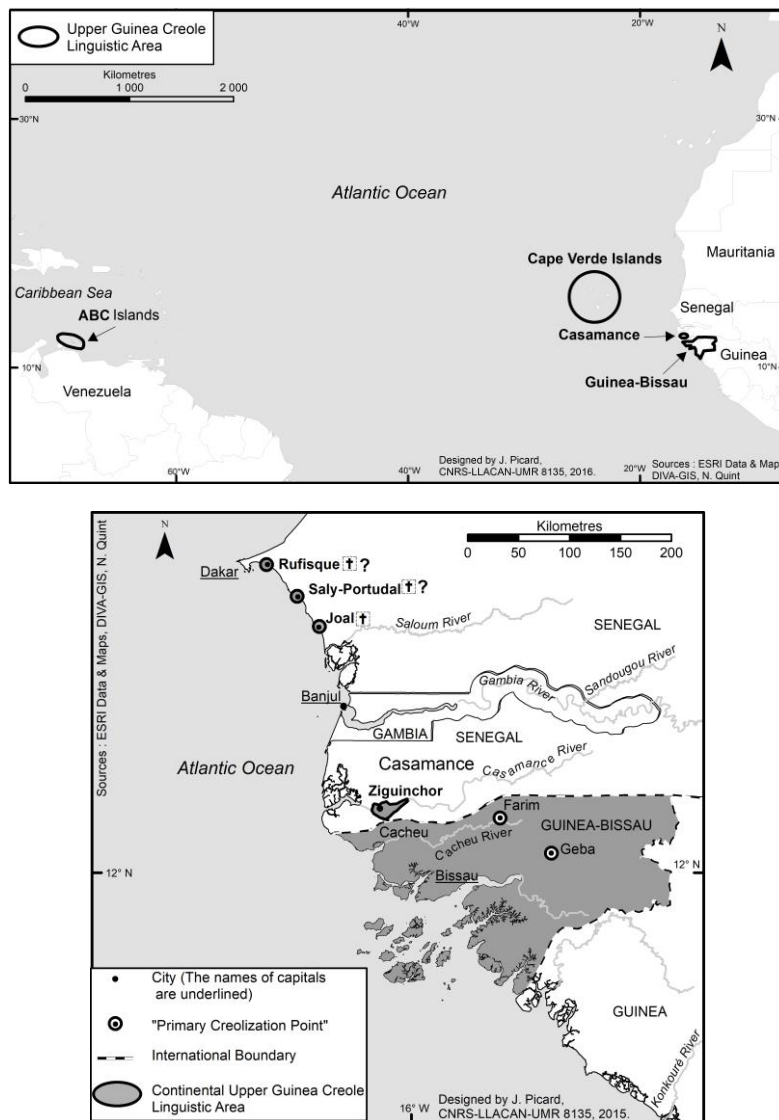
This study provides a systematic comparison of the African lexical component in seven different varieties belonging to the three branches² of the Upper Guinea

¹ We are much indebted to our various language consultants for the different creoles studied herein, as well as to the following people: Richenel Ansano, Sokhna Bao Diop, Gerard van Buurt, Denis Creissels, El-Hadji Dièye, Tjerk Hagemeijer, Bart Jacobs, Ange Jessurun, Braima Sam Mendes, Aires Semedo, Celina Tromp, who helped us refine our data and knowledge of the various languages considered in this study. The usual authorial disclaimer applies. This research was supported by the research program of the French Agence Nationale de la Recherche, *Sénélangues* (ANR-09-BLAN.0326-02), supervised by Stéphane Robert.

² For the nomenclature of the three branches of UGPCs, see Biagui, Nunez & Quint (in press).

Portuguese Creoles (UGPC): Insular (Capeverdean), with two varieties (Fogo (FG) and Santiago (ST)); Continental (Guinea-Bissau and Casamance), with four varieties (Bissau (BI), Cacheu (CU), Casamance (CS) and Geba (GE)); and finally ABC (Papiamentu), represented by one variety (Curaçao (ABC)). The geography of these branches and varieties are shown in the two maps shown in Figure 1.

Figure 1: Maps of the branches and varieties of UGC used in this study



The role and importance of substrate languages is one of the most debated questions in the field of Creole studies (Arends, Kouwenberg & Smith 1995: 99-109; Holm 1988: 65-68, 79-89; Parkvall 2000; Todd 1990: 53-55). As far as

the UGPCs are concerned, the substrate is essentially represented by several Niger-Congo languages (Quint 2008; Rougé 1999a), belonging mainly to Mande (Manding, especially Mandinka) and Atlantic (Wolof and Temne). Several researchers have published extensively on the African component present in the various branches of UGPC; a list of key references includes Biagui, Nunez & Quint (in press), Lang (2009), Quint (2000b, 2006, 2008, 2012) and Rougé (1988, 1994, 1999a, 1999b, 2004, 2005, 2006) for the Cape Verde and Guinea-Casamance areas, and van Buurt (2001), Jacobs (2014) and Maurer (1988, 1994, 1998) for Papiamentu. Other works worthy of mention are Dieng (1985), Holm & Intumbo (2009), Martinus (1997), Santos (1979) and Scantamburlo (2002). Following Quint (2000b and elsewhere) and Jacobs (2012 and elsewhere), this paper considers Papiamentu a member of the UGPC group.

The present paper is organized as follows. In section 2, the choice of each creole variety considered in this paper is explained and justified. Section 3 contains the methodology and implications of this study and discusses (i) the standard used for comparison and the data provenance, (ii) the criteria used to compare the different UGPC varieties at stake, and (iii) the significance and scientific interest of this comparison. Section 4 is the core of the paper: it provides an analysis of the results and deals with the existence of an African core common to most UGPC varieties and of an Insular vs. Continental split. On the basis of African-derived lexical items, section 4 additionally discusses the relationships between the UGPCs included in the comparison, namely Papiamentu (ABC), Fogo and Santiago Capeverdean, Casamance Creole and the northern Continental sub-branch, and Geba Creole and the eastern Continental branch. Section 5, the conclusion, summarizes the main findings and results and suggests possible directions for future research. The data used in this study and their statistical treatment are presented in Appendices 1 to 3.

2. Justification of the choice of the varieties

The Insular varieties of Santiago and Fogo appear to be the oldest Portuguese Creoles (PCs) to have historically developed in Cape Verde (Andrade 1996; Moreira *forthc.* a, b; Quint 2008), since these two islands were both settled before the second half of the 16th century. In all likelihood, Santiago is the Capeverdean variety displaying the highest number of African (or non-Portuguese) features at all levels of its grammar, including lexicon (Quint 2000b: 67-97, 2008: 18; Rougé 2004: 1-3). Fogo African-derived lexicon has hitherto not been the subject of any detailed study. However, first-hand data

collected by Moreira (forthc. b) allow us to shed new light on the peculiarities of this original, understudied Creole variety. Note that both Fogo and Santiago belong to the *Sotavento* (Leeward) or southern group of Insular UGPCs. The northern group of islands (*Barlavento* or Windward) is not taken into consideration here for two reasons: first, the lexical data available on these varieties remains scanty, in particular as regards African-derived items; second, the Capeverdean Barlavento islands seem to have been settled by two main types of population (Quint 2000b; Andrade 1996), coming, respectively, from Capeverdean Sotavento islands and the Portuguese-speaking islands Madeira and Azores. Therefore, there does not seem to have been any significant African influx during the settlement period of Barlavento and, in the absence of specific lexical studies dealing with Capeverdean Barlavento varieties, it appears safe to assume that the African-derived items these varieties present made their way into the Barlavento group mainly due to the presence of Sotavento settlers.

With respect to Continental varieties, especially since the second half of 20th century, the Bissau variety – originally spoken in the city of Bissau since the 17th century – has become the dominant language of Guinea-Bissau, being spoken as a first or second language by at least 1,000,000 people (Jacobs & Quint 2016: 69). Other specific creole varieties are historically attested in several other Bissau-Guinean cities, in particular Cacheu, the first Portuguese trading post in Guinea Bissau, founded ca. 1590 (Boulègue 2006: 49; Carreira 1984: 19-20), and Geba, founded, like Bissau, in the mid-17th century (Carreira 1984: 20). These are respectively considered to illustrate the northern and eastern varieties of Continental UGPCs, contrasting with the Bissau variety, which can be considered the central (=southern) variety (Intumbo, Inverno & Holm 2013: 31; Wilson 1962: VII cited in Couto & Embaló 2010: 35). In practice, the Cacheu and Geba UGPCs have never been the object of any specific descriptive work. Casamance Creole, like Cacheu Creole, belongs to the northern group of Continental UGPCs. However, its sociolinguistic and historic profile is quite distinct given that (i) since the cession of Ziguinchor to France in 1886 (Biagui 2018: 18; Biagui & Quint 2013: 41), it is the only Continental UGPC to be spoken in a place where Portuguese is not the official language; (ii) since 1645, the foundation of Ziguinchor by Afro-Portuguese settlers coming from Cacheu, Casamance Creole has developed (Biagui, Nunez & Quint, in press) in close contact with various local adstrates (in particular Nyun), which are not present in the other ancient trading-posts where Continental UGPCs have historically developed. Casamance Creole lexicon, which was already studied by Rougé (1988), is now being described in depth by Biagui (in progress).

Concerning the third group, the ABC islands, there is an ongoing debate about the Portuguese component of Papiamentu, an Iberian-based Creole whose contemporary form displays a majority of Spanish elements. Several studies point towards the existence of a strong relationship between Papiamentu and the remaining UGPCs (Martinus 1997; Quint 2000b; Jacobs 2012). The present paper accepts that the evidence provided by these studies justifies the inclusion of Papiamentu in the family of UGPCs. As already shown in Quint (2000b: 159-160), and as will be discussed below, Papiamentu exhibits a few African-derived lexical elements that are shared with other UGPCs. The Papiamentu variety considered herein for the sake of comparison is Curaçaoan Papiamentu, the best-known variety of the ABC islands. Unfortunately, detailed lexical information is not available for the remaining two Papiamentu varieties of Aruba and Bonaire.

3. Methodology and implications of the present paper

3.1. Standard used for the comparison and data provenance

The present comparative study is based on a standard, i.e. a list of African-derived terms most of which were already identified by Quint (2000b, 2006, 2008) and Rougé (2004) for Santiago Capeverdean Creole.³ The choice of this standard is related to the following: (i) the African lexical element in Santiago Capeverdean has already been studied in detail and therefore offers a good starting point for a comparative study among UGPCs; (ii) quantitatively, the size of the African-derived lexicon in Santiago (96 items are discussed in this paper) is somewhat halfway between Papiamentu (less than 20 known African-derived items, of which 4 have cognates in other UGPCs, see Table 1.8) and Continental Creoles, in which the African-derived items represent probably over 10% of the total lexicon in conservative varieties and are far from being fully documented or classified today. At present, it is therefore impossible to take Papiamentu as a standard for comparison since the African-derived items are too few to allow any quantitative study. Conversely, the abundance and diversity of African-derived items in the Continental Creoles is not yet fully described.

³ The etyma of words, which were not already discussed by Quint or Rougé or for which a new origin is proposed, are discussed in Appendix 2 (Table 2.1).

Note also that, among the Insular UGPCs, Santiago Creole qualifies better than Fogo as a suitable standard for comparison, given that the present paper shows that the African-derived lexical element in Fogo is noticeably more reduced than in Santiago. Furthermore, research on the African lexical element in Fogo is incipient and to date the main available resources are the data published in this paper.

A further reason why it makes sense to take the Santiago variety as the standard is that the socio-historical and linguistic data on the UGPCs strongly suggest that Santiago was the place where Proto-UGPC nativized and from where it spread to the other regions where an UGPC developed (see Jacobs 2010 and Quint 2000b for discussion).

The standard sample of Santiago Capeverdean African-derived lexemes comprises 96 items, coming from various languages, as shown in Table 1.

Table 1: Origin of African-derived items in the Santiago Capeverdean standard sample

Origin	Number	%
Manding (Mandinka) ⁴	42	44
Wolof	22	23
Manding and/or Wolof ⁵	9	9
Temne	4	4
Other Atlantic languages ⁶	5	5
Bantu languages	3	3
Unidentified non-European items ⁷	11	12
TOTAL	96	100

⁴ Mandinka is the westernmost variety of the Manding cluster. In practice, most Manding-derived terms attested in UGPCs can be traced back to Mandinka (Quint 2008: 34-38, 2000b: 25-28).

⁵ A significant number (9/96 = 10%, see above) of Santiago African-derived items cannot be clearly attributed either to Manding or Wolof. This is due to the fact that both languages have been in contact for centuries and therefore have extensively borrowed from each other (Rougé 2006: 64-65). Furthermore, items shared by both Manding and Wolof must have made their way more easily into incipient UGPC (from which Santiago Creole descends) since these items were easily understandable by most African ancestors (whether Manding- or Wolof-speaking) of today's UGPC speakers.

⁶ Five languages that each contributed one lexical item are Balanta, Biafada, Fula, Manjaku/Mankanya, Nyun.

⁷ The items that are discussed under the label 'unidentified non-European items' could not be traced back to any known African language or variety. However, they are included here because (i) they are clearly not derived from Portuguese (or another European language), which almost certainly implies that they have an African origin, and (ii) they are shared by at least three different varieties of UGPC, which allows for reconstruction of proto-forms.

This sample was checked in turn in each of the remaining six UGPCs (Fogo, Bissau, Cacheu, Casamance, Geba, and Curaçao Creoles) included in the comparison. The data provenance is indicated in Table 2.

Table 2: Provenance of the data used for the present comparative study⁸

UGPC variety	data collection	researcher
Santiago	available publications (mainly Quint and Rougé – see above) + additional research	Quint
Fogo	fieldwork – 2009, 2017, and 2018	Moreira + Quint
Cacheu	fieldwork – 2012/11/13-16	Quint
Geba	fieldwork – 2010/02/02-05	Quint + Biagui
Bissau	perusal of various references (Biasutti 1987; Rougé 1988, 2004; Scantamburlo 1981, 2002) + additional research	Quint
Casamance	fieldwork – 2010/02/08-16 + Biagui’s unpublished Casamance lexical database + additional research	Quint + Biagui
Curaçao	fieldwork – 1997 + various references (Jacobs 2014; Maurer 1988, 1994, 1998; Quint 2000b)	Quint

In most cases ($83(=69+14)/96=86\%$), the standard list was checked in 5 or 6 UGPC varieties and in all cases it was checked in at least 3 varieties (Table 3).

⁸In 2010/11/20-21, Quint and Biagui conducted a linguistic survey in Joal (see Map 2), an ancient creole-speaking place in Senegal Petite Côte (ca. 100 km South of Dakar): several UGPC-speakers were found but all of them had learnt an UGPC variety through contact with either Capeverdean or Bissau speakers (mostly in Dakar). This survey confirms that no native speaker of Joal UGPC is left today (see also Boulègue (1972) and references therein), although several local Christians told us that they had heard some old people still using a specific form of UGPC in Joal in their childhood. In 2012/11/17-18, Quint also collected data in Farim (Guinea-Bissau) among the *kristonj* (literally ‘Christian’) community, supposed to represent the most ancient layer of local Creole speakers. As a matter of fact, in several places in Guinea-Bissau, the label *kristonj* refers to people who, after the arrival of the Portuguese in the Renaissance, chose to identify themselves principally as (catholic) Christians and to adopt the incipient proto-UGPC as an in-group language, thereby giving up their traditional ethnic labels such as Balanta, Fula, Pepel, etc. However, from talks with local clerics, it seems that today’s Farim *kristonj* community is relatively recently settled in the city (since 1947) and that there isn’t probably any continuity between this community and the old Portuguese factory of Farim dating back to the 17th century (Carreira 1984: 20). Therefore we have decided not to include Farim data in this comparative study, except for one feature (see Table 18 below).

Table 3: Number of UGPC varieties (including Santiago and excluding Papiamentu⁹) for which the standard list of African-derived items was checked

Number of UGPC varieties	6	5	4	3	<3	TOTAL
Number of items checked	69	14	10	3	0	96

The results of the comparison are given in Appendix 1 and a Proto-UGPC form is proposed for each item attested in at least one variety other than Santiago.¹⁰

3.2. Criteria for comparison

Three main criteria were retained in order to refine the analysis of the results of this comparative study (see tables in Appendix 1 and 3), namely (i) presence/absence (PRA), (ii) phonetic similarity (PHON), and (iii) semantic similarity (SEM).

- Presence/Absence: when a form was not found in one or more varieties, the name of these varieties is preceded by 0, for example:

(1) 0FG means that no reflex of a given African-derived item was found in Fogo Capeverdean;

(2) 0FG+BI means that no reflex of a given African-derived item was found either in Fogo or Bissau varieties.

- Phonetic similarity: the phonetic segments which allow to distinguish one or more UGPC varieties with respect to other varieties are given in bold, for example:

(3) for Proto-UGPC */**joŋ**'go/ 'doze' (Table 1.1), the vowels of the Bissau and Geba forms, respectively [**juŋ**gu] and [**juŋ**'gu], are in bold, since they are distinct from all other UGPCs and allow to establish a contrast between two clusters, namely BI+GE vs. all [other varieties].

- Semantic similarity: whenever relevant, the differences in meaning are indicated in the column 'Meaning' of Tables 1.1 to 1.8.

⁹ Since the number of Papiamentu African-derived items matching the other UGPCs is very low (see above), this was not taken into consideration in Table 3.

¹⁰ For a first list of Proto-UGPC forms (based on a smaller sample of UGPC varieties), see also Quint (2000b: 197-208, 307-318).

The results of the comparison between the UGPC varieties at stake for each of the three criteria are given in Appendix 3 (Tables 3.1 to 3.3).

3.3. Significance and scientific interest of this comparison

The African lexical component is a key element to explore the historical relationships between UGPCs for at least two reasons. First, since all UGPCs have evolved and developed in a situation of diglossia with Portuguese as the dominant language for the first generations of speakers, followed in some cases by French (for Casamance Creole since the 19th century) and Spanish and Dutch (for Papiamentu since the 17th century), it is often not easy to know whether a given Romance-derived lexical item found in one (or several) of these languages was present since the settlement/formation period or whether it only entered at some point in time between the emergence of UGPCs and the present.¹¹ Regarding the African lexical component of UGPCs, in particular for Insular and ABC varieties, where there has not been any regular contact with African languages since the emergence of these varieties, we can be almost sure that a given African-derived item shared by several of these varieties can be traced back to the period of their emergence. Second, the above-mentioned diglossia between UGPCs and Romance languages such as Portuguese and Spanish entails that the form of many Romance-derived items must have changed over time. These changes are due to processes of hypercorrection, adjustment or convergence, which tend to match the pronunciation and usage of the UGPC term with the pronunciation and usage attested for the cognate of this term in the dominant Romance language.¹² Regarding African-derived items, especially for Insular and ABC varieties, such processes of hypercorrection cannot have taken place, given that African-derived items do not have any cognate in the dominant European languages. Since the African component of UGPCs is less subject to the influence of dominant languages, this component arguably suffered fewer contact-induced changes since the emergence of UGPCs.¹³ Therefore, African-derived items can be considered a particularly conservative lexical layer of the UGPCs and their study give us privileged access to the

¹¹ For an account of the diachrony of the Romance component of UGPCs, see Quint (2000a, 2001, 2009).

¹² For the tendency to calque Romance-derived items on the contemporary forms of the main lexifiers, see Quint (2009) for Capeverdean (calquing on modern Portuguese) and Quint (2000b: 190-192) for Papiamentu (calquing on modern Spanish).

¹³ In 4.6.1, however, we present a case of contact-induced change in Casamance Creole under the influence of Mandinka.

historical period of emergence and development of the varieties and branches of the UGPCs.

4. Analysis of the results

The results obtained for each criterion retained in this comparative study (see 3.2 above) are summarized in Table 4, which is discussed in detail in the following subsections.

Table 4: Values observed for the three main criteria

Criteria		Phonetic (PHO) and Semantic (SEM) similarities			PRA+PHON+SEM together	
Presence/Absence (PRA)		PHON/SEM	PHON	SEM		
Main profiles (MP)	% MP	Main profiles	% MP	% MP	Total%	Main profiles
All	41	0 [=NSD]	31	67	139	UGPC African Core
IN vs. CO	6	IN/ST vs. CO	26	20	52	IN vs. CO
0FG (+GE/BI/CS) vs. all	25	FG vs. all	3	3	31	FG vs. all
0CS vs. all	2	CS (+IN/ST/CU) vs. all	15	IRR	20	CS vs. all
CS+ST vs. 0BI(+FG)	3					
ST vs. 0all	15	ST vs. 0all	IRR	IRR	15	ST vs. all
0GE vs. all	2	0GE vs. all	IRR	IRR	2	GE vs. all
Others	6	Others	18	6	30	others
Unclear	IRR	Unclear	7	4	11	unclear
TOTAL	100		100	100	300	

4.1. A common African core

We found that 41% (see PRA in Table 4) of all Africanisms identified in Santiago Capeverdean occur in all the other UGPC varieties, except for Papiamentu (see 4.3 for a discussion of this language). Moreover, 31% of the Africanisms found in at least Santiago and two other varieties do not present any significant difference regarding their phonetic shape and two thirds (67%)

of these items have roughly the same meaning in the UGPC where they are attested.

Note also that the proportion of the different African substrates does not show any significant variation between the Santiago Capeverdean (ST) sample and the items shared by all UGPCs (see Table 5). As for Santiago Capeverdean, three languages qualify as the main substrates of UGPCs: Manding, Wolof, and Temne (in order of decreasing importance).

Table 5: Origin of the African-derived items in the Santiago Capeverdean standard sample compared with the African-derived items present in all UGPCs but Papiamentu

Origin	ST sample		Present in all UGPCs	
	Number	%	Number	%
Manding (Mandinka)	42	44	20	51
Wolof	22	23	8	21
Manding and/or Wolof	9	9	2	5
Temne	4	4	2	5
Other Atlantic languages	5	5	2	5
Bantu languages	3	3	0	0
Unidentified non-European items	11	12	5	13
TOTAL	96	100	39	100

These results clearly point towards the existence of an African lexical core common to all UGPCs. Given that (i) the Insular UGPCs have not had significant contact with any African language for a long time¹⁴ (probably several centuries) and that (ii) the various Continental UGPCs are nowadays in contact with different African languages (see Table 6), this common African lexical element can only be explained through the existence of an ancient Proto-UGPC stage (Quint in press, Jacobs 2012), i.e., a first Portuguese-based contact language (either a pidgin or a nativized language) which would have included a mainly Manding/Wolof/Temne African lexical component. The predominance of these three languages can be satisfactorily accounted for by their vehicular use at the time of the arrival of the Portuguese in West-Africa (Lang 2009; Quint 2008; Rougé 1994: 94, 1999a: 61-62).

¹⁴ This also applies to Papiamentu (i.e. the ABC branch of UGPCs).

Table 6: Main African languages currently in contact with continental UGPCs

UGPC variety	African contact language(s)
Geba	Manding
Cacheu	Manjaku
Casamance	Manding + Nyun + Jola (Diola) [+ Wolof (recent)]
Bissau	Pepel + Balanta [+ various (recent)]

The fact that Manding, Wolof, and Temne seem to represent the most ancient African lexical contribution to the UGPCs motivates their designation as the main African *substrate languages* of this group.

By contrast, languages such as Nyun (for Casamance, see Biagui, Nunez & Quint in press; Quint in press), Pepel or Balanta (for Bissau, see Holm & Intumbo 2009) should be viewed as *adstrate languages*, i.e. languages which have recently (or at least after the formative period of Proto-UGPC) contributed African lexical items generally limited to one (or a subgroup of) UGPC(s).

In some cases the same language can be both a substrate and an adstrate for a given UGPC: this is the case of Manding with respect to Geba UGPC.

4.2. The Insular vs. Continental split

As shown in Table 4, the most frequently recurring contrast throughout the data is found between Insular and Continental UGPCs. The Insular vs. Continental split is particularly clear in the domains of phonetics and semantics, in which respectively 26% and 20% of the 82 African-derived items shared by ST and at least another UGPC variety display significant differences between the Insular and Continental UGPCs (see Tables 3.2 and 3.3 in Appendix 3).

Table 7: Some Insular and Continental UGPC African-derived items compared with their respective etyma

Insular UGPCs		Continental UGPCs		African substrate		
Form	Meaning	Form	Meaning	Language	Form	Meaning
FG+ST [bõbu'rõ]	'saddle sp.'	BI+GE+CS [bamba'raŋ]	'cloth used to carry s.o. on back'	Manding (Mandinka)	<i>bàmbùràŋ</i>	= CO
ST [dʒɛ'gesi] + FG [dʒɛgɛ'si]	'mix (alcohol, food)'	BI [jagasi] ~ [jakasi] + GE+CU +CS [jaka'si]	'mix (general)'	Wolof	<i>jaxase</i> [jaxase]	= CO
FG+ST [tɛ'bākɛ]	'carnival sp.'	BI+GE+CU+ CS [ta'baŋka]	'village'	Temne	<i>ka-banka</i>	'fortifica- tion, palisade'

The African-derived items attested in the Continental varieties generally remain semantically closer to the African etymon than their Insular cognates, as shown in Table 7 (see also Rougé 1999a, 2004), which is quite understandable if we consider that Continental UGPCs have always remained in close contact with Atlantic and Mande languages (see Table 6 above), with semantic structures quite similar to the three main UGPC substrates (Wolof, Mandinka, and Temne). In all likelihood, this contact has reduced the possibility of semantic divergence between UGPC African-derived items and these substrates. On the other hand, Insular UGPCs, which were almost exclusively in contact with Portuguese, were able to develop more idiosyncratic (or non-African) semantic patterns.

Regarding the phonetic shape of the African-derived items, the situation is less clear-cut. For instance, if we consider the UGPC reflexes of Manding *bàmbùràŋ* in the table above, FG+ST have maintained the medial /u/ of Manding, which was replaced by /a/ in Continental UGPCs, while Continental varieties have kept the initial and final /a/ of Manding, which were replaced by a nasalized /o/ in FG and ST (Insular varieties).

In all likelihood, the sharp divide between Insular and Continental varieties is due to two main reasons. First, although Cape Verde and Guinea-Bissau (including Casamance until the end of the 19th century – see 2 above) have been part of the Portuguese Empire (i.e. of the same macro-political unit) for several centuries, the economical exchanges and human flows between the Archipelago and the Continent must have remained limited after Cape Verde's settlement period. Second, contrary to what happened in Cape Verde, there has been ongoing linguistic contact between Continental UGPCs and local African languages. Therefore, the limited relationships between the Insular and the Continental region, on the one hand, and the continuous contact of Continental UGPCs with local African languages, on the other hand, favoured the progressive differentiation of the two subgroups, i.e. Insular and Continental varieties.

4.3. Papiamentu (ABC) and UGPCs

In spite of systematic research on further African-derived items, we were able to identify only four African items (see Quint 2000b: 159f., 311). These items are shown in Table 1.8 in Appendix 1.

Unsurprisingly, all four African-derived elements shared by Papiamentu and other UGPCs are Manding-derived (with a possible influence of Wolof in one case, namely Papiamentu [kalem'be] < Proto-UGPC */kal'ma(ŋ)/ ~

*/**kalam'ba(ŋ)**/). As Manding is the main contributor of lexical items to UGPCs (see Tables 1 and 5), Manding-derived items were more likely to survive in a more reduced sample of African-derived items inherited from Proto-UGPC, such as the one attested in contemporary ABC varieties. Two of these items deserve further comments:

- [ka'ranga] 'louse' (only attested in old Papiamentu (see Quint 2000b: 160), nowadays replaced by *piew* < Spanish *piojo* in mainstream Papiamentu) is shared with all Continental UGPCs, but not with the Insular branch, which means that a similar Manding-derived item may well have existed in Insular varieties (i.e. Capeverdean) but was lost under the influence of Portuguese in a similar way to what happened more recently in Papiamentu.
- Mandinka *jòŋkótó* 'squat', the etymon of Proto-UGPC */jɔŋgo'to/ (including Papiamentu - [jɔŋgo'ta]), also has reflexes in several Afro-European Creoles spoken in the New World (Quint 2000b: 159), in particular in the varieties spoken in Surinam (English-based Creoles) and French Guyana (French-based Creole). Further research is needed to ascertain whether this Manding-derived item spread in the Caribbean area due to the presence of an important contingent of Manding-speaking slaves or through the use of an Afro-Portuguese variety influenced by Proto-UGPC. Note that at least another Manding item in the UGPC African lexical core, Mandinka *jìŋkóo* 'doze' > Proto-UGPC */jɔŋ'go/, also has a very similar reflex in an American Afro-European Creole: Saramaccan *djònkõ* 'to nod in sleep' (SIL: 2003). Since a significant part of the Saramaccan lexicon is Portuguese-derived and since this Creole language has already been shown to present some commonalities with the UGPCs (Jacobs & Quint 2016), the similarity between *djònkõ* and the reconstructed Proto-UGPC form */jɔŋ'go/ could be due to contact with an ancient UGPC variety rather than to chance.

Be that as it may, the scrutiny of the African-derived lexicon in Papiamentu highlights the outlier status of Papiamentu among UGPCs, not only because of the paucity of the lexical items shared with the rest of the group, but also because

of the existence of a Bantu/Kwa lexical component that is unattested in the other UGPC varieties (Quint 2000b: 181-182).¹⁵

4.4. Fogo

Fogo Creole can easily be set apart from the other (Insular and Continental) UGPC varieties. Indeed, regarding the first criterion considered in our comparison, i.e. the presence/absence of African-derived items attested in the standard Santiago sample, Fogo is the only variety to lack as much as 25% of the items attested in the other UGPCs (see Table 4 and Table 3.1 in App. 3).

Table 8: Origin of African-derived items unattested in Fogo compared with those found in the ST sample and in other UGPCs varieties

Origin	ST sample		Present in all UGPCs		Present in all UGPCs but Fogo		0FG vs. all	0FG+GE/BI/CS vs. all ¹⁶
	Nr.	%	Nr.	%	Nr.	%	Nr.	Number
Manding/Mandinka	42	44	20	51	27	46	7	8
Wolof	22	23	8	21	12	21	4	5
Manding and/or Wolof	9	9	2	5	4	7	2	5
Temne	4	4	2	5	3	5	1	1
Other Atlantic languages	5	5	2	5	3	5	1	1
Bantu languages	3	3	0	0	1	2	1	1
Unidentified non-European items	11	12	5	13	8	14	3	3
TOTAL	96	100	39	100	58	100	18	24

As can be seen in the last three columns of Table 8, the origin of UGPC or ST African-derived items absent in Fogo does not seem to be linked with any

¹⁵ This Papiamentu idiosyncrasy is most likely due to historical circumstances, namely the early separation of Papiamentu from the rest of the UGPC family and subsequent differential language contact. In fact, the ancient UGPC variety that gave rise to modern Papiamentu was presumably transplanted to the ABC islands in the second half of the 17th century. However, since the beginning of the 18th century, the slaves imported to Curaçao were predominantly Bantu- and Kwa-speakers, which accounts for the presence of Bantu-/Kwa-derived items in today's Papiamentu (see Jacobs (2012: Ch. 7) for further details).

¹⁶ Here (and in Table 4), the cases in which a Santiago African-derived item is unattested in Fogo plus another variety (0FG+GE/BI/CS vs. all) have been merged with the cases when a Santiago African-derived item is unattested in Fogo only. Indeed, whenever the item is missing in only one continental variety, this lack may be due to an incomplete data collection rather than to a genuine dialectal peculiarity.

particular substrate. Fogo is simply the least African of all UGPC considered in this comparison. However, Fogo African-derived lexical component also displays some peculiarities, such as the existence of at least one Wolof-derived item attested only in Fogo, while all other UGPC varieties resort to Romance-derived items for the same meaning (see Table 9).

Table 9: A Wolof-derived item exclusive to Fogo and its counterparts in other UGPCs

Meaning	UGPC variety	Etymon
‘winnow (harvested or husked – not pounded – grain)’	FG [bi’ʃi] ~ [’biʃi]	Wolof <i>bees</i> /be:s/ (JLD 2003: 65, AF 1990: 43)
	ST [’bētjɛ], ABC [’bentʃa]	Pt. <i>ventear</i> ¹⁷
	BI [balia], CU/CS [bali’ja]	Pt. <i>balaio</i> X -ar
	GE: X (unattested)	X

Furthermore, FG also presents some distinctive phonetic treatments of pan-UGPC African-derived items, as suggested by the two elements presented in Table 10.

Table 10: Two African-derived items whose phonetic treatment in Fogo contrasts with all other UGPC varieties

Origin	Meaning	Proto-UGPC	Fogo	Fogo vowels	Other UGPCs vowels
Wolof	‘pinch’ ~ ‘peck/pick up (food) with fingers’	*/cupu’ti/	[tʃɛpu’ti]	/ɛ-u-i/	/u-u-i/
unknown	‘scrape (with spoon)’ ~ ‘slander’	*/korko’ti/	[kɛrku’ti]	/ɛ-u-i/	/o-o-i/

In other cases, Fogo Capeverdean is similar to the other UGPCs, except for Santiago Capeverdean. This tendency is illustrated by a single African-derived word and several items of Portuguese provenance (see Table 11). The elements presented there are highly significant concerning the history of UGPCs. They show that phonetically (see the vowels of the reflexes of Balanta *rikiti*) and lexically, the Insular vs. Continental split discussed above in 4.2 does not always hold. The fact that some Fogo traits are shared with Continental varieties but not with Santiago suggests that creolization (or the development of a local variety of UGPC) in Fogo took place at least partly independently from

¹⁷ Regarding Papiamentu, Latino-American Spanish *ventear* (contrasting with standard peninsular *aventar*) is also a plausible source.

Santiago¹⁸ and that regarding Insular UGPCs no comparison with Continental varieties can be based on Santiago only, given that some non-Santiago phonetic and lexical features are also found in Fogo and possibly in other Insular varieties.

Table 11: Several African- and Portuguese-derived items whose forms in Fogo match with Continental UGPCs only

Meaning	UGPC forms associated with the meaning			Etymon	
	FG	Continental UGPCs	ST	ST	FG+CO
‘take off a bit of’ ~ ‘pinch’	[ɓiki’ti]	BI/GE/CU [riki’ti]	[ru’kuti]	Balanta <i>rikít</i>	
‘small [particles of pounded grain]’	[dɛr’gadu]	CS [dal’gadu]	[’mjodu]	Pt. <i>miúdo</i>	Pt. <i>delgado</i>
‘machete’	[(ɐ)trɛ’zadu] ¹⁹	BI [ter’sadu] ~ [tar’sadu], GE/CU [tar’sadu], CS [tor’sadi] ~ [tor’sadu]	[mɛ’fɪ]	Pt. <i>mach(ad)i-nho</i> or <i>manchil</i> (?)	Pt. <i>terçado</i>
‘foot sole’	[’pe di ’bafu]	CS [bas di ’pe]	[’sɔləl ’pe]	Pt. <i>sola</i>	Pt. <i>baixo</i>

4.5. Santiago Creole

Santiago Creole is not as clearly an outsider as Fogo in the realm of UGPCs. In the present comparison, 14 Santiago items, i.e. 15% of the items in the sample, lack cognates in any other UGPC (see Tables 4 above and 3.1 in Appendix 3). Since Santiago is the standard for the comparison (see 3.1), it is expected that some Santiago African-derived items lack a counterpart in other UGPCs. A similar issue would arise if another UGPC variety were the standard (see FG [bi’fɪ] ~ [’bɪfɪ] discussed in Table 9).

¹⁸ In some cases, the similarities between Fogo and Continental varieties may be due to the fact that Fogo could have retained a Proto-UGPC item, which was lost or replaced in Santiago. However, this explanation probably does not hold for all the cases presented above.

¹⁹ The most commonly used item in Fogo is [’spadɐ], which is also sporadically attested in Santiago. However, [(ɐ)trɛ’zadu] ‘machete’ was consistently found in the four localities of *Kemada Gintxu*, *Fajanzinha*, *Sunbangu* and *Otrubanda*. In the latter, [ɐtrɛ’zadu] was considered an old form by our informant during a survey carried out by Quint (2009/08/15).

Table 12 shows that the origin of African-derived items specific to Santiago does not seem to display any significant bias when compared with the total sample of African-derived items in Santiago.

Table 12: Origin of the African-derived items in the Santiago Capeverdean standard sample and among the items specific to Santiago Capeverdean

Origin	ST sample		Items specific to ST
	Nr.	%	Nr.
Manding (Mandinka)	42	44	7
Wolof	22	23	5
Manding and/or Wolof	9	9	0
Temne	4	4	0
Other Atlantic languages	5	5	1
Bantu languages	3	3	1
Unidentified non-European items	11	12	0
TOTAL	96	100	14

Table 11 above shows that, for some lexical and phonetic features, Santiago stands apart both from Fogo and from the Continental UGPCs. Furthermore, the phonetic treatment of the African substrate non-final /i/ > ST /u/ is attested in at least one other African-derived item in Santiago, which supports the idea that this feature displayed a certain regularity in the history of Santiago Capeverdean (see Table 13).²⁰

Table 13: Two African-derived items displaying a parallel phonetic treatment exclusive to Santiago Capeverdean

Meaning	Etymon	Proto-UGPC	ST
‘take off a bit of’ ~ ‘pinch’	Balanta <i>rikít</i>	*/riki’ti/	[ru’kuti]
‘thrash about’	Md. <i>fitífití</i>	Unattested	[’futi’futi]

Rougé (1999a: 56) asserts that “80% of Santiago African-derived items are also found in Guinea-[Bissau], while 75% of the African-derived items present in Continental [UGPC] are specific [to these varieties, i.e. not found in

²⁰ See however some counterexamples within the Santiago African-derived lexical component itself, e.g. [bībī’rī] ‘millet’ < Md. (Maninka) *binbiri*, and [di’gigi] ‘shake, be shaky’ < Md. (Mandinka) *jūjū* (Mande varieties other than Mandinka are more similar to Santiago, e.g. Maninka *dyigidyigi*, Bambara *yigiyigi* (Bailleul 1996: 431, 1998: 284)).

Santiago].”²¹ Our data confirm the first part of Rougé’s statement: out of a total of 96 Santiago African-derived items, as said above, 14 are not found in any UGPC variety and 6 are found in both Insular varieties (including Santiago), but not in Continental varieties (see Main Profile IN vs. CO in Table 3.1 of Appendix 3, or the cell IN vs. CO in the column PRA of Table 4). The percentage of Santiago Africanisms also found in Continental UGPCs is therefore $(96-14-6)/96=79\%$, which nicely matches Rougé’s result.

We are presently unable to corroborate the second part of Rougé’s claim. As a matter of fact, the African-component of an UGPC such as Casamance Creole seems to be far larger than what is found in Insular varieties. However, among the African-derived items of Casamance Creole, many have been borrowed from adstrate languages (see Table 6 and discussion in 4.1 and also Biagui, Nunez & Quint (in press)) and are not found in other Continental UGPCs. Hence, in order to be able to provide a statistically-based comparison of the African-component of Continental UGPCs as opposed to Insular varieties, we should first build a database of the African-derived items shared by all (or most of) Continental varieties. Under the hypothesis that these shared items go back to Proto-Continental UGPC, we would have an idea of what the African substrate (as opposed to the various local adstrates) of the Continental UGPCs looks like and we would be able to actually assess the second part of Rougé’s statement. However, such an undertaking goes well beyond the scope of the present paper.

4.6. Casamance Creole and the northern Continental sub-branch

As shown in Table 4, after Fogo, Casamance Creole (CS) is the variety that displays the highest number of peculiarities with respect to its African-derived items in relation to some or all of the remaining UGPCs. CS distinctive features can ultimately be related to three main patterns, which will be discussed below.

4.6.1. Casamance Creole contrasts with all other UGPCs (0CS vs. all)

As exemplified in Table 14, CS is sometimes at variance with all other UGPCs either by lacking an African-derived item attested elsewhere (e.g. the reflex of Balanta *rikít*) or by showing a peculiar phonetic feature (e.g. retention of the voiceless plosive [k] of (Mandinka) Manding *móñonko*).

²¹ In Rougé’s (1999a: 56) original text: “Estimamos que 80% das palavras de origem africana do crioulo de Santiago existem também na Guiné, enquanto, no crioulo do continente, 75% dos lexemas cujo étimo é africano lhe são específicos.”

Table 14: African-derived items for which Casamance Creole contrasts with all other UGPCs

CS peculiarity	Meaning	Proto-UGPC	Cognate in CS	Etymon
different phonetic feature	‘crumple (v.)~ ‘shell, (maize)’ ~ ‘grind, crumble (v.)’	*/moŋoŋ’go/ ~ */moŋoŋ’gi/	[muŋoŋ’ka]	Md. <i>móñonko</i> (Creissels & al. 1982: 121)
missing item	‘take off a bit of’ (IN) vs. ‘pinch’ (CO)	*/riki’ti/	Unattested	Balanta <i>rikit</i>

Regarding the different phonetic forms of African-derived items in CS, the three cases we have found in the sample are all Manding-derived and in all cases the CS form appears to be closer to Mandinka, which is one of the main adstrates of CS (see Table 6). Here the divergent tendency observed in CS vis-à-vis other UGPCs is therefore probably related to the influence of contemporary Mandinka, which acted as an adstrate and triggered a modification of the Manding-derived Proto-UGPC form in CS in order to more closely resemble the Mandinka variety CS-speakers are daily confronted with.

As for the cases of missing items, they may well be accounted for through the relative isolation of today’s CS speech community. The political boundary established between Guinea-Bissau and Senegal more than 130 years ago had the effect of limiting linguistic contact with the other Continental UGPC varieties (see section 2). Some Proto-UGPC items in CS were probably lost for that reason.

4.6.2. Casamance Creole clusters with Insular UGPCs only

This pattern is mostly instantiated in the domain of phonetic similarities (see PHO in Table 4).

As shown in Table 15, in many cases, Proto-UGPC vowels /e/ and /o/ exhibit the form /i/ and /u/, respectively, in Continental UGPCs other than CS. This phonetic change seems to have been historically initiated by Bissau UGPC and also regularly applies to Romance-derived items (see Doneux & Rougé 1988: 75-76 and 4.6.3 below). Given that CS is the northernmost variety of all Continental UGPCs and that it is separated from the remaining Continental varieties by a political border (see 4.6.1 above), CS has remained largely unaffected by the vowel changes observed in southern Continental UGPCs. What makes CS cluster with Insular UGPCs and not with other Continental varieties is its conservative nature, which led it to maintain the Proto-UGPC vowel system. At any rate, this clustering pattern clearly illustrates the need to

take into account as many varieties as possible for a comparison between Insular and Continental Creoles.²²

Table 15: Vowels of three African-derived items for which Casamance Creole shows the same phonetic treatment as Insular UGPCs

Meaning	Insular+CS			Continental UGPCs other than CS	Vowels	
	Proto-UGPC	Insular	CS		IN+CS	CO other than CS
‘incline’	*/je’ɲje/ ~ /ge’ɲge/	ST [‘dʒɛ̃dʒe]	[je’ɲje]	BI [jɪŋgi], GE+CU [‘gɪŋgi]	/e-e/	/i-i/
‘look (v.)’	*/jo’be/	ST [‘dʒobe], FG [dʒo’be]	[jo’be]	BI [jubi], GE [‘jubi], CU [ju’bi]	/o-e/	/u-i/
‘curl up, triturate’ ~ ‘shape food in one’s hand’	*/mon’do/	ST [‘mõdo], FG [mõ’do]	[mo’nd o]	BI [mundu], GE+CU [mu’ndu]	/o-o/	/u-u/

4.6.3. Casamance and Cacheu Creole: the northern branch of Continental UGPCs

In at least four cases (see Table 16 below and Tables 3.1 and 3.2 in Appendix 3) involving either the criteria of presence/absence (PRA) or phonetic similarity (PHO), Casamance Creole (CS) clearly clusters with Cacheu Creole (CU) against all other UGPC varieties.

Table 16: African-derived items for which Casamance Creole clusters with Cacheu

Meaning	UGPC forms associated with the meaning		Etyma	
	Proto-UGPCs	CU+CS	CU+CS	Proto-UGPC
‘squat’	*/jonŋo’to/	[joko’ni]	Wf. <i>jonkon</i>	Md. <i>jòŋkótó</i>
‘silent fart’	*/fus/	[ka’fus]	Md. <i>fùusí</i>	
‘make use of any resource’ ~ ‘glean’	*/ferfe’ri/	[fefe’ri]	Md. <i>fèefé</i>	
‘stick (n.)’	*/man’duku/	[ma’nduk]	Md. <i>dòkó</i> ²³	

²² Most comparative works hitherto published on Insular and Continental Creoles (e.g. Baptista, Mello & Suzuki 2007; Hagemeyer & Alexandre 2012; Marques Barros 1887-1908; Quint 2000b: 99-117) only contrast one variety of each group, generally Santiago (Insular) and Bissau (Continental), which significantly limits the range of their results.

²³ For an account of the /ma/- prefix developed by UGPC forms, which is missing from the Manding etymon, see Quint (2008: 50-53, 2012: 12).

These similarities between Casamance and Cacheu are striking and cannot be due to chance. They show that there exists a particularly strong relationship between the two varieties, which derives from the fact that Ziguinchor, the first Portuguese trading post in Casamance, where CS began to develop, was reportedly founded by Cacheu settlers in 1645 (see above section 2). At least two items deserve further discussion, namely the terms that translate as ‘to squat’ and the prefix [ka] of [ka’fus].

From a comparative and historical point of view, it is highly significant that CS and CU share a Wolof-derived item [joko’ni] for ‘squat’, whereas all other UGPC varieties (including Papiamentu, see 4.3 and Table 1.8 in Appendix 1) share a Manding-derived item, reconstructed as Proto-UGPC */jɔŋgo’to/. In fact, Wolof is not an adstrate of CU and only recently came in direct contact with CS-speakers (mainly in the second half of the 20th century). Therefore, CS+CU [joko’ni] cannot be assigned to a recent influence of Wolof and probably goes back to the formative period of Cacheu Creole at the end of the 16th century (see section 2), from which it must have been transferred to Casamance (see above). Since reflexes of Proto-UGPC */jɔŋgo’to/ are found in all three branches of UGPCs, the Manding root is clearly a part of the most ancient core-vocabulary common to all UGPCs. This being said, two plausible explanations can account for the existence of Wolof-derived [joko’ni] in Cacheu and Casamance UGPCs. First, there was a particularly strong Wolof-speaking community among the first settlers of Cacheu. This ‘Wolof-settlers’ hypothesis (WSH) does not seem to be borne out by the existence of many other Wolof-derived items specific to both CS and CU. However, without an exhaustive study of the African-derived component in the various UGPC varieties (see above 4.5.), the WSH cannot be ruled out. Second, two different roots - a Manding and a Wolof one - could have been used concurrently in the incipient Proto-UGPC. The Wolof root would have been favored in Proto-UGPC varieties used along the northern shore of the captaincy of Cape Verde (Quint 2012), i.e. the mouth of the Gambia River, where Wolof must have been the main vehicular language, whereas the Manding root would have dominated in the southern part of the same area. As Cacheu (and Ziguinchor) are situated to the south of the Gambia river’s mouth but closer to it than any attested Continental UGPC variety, the incipient varieties of UGPCs spoken in and around Cacheu (and later in Ziguinchor) could have been in a region where both the southern Manding root and the northern Wolof root coexisted. If this ‘North vs. South Continental divide hypothesis’ is accurate, Wolof-derived [joko’ni] ‘squat’ found in Casamance and Cacheu bears testimony of the existence of more Wolofized lects of UGPCs spoken north of the Gambia river’s mouth, in

places such as Joal, whose UGPC variety unfortunately disappeared undocumented (see footnote 8 above and references therein).

Prefix [ka], found in CU and CS [ka'fus] 'silent fart', was probably added to this Manding-derived root due to the influence of the various Atlantic noun-class²⁴ languages still spoken in the vicinity of Cacheu (mainly Manjaku²⁵) and Casamance (mostly Nyun and Jola (Diola), see Table 6). Since history shows that Casamance Creole was an offshoot of Cacheu UGPC, the specific CU+CS form must have appeared first in CU and then spread to CS: Manjaku is, therefore, the most plausible source for [ka] and the attachment of an Atlantic prefix to a Manding-derived root shows that, at least in Cacheu, the influence of the Manding substrate on UGPC was not direct. Instead, it was mediated by the speakers of other African languages.

As will be shown below, the Portuguese lexical component also confirms the existence of a special link between Cacheu and Casamance Creole. First, we have been able to find at least one frequent Portuguese-derived item which seems unique to Cacheu and Casamance (see Table 17).

Table 17: A Portuguese-derived item for which Casamance Creole clusters with Cacheu

Meaning	UGPC variety	Etymon
'(earthenware) jar to store water inside a house'	FG/ST ['pɔti], BI ['puti] ~ ['pute] ~ ['poti] ~ ['pote], GE ['puti]	Pt. <i>pote</i>
	CU [kan'trera], CS [kan'tilera]	Pt. <i>cantareira</i>

Second, in several items displaying the afore-mentioned phenomenon of vowel-raising in Bissau UGPC and other southern UGPC varieties (see Table 15 and the following comments in 4.6.2), some Cacheu speakers interviewed in 2010 still exhibit the /e, o/ vowels of Proto-UGPC. Table 18 provides some instances of this conservative behavior.

²⁴ Contrary to Atlantic languages, the different varieties of Manding (such as Mandinka, Maninka or Bambara) have neither class-prefixes nor any other type of noun-class morphology.

²⁵ Note that the singular noun-class prefix [ka] is attested in Manjaku with a singulative/deverbal value (Basso Marques 1947: 80; Buis 1990: 18-19, 21; Carreira & Basso Marques 1947: 29, 34; Doneux 1967: 262-263; Karlik 1972: 259).

Table 18: Examples of Portuguese-derived items for which Casamance Creole clusters with Cacheu and Insular varieties

Mean.	Proto-UGPC	Insular	CS	CU	Farim	BI ²⁶ +GE	Pt. etymon
‘bite’	*/mor’de/	ST [ˈmorde], FG [mor’de]	[mor’de]	[mor’de] ~ [mur’di]	[mur’dĩ]	BI [murdi], GE [ˈmurdi]	<i>morder</i>
‘can (v.)’	*/po’de/	ST [ˈpo de], FG [po’de]	[po’de]	[po’de] ~ [pu’di]	[po’de]	BI [pudi], GE [ˈpudi]	<i>poder</i>
‘fit’	*/ke’be/	ST [ˈkebe] FG [ke’be]	[ke’be]	[ke’be]	[ki’bi]	BI [kibi]	<i>caber</i>
‘bring’	*/ter’sē/	ST [ˈtɛrse], FG [trɛ’ze]	[tɛ’sē]	[tɛ’sē] ~ [tĩ’si]	[tisi] ²⁷	BI [tisi], GE [ˈtisi]	<i>trazer</i>
‘know’	*/sɛ’bɛ/	ST [ˈsɛbɛ], FG [sa’bɛ]	[sɛ’bɛ]	[sɛ’bɛ] ~ [si’bi]	[si’bi]	BI [sibi]	<i>saber</i>
‘hyena’	*/ˈlobu/	ST+FG [ˈlobu]	[ˈlobu]	[ˈlobu]	X	BI [ˈlubu]	<i>lobo</i>
‘twist (v.)’	*/tor’sɛ/	ST [ˈtɔrse], FG [tor’sɛ]	[tɔr’sɛ]	[tur’si(du)] ²⁸	[ˈtursi]	BI [tursi], GE [ˈtursi]	<i>torcer</i>
‘cough (v.)’	*/to’si/	ST [ˈtose] ~ [ˈtosi] FG [to’sɛ]	[tɔr’sɛ]	[ˈtorse]	[ˈtorse]		<i>tossir</i>

As can be seen there, for many items, CU shows variation between /e, o/ and /i, u/, while CS regularly maintains /e, o/ of Proto-UGPC. These variations are probably due to dialect levelling presently at work throughout Guinea-Bissau, where the Bissau variety has become the main vehicular of the whole country and tends to superimpose its characteristics (here the /i, u/ forms) upon the remaining local varieties. CU /e, o/ forms most certainly represent older variants of CU /i, u/ forms.²⁹ Farim UGPC also retains /e, o/ forms, but to a lesser extent than CU.

²⁶ Biasutti (1987) also mentions some /e, o/ forms in Bissau, e.g. *mordi* ‘bite’ (p. 159) or *sebɛ* ‘know’ (p. 205), alongside *murdi* (p. 161) and *sibi* (p. 206) respectively. However, such /e, o/ forms are not confirmed for Bissau by authors such as Rougé (1988) or Scantamburlo (2002) nor by Quint’s personal observations *in situ*: their presence in Biasutti’s dictionary could be motivated by the author’s desire to take into account the different Creole varieties (other than Bissau UGPC) spoken in Guinea-Bissau. In view of the above and for the sake of clarity, we have therefore chosen not to mention Biasutti’s /e, o/ forms.

²⁷ In some transcriptions, the stressed syllable is not indicated, as it was missing in the sources consulted or omitted in the notes taken on the field.

²⁸ For ‘twist’ in Cacheu and ‘bite’ in Farim (next column), Quint was only able to collect inflected forms, [tur’sidu] ‘twisted’ and [mur’dĩ] ‘bite.1SG’.

²⁹ Doneux & Rougé (1988: 75-76) only mention /e, o/ forms for both Cacheu and Casamance Creoles. Although Quint interviewed older speakers who spent most of their lives in Cacheu, none of these speakers made a regular use of /e, o/ forms in 2010. Some used exclusively /i, u/ forms while some others mixed /i, u/ and /e, o/ forms. While it is possible that the present

The contrast between ‘twist’ and ‘cough’ is worthy of discussion. While Insular varieties have remained closer to their Portuguese and Proto-UGPC etyma, BI and GE have merged both verbs with a unique form in /i, u/. However, CS, CU, and Farim all maintain the two verbs distinct resorting to a vocalic contrast (/ɔ-ɛ/ vs. /o-e/ in CS, /u-i/ vs. /o-e/ in CU+Farim). Here again, Cacheu and Casamance varieties cluster together and Farim also displays the same specific phenomenon.

The Romance lexical component confirms what was already apparent from the scrutiny of the African-derived items: Casamance and Cacheu Creole belong to the same genetic unit, i.e. the northern sub-branch of Continental UGPCs. Irrespective of the question of its historical continuity (see fn. 8), Farim seems to be more closely related to this northern sub-branch than any other Continental UGPC variety.

4.7. Geba Creole and the eastern Continental branch³⁰

As mentioned previously, Geba is commonly considered as the most characteristic variety of the eastern sub-branch of Continental UGPCs (see section 2) and many Bissau-Guineans believe that Geba was the actual cradle of all UGPCs presently spoken in the country. However, the material collected *in situ* by Quint and Biagui among old, conservative speakers of Geba Creole, members of the ancient local Christian community, has not revealed any significant difference with respect to Bissau Creole.

Only two items of the standard sample are found in all UGPCs but Geba (see Table 19).

Table 19: The two African-derived items found in all UGPC varieties except Geba

Meaning	Proto-UGPC	Geba
‘corn husk’ ~ ‘bean pod’ ~ ‘rice bran’	*/ ɲaɲa /	[‘soka] < Pt. <i>soca</i>
‘scramble, manage’ ~ ‘jump’	*/ jugu’ta /	[ci’rimpa] < unknown etymon

Geba [‘soka] probably results from a misunderstanding while eliciting data with informants, since the survey was carried out during the first fieldtrip mission of Quint and Biagui outside of their previous areas of expertise (respectively Cape

situation is only due to the recent arrival of Bissau forms in Cacheu, we find it more likely that alternations between /i, u/ and /e, o/ forms must have existed there for at least several decades, since even the most conservative speakers of Cacheu Creole were not consistent in this respect.

³⁰ Bissau UGPC is not specifically discussed in any particular subsection, since it is already systematically contrasted with the other varieties in section 4.

Verde and Casamance). The form was obtained by asking the equivalent of Santiago ‘corn husk’, i.e. ‘a grainless ear of maize’. The cognate of Geba [ˈsoka] in Santiago is [ˈsøkə] ‘maize stalk’, i.e. ‘a dried maize stem deprived of any ear’.

As for [ciˈrimpa], this item (meaning both ‘jump’ and ‘spurt’) is clearly a Geba shibboleth, recognized as such by many Creole speakers outside Geba as we have observed through conversations with various Guinea-Bissau citizens. However, we cannot make Geba UGPC a genetic group of its own on the basis of a single lexical item. Geba Creole does not seem to present any significant or recurring phonetic or semantic peculiarity vis-à-vis other UGPCs and in particular the Bissau variety. The same also generally applies to the Portuguese-derived elements we were able to collect. We only found the items in Table 20 to be at variance with Bissau forms.

Table 20: Portuguese-derived items having a form exclusive to Geba

Meaning	IN	CS	BI	GE	Pt. etymon
‘woman’	ST [muˈdʒer], FG [ˈmudʒe]	[miˈɲjer]	[miˈɲjer]	{munjer} ³¹	<i>mulher</i>
POSS.PRO.3SG ‘his/her/its’	ST [diˈsel] ~ [diˈsew], FG [diˈsel]	[diˈsol]	[diˈsil]	[diˈsiw]	<i>de + seu + (ele)</i>
POSS.ADJ.3SG ‘his/her/its’+noun	ST [si] ~ [se], FG [se]	[si]	[si]	[su]	<i>seu</i>

During the Geba data collection, we thought that the scarcity of distinctive linguistic features in Geba could be explained by the recent expansion of Bissau Creole throughout Guinea-Bissau (see 4.6.3). However, after the fieldtrip, we were also able to hear recordings collected from Geba speakers by a Senegalese linguist in the 1980s. These recordings, which were checked by Biagui, do not seem to display any additional traits that are exclusive to Geba. In view of these elements, it appears that the eastern sub-branch of Guinea-Bissau Creole does not exist as such anymore, if it ever existed. Two hypotheses can be envisaged. Either Geba Creole merged quite early into Bissau Creole or the Geba Creole never used to be very different from Bissau Creole. The first scenario does not seem attractive in view of the social reality in Geba, a small city situated far inland, where a tiny (some few hundred people at most), isolated Christian community coexists with a majority of Muslim Manding-speakers who, at least until the mid-20th century, only learnt Creole from their Christian neighbors,

³¹ This form is only mentioned by Rougé (1995: 86). We have maintained his transcription in Table 20.

according to an elder of the Manding community interviewed during our 2010-survey. In this scenario, it is difficult to explain how this secluded Christian group gave up their Creole variety before the population of the port-city of Cacheu. Therefore, the available evidence seems to favor the second hypothesis. The similarity between Geba and Bissau UGPCs may be due to intense economic links between the two cities during and after the founding period or to the fact they were settled by related communities, as appears to have been the case for Cacheu and Ziguinchor in Casamance (see 4.6.3). It follows that the traditional threefold distinction between an eastern, a central (or southern) and a northern sub-branch of Continental UGPCs should probably be abandoned in favor of a twofold distinction between a southern sub-branch (including Bissau and Geba) and a northern sub-branch (including Cacheu, Casamance and, possibly, Farim).

5. Conclusion

This paper is the first study that compares the African element in a wide array of UGPC varieties. In addition to the identification of additional African etyma, this comparison provides new evidence for the genetic classification of the UGPCs. First, the lexicon confirms that all these languages share a common African core, whose existence can probably be traced back to the African substrates (Manding, Wolof, and Temne), which contributed to the formation of a first Proto-UGPC that came into existence at the turn of the 15th and 16th centuries. Second, the three criteria investigated for the African component, namely (i) presence/absence, (ii) phonetic similarities, and (iii) semantic similarities, reveal a sharp divide between Insular (Capeverdean) and Continental (Guinea-Bissau and Casamance) Creoles. Third, the present study shows that some African-derived items in Papiamentu can be traced back to Proto-UGPC, hereby strengthening the hypothesis of a genetic relationship between ABC and other UGPCs (Martinus 1997; Quint 2000b; Jacobs 2012). However, we were not able to increase the small sample of African roots shared between Papiamentu, Insular, and Continental Creoles. Fourth, among the six non-Caribbean UGPCs we studied, the Fogo lexicon appears to be the most divergent and the least African. At the same time, this variety presents some striking commonalities with Continental UGPCs but not with Santiago, a fact that seems to be at odds with the Insular vs. Continental UGPC division. Fifth, regarding Santiago Creole, our results confirm Rougé's (1999a) assertion that 80% of the African-derived items found in Santiago are also found in

Continental UGPCs. Sixth, Casamance Creole is an outlier among Continental UGPCs. It exhibits some exclusive features and shares others with Insular varieties. However, Casamance Creole can also be shown to have a particularly close relationship with Cacheu Creole, which underlies the inclusion of these varieties, and possibly also Farim UGPC, into the same genetic subgroup, namely the northern sub-branch of Continental UGPCs. Seventh, the Geba variety shows almost no difference with Bissau, which prompts us to propose the inclusion of these UGPCs into the same genetic subgroup, i.e. the southern sub-branch of Continental UGPCs and to discard the existence of a specific eastern branch for the Geba variety.

This paper also emphasizes the importance of including a wide array of language varieties in comparative studies in order to improve linguistic and historical reconstructions. It further highlights the next challenge for a good understanding of the African-derived lexicon of UGPCs, i.e. the need to collect, classify and compare as many Africanisms as possible in each of the various local varieties which have not been fully studied to-date (in particular Cacheu, Geba, and Casamance UGPCs). Only by doing so can we grasp the full significance of the influence African languages have exerted upon Upper Guinea Portuguese Creoles for over half a millennium.

The scope of the present study would certainly benefit from pursuing two other lines of investigation. One line of research would be to raise the number of criteria considered, for example by taking into account the respective word-category (or part of speech) of the Africanisms in both the substrate languages and the UGPCs at stake. A second line of research would be to correlate the findings of this paper, which mostly focuses on the lexical component of UGPCs, with the available work on African-related grammatical features in the UGPCs.

Abbreviations

ABC=Papiamentu (Aruba-Bonaire-Curaçao); **ADJ**=Adjective; **AF**=Arame Fal & al.; **BI**=Bissau; **CO**=Continental Creoles; **CS**=Casamance; **CU**=Cacheu; **DC**=Denis Creissels; **EX**=Exogenous (known by speaker but considered to be item of a neighboring non-Creole language); **FG**=Fogo Capeverde; **GE**=Geba; **IN**=Insular Creoles (Capeverdean); **IRR**=Irrelevant; **JLD**=Jean-Léopold Diouf; **JLR**=Jean-Louis Rougé; **Md.**=Manding; **Mdka**=Mandinka; **MP**=Main profile; **NC**=Not checked; **NSD**=No significant difference; **PC**=Portuguese Creole; **PHON**=Phonetic similarity; **POSS**=Possessive; **PRA**=Presence/absence; **PRO**=Pronoun; **Pt.**=Portuguese; **SEM**=Semantic similarity; **SG**=Singular;

ST=Santiago Capeverdean; **UGPC**=Upper Guinea Portuguese Creole; **v.**=Verb; **Wf.**=Wolof; **WSH**=Wolof settlers' hypothesis

References

- Andrade, Elisa Silva. 1996. *Les Îles du Cap Vert de la «Découverte» à l'Indépendance Nationale (1460-1975)*. Paris: L'Harmattan.
- Arends Jacques, Sylvia Kouwenberg & Norval Smith. 1995. Theories focusing on the non-European input. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and Creoles: An Introduction*. Amsterdam: Benjamins.
- Bailleul, Charles. 1996. *Dictionnaire Bambara-Français*. Bamako: Donniya.
- Bailleul, Charles. 1998. *Dictionnaire Français-Bambara*. Bamako: Donniya.
- Baptista, Marlyse, Heliana Mello & Miki Suzuki. 2007. Kabuverdianu, or Cape Verdean, and Kriyol, or Guinea-Bissau (Creole Portuguese). In John Holm & Peter L. Patrick (eds.), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars*, 53-82. London: Battlebridge.
- Basso Marques, João. 1947. Aspectos do problema da semelhança da língua dos papéis, manjacos e brames. *Boletim Cultural da Guiné Portuguesa* 5(2). 77-109.
- Biagui, Noël B. 2018. *Description Générale du Créole Afro-Portugais de Ziguinchor (Sénégal)*. Paris: Karthala.
- Biagui, Noël B. in progress. *Dictionnaire Créole Casamançais-Français*.
- Biagui, Noël B., Joseph J.-F. Nunez & Nicolas Quint. in press. *Casamance Creole*. In Friederike Lüpke (ed.), *Oxford Guide to the World's Languages: Atlantic*. Oxford: Oxford University Press.
- Biagui Noël B. & Nicolas Quint. 2013. Casamance Creole. In Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Survey of Pidgin and Creole languages*, vol. II (*French-based, Portuguese-based and Spanish-based Languages*), 40-49. Oxford: Oxford University Press.
- Biasutti, Artur. 1987. *Vocabulário Kriol-Português*. Bubaque: Missão Católica de Bubaque (Guiné-Bissau).
- Boulègue, Jean. 1972. *Les Luso-Africains de Sénégambie* (Travaux et documents, vol. 1). Dakar: Université de Dakar.
- Boulègue, Jean. 2006. Présence portugaise et sociétés africaines sur la côte de la "Guinée du Cap-Vert". In Jürgen Lang, John Holm, Jean-Louis Rougé & Maria João Soares (eds.), *Cabo Verde: Origens da sua Sociedade e do seu Crioulo*, 45-52. Tübingen: Günter Narr.
- Buis, Pierre. 1990. *Essai sur la Langue Manjako de la Zone de Basserel*. Bissau: INEP (Instituto Nacional de Estudos e Pesquisa).
- Buurt, Gerard van. 2001. Afrikaanse woorden in het Papiamentu. *Kristòf* 11(3). 1-24.
- Carreira, António. 1984. *O Crioulo de Cabo Verde - Surto e Expansão*. Lisbon: author's edition.
- Carreira, António & João Basso Marques. 1947. *Subsídios para o Estudo da Língua Manjaka*. Lisbon: Centro de Estudos da Guiné Portuguesa 3.
- Couto, Hildo Honório do & Filomena Embaló. 2010. Literatura, língua e cultura na Guiné-Bissau. *Papia* 20.

- Creissels Denis, Sidia Jatta & Jobarteh Kalifa. 1982. Lexique mandinka-français. *Mandenkan* 3.
- Creissels Denis. 2012. *Lexique mandinka-français* (unpublished), http://www.deniscreissels.fr/public/Creissels-lexique_mandinka_2012.pdf. (20 August, 2018.)
- Dieng, Sady. 1985. *Étude contrastive des lexies du créole endogène "portugais" et du wolof*. Lyon: Université Lumière Lyon 2, PhD dissertation.
- Diouf, Jean-Léopold. 2003. *Dictionnaire Wolof-Français et Français-Wolof*. Paris: Karthala.
- Doneux, Jean L. 1967. Le manjaku: Classes nominales et questions sur l'alternance consonantique. In Gabriel Manessy (ed.), *La Classification Nominale dans les Langues Négro-Africaines*, 261-276. Paris: CNRS.
- Doneux, Jean L. & Jean-Louis Rougé. 1988. *En Apprenant le Créole à Bissau ou Ziguinchor*. Paris: L'Harmattan.
- Fal Arame, Rosine Santos & Jean L. Doneux. 1990. *Dictionnaire Wolof-Fançais, Suivi d'un Index Français-Wolof*. Paris: Karthala.
- Hagemeyer, Tjerk & Nélia Alexandre. 2012. Os crioulos da Alta Guiné e do Golfo da Guiné: uma comparação sintáctica. *Papia* 22(2). 233-251.
- Holm, John. 1988. *Pidgins and Creoles*. Cambridge: Cambridge University Press.
- Holm, John & Incanha Intumbo. 2009. Quantifying superstrate and substrate influence. *Journal of Pidgin and Creole Languages* (JPCL) 24(2). 218-274.
- Intumbo, Incanha, Liliana Inverno & John Holm. 2013. Guinea-Bissau Kriyol. In Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole Language Structures*, 31-38. Oxford: Oxford University Press.
- Jacobs, Bart. 2010. Upper Guinea Creole: Evidence in favor of a Santiago birth. *Journal of Pidgin and Creole Languages* 25(2). 289-343.
- Jacobs, Bart. 2012. *Origins of a Creole. The History of Papiamentu and its African Ties*. Berlin: De Gruyter.
- Jacobs, Bart. 2014. Lexico-semantic traces of kinship between Papiamentu and Upper Guinea Portuguese Creoles. *Papia* 24(1). 133-147.
- Jacobs, Bart & Nicolas Quint. 2016. Relevance of classical Portuguese features in four Atlantic creoles. In Armin Schwegler, Liane Stroebel & John McWhorter (eds.), *The Iberian Challenge: Creole Languages Beyond the Plantation Setting*, 67-83. Madrid/Frankfurt am Main: Klaus Vervuert/Iberoamericana.
- Karlik, Jan. 1972. *A Manjako grammar with special reference to the nominal group*. London: University of London (School of Oriental and African Studies), PhD dissertation.
- Lang, Jürgen. 2009. *Les Langues des Autres dans la Créolisation. Théorie et Exemplification par le Créole d'Empreinte Wolof à l'île de Santiago du Cap Vert*. Tübingen: Günter Narr.
- Marques Barros, Marcelino. 1887-1908. O Guineense. *Revista Lusitana* 5, 6, 7 & 10.
- Martinus, Efraïm F. 1997. *The Kiss of a Slave, Papiamentu's West-African Connections*. Curaçao: De Curaçaosche Courant.
- Maurer, Philippe. 1988. *Les Modifications Temporelles et Modales dans le Papiamentu de Curaçao (Antilles Néerlandaises)*. Hamburg: Helmut Buske Verlag.
- Maurer, Philippe. 1994. L'influence des langues africaines sur la signification des lexèmes du papiamentu de Curaçao. In Ramón Lorenzo (ed.), *Actas do XIX Congreso Internacional de Lingüística e Filoloxía Románicas*, 771-779. Santiago de Compostela: Universidad de Santiago de Compostela.

- Maurer, Philippe. 1998. El papiamentu de Curazao. In Matthias Perl & Armin Schwegler (eds.), *América Negra: Panorámica Actual de los Estudios Lingüísticos sobre Variedades Hispanas, Portuguesas y Criollas*, 139-217. Frankfurt am Main: Vervuert.
- Moreira, Ana Karina Tavares. forthcoming (a). Le créole afro-portugais de l'île de Fogo: Brève présentation historique et linguistique. *Annales de la Fondation Martine Aublet*.
- Moreira, Ana Karina Tavares. forthcoming (b). *Documentation et description grammaticale et lexicale du créole afro-portugais de l'île de Fogo (République du Cap-Vert, Afrique de l'Ouest)*. Paris: INALCO, Phd dissertation.
- Parkvall, Mikael. 2000. *Out of Africa; African Influences in Atlantic Creoles*. London: Battlebridge.
- Quint, Nicolas. 2000a. *Grammaire de la Langue Cap-verdienne*. Paris: L'Harmattan.
- Quint, Nicolas. 2000b. *Le Cap-verdien: Origines et Devenir d'une Langue Métisse*. Paris: L'Harmattan.
- Quint, Nicolas. 2001. Langues créoles, diachronie et procédés de reconstruction. *Bulletin de la Société Linguistique de Paris* XVI(1). 265-284.
- Quint, Nicolas. 2006. Un bref aperçu des racines africaines de la langue capverdienne, In Jürgen Lang, John Holm, Jean-Louis Rougé & Maria João Soares (eds.), *Cabo Verde - Origens da sua Sociedade e do seu Crioulo*, 75-90. Tübingen: Günter Narr.
- Quint, Nicolas. 2008. *L'Élément Africain dans la Langue Capverdienne / Africanismos na Língua Cabo-Verdiana*. Paris: L'Harmattan.
- Quint, Nicolas. 2009. As formas divergentes em cabo-verdiano. In Ana Carvalho (ed.), *Português em Contacto*, 67-85. Madrid/Frankfurt am Main: Iberoamericana/Vervuert.
- Quint, Nicolas. 2012. African words and calques in Capeverdean Creole (Santiago dialect). In Philip Baker & Angela Bartens (eds.), *Black Through White. African Words and Calques in Creoles and Transplanted European Languages*, 3-26. London/Colombo: Battlebridge.
- Quint, Nicolas. in press. Djifanghor Nyun (Baïnouck). In Friederike Lüpke (ed.), *Oxford Guide to the World's Languages: Atlantic*. Oxford: Oxford University Press.
- Rougé, Jean-Louis. 1988. *Petit Dictionnaire Étymologique du Kriol de Guinée Bissau et de Casamance*. Bissau: INEP.
- Rougé, Jean-Louis. 1994. À propos de la formation des créoles du Cap Vert et de Guinée. *Papia* 3(2). 137-149.
- Rougé, Jean-Louis. 1995. A propósito da formação dos Crioulos de Cabo Verde e da Guiné. *Soronda* 20. 81-98.
- Rougé, Jean-Louis. 1999a. Apontamentos sobre o léxico de origem africana dos crioulos da Guiné e de Cabo Verde (Santiago). In Klaus Zimmermann (ed.), *Lenguas Criollas de Base Lexical Española y Portuguesa*, 49-67. Frankfurt am Main: Vervuert.
- Rougé, Jean-Louis. 1999b. A propos du vocabulaire africain dans les relations de voyages des Portugais en Guinée du Cap Vert (1594-1697). In Anne-Marie Quint (ed.), *Cahier du CREPAL N°6: Le Conte en Langue Portugaise*, 177-201. Paris: Presse de la Sorbonne Nouvelle.
- Rougé, Jean-Louis. 2004. *Dictionnaire Étymologique des Créoles Portugais d'Afrique*. Paris: Karthala.
- Rougé, Jean-Louis. 2005. A formação do léxico dos crioulos portugueses da África. *Papia* 15. 7-17.
- Rougé, Jean-Louis. 2006. L'influence manding sur la formation des créoles du Cap-Vert et de Guinée-Bissau et Casamance. In Jürgen Lang, John Holm, Jean-Louis Rougé & Maria

- João Soares (eds.), *Cabo Verde - Origens da sua Sociedade e do seu Crioulo*, 63-74. Tubingen: Gunter Narr.
- Santos, Rosine. 1979. Comparação entre o Crioulo de Cabo Verde e as línguas africanas. *I Colóquio linguístico sobre o crioulo de Cabo Verde*, 121-44. Praia/Mindelo: Ministério da Educação e Cultura, Direcção Geral da Cultura.
- Scantamburlo, Luigi. 1981. *Gramática e Dicionário da Língua Criol da Guiné-Bissau (GCr)*. Bologne: Missionaria Italiana.
- Scantamburlo, Luigi. 2002. *Dicionário do Guineense*, vol. II (*Dicionário Guineense-Português. Disionariu Guinensi-Portuguis*). Lisbon: Edições Colibri, Faspebi.
- SIL. 2003. *Saramaccan-English Dictionary*, <http://www.suriname-languages.sil.org/Saramaccan/English/SaramEngDictIndex.html>. (20 September, 2018.)
- Todd, Loreto, 1990. *Pidgins and Creoles*. London/New York: Routledge.
- Wilson, William A. A. 1962. *The Crioulo of Guiné*. Johannesburg: Witwatersrand University Press.

APPENDIX

Appendix 1. African-derived items attested in ST and/or one or several other UGPC varieties

Table 1.1: UGPC lexical items derived from Manding

Proto-UGPC	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/ baga'baga / (+ABC) ³²	'ant sp. (IN) vs. termit (CO+ABC)'	['bagɛ'bagɛ]	['bage'bage]	['baga'bagaga]	['baga'bagaga]	['baga'bagaga]	['baga'bagaga]	6	all	0	IN vs. CO
*/ bam'bu /	'carry on one's back'	['bõbu] ~ ['bẽbu]	[bẽ'bu]	['bambu]	[bam'bu]	[bam'bu]	[ba'mbu]	6	all	0	0
*/ bambu'raŋ /	'saddle sp.' (IN) vs. 'cloth used to carry s.o. on one's back' (CO)	[bõbu'rõ]	[bõbu'rõ]	[bamba'raŋ]	[bamba'raŋ]	NC	[bamba'raŋ]	5	all	IN vs. CO	IN vs. CO
*/ bimbi'riŋ / (IN)	'millet'	[bĩbi'rĩ]	[bĩbi'rĩ]	X	X	X	X	6	IN vs. 0CO	0	0
*/ bo'li /	'gourd sp.'	['boli]	[bo'li]	['buli] ~ ['bule]	['buli]	['buli]	['boli]	6	all	IN+CS vs. all	0
*/ bu'su /	'graze (v.) (=damage one's skin)' (IN + CS) vs. 'unsheathe, unroot' (BI) vs. 'lose straw (broom)' (GE)	['busu]	[bu'su]	[bu'si]	['busi]	[bu'si]	[bu'si]	6	all	IN vs. CO	unclear
*/ bur'bur /	'non-sticky, powdery' (ST) vs. 'crumbs, grinding residues' (CO)	bur-bur [bur'bur]	X	[burbur]	[bur'bur]	[bur'bur]	X	6	0FG+ CS vs. all	0	ST vs. CO

³² The label '+ABC' means that the Proto-UGPC reconstructed form is also based on the Papiamentu (ABC) cognate.

* coro'ti/ ~ * poro'ti/	'drip' (ST+FG) vs. 'be stingy towards s.o.' (ST+GE)	[tʃo'roti]	[tʃoro'ti]	X	[ɲoro'ti]	NC	NC	4	0BI vs. all	IN vs. CO	uncle ar
* cu'ki/ ~ * co'ki/	'stick into, pierce' (IN+BI /u/) vs. 'pack (down), cram sth. into (v.)' (CS+CU+GE +BI /o/)	[tʃuki]	[tʃu'ki]	[cuki ~ coki]	['coki]	[co'ki]	[co'ki]	6	all	IN+BI vs. all+BI	IN+B I vs. all+B I
* da'bi/	'bedbug'	[de'biw]	([perse'be fu]) < Pt. percevejo)	[da'bi ~ de'bi]	[da'bi]	[da'bi]	[de'bi]	6	0FG vs. all	ST vs. CO	0
* disdaŋ'gu/	'neglected' (FG) vs. 'refuse to answer, to sulk at' (all) vs.	[diz'dɛgu]	[dizdɛ'gud u]	[dis'daŋ gu]	[dis'daŋ u]	[dis'daŋgu]	[disda'ŋgo]	6	all	CS vs. all	FG vs. all
* ferfe'ri/	'make use of any resource' (ST) vs. 'glean' (CO)	[fer'feri]	X	[ferferi]	['ferfer] ~ [fer'feri]	[fefe'ri]	[fefe'ri]	5	0FG vs. all	CU+CS vs. all	ST vs. CO
* finu/ (IN) vs. * nok/ (CO)	'very (black)'	['finu]	['finu]	['nok]	['nok]	['nɔk]	['nok]	6	IN vs. CO	0	0
* fus/	'silent fart'	['fus]	X	['fus]	['fus]	[ka'fus]	[ka'fus]	6	0FG vs. all	CU+CS vs. all	0
* jamba/	'big locust sp.' (ST) vs. 'bird sp. (CO)	[dʒãbɛ]	X	['jamba]	['jamba]	['jamba]	['jamba'tut u]	6	0FG vs. all	CS vs. all	ST vs. CO
* jo'be/	'look (v.)'	[dʒobe]	[dʒo'be]	[jubɪ]	['jubɪ]	[ju'bi]	[jo'be]	6	all	IN+CS vs. all	0
* joŋ'go/	'doze (v.)'	[dʒɔŋgo]	[dʒɔ'go]	[juŋgu]	[juŋ'gu]	[joŋ'go]	[jo'ŋgo]	6	all	BI+GE vs. all	0
* joŋgo'to/ (+ABC)	'squat (v.)'	[dʒɔ'goto]	[dʒɔŋgo'to]	['juŋgut u]	['juŋkutu]	([joko'ni] < Wf. jonkon)	([jokoni] < Wf. jonkon)	6	0CU+ CS vs. all	IN vs. CO	0

The common African lexical core of Upper Guinea Creoles and its historical significance

*/ju'tu/ ~ */ju'ti/	'underestimate'	[ɲdʒutu]	[dʒu'tu]	[juti] ~ [ɲjuti]	[ɲdʒuti]	[ju'ti]	[ju'ti]	6	all	IN vs. CO	0
*/ka'raŋga/ (+ABC)	'louse'	([ɲ'pjoɖʒu ɲ < Pt. pioɻho)	([ɲ'pjoɖʒu < Pt. pioɻho)	[ka'raŋ a]	[ka'raŋ a]	[ka'raŋ a]	[ka'raŋ a]	6	0IN vs. CO	0	0
*/kaŋk(a)ra (ŋ)/	'bed sp.' (IN) vs. 'framework /roof' (CO)	[kɛkɛ'rɛ]	[kɛkɛ'rɛ]	['kaŋkra]	['kaŋkra]	['kaŋkra]	['kaŋkra]	6	all	IN vs. CO	IN vs. CO
*/ko'tʃi/	'pound grain (v.)'	[kotʃi]	[ku'tʃi]	[koci] ~ [koce]	['koci]	['koci]	[ko'ci]	6	all	0	0
*/ko'de/	'last-born child'	[ko'de]	[ko'de]	[ko'de]	[ko'de]	[ko'de]	[kɔ'dɛ]	6	all	0	0
*/koŋ'ki/ ~ */koŋ'ko/	'knock (door...)'	[kɔko]	[kɔ'ko ~ kɔ'ki]	[koŋki]	[koŋ'ki]	[koŋ'ki]	[ko'ŋki]	6	all	ST+FG vs. CO+FG	0
*/kundin'diŋ /	'coccyx'	[kūdi'dē]	[kūdi'dē]	[kundin' diŋ]	[kundin'd iŋ]	[kundin'diŋ]	[kundi'ndiŋ]	6	all	IN vs. CO	0
*/'loti'loti/	'soft, flaccid'	[lɔti'loti]	X	X	[(loti)lo'ti du]	X	X	4	ST+G E vs. Oall	unclear	0
*/man'duku/	'stick (n.)'	[mɛ'duku]	[mɛ'duku]	[man'du ku]	[man'duk u]	[man'duk u]	[ma'nduk]	6	all	CU+CS vs. all	0
*/mon'do/	'curl up, triturate' (IN) vs. 'shape food in one's hand' (CO)	[mōdo]	[mō'do]	[mundu]	[mun'du]	[mun'du]	[mo'ndo]	6	all	IN+CS vs. all	IN vs. CO
*/moŋoŋ'go/ ~ */moŋoŋ'gi/	'crumple (v.), rumple (v.)' (IN+GE+CU /i/) vs. 'shell, shuck (maize)' (BI +CS + CU /u/) vs. 'grind, crumble (v.)' (BI)	[mo'ŋɔŋgi]	[moŋɔ'go]	[mu'ŋuŋ gu]	[mu'ŋuŋ i]	[muŋuŋ'gi ~ muŋuŋ'gu]	[muŋoŋ'ka]	6	all	CS vs. all	uncle ar
*/ɲe'me/	'chew'	[ɲeme]	[ɲɛ'mi]	[ɲeme] ~ [ɲemi]	[ɲe'me]	[ɲe'me]	[ɲe'me]	6	all	FG vs. all	0

*/ ɲoli /	‘look daggers at s.o.’	['ɲuli] ~ ['ɲoli]	X	[ɲuli]	['ɲoli]	['ɲuli]	[ɲo'li]	6	0FG vs. all	CS+GE+ST vs. BI+CU+S T	0
*/ so'li /	‘fill’	['nsoli]		[su'li]	[suli]	['suli] ~ ['soli]	[so'li]	6	all	ST+CS+CU vs. all+CU	0
*/ so'ti /	‘insert (v.)’	['soti]	X	[suti]	['suti]	['suti] ~ ['soti]	[su'ti]	6	0FG vs. all	ST+CU vs. all	0
*/ ten'te ³³ /	‘winnow (v.)’	['tête]		[tê'te]	X	(EX [ten'te])	X	6	IN vs. OCO	0	0
*/ wan'daŋ /	‘[swing] wide open’	[wẽ'dê]	X	[wan'daŋ]	[wan'daŋ]	NC	NC	4	0FG vs. all	0	0

Table 1.2: UGPC lexical items derived from Wolof

Proto-UGPC	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/ ba'ci /	‘pound a second time (husked corn/rice)’	['pɛtʃi]	([so'tɛ] < Pt. <i>açoitar</i>)	([limpsa] < unknown, Pt. <i>limpo</i> ?)	NC	NC	[wa'ci]	4	0FG+Bi vs. ST+CS	ST vs. CS	0
*/ bindi /	‘pot with holes’	['bīdi]	['bīdi]	[bindi]	['bindi]	['bindi]	['bindi]	6	all	0	0
*/ cupu'ti /	‘pinch (v.)’ (IN+GE+CU) vs. ‘peck/pick up (food) [with fingers]’ (BI+CS) ³⁴	[tʃu'puti]	[tʃɛpu'ti]	[cuputi] ~ [copoti] ~ [cupiti]	['cuputi]	[cupu'ti]	[copo'ti]	6	all	unclear	BI+CS vs. all

³³ In Continental UGPCs, the most common semantic equivalent of (Proto-)Insular UGPC */**tente**/ is an item derived from Proto-UGPC */**fe'ki**/, also attested in ST, but with a different meaning (see Table 1.7 below).

³⁴ In both BI and CS, another verb is associated with the meaning ‘pinch’, respectively [ʃɔpoti] (BI) and [ɲopo'ti] (CS). These forms are both derived from Mdk. *ñópótí* (DC 2012:199) or Wf. *ñoppati* (AF 1990:161) ‘pinch’, of which the latter is possibly related to Wf. *coppati* [cɔ:ppati] ‘pick up (food) with fingers’ (AF1990:53, JLD 2003:84). The meaning of GE and CU forms was not checked in details: it may well be the case that these two other continental forms display the same meaning as BI and CS and that we have another case of semantic contrast between Insular and Continental UGPCs.

The common African lexical core of Upper Guinea Creoles and its historical significance

*/'feti'feti/	'rub linen'	['feti'feti]	[fi'ti fi'ti]	[fet feti]	['feti'feti]	['feti'feti]	[fet fe'ti]	6	all	unclear	0
*/'jaga'si/	'mix (alcohol, food)' vs. 'mix (general)'	[dʒɛ'gɛsi]	[dʒɛgɛ'si]	[ʒagasi] ~ [ʒakasi]	[ʒaka'si]	[ʒaka'si]	[ʒaka'si]	6	all	IN+ BI vs. all+BI	IN vs. CO
*/'joŋgor'ni / ~ */'joŋjor'ni /	'overlapping (teeth)' (ST) vs. 'overlap' (CO)	[dʒɔŋgor'ni dʒɔŋjor'ni]	([mũ'tadu] < Pt. <i>montar</i>)	X	NC	NC	[joŋjor'ni]	4	0FG+ BI vs. ST+C S	ST vs. CS	ST vs. CS
*/'lam'bu/	'wrap up' (IN) vs. 'collect/steal/take/keep' (CO)	['lɔbu] ~ ['lɛbu]	[lɛ'bu]	[lambu]	[lam'bu]	[lam'bu]	[la'mbu]	6	all	0	IN vs. CO
*/'loko'ti/	'scrape with finger'	[lo'koti]	([gɛrba'tɛ] < Pt. <i>esg(a)ravatar</i>)	[lokoti] ~ [lokati]	(['ɲori]) ³⁵	[loko'ti]	[loko'ti]	6	0FG+ GE vs. all	0	0
*/'ma'fe/	'fat (Tarrafal area)' (ST) vs. 'garnish, accompaniment' (CO)	[mɔ'fe]	X	[ma'fe]	[ma'fe]	[ma'fe]	[ma'fe]	6	0FG vs. all	ST vs. CO	ST vs. CO
*/'moku/	'completely (pounded)'	['moku]	X	['moku]	NC	NC	['moku]	4	0FG vs. all	0	0
*/'mu'ga/ (IN) vs. */'mu'fa/ (CO)	'eat with hand' (BI) vs. 'eat powdered food' (all)	['mugɛ]	[mu'gɛ]	[mufa]	NC	[mu'fa]	[mu'fa]	5	all	IN vs. CO	BI vs. all
*/'ndor/ (IN)	'triggerfish'	['ndɔr] ~ [kɛ'gul]	[do'rotʃi] ~ [fɛ'bil] ~ [fɛ'bi]	([du'tur] < Pt. <i>doutor</i> ~ ['pis- ga'lija] < Pt. <i>peixe</i> + <i>galinha</i>)	NC	X	X	5	IN vs. OCO	ST vs. FG	0
*/'ɲapa'ti/	'bite off (v.)'	[ɲɛ'peti]	X	['ɲapati] ~ ['napati]	['ɲapati]	[ɲapa'ti]	[ɲapa'ti]	6	0FG vs. all	ST vs. CO	0
*/'ɲgor(i)/ (x2)	'potter wasp'	['ɲgori]	([kɔs'tansjɛ gɛr'siɛ] > Pt. <i>anthroponyms</i>)	[ɲgor ɲgor]	NC	[ɲgɔriŋgɔ r]	[ɲgoruŋgor]	5	0FG vs. all	ST vs. CO	0
*/'ori/	'wari (game)'	['ori] ~ ['oru]	['ori]	[uri]	NC	['oro] ~ ['oro]	([awa'le])	5	0CS vs. all	CU+ ST vs. FG+BI+S	0

³⁵ A cognate also exists in Cacheu, namely [ɲu'ri] with the meaning 'scrape (plate) with fingers'.

										T	
*/si'ki(du)/	'stand'	[sɐ'kɛdu]	[sɐ'ke]	[si'kidu]	NC	NC	[si'kidu]	4	all	IN vs. CO	0
*/(ŋ)u'pi/	'show one's butt to'	[ʔuɲi] ~ [ʔupe]	[u'pi]	[ŋuɲi]	[ʔuɲi]	[ŋu'pi]	[wo'ŋi] ~ [wo'pi]	6	all	IN+CS vs. all	0

Table 1.3: UGPC lexical items derived from Manding and/or Wolof

Proto-UGPC	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/fɛp/	'completely, totally'	[fɛpu]	[fɛpu]	[fɛp]	[fɛp]	[fɛp]	[fɛp]	6	all	IN vs. CO	0
*/fVkvVt/ ~ /*pirkit/	'[raise] at once'	[fɛ'kɛti] ~ [pri'kiti]	X	X	[fikiti]	NC	[pir'kit]	5	0FG+BI vs. all	unclear	0
*/jam(b)u/	'mourn, pity (v.)'(ST) vs. 'praise' (BI)	[dʒɛbu]	X	[jamu]	X	X	X	6	ST+BI vs. 0all	ST vs. BI	ST vs. BI
*/jɛ'ɲɛ/ ~ /ge'ŋge/	'incline'	[dʒɛdʒɛ]	X	[jɲɲi]	[gɲɲi]	[gɲɲi]	[jɛ'ɲɛ]	6	0FG vs. all	ST+CS vs. all	0
*/kal'ma(ŋ)/ ~ */kalam'ba(ŋ) / (+ABC)	'ladle sp.'	[kɛl'mɛ] ~ [kɛr'mɛ]	[kɛl'mɛ]	[kal'mon] ~ [kalma]	[kal'ma]	[kal'mon] (old) ~ [kal'ma]	[kal'mon]	6	all	unclear	0
*/lonki/ ~ */longi/ ~ */londi/ ~ */lonti/	'swing (v.) on a stick tied to a horizontal branch, oscillate'	[dʒo'lɔŋgi] ~ [lo'lɔdi]	([dɛ bɔ'lɛ] < Pt. 'dar' + unknown)	[lonki] ~ [lonti- lonti]	([indoli] < unknown)	[lonti'lonti]	[gle'randu] ~ [lelo'nte] ~ [fa'si le'lo]	6	0FG+GE vs. all	ST vs. CO	0
*/me'sɛs/	'poultry's louse'	[me'sɛs]	([pʒodʒu gɐ'liɲɛ] < Pt. 'piolho + galinha')	[mesis]	([ka'ranɲa di ga'liɲɛ] < Md. karanga 'louse' + Pt. de + galinha)	[mi'sis]	[mi'sis]	6	0FG+GE vs. all	ST+BI vs. CU+CS+BI	0
*/se'reŋ/	'dish made of pounded cereals'	[ʃɛ'rɛ]	[ʃɛ'rɛ]	[se'reŋ]	[se'reŋ]	X	NC	5	0CU vs. all	IN vs. CO	0
*/sibi/	'fan palm'	[sibi]	X	[sibi]	[sibi]	[sibi]	[sibi]	6	0FG vs. all	0	0

Table 1.4: UGPC lexical items derived from Temne

Proto-UGPC	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/fũḡku/ (IN)	'hut sp.'	[fũku]	[fũku]	X	X	X	X	6	0CO vs. IN	0	0
*/i'raŋ/	'genius, spirit'	[i'rẽ]	X	[i'raŋ]	[i'raŋ]	[i'raŋ]	NC	5	0FG vs. all	0	0
*/po'loŋ/	'kapok tree'	[po'lõ]	[po'lõ]	[po'loŋ]	[po'loŋ]	[po'loŋ]	[po'loŋ]	6	all	0	0
*/ta'baŋka/	'carnival sp.' (IN) vs. 'village' (CO)	[tẽ'bãkẽ]	[tẽ'bãkẽ]	[ta'baŋka]	[ta'baŋka]	[ta'baŋka]	[ta'baŋka]	6	all	0	IN vs. CO

Table 1.5: UGPC lexical items derived from Other Atlantic Languages

Proto-UGPC	Origin	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
			Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/ka'mati/	Manjaku/ Mankanya	'wild tomato' (ST) vs. 'tomato' (CO)	[kẽ'mati]	X	[ka'mati] ~ [ka'mate] ~ [ka'matu]	[ka'mati]	[ka'mati]	[ka'mati]	6	0FG vs. all	BI+CU vs. all	ST vs. CO
*/ma(n)'fafa/	Biafada	'yam sp.'	[mẽ'fafẽ]	[mẽ'fafẽ]	[man'fafa]	[man'fafa]	[man'fafa]	[man'fafa]	6	all	IN vs. CO	0
*/maŋ'kara/	Nyun	'peanut'	[mẽ'kare]	[mẽ'kare]	[maŋ'kara]	[maŋ'kara]	[maŋ'kara]	[maŋ'kara]	6	all	0	0
*/riki'ti/	Balanta	'take off a bit of' (IN) vs. 'pinch' (CO)	[ru'kuti]	[riki'ti]	[riki'ti]	[riki'ti]	[riki'ti]	X	6	0CS vs. all	ST vs all	IN vs. CO

Table 1.6: UGPC lexical items derived from Bantu Languages (Kikongo and Kimbundu)

Proto-UGPC	Origin	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
			Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/bunda/	Kimbundu	'ass, buttocks'	[bũde]	X	[bunda]	[bunda]	[bunda]	[bunda]	5	0FG vs. all	0	0
*/ma'laŋka/	Kikongo	'yam sp.'	[mẽ'lãkẽ]	[mẽ'lãkẽ]	X	[ma'laŋka]	([jãmbi] < Wf. ñambi 'cassava')	NC	5	0BI+CU vs. all	0	0

Table 1.7: Unidentified putative African-derived UPGC lexical items

Proto-UGPC	Meaning	Insular Creoles		Continental Creoles				Data	Criteria		
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		PRA	PHON	SEM
*/fe'ki/	'shake (pot/pan) so that the food does not burn' (ST) vs. 'winnow' (CO)	['fetʃi] ~ ['feki 'feki]	X	[feki]	['feki]	['feki]	[fe'ki]	5	0FG vs. all	ST vs. CO	ST vs. CO
*/futu'tu(t)/	'strongly (smell)'	[futu'tu]	NC	X	NC	NC	[futu'tut]	3	0BI vs. ST+CS	0	0
*/(p)joja/	'fix/repair (rope)'	['dʒõdʒõ] ~ ['dʒõdʒõ] ³⁶	['dʒõdʒõ]	['ɲɲɲɲ]	[ɲɲɲɲ]	['ɲɲɲɲ]	[ɲɲɲɲ]	6	all	IN vs. CO	0
*/jorom'bo/	'shake basket horizontally to separate pounded grain according to size'	[dʒõ'rõbo] ~ [õ'rõbo]	[ɛrõ'bo]	[jurumbu] ~ [jurumbu]	[jurum'bu]	[jurum'bu]	[joro'mbo]	6	all	IN+CS vs. all	0
*/jugu'ta/	'scramble, manage' (IN) vs. 'jump' (CO)	[dʒu'gutɛ]	[dʒugu'tɛ]	['juguta ~ 'jukuta ~ 'jukta]	[ci'rimpa] < unknown	[jugu'ta]	[jugu'ta] ~ [juk'ta]	6	0GE vs. all	0	IN vs. CO
*/korko'ti/	'scrape (with spoon)' (ST) vs. 'slander' (FG)	[kor'koti]	[kɛrku'ti]	[kokori]	[koko'ri]	[kokor'ti]	[koko'ri]	6	all	IN vs. CO	FG vs. all
*/lalu/	'slip, be slippery (v.)'	['lolo]	[skõkõ'gɛ] < Pt. <i>escorregar</i>	[lalu]	['lalu]	NC	[la'li]	5	0FG vs. all	ST vs. CO	0
*/lo'fi/	'pound (plant leaves)'	['lofi]	X	[lofi]	NC	NC	[lo'fi]	4	0FG vs. all	0	0
*/'ɲana/	'corn husk' (ST+CS) vs. 'bean pod' (FG) vs. 'rice bran'	['ɲanɛ]	['ɲganɛ]	['ɲana]	['soka] < Pt. <i>soca</i>	['ɲana]	['ɲana]	6	0GE vs. all	FG vs. all	CU+C S vs. ST + BI vs. FG

³⁶ The forms ['dʒõdʒõ] and ['dʒõdʒõ] also occur in ST, but were only attested in less than 10 out of 300+ points of inquiry.

The common African lexical core of Upper Guinea Creoles and its historical significance

	(CU+CS)												
*/re'be/	'leak (pot)' (IN) vs. 'melt' (CO)	[ˈrebe]	[ʁe'be]	[ri'bi] ~ [re'be]	[ˈri'bi]	[ri'bi]	[re'be]	6	all	IN+CS + BI vs. GE+ CU+BI	IN vs. CO		
*/CV(C)CV/ + /tutu/	'antlion'	[kutübē'bɛ] ~ [tu'tu]	[ma'ma tu'tu] ~ [tu'tu bɛ'relə]	[jamba'tut u]	[jamba'tutu]	[jamba'tutu]	[ma'ma ku'ti] ~ [koko'le-ko'le] ~ [ma'me di tuk-'tuk]	6	all	unclear	0		

Table 1.8: African-derived lexical items common to Papiamentu and at least one other branch of UGPCs

Proto-UGPC	Origin	Meaning	Insular Creoles		Continental Creoles				ABC	Data	Criteria		
			Santiago	Fogo	Bissau	Geba	Cacheu	Casamance	Curaçao		PRAB	PHON	SEM
*/'baga'baga/	Md.	'ant sp. (IN) vs. 'termite' (CO+ABC)	[ˈbagɛ'bagɛ]	[ˈbagɛ'bagɛ]	[ˈbaga'bagɛ]	[ˈbaga'bagɛ]	[ˈbaga'bagɛ]	[ˈbaga'bagɛ]	[ˈbadʒa'ga]	7	all	ABC vs. all	IN vs. all
*/jongo'to/	Md.	'squat (v.)'	[dʒɔ'gɔto]	[dʒɔ'gɔ'to]	[ˈjuŋgut u]	[ˈjuŋ kutu]	([joko'ni] < Wf. jonkon)	([joko'ni] < Wf. jonkon)	[jɔŋgo'ta]	7	OCU+CS vs. all	ABC vs. IN vs. CO	0
*/kal'ma(ŋ) ~ /kalam'ba(ŋ)/	Md./Wf..	'ladle sp.'	[kəl'mɛ] ~ [kɛr'mɛ]	[kəl'mɛ]	[kal'moŋ] ~ [kalma]	[kal'ma]	[kal'moŋ] (old) ~ [kal'ma]	[kal'moŋ]	[kalɛm'bɛ]	7	all	ABC vs. all	0
*/ka'raŋga/	Md.	'louse'	([ˈpjoʒu] < Pt. piolho)	([ˈpjoʒu] < Pt. piolho)	[ka'raŋga]	[ka'raŋga]	[ka'raŋga]	[ka'raŋga]	[ka'raŋga]	7	IN vs. all	0	0

Table 1.9: African-derived lexical items found only in Santiago Capeverdean

Origin	Meaning	Insular Creoles		Continental Creoles				Data	Criterion
		Santiago	Fogo	Bissau	Geba	Cacheu	Casamance		
Md.	‘shake, be shaky’	[di'gigi]	([sɐku'di] < Pt. <i>sacudir</i>)	X	X	X	X	6	ST vs. 0all
Md.	‘black ant sp.’	[dũdu]	([fɨrmi'gijnɐ] < Pt. <i>formiga</i> + <i>-inha</i>)	X	X	([fir'miŋga] < Pt. <i>formiga</i>)	X	6	ST vs. 0all
Wf.	‘bull calf, young bull’	[dʒaki]	([ʔɛs] < Pt. <i>rés</i>)	X	([ʔuru'siju] < Pt. <i>touro</i> + <i>-zinho</i>)	([ʔuru'siju] < Pt. <i>touro</i> + <i>-zinho</i>)	([ʔuru'siju] < Pt. <i>touro</i> + <i>-zinho</i>)	5	ST vs. 0all
Fula	‘hurried, awkward’	[dʒur'dʒur]	X	X	X	X	X	6	ST vs. 0all
Md.	‘thrash about’	[futi'futi]	X	X	X	X	X	6	ST vs. 0all
Md.	‘very (old)’	[gɔbu'gɔbu]	X	X	NC	NC	NC	3	ST vs. 0all
Md.	‘very (old)’	[kɔti'kɔti]	X	X	NC	NC	NC	3	ST vs. 0all
Wf.	‘sound of pestle on mortar’	[kũ'dũ kũ'dũ]	X	X	NC	NC	[cum cum]	4	ST vs. 0all
Wf.	‘big tree sp.’	[lɛmbɛ lɛmbɛ]	X	X	X	X	X	6	ST vs. 0all
Kikongo/ Kimbundu	‘snare (n.)’	[mɛtɛ'bu]	X	X	X	X	X	6	ST vs. 0all
Wf.	‘trap sp.’	[mbipu]	X	X	NC	NC	X	4	ST vs. 0all
Md.	‘scintillation’	[mɛni'mɛni]	X	X	X	X	X	6	ST vs. 0all
Wf.	‘gnaw’	[ɲɛpi]	X	X	X	X	X	6	ST vs. 0all
Md.	‘chewing tobacco’	[si're]	(<i>bora</i> < Pt. ?)	X	(EX [si're])	X	X	6	ST vs. 0all

Appendix 2. UGPC African-derived items whose etyma have not been previously published or discussed in Quint (2000b, 2006, 2008)

Table 2.1: UGPC African-derived lexical items and their etyma not discussed in previous work by Quint

Proto-UPGC *// or ST [] form	Origin	Meaning	Etymon
*/bambu'raŋ/	Md.	'saddle sp.' (IN) vs. 'cloth used to carry s.o. on one's back' (CO)	Mdk. <i>bàmbùràŋ</i> 'cloth used to carry s.o. on one's back' (DC 2012:14)
*/coro'ti/ ~ */ɲoro'ti/	Md.	'drip' (ST+FG) vs. 'be stingy towards s.o.' (ST+GE)	Mdk. <i>córóti</i> 'drip' (DC 2012:34)
*/fep/	Md./Wf .	'completely, totally'	Mdk. <i>fěw</i> '[finish] completely' (DC 1981:47, JLC 2004:308) X Wf. <i>-épp</i> 'completely'; <i>fěpp</i> [fɛ:ppə] 'complete, every(where)' (JLD 2003:125)
*/'finu/ (IN) vs. */'nok/ (CO)	Md.	'very (black)'	1) Mdk. <i>fīŋ</i> (IN) 'black' (DC 2012:67; JLR 2004:309) X Port. <i>fino</i> 'fine, refined' 2) Mdk. <i>nót</i> (CO) 'very (black and dirty)' (DC 2012:188)
*/fVkvVt/ ~ */'pirkit/	Md./Wf .	'[raise] at once'	Mdk. <i>cókót</i> '[raise] at once' (DC 2012:33) X Wf. <i>ñokket</i> [ɲo:kket] '[raise] at once' (JLD 2003:269) + Wf. <i>fojjet</i> [fɔ:ʝet] '[raise] immediately' (JLD 2003:128)
*/jam(b)u/	Md./Wf .	'mourn, pity (v.)' (ST) vs. 'praise' (BI)	Mdk. <i>jàmúŋ</i> 'praise' (DC 2012:88) X Wf. <i>jàmbat</i> [ja:mbat] 'complain/lament constantly' (JLD 2003:158). The meanings of Mdk. and Wf. differ considerably and match respectively with BI and ST. Note however that ST ['dʒɛbu] seems to display similarities with the Mdk. etymon.
*/je'ɲje/ ~ /ge'ŋge/	Md./Wf .	'incline'	Mdk. <i>ɲèkè</i> 'be twisted/bent' (DC 2012:91) X Wf. <i>jeng</i> [jɛ:ŋgə] 'tilt, incline' (JLD 2003:168)
*/joŋgor'ni/ ~ */joŋgor'ni/	Wf.	'overlapping (teeth)' (ST) vs. 'overlap' (CO)	Wf. <i>jaŋar</i> [jaŋar] 'have overlapping teeth' (JLD 2003:160); <i>jang</i> 'overlap (of teeth)' (AF 1990:92)
*/kal'ma(ŋ)/ ~ */kalam'ba(ŋ)/ (+ABC)	Md./Wf .	'ladle sp.'	Mdk. <i>kàlàmaa</i> 'big spoon' (DC 2012:105) X Wf. <i>kalamba</i> 'wooden recipient used to measure millet' (JLD 2003:180)
*/loŋki/ ~ */loŋgi/ ~ */londi/ ~ */lonti/	Md./Wf .	'swing (v.) on a stick tied to a horizontal branch, oscillate'	Mdk. <i>jòlóng</i> 'drop, fall down'; <i>lènléŋ</i> 'hang on, cling' (DC 2012:95, 156) X Wolof <i>lonk</i> [lɔ:ŋkə] 'hang on, cling' (JLD 2003:205)
*/me'ses/	Md./Wf .	'poultry's louse'	Mdk. <i>mèséesi</i> 'poultry's louse' (DC 2012:171) X Wf. <i>meesees</i> [mɛ:sɛ:s] 'poultry's louse' (JLD 2018:223)
*/'moku/	Wf.	'completely (pounded)'	Wf. <i>mokk</i> [mɔ:kkə] 'be ground/crushed' (Dieng 1985:247; JLD 2018:226)
*/riki'ti/	Balanta	'take off a bit of' (IN) vs. 'pinch' (CO)	Balanta <i>rikít</i> '[pinch] fiercely/strongly' (JLR 2004:348)
*/se'reŋ/	Md./Wf .	'dish made of pounded cereals'	Mdk. <i>sèréŋ</i> 'cooked/boiled millet' (DC 2012:219) X Wf. <i>cere</i> 'millet couscous'; <i>sereŋ</i> 'millet couscous' (JLD 2003:82, 312)
*/'sibi/	Md./Wf .	'fan palm'	Mdk. <i>sibi</i> 'fan palm' (DC 2012:219) X Wf. <i>sibi</i> 'fan palm' (AF 1990:199)

*/si ki(du)/	Wf.	‘stand’	Wf. <i>siggì</i> ‘stand up’; Rougé (2004:242) proposes a Pt. etymon <i>quedo</i> ‘calm, still’ but the Wolof etymon (already proposed by Dieng (1985:248) seems much more plausible especially if one considers the formal similarity between Wf. and Continental UGPC forms [si’kidu].
*/wan’dan/	Md.	‘[swing] wide open’	Mdk. <i>wáránj</i> ‘[swing] wide open’ (DC 2012:264)
ST [‘gɔbu’gɔbu]	Md.	‘very (old)’	Mdk. <i>kób</i> ‘very (old)’ (DC 2012:264)
ST [‘kɔti’kɔti]	Md.	‘very (old)’	Mdk. <i>kòtó/kòtóo</i> ‘(be) old, elder’ (DC 2012:137)
*/ba’ci/	Wf.	‘pound a second time (husked corn)’	Wf. <i>bàcc</i> [ba:ccə] ‘pound (grain)’ (JLD 2003: 58)

Appendix 3. Comparison of UGPC African-derived lexical items according to three criteria

Table 3.1: African-derived lexical items classified according to presence/absence in one or several UPGCs

Presence/absence (PRA)	Origin							TOTAL	%	% MP	Main profiles (MP)
	Md.	Wf.	Md./Wf.	Temne	Other Atlantic	Bantu	Unidentified				
all	20	8	2	2	2	0	5	39	41	41	all
0FG vs. all	7	4	2	1	1	1	3	18	20	25	0FG (+GE/BI/CS) vs. all
0FG+GE vs. all	0	1	2	0	0	0	0	3	3		
0FG+BI vs. all	0	0	1	0	0	0	0	1	1		
0FG+CS vs. all	1	0	0	0	0	0	0	1	1		
ST vs. 0all	7	5	0	0	1	1	0	14	15	15	ST vs. 0all
0GE vs. all	0	0	0	0	0	0	2	2	2	2	0GE vs. all
0CS vs. all	0	1	0	0	1	0	0	2	2	2	0CS vs. all
0CO vs. IN	2	1	0	1	0	0	0	4	4	6	IN vs. CO
0IN vs. CO	1	0	0	0	0	0	0	1	1		
IN vs. CO	1	0	0	0	0	0	0	1	1		
0FG+BI vs. ST+CS	0	2	0	0	0	0	0	2	2	3	ST+CS vs. 0BI(+FG)
0BI vs. ST+CS	0	0	0	0	0	0	1	1	1		
0BI vs. all	1	0	0	0	0	0	0	1	1	6	Others
0BI+CU vs. all	0	0	0	0	0	1	0	1	1		
0CU vs. all	0	0	1	0	0	0	0	1	1		
0CU+CS vs. all	1	0	0	0	0	0	0	1	1		
ST+BI vs. 0all	0	0	1	0	0	0	0	1	1		
ST+GE vs. 0all	1	0	0	0	0	0	0	1	1		
TOTAL	42	22	9	4	5	3	11	96	100	100	

Table 3.2: UPGC African-derived lexical items classified according to phonetic similarity

Phonetics (PHON)	Origin							TOTAL	%	% MP	Main profiles (MP)		
	Md.	Wf.	Md./Wf.	Temne	Other Atlantic	Bantu	Unidentified						
0 [=NSD]	10	4	1	4	1	2	3	25	31	31	0		
IN vs. CO	7	2	2	0	1	0	2	14	17	26	IN/ST vs. CO		
ST vs. CO	1	3	1	0	0	0	2	7	9				
unclear	1	2	2	0	0	0	1	6	7	7	unclear		
IN+CS vs. all	3	1	0	0	0	0	1	5	6	15	CS (+IN/ST/CU) vs. all		
ST+CS vs. all	0	0	1	0	0	0	0	1	1				
CU+CS vs. all	3	0	0	0	0	0	0	3	4				
CS vs. all	3	0	0	0	0	0	0	3	4				
FG vs. all	1	0	0	0	0	0	1	2	3	3	FG vs. all		
ST vs. FG	0	1	0	0	0	0	0	1	1	18	Others		
IN(+BI) vs. all(+BI)	1	1	0	0	0	0	0	2	3				
ST vs. CS	0	2	0	0	0	0	0	2	3				
ST+CU vs. all	1	0	0	0	0	0	0	1	1				
ST+CS(+CU) vs. all(+CU)	1	0	0	0	0	0	0	1	1				
ST(+FG) vs. CO(+FG)	1	0	0	0	0	0	0	1	1				
ST(+BI) vs. CU+CS(+BI)	0	0	1	0	0	0	0	1	1				
ST vs. BI	0	0	1	0	0	0	0	1	1				
ST vs all	0	0	0	0	1	0	0	1	1				
IN+CS(+BI) vs. GE+CU(+BI)	0	0	0	0	0	0	1	1	1				
CU(+ST) vs. FG+BI(+ST)	0	1	0	0	0	0	0	1	1				
CS+GE(+ST) vs. BI+CU(+ST)	1	0	0	0	0	0	0	1	1				
BI+GE vs. all	1	0	0	0	0	0	0	1	1				
BI+CU vs. all	0	0	0	0	1	0	0	1	1				
TOTAL	35	17	9	4	4	2	11	82	100			100	

The common African lexical core of Upper Guinea Creoles and its historical significance

Table 3.3: UPGC African-derived lexical items classified according to semantic similarity

Semantics (SEM)	Origin							TOTAL	%	% MP	Main profiles (MP)
	Md.	Wf.	Md./Wf.	Temne	Other Atlantic	Bantu	Unidentified				
0 [=NSD]	23	11	8	3	2	2	6	55	67	67	0
IN vs. CO	4	2	0	1	1	0	2	10	12	20	IN/ST vs. CO
ST vs. CO	3	1	0	0	1	0	1	6	8		
unclear	3	0	0	0	0	0	0	3	4	4	unclear
FG vs. all	1	0	0	0	0	0	1	2	3	3	FG vs. all
CU+CS vs. ST + BI vs. FG	0	0	0	0	0	0	1	1	1	6	others
IN(+BI) vs. all(+BI)	1	0	0	0	0	0	0	1	1		
ST vs. BI	0	0	1	0	0	0	0	1	1		
ST vs. CS	0	1	0	0	0	0	0	1	1		
BI vs. all	0	1	0	0	0	0	0	1	1		
BI+CS vs. all	0	1	0	0	0	0	0	1	1		
TOTAL	35	17	9	4	4	2	11	82	100		