

The Gulf of Guinea creoles: A case-study of syntactic reconstruction

Tjerk Hagemeijer¹

Universidade de Lisboa

This paper argues that creole languages do not face some of the typical problems that have been discussed with respect to syntactic reconstruction of older languages. Creoles often belong to young language families and are therefore expected to show a significant amount of syntactic identity among sister languages. Other factors, such as their isolating typology and geographical isolation, may be additional advantages in the success of syntactic reconstruction. This hypothesis is tested on the four Portuguese-related Gulf of Guinea creoles, where a high degree of identity and the use of other processes, such as directionality, prove to provide good insights into the syntactic features of the proto-language.

Keywords: Gulf of Guinea creoles, creolization, reconstruction, syntax, directionality

1. Introduction

The Comparative Method (CM) has a long-standing tradition in the classification of languages as genetic units and in the reconstruction of proto-languages. It was designed to identify regular sound changes between cognates in sister languages. Since it is often a short leap from phonology to morphology, the CM is frequently extended to the reconstruction of morphological properties and functional material. Taking reconstruction in the realm of the CM at the level of syntax, on the other hand, is *a priori* more problematic because we do not think of correspondences between sentences, i.e. cognate sentences, in the same sense we think of lexical and sound correspondences. This paper does not engage so much in this epistemological debate itself, but means to present what the contribution of creole languages may be. It will be argued that these languages are

¹ This work was supported by the Portuguese Foundation of Science and Technology (FCT) as part of the project *The origins and development of creole societies in the Gulf of Guinea: An interdisciplinary study* (PTDC/CLE-LIN/111494/2009). Unlike indicated otherwise, the data from the GGCs were drawn from the corpora of these languages that are being build within the project (e.g. Hagemeijer *et al.* 2014). I'm grateful to Hugo Cardoso, Matthias Jenny, Philippe Maurer, and Carlotta Viti for their precious comments on this paper and to Armando Zamora for generously helping me with some of the Fa d'Ambô data.

interesting for syntactic reconstruction due to the particular sociohistorical context in which they arose and due to their generally isolating typology.

The focus will be on the four Gulf of Guinea creoles (GGCs) in West Africa, which arguably branched from a single proto-language that started arising on the island of S. Tomé at the end of the 15th century as the result of contact between Portuguese, the lexifier language, and continental African languages from the Benue-Congo family (e.g. Ferraz 1979; Hagemeyer 2011). The contemporary GGCs share a great amount of lexical and linguistic features, but largely lack mutual intelligibility. The oldest written records of these creoles date back to the late 19th century and in-depth and well-informed descriptions of these languages only became available from second half of the 20th century on, which means that any attempt to reconstruct their linguistic history must be undertaken mainly from the contemporary languages.

Section 2 briefly addresses the classificatory problem creoles pose for historical linguistics and is followed by a short introduction to the GGCs in section 3. The next two sections are concerned with syntactic reconstruction. Section 4, in particular, focuses on the place of creoles in the debate on syntactic reconstruction. It will be argued that creoles, as young languages, do not face many of the challenges that come with increasing time-depth. The main section of this paper, section 5, provides a case-study of the GGCs, showing that many syntactic features can be comfortably reconstructed on the basis of identity or by taking into account directional processes or other types of information.

2. The classification of creole languages

For a few decades now, especially since the work of Bickerton, there has been a fertile debate on how creole languages come about and how they relate to each other, to the lexifier and to the putative substrate languages. The idea that all creoles may have a common genetic origin as expressed in the monogenesis theory is of course long gone, but in more recent times several attempts have been made to group these languages together as a typological class. Bickerton (1981, 1984), for instance, claimed that creoles

share many linguistic properties because in context of radical language change children use their universal, biological language devices to transform pidgins into native languages. While many of Bickerton's claims were shown to be problematic,² the search for a specific creole typology has continued, but focusing in particular on shared linguistic properties related to (lower) creole complexity (e.g. McWhorter 2001; Bakker *et al.* 2011). In sum, for these scholars the claim that creoles form a typological class boils down to their alleged pidgin past.

The exact place of creoles in language typology is also connected to their problematic classificatory status. Thomason & Kaufman (1988), for instance, claim that the sharp break in transmission that characterizes some creoles (including the GGCs) implies that do not have a place in the Stammbaum tradition. This tradition has perhaps influenced the way creoles are organized in places such as Ethnologue (Lewis 2009), where creoles are lumped together as if they were one big language family whose internal branching is based on the lexifier language (Portuguese, English, French, etc.). Paradoxically, they are not embedded in the larger family where their lexifier belongs to, which would be Indo-European in most cases. This classification thus reflects the view that creole languages somehow represent a special – and controversial – class of languages.

Of course there is also the view that creoles are not the result of a break in transmission and thus more closely related to their lexifier (e.g. Chaudenson 1992, Mufwene 2001). DeGraff (2009) points out that, strictly speaking, creole languages should have a place somewhere in the language tree of their lexifier language, since the main portion of their lexicon is regularly derived from the lexifier. Creolization and Romanization are not fundamentally different for this author. Even though the transition from (classic) Latin to the Romance languages is generally considered a process of regular transmission over time and creolization an abrupt type of contact leading to the formation of new languages, both cases rely on Second Language Acquisition and

² See, for instance, Veenstra (2008), for an overview and discussion of Bickerton's claims with respect to creolization.

significant restructuring.³ Here I concur with DeGraff's (2009) assumption that creoles are genetically related to their lexifier and that the outcomes are expected in the light of theories of language contact and SLA. Historical, linguistic, and social facts will then explain why some creoles are claimed to be closer or less close to the lexifier. The fact that creoles are generally thought of as being at some typological distance from their lexifiers is not a sufficient argument to automatically rule them out from the lineage of their lexifier. In many genetic units beyond creoles, quite different typologies (e.g. analytic and synthetic) co-exist peacefully. From this perspective, the Portuguese-related creoles can be said to form relatively young genetic units, such as the Upper Guinea creoles and the Gulf of Guinea creoles, which branched from Portuguese, their lexifier, in the 15th or 16th century. Despite the fact that both these units are ultimately connected through the Portuguese lexicon, they constitute distinct, independently born branches with many distinct linguistic features (Ferraz 1987).

3. The Gulf of Guinea creoles

The GGCs form a cluster of four languages spoken on three West-African islands: Santome (ST) and Angolar (ANG), spoken on the island of S. Tomé, Principense (PR) on the island of Príncipe, and Fa d'Ambô (FA) on the island of Annobón.⁴ Although the mutual intelligibility between these languages is limited, it is usually assumed that they are genetically related. The only scholar who expressed a slightly different view is Ferraz (1987), a pioneer of the studies of the GGCs, who claims that the similarities between the GGCs can be explained by the introduction of slaves through the central administration on S. Tomé (the main island) and not as the result of a branching proto-language. In-

³ Recent work has shown that nativization of a creole language requires at least two generations (Roberts 2000) but under certain historical conditions this process may arguably take longer (Arends 1995). On the other hand, language change related to Romanization might have been less gradual than is generally proposed (Adams 2007).

⁴ I will use the language names that are used in Michaelis *et al.* (2013).

depth studies and new data produced in the last two decades, however, do not speak in favor of this scenario.⁵

According to Ferraz (1979), ST, the continuation in time and space of the putative proto-language, shares 67% of its lexicon ANG, 82% with FA, and 88% with PR. The main portion of the lexicon⁶ is derived from Portuguese but the most significant lexical differences between these languages are found in the much smaller, African portion of the lexicon: PR shows stronger impact from Niger-Delta languages, in particular from the Edoid family (Maurer 2009), whereas ANG exhibits a unusual high portion of Western Bantu lexicon, in particular from Kimbundu (Lorenzino 1998; Maurer 1992, 1995). ST and FA show a relatively proportional mix of substrate-derived lexical items from Edo and Kikongo (Ferraz 1979; Granda 1985). From a typological perspective, the GGC are strongly isolating languages with a very substantial number of substratal features in different domains of their grammar, such as labial-velars (PR), implosives, prenasalization, extensive reduplication, verb serialization, discontinuous or final sentence negation, word order (N-POS, N-DEM, ADJ-ADV, etc.) and ideophones.

Ferraz (1979) argued that the African features in ST (and thus by analogy in the other GGCs as well) are halfway between Nigerian and Bantu typology. In Hagemejjer (2011) I argued that many lexical and grammatical features shared by the GGCs are typically related to the Nigerian typology, leading to the interpretation that the new contact language that arose on the island of S. Tomé resulted mainly from the contact between Portuguese and languages from this area.⁷ In this scenario, the contribution of Western Bantu languages is one of secondary contact that set in at a point of time when the proto-language was already gaining shape. Historical work on the slave trade to S. Tomé shows that during the first decades after the permanent settlement of S. Tomé, in

⁵ See Hagemejjer (2011) for discussion.

⁶ Although the percentages of Portuguese and African lexicon in the GGCs have not been studied and compared beyond the basic lexicon, it can be estimated that the Portuguese lexicon in each creole represents between 80 and 90% in ANG (Lorenzino 1998) and over 90% in the remaining three GGCs.

⁷ The lexicon is a good example. African-derived cognates in the basic vocabulary of the four GGCs are Edo(id)-related. In fact, ANG is the only GGC that exhibits Bantu-derived basic vocabulary. In addition, several shared function words in the GGCs are from Edo(id) but typically not from Bantu.

1493, slaves were in fact almost exclusively taken from the Niger delta area (e.g. Caldeira 2008; Thornton 1992; Vogt 1973) which emphasizes the linguistic founder effect of this area in the making of a new language. The import of slaves speaking western Bantu languages only became significant and quickly dominant when the sugar cycle started blossoming, some two decades after the permanent settlement of S. Tomé.

4. Creoles and syntactic reconstruction

Reconstruction of phonology and morphology within the realm of the CM has a long-standing tradition and has widespread acceptability. Cognacy can be established on the basis of regular correspondences between sounds and therefore warrants the establishment of proto-sound and, by extension, proto-morphemes and proto-words. Syntactic reconstruction within the CM, on the other hand, has failed to reach a consensus among historical linguists, in particular because it is not clear what the notion of correspondences would be in syntax and how directionality operates (e.g. Lightfoot 2002). Campbell & Harris (2002) and Harris & Campbell (1995) are more optimistic about syntactic reconstruction but emphasize that cognate sentences are not direct descendents from some shared sentence but cases of pattern sharing through the proto-language. Walkden (2013:109) reviews the different positions and tries to reach a compromise between them by arguing for some degree of isomorphism between phonology and syntax in the following way: “(...) whereas in phonology we might reconstruct the lower level unit, sounds, through their context of appearance in lexical items attested in the daughter languages, in syntax we might reconstruct the lower level unit, lexical items, through their context of appearance in sentences attested in the daughter languages.” According to Walkden (2013), much then boils down to the plausibility of syntactic reconstructions. In what follows it will be argued that creoles offer a good working place for syntactic reconstruction which has not been sufficiently explored.

Despite the interest of creole languages for syntactic reconstruction, the intense debates on their typological classification (e.g. Bakker *et al.* 2011; Bickerton 1984;

McWhorter 2005) and their importance for the field of language acquisition and creation (e.g. DeGraff 1999; Plag 2008, Siegel 2008) have somewhat overshadowed the study of genetic relations between certain creole languages. Nevertheless, the assumption that creole languages form diffused genetic units has a long-standing tradition⁸ and there are quite a few studies that investigate the relation between clusters of creole languages. With respect to the Portuguese-related creoles, recent comparative work by Jacobs (2009, 2011), for instance, renews the proposal to include Papiamentu in the Portuguese-related Upper Guinea creoles (the varieties of Capeverdean, Guinea-Bissau and Casamance creole), a connection that had already been hinted at in earlier work (e.g. Martinus 1996; Quint 2000), but not worked out in a more detailed way.⁹ Apart from the need to obtain a better understanding how creoles within specific units relate to each other, creoles may provide a good testing ground for syntactic reconstruction due to their historical and linguistic specificity.

In the first place, creoles are young languages and thus constitute young language families, irrespective of the more general classification problem. The languages we call creoles are at most five centuries old (the Portuguese-related creoles being the oldest), and often much younger (e.g. Hawaiian Creole came about in the 19th century). Therefore, the birth of the proto-language and subsequent branching, whenever relevant, have a shallow time-depth. Moreover, in the case of many creoles insularity has contributed to geographic isolation. Of course not all creoles are spoken on islands and they have often been in contact with either the lexifier (particularly among exogenous creoles in the Atlantic world) or substrate languages (particularly among endogenous creoles in Asia, but also in the case of Guinea-Bissau and Casamance creole in West-Africa). If there is substantial evidence that creoles are part of a larger genetic unit (e.g. GGCs, Suriname creoles), shared features are likely to represent inherited material. Finally, creoles are typically isolating languages with no inflectional morphology, no overt case marking and,

⁸ In the case of the GGCs, for instance, historian Matos (1842:107) writes that ST, FA and PR are very similar languages (dialects in the original). This claim was made decades before the publication of the first descriptions and data of these languages.

⁹ This is not a case closed, however. Maurer (2013), for instance, highlights a number of features that require a different explanation, possibly involving the GGCs.

therefore, are likely to exhibit more rigid syntax.¹⁰ In other words, word order patterns are expected to be more stable over time under these conditions.

Altogether, it is thus expected that creoles are likely to exhibit a high degree of diachronic stability and substantial identity between sister languages within well identified genetic units. Creoles thus meet a number of criteria that, at first glance, favor successful reconstruction of their proto-language. Although the debate on syntactic reconstruction generally focuses on features of older genetic units (e.g. Indo-European word order patterns), it is frequently mentioned – even among the most skeptic – that full identity between sister languages and less time-depth are success factors in doing reconstruction (e.g. Lightfoot 2002; Pires & Thomason 2008; Harris 2008)

5. The GGCs and syntactic reconstruction

The purpose of this section is to investigate several syntactic constructions in the GGCs in order to show that syntactic comparison can be successfully used to establish or reinforce the genetic connection between these languages and to reconstruct patterns in the proto-language. It is expected that cases of identity will lead us directly to the proto-structure, whereas cases of non-identity allow reconstruction by taking into account directional processes and other types of knowledge about the GGCs, such as language-internal variation.

5.1. NP locative interrogatives

Standard Wh-interrogatives in the GGCs are typically fronted to the clause or sentence-initial position, but there is one interrogative construction in these languages that fails to exhibit this property, NP locative interrogatives (*bô* or *ba*). In these constructions only

¹⁰ Note that this is of course not an absolute statement: several Indo-Portuguese creoles, such as Diu (Cardoso 2009) and Korlai (Clements 1996), exhibit some inflectional morphology and case-marking related to syntactic and semantic properties of arguments.

NPs can be questioned and the interrogative word obligatorily occurs to their right, in final position. This is illustrated in the following examples.

- (1) *Lomba ê, moto bô bô?* (ST)
 Lomba VOC motorbike POS.2SG INT
 Hey Lomba, where is your motorbike?’
- (2) *Kasô tê ba?* (PR, Maurer 2009:149)
 dog POS.2SG INT
 ‘Where is your dog?’
- (3) *Bo bô?* (FA, Zamora 2010:94)
 2SG INT
 ‘Where are you?’
- (4) *Lêlu ô bô?* (ANG, Maurer 1995:139)
 Money POS.2SG INT
 ‘Where is your money?’

The syntax and semantics of this interrogative pattern, as well as the phonetic shape of the interrogative element¹¹ can therefore be safely reconstructed to the proto-GGC. The origin of this pattern and the interrogative marker itself are arguably derived from Edoid language Edo, where *vbòó*, [βòó], exhibits identical properties, as illustrated in (5).

- (5) *Rèn vbòó?* (Edo, Melzian 1937: 218)
 3SG INT
 ‘Where is s/he?’

Since lexical idiosyncrasies, such as irregular degrees of comparison (‘good – better – best’), are often taken as the type of evidence that reinforces a genetic connection between languages, syntactic idiosyncrasies, such as the one described here in the domain of interrogation, contribute to the establishment of a historical relation.

¹¹ The phonetic form *ba* in PR results from a sandhi rule whereby *bô* contracted with final interrogative particle *a*.

5.2. Body-reflexives

Yet another crystal-clear case of identity are the so-called body-reflexives, which occur with a number of verbs in the GGCs.

- (6) Ê mata **ubwê** dê. (ST)
(7) Ê mata **igbê** sê. (PR, Maurer 2009: 152)
(8) Ê mata **ôngê** rê. (ANG, Maurer 1995: 145)
(9) Ê mata **ôgê** dêli. (FA, Zamora [p.c.])
 3SG kill body POS.3SG
 ‘S/he committed suicide.’

Once again, these structures strongly resemble the syntax and semantics of a language such as Edo, although body-reflexives of this type also occur in many other Benue-Congo languages, but crucially not in Bantu (e.g. Parkvall 2000).

- (10) Ò gbé-ègbé èré ruà. (Edo, Hagemeyer 2011: 129)
 3SG kill -body POS.3SG PRT
 ‘S/he committed suicide.’

Note that this reflexive strategy is also found in other creole languages with the difference that, in these cases, the body(part) item is typically derived from the lexifier language and not from substrate languages (e.g. Heine 2005; Muysken & Smith 1994).

On a side note, the examples above show that the body-reflexive in PR is closest to the proto-item due to the survival of labial-velar / \widehat{gb} /, which is found almost exclusively in languages from the Sudanic belt (e.g. Güldemann 2008).¹² PR voiced labial-velar / \widehat{gb} / and its unvoiced pair / \widehat{kp} / are absent from the other three GGCs, which do often exhibit cognate items but with other solutions, such as labialized velars (ex. PR

¹² Although the Sudanic stretches over a large area of West and Central Africa, including families such as Kwa, Yoruboid or Edoid, it does not include narrow Bantu languages.

ukpaku vs. ST *kwaku* ‘traditional toothbrush’ or PR *ikpe* vs. ST *ukwê* ‘seed, grain’). Due to the marked status of labial-velars crosslinguistically, I assume that directionality operated from labial-velars to labializations and not the other way around. Therefore it can be concluded that labial-velars constitute a relic in PR, whereas the other GGCs innovated.


5.3. Serial verb constructions

Research on the GGCs in the past two decades has shown that serial verb constructions (SVCs) are a prominent feature of these languages (Post 1992; Maurer 1995, 1999, 2005, 2009; Hagemeijer 2001; Hagemeijer & Ogie 2011). On the African continent, SVCs constitute an areal feature whose core comprises the Kwa family and several Nigerian language families (e.g. Dimmendaal 2005). Although different definitions of SVCs are around, they should crucially present more than one verb in an asyndetic construction, exhibit a mono-eventive interpretation, bear at most one negation and tense marker with scope over the whole construction, and lack discourse pauses. Since it is beyond the scope of this paper to discuss all the types of SVCs, I have selected a few structures where the GGCs show identity.¹³ The following examples depict the locative serial, where the verb *pê* ‘to put’ combined with transitive verbs in the first position with a dynamic or motion feature (such as *saya*, etc. below) indicates the end point of the event.

- (11) Ê *saya kanwa pê matu.* (ST, Maurer 1999: 3)
 (12) Ê *saa kanwa pwê umatu.* (PR, Maurer 1999: 3)
 (13) Ê *thaa kanua pê matu.* (ANG, Maurer 1999: 6)
 (14) Ê *saa batelu pê matu.* (FA, Zamora [p.c.])
 3SG pull canoe put bushes
 ‘S/he pulled the canoe to the bushes.’

¹³ Not all SVCs in the GGCs show identity, however, and more comparative research is required in this domain.

Directional SVCs are another highly productive type in the GGCs and involve verbs of movement, combining manner of motion and directed motion. The verb corresponding to ‘to go’, for instance, can be used as a main verb, as an auxiliary and, in SVCs, to indicate the goal of movement (15-17) or as a purposive connector introducing nonfinite clauses (18-20).

- (15) *So ê kôlê ba palaxu ê.* (ST)
 then 3SG run go palace PRT
 ‘Then he ran to the palace.’
- (16) *S’ê khôlê ba Palea.*¹⁴ (FA)
 then-3SG run go Palea
 ‘Then he run to Palea.’
- (17) *Ê kwê we umatu.* (PR, Maurer 2009: 113)
 3SG run go forest
 ‘He ran into the forest.’
- (18) *Mese kôlê ba koplâ kwa.* (ST)
 master run go buy thing
 ‘The master ran to buy it.’
- (19) *M ba nha ba lël’ê.* (ANG, Maurer 1995: 107)
 1SG go there go accompany-3SG
 ‘I went down to pay him a visit.’
- (20) *S’ê khôlê ba tokha mina dêli.* (FA) 
 then-3SG run go touch child POS
 ‘Then he ran to touch his child.’

In Hagemejijer (2001), it was shown that in ST verbs in the second position, such as *ba* and *pê* above, are hybrid items sharing properties with both verbs and prepositions.¹⁵ The

¹⁴ Graphemes <kh> stands for velar fricative /x/.

following examples show a type of commitative SVC which expresses the idea of togetherness.

- (21) *Nen zunta kume lôsô.* (ST)
 3PL join eat rice
 ‘They ate rice together.’
- (22) *Inen zunta we posan.* (LU, Maurer 2009:118)
 3PL join go city.
 ‘They went to town together.’
- (23) *Inen zunta kumu.* (FA, Armando Zamora [p.c.])
 3PL join eat
 ‘They ate together.’

In some cases, a verb sequence has (partially) lexicalized, as in the following examples where the two verbs are highlighted.

- (24) *Ê pô sa mina kwali, a na ka **zug’e** **buta** fa.* (ST)
 3SG can be small basket IMP NEG TAM throw-3SG throw NEG
 ‘Even if it is a small basket, one doesn’t throw it away.’
- (25) *Ê **zuga** kwisê **bota.*** (PR, Maurer 2009: 113)
 3SG throw this throw
 ‘S/he threw this away.’

¹⁵ Example (i) shows property V2 shares with verbs, namely the possibility to take an aspect marker (*ka*) to derive iterative interpretations. Example (ii) shows a property shared with prepositions. Since ST is a non *pro*-drop language, the answer *pê lata* ‘in the can’, without an over subject, indicates a non-verbal property.

- (i) *Ê ka bloka awa ka pê lata.* (ST)
 3SG TAM poor water TAM put can
 ‘S/he always poor water in the can.’
- (ii) *Andji ku ê bloka awa pê? / Pê lata.* (ST)
 Where COMP 3SG poor water put / Put can
 ‘Where did s/he poor the water in? In the can.’

(26) *N dhuga ta pê awa.* (Maurer, 1995: 108)
 1SG throw throw put water
 ‘I threw it in the water.’

(27) *Ê zugwa ta pê a saaga.* (FA)
 3SG throw throw put water salty
 ‘He threw it in the salty water.’

In these cases, *zuga buta* (ST) and its equivalents in the other creoles are constructed from the Portuguese verbs ‘jogar’ and ‘botar’. This second verb, however, lacks verbal properties in these constructions. It cannot, for instance, select for an internal argument or be preceded by TMA markers, which can be seen in the ANG and FA examples in (26-27), where *pê* is required to Case-mark the internal argument.

A final example comes from the use of the lexical item *da*, whose primary meaning as a main verb is ‘to give’. This item, however, also occurs in a structure that is often classified as a SVC, the so called ‘give’-serials, which typically occur in serializing languages in West-Africa and Southeastern Asia (Lord 1993) and is also commonly found in Atlantic creoles. Consider the following examples from the GGCs:

(28) *Sama Dêsu da non ê.* (ST)
 call God give 1PL PRT
 ‘Call God for us.’

(29) *Ê kyê na ubanku da usan.* (PR, Maurer 2009: 109)
 3SG fall LOC chair give ground
 ‘He fell from the chair to the ground.’

(30) *No tega kikiê ra pato.* (ANG, Maurer 2009: 111)
 1PL hand.over fish give boss
 ‘We handed the fish over to the boss.’

(31) *Andêlê kumpa an lavulu da’l.* (FA, Zamora 2010: 331)
 Andrés buy a book give-3SG
 ‘Andrés bought a book for him/her.’

In these constructions, *da* takes on a prepositional meaning (‘for, to’) introducing typically benefactive, recipient and goal arguments and lacks verbal properties. In other words, *da* is phonetically derived from a verb but, unlike for instance *pê* and *ba*, shares its properties with prepositions.¹⁶ There is, however, one property that distinguishes *da* from other prepositions, such as *ku* ‘with’ or *di* ‘of’. When the argument introduced by this item is extracted, *da* is stranded, whereas prepositions are stranded with an invariable trace corresponding to the 3sg pronoun, as follows from the contrast between the Wh-interrogatives in (32) and (33).

(32) *Kê ngê ku ê tlaba da?* (ST)

what person COMP 3SG work give

‘Who did he work for?’

(33) *Kê ngê ku ê tlaba {ku ê/*ku}?* (ST)

what person COMP 3SG work with/with-3SG

‘Who did he work with?’

The other GGCs also exhibit this specific property, as illustrated by the following relative clauses in PR and ANG.

(34) *Txi sêbê ningê ki mosu sê sa xivi da a?*

2SG know person REL man DET TAM work give PRT

‘Do you know the person this young man is working for?’ (PR, Maurer 2009: 52)

(35) *ome si ma n ga taba da* (ANG, Maurer 1995: 56)

man DEM REL 1SG TAM work give

¹⁶ Here I concur with Bruyn (1995), who argues that creoles may exhibit structures that were taken directly from the substrate languages and not necessarily the result of a process of language-internal grammaticalization. This hypothesis seems to be very effective for the discussion on *da* in serial-like constructions. Lord (1993) provides crosslinguistic evidence of similar constructions where a verb lost part or all of its verbal features and is better dealt with as a transitive preposition. One of the main substrate languages, Edo, is one such case (e.g. Agheyisi 1990).

‘The man I work for’

In this section it was shown that the GGCs frequently show identity not only in the typology of verb serialization but also with respect to fine-grained features. In the remaining sections the discussion focuses on cases where the four GGCs lack identity.

5.4. Sentence negation

The GGCs are well known for their typologically marked negation patterns. ST, ANG, and FA exhibit a discontinuous sentence negation pattern, whereas PR exhibits canonical final negation (Maurer 1995, 2009; Post 1997; Hagemeyer 2007, 2009b; Zamora 2010). These patterns are illustrated in examples (36-39).

(36) *A na kuvida non fa.* (ST; Ferraz 1979: 68)

IMP NEG invite 1PL NEG

‘They didn’t invite us.’

(37) *N sa podi da txi fa.* (PR; Günther 1973: 78)

1SG TAM can give 2SG NEG

‘I can’t give it to you.’

(38) *Bô pô na f'e wa.* (ANG; Maurer 1995: 131)

2SG can NEG do-3SG NEG

‘You may not do it.’

(39) *Amu na pô fe f.* (FA; Post 1997: 303)

1SG NEG can do NEG

‘I can’t do it.’

The standard negation patterns in the four GGCs are summed up in Table 1 below.

Table 1. Default negation patterns in the GGCs.

	Santome	Principense	Angolar	Fa d’ Ambô
Default sentence	na...fa	...fa	(a~na)...wa ~	na...f

negation			va	
----------	--	--	----	--

According to Maurer (1995:131), the preverbal negation marker in ANG is generally optional. Despite this optional status in ANG and the absence of the preverbal marker in PR, there is strong evidence that a discontinuous negation can be reconstructed for all the GGCs, since certain specific environments, such as purpose clauses, require the preverbal marker (*na*) only.

(40) *Soku n sa tlaxi san ku sustu pa san na da mu sôtxi.*
 so 1SG be behind lady with fear PURP lady NEG give me whip
 ‘So I stood frightened behind the lady so she wouldn’t hit me.’ (ST)

(41) *Mene bê kukunu kôndê pa uwê jingantxi na vê li.*
 Mene also crouch hide PURP eye ogre NEG see 3SG
 ‘Mene crouched and hid so that the ogre’s eyes would not see him.’
 (PR, Maurer 2009:133)

(42) *Ê lôkê kwa bisi rê pa ê na nana.*
 3SG clean.up thing dress POS.3SG PURP 3SG NEG spoil
 ‘He cleaned up his clothes so they wouldn’t spoil.’ (ANG; Maurer 1995: 132)

(43) *Osyi ku eli pe olemu sa pa batelu na fo buka.*
 when that 3SG put paddle be for canoe NEG go turn
 ‘When he puts the paddle, it is to avoid that the canoe turns.’ (FA; Post 1997: 308)

In Hagemejijer (2007) I claimed that a discontinuous negation pattern (*na...fa*) should therefore be reconstructed for the four GGCs and that this pattern existed prior to diffusion in time and space of the proto-language. Gûldemann & Hagemejijer (2006) discuss how a discontinuous pattern may have developed in the proto-GGC using both linguistic material from Portuguese and African substrate languages. Directionality is also relevant to our understanding of sentence negation in the GGCs (Hagemejijer 2007). Jespersen’s cycle arguably underlies the earlier stages of the formation of the proto-language: Initially, a final, intensifying element (*f(a)/va~wa*) became associated with an unmarked preverbal negation (*na*) pattern. Upon diffusion of the proto-language and

subsequent internal development, the cycle was taken a step further in ANG, where the preverbal negation marker became optional, and even further in PR, where *na* only survived in highly marked environments, in particular purpose clauses. Standard negation in PR thus reflects the most innovative state of affairs, whereas ST and FA appear to be more conservative. In sum, while the negation patterns in the contemporary GGCs do not show identity, they derive from a common, discontinuous proto-pattern. Directionality is responsible for the innovating pattern that can be observed in ANG and PR.

5.5. *The Noun Phrase*

As in other domains of their grammar we have seen so far, the GGCs are also relatively homogenous with respect to word order in the DP, i.e. the relation between modifiers and the head noun. In all the GGCs, demonstratives, possessives, adjectives, relative clauses, and genitives occur to the right of the head noun in all GGCs, as shown in (44-45).

(44) *mina se mu glavi ku bi nai* (ST)
 child DEM POS.1SG beautiful REL come here
 ‘the/that beautiful child of mine who came here’

(45) *ke pe dê*
 house father POS.3SG
 ‘his father’s house’

Several other modifiers, such as indefinite singular (46) and plural marker, as well as diminutive and augmentative modification (46-47), on the other hand, typically occur to the left of the head noun.

(46) *ũa n'na parô* (ANG, Maurer 1995: 40)
 ART DIM basket
 ‘a small basket’

(47) *memen pe vapô* (ST)
 AUG AUG ship

‘a huge ship’

Despite the many similarities between NP word order patterns in the four GGCs, some differences or specific patterns can be found at a more fine-grained level. Three of these cases, namely the syntax of demonstratives, numerals and adjectives, will be addressed below.

5.5.1. Demonstratives

Here I will focus on the more fine-grained syntax of the (proximal) demonstrative modifier *se* in the examples above and its counterparts in the other three GGCs. I will use the label ‘demonstrative’, being aware that other labels have been used for these items in the GGCs, such as determinant or specific marker (Alexandre & Hagemeyer 2007). The following sentences exemplify the demonstrative in the remaining GGCs.

(48) *Ê kume inhemî sê pe.* (PR, Maurer 2009: 34)

3SG eat yam DEM completely

‘He ate the whole yam.’

(49) *Tuu ane ale e vitxa nha.* (ANG, Maurer 1995: 41)

all PL king DEM arrive there

‘All the kings arrived there.’

(50) *Ta ku nan pe nen se sa kha bay,...* (FA)

when COMP PL man PL DEM be TAM go

‘When these men left,...’

In ST, *se* can be reduced to *e* in certain contexts (*kwa se* > *kw’e* ‘that, this/that thing’, lit. ‘thing DEM’). In ANG *e* varies with *dhe* and very scarcely with *the*.¹⁷ In FA, *se* is found in complementary distribution with the form *say*, as illustrated in (51).¹⁸

¹⁷ Graphemes <dh> and <th> stand for voiced and unvoiced interdental sibilants respectively.

¹⁸ There are two possible sources for these demonstratives:

- (51) *S'ê ten ba khay say.* (FA)
 Then-3SG also go house DEM
 'Then he also went to that place.'

Although additional research on long and short forms in FA is necessary, the long forms occur in contexts where there is a discourse break, such as a full stop. Note that long and short forms also affect other word categories, including nouns and verbs. Otherwise the short forms occur. Moreover, FA is the only GGC where the demonstrative can be overtly pluralized by a postnominal plural marker, *nen* in (50) above.

In Alexandre & Hagemeyer (2007), it is shown that demonstrative determiner *se* in ST has the properties of a clitic which occurs to the immediate right of nouns, as in (44) above. Other lexical material, such as possessives or adjectives, cannot intervene and have to occur to the right of *se*. The rigid syntax of the demonstrative in ST contrasts with the data from the other creoles. In ANG, *e/dhe* follows adjectives and may either follow or precede possessives.


- (52) a. *moto ngai e/dhe* (ANG, Maurer 1995: 43)
 motorbike big DEM
 'This big motorbike'
 b. **moto e~dhe ngai*

- (53) a. *n'na dhe m ...* (ANG, Maurer 1995: 45)
 child DEM POS.1SG
 'My child...'
 b. *N tua n'kila rê dhe, ...* (ibidem)
 1SG take tail POS DEM

These forms are derived from Portuguese demonstrative *esse* 'that' or *este* 'this'. In the latter case with the standard reduction of the cluster /st/ to /s/ (e.g. ST *fesa*<festa 'party', *sa*<estar 'to be', *mêê*<mister (old Ptg.) 'to want, to love, to need', etc.).

These forms are derived from periphrastic construction *sa ai* 'be here', following a sandhi rule whereby the Portuguese diphthong /aj/ becomes /ɛ/ (e.g. ST *pe*<pai 'father', *be*<vai 'to go').

‘I took his tail...’

PR and FA are different from ST in the sense that demonstrative may precede or follow simplex possessives but is also different from ANG because the demonstrative precedes adjectives. 

(54) a. *Ine ufaka tê sê pêdê.* (PR; Maurer 2009: 37)

PL knife POS.2SG DEM lose

‘These knives of yours got lost.’

b. **Ine ufaka sê tê pêdê.*

(55) *khadji mu se...* (FA, Zamora [p.c.])

house POS.1SG DEM

‘my house’ / ‘this house of mine...’

Maurer shows that the POS-DEM word order between modifiers in (54a) is already attested in the oldest known sketch of PR from the late 19th century. Note further that demonstratives in the GGCs may also follow clauses emphasizing the situation the verb refers to (Maurer 2009: 35). Although this property is not yet well understood, it appears to be particularly common in progressive-marked relative clauses (56-57) and temporal clauses (58).¹⁹

(56) *Tubulo ki thêka bi e.* (ANG, Maurer 1995: 42)

shark REL TAM come DEM

‘The shark that is coming.’

¹⁹ Note that temporal clauses may be related to relative clauses because the temporal conjunction is usually derived from a noun+complementizer construction: *ola ku* ‘when’ (lit. ‘(the) hour that’). In this respect, note that at least PR the demonstrative is also found in interrogative structures with a relative-like Wh (*ki* = COMP).

(i) *Kwa ki txi sa fêzê sê?* (Maurer 2009: 25)

thing COMP 2SG TAM do DEM

‘What are you doing?’

(57) [...] *a sa buka lavin ki sa ke sê*, ... (PR; Maurer 2009: 35)
 INDEF TAM search ship REL TAM go DEM

‘[...] they were looking for the boat that was passing by [...].’

(58) *Ola san xka pali se, san ka glita...* (ST)

when she TAM give.birth DEM, she TAM scream...

‘When she was giving birth, she screamed...’

In this subsection it was shown that the phonetic shape of the demonstratives under discussion can be readily assigned to a common source. Syntactically, these forms occur in postnominal position but their relative position to the noun is variable across the GGCs. The fact that these demonstratives also have eventive modification properties contributes to the idea that these demonstratives were originally free morphemes which could attach to different word categories or clausal structures. However, the facts show that in the domain of nominal modification they are clearly bound morphemes in ST (Alexandre & Hagemeyer 2007). In PR, their status is less bound but more restricted than in ANG, where they may even follow adjectives. In fact, it could well be that in PR, ST and FA demonstratives and possessives form clitic clusters that differ with respect to the ordering pattern of the clitics. I am therefore assuming that the demonstratives under discussion are showing effects of grammaticalization toward clitics: ANG exhibits the most conservative patterns whereas ST is further down the road of innovation.

5.5.2. Numerals

The GGCs do not exhibit homogeneous behaviour with respect to the position of numerals in the NP. According to Maurer (1995), cardinal numerals always precede the noun in ANG.²⁰ This is also the typical pattern in ST.

(59) *tano litu awa* (ANG, Maurer 1995: 47)

five liter water

‘five liters of water’

²⁰ Note that the ANG numerals from one to three have a Portuguese etymology, whereas those from four to teen are Bantu-derived.

(60) *Inen dôsu manu se sa jingantxi.* (ST)
PL two brother DEM be giant
'The two brothers in question are giants.'

While the prenominal position is by and large dominant in ST, postnominal numerals occur sporadically in spontaneous data and are considered grammatical by native speakers in elicitation tasks.

(61) *inen sun se dôsu...*
PL man DEM two
'these two men...'

In FA, except for 'one', which is typically prenominal, non-composed numerals are typically postnominal, although some variation can be observed, as illustrated in (62-63).

(62) *Lavulu dôs sa beetu.* (FA, Zamora 2010: 234)
book two be open
'The two books are open.'

(63) *dôs lavulu* (ibidem, 238)
'two books'

In contemporary PR, 'one' is always postnominal and all other numbers are normally prenominal, but Maurer (2009: 42) mentions that older speakers accept postnominal numerals in general. On the whole then, it can be concluded that prenominal numerals are dominant in the GGCs but ANG is the only GGC where the postnominal pattern has not been attested. Against the rule of majority, however, there are reasons to believe the pattern N-NUM should be reconstructed, on the assumption that the two creoles in the periphery, but especially FA, which was more strongly isolated from an early period on, exhibit an archaic pattern. Prenominal numerals would then represent an innovative pattern that spread on the island of S. Tomé, where ANG and ST are spoken. Historical language contact between these two languages, and between ST and PR, would then be at

the basis of this innovation. A reconstructed N-NUM pattern also fits the general tendency toward head-initial DPs in the GGCs. A similar split between ST and ANG, on the one hand, and PR and FA, on the other hand, can also be observed for the position of universal quantifier *tudu* ‘all’, which is prenominal in the former two and postnominal in the latter two creoles, and for a few other grammatical features (cf. Hagemeyer 2009a: 43). Since the two creoles in the periphery were the least likely to be in contact with each other, chances are that shared patterns in these two creoles may well be a glimpse into the proto-language.

5.5.3. Adjectives

While adjectives in the GGCs typically follow the noun (see for instance example (44) above), there are a very few exceptions to this rule, which come in two types: adjectives that are exclusively prenominal and adjectives that can occur both pre and postnominally. The items corresponding to the first type are *bon* ‘good’ (also *bwa* in ANG) and *ma* ‘bad’, which are derived from Portuguese *bom* (masc., sing.) and *má* (fem., sing.):

- (64) a. *bon/ma ngê* ‘good/bad person’ (ST)
 b. **ngê bon/ma*
- (65) a. *bon/ma kumê* ‘good/bad food’ (PR, adapted from Maurer 2009: 47)
 b. **kumê bôn/ma*.
- (66) a. *ũa bwa/bo ngê* ‘a good person’ (ANG, Maurer 1995: 50)
 b. *ũa ma ngê* ‘a bad person’ (ANG, Maurer 1995: 50)

One of the very few items that is able to occur in both positions is *ve* ‘old’, as illustrated in the examples below.

- (67) a. *panu ve* ‘old cloth’ (ST)
 b. *ve panu* ‘rag’
- (68) a. *ve jinku* ‘old zinc’ (PR, Maurer 2009: 47)
 b. *jinku ve* ‘old zinc’
- (69) a. *panu veyu* ‘old cloth’ (FA, Zamora 2010: 208)

b. *veyu panu* ‘rag’

While *bon* and *ma* above are productively used in prenominal position, *ve* is heavily restricted in its prenominal use, which explains why at least in the FA and ST example the translation ‘rag’ suggests that we are dealing with a lexicalized chunk. Differently from Portuguese, where a much larger array of adjectives may occur in prenominal position typically yielding specific semantic readings, prenominal adjectives in the GGCs are therefore truly exceptional in these languages and constitute the type of shared syntactic idiosyncrasy that must have spread from the proto-language.

6. *Final remarks*

In this preliminary case-study on the GGCs we argued that creole languages may constitute fertile ground with respect to syntactic reconstruction. The fact that we are dealing with young isolating language families which are often spoken in relatively isolated geographical contexts favors the success of syntactic reconstruction. These factors are expected to warrant a high degree of full structural identity between sister languages but also facilitates the insight into directional processes, with the additional advantage that it will be easier to tell apart inherited and borrowed features. Comparative research on the GGCs in particular shows that many shared syntactic properties must have been inherited and diffused from the proto-language, hereby providing not only important insights into language-internal development but also into the relevance of the different strata that contributed to the formation of the proto-language. From the shared common ground between the four GGCs we conclude that they did not undergo heavy structural changes in the course of their young history and that many shared features highlight the importance of the typology of the Niger delta substrate area in the formation of the proto-GGC. Reconstruction of specific syntactic features may also contribute to the debate on the typology of creole languages. It was shown that some typologically marked features, such as discontinuous sentence negation or labial-velars, must have spread from the proto-GGC, countering the idea that there is such a thing as an “unmarked” creole

prototype. Finally, future research in the field of creoles is also expected to shed light on issues related to the rate of change in the lexicon and in different domains of grammar (e.g. Longobardi & Guardiano 2009; Parkvall 2000). The ANG lexicon, for instance, is clearly further apart from the other GGCs but many syntactic features of ANG show stronger resemblance with ST than with the other two creoles spoken in the periphery. Altogether, however, this brief case-study showed that comparing the syntax of the GGCs reinforces their genetic relatedness as a complex network of innovative and conservative features and not as a monolithic family tree.

References

- Adams, James N. 2007. *The regional diversification of Latin, 200 BC–AD 600*. Cambridge, UK: Cambridge University Press.
- Agheyisi, Rebecca. 1990. *A grammar of Edo*. Unesco.
- Alexandre, Nélia & Hagemeijer, Tjerk. 2007. Bare nouns and the nominal domain in Santome. In Marlyse Baptista & Jacqueline Guéron (eds), *Noun phrases in creole languages: a multi-faceted approach*, 37-59. Amsterdam, Philadelphia: John Benjamins.
- Arends, Jacques. 1995. Demographic factors in the formation of Sranan. In Jacques Arends (ed.), *The early stages of creolization*, 203-232. Amsterdam, Philadelphia: John Benjamins.
- Bakker, Peter; Daval-Markussen, Aymeric; Parkvall, Mikael; Plag, Ingo. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1): 5-42.
- Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma.
- Bickerton, Derek. 1984. The Language Bioprogram Hypothesis. *Behavioral and Brain Sciences* 7, 173–231.
- Bruyn, Adrienne. 1995. *Grammaticalization in creoles: The development of determiners and relative clauses in Sranan*. Amsterdam: IFOTT.

- Caldeira, Arlindo. 2008. Tráfico de escravos e conflitualidade: O arquipélago de São Tome e Príncipe e o reino do Congo durante o século XVI. *Ciências & Letras* 44: 55-76.
- Campbell, Lyle. & Harris, Alice C. 2002. Syntactic reconstruction and demythologizing 'Myths and the prehistory of grammars'. *Journal of Linguistics* 38(3): 599-618.
- Cardoso, Hugo C. 2009. *The Indo-Portuguese language of Diu*. Utrecht: LOT.
- Chaudenson, Robert. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.
- Clements, Joseph C. 1996. *The genesis of a language*. Amsterdam, Philadelphia: John Benjamins.
- DeGraff, Michel. 1999. *Language creation and language change: Creolization, diachrony and development*. Cambridge, MA: The MIT Press.
- DeGraff, Michel. 2009. Language acquisition in creolization and, thus, language change: Some Cartesian-Uniformitarian boundary conditions. *Language and Linguistics Compass*, 3/4: 888-971.
- Dimmendaal, Gerrit. 2005. Areal diffusion versus genetic inheritance: An African perspective. In Alexandra Aikhenvald & R.M.W. Dixon (eds.), *Areal diffusion and genetic inheritance*. Oxford: Oxford University Press.
- Ferraz, Luiz I. 1979. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press.
- Ferraz, Luiz I. 1987. Portuguese creoles of West Africa and Asia. In Gilbert, Glen (ed.), *Pidgins and creole languages: essays in memory of John E. Reinecke*, 337-60. Honolulu: University of Hawaii Press.
- Granda, German de. 1985. El 'vocabulario fundamental' del criollo portugués de Annobón. Rasgos caracterizadores. In German de Granda, *Estudios de lingüística afro-románica*. 157-196. Valladolid: Universidad de Valladolid.
- Güldemann, Tom. 2008. The macro-Sudan belt: Towards identifying a linguistic area in northern sub-Saharan African. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151-185. Cambridge: Cambridge University Press.
- Güldemann, Tom & Hagemeyer, Tjerk. 2006. Negation in the Gulf of Guinea creoles: typological and historical perspectives, *Encontro Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola*, Coimbra.

- Günther, Wilfried. 1973. *Das Portugiesische Kreolisch der ilha do Príncipe*. Marburg an der Lahn: Marburger Studien zur Afrika- und Asienkunde.
- Hagemeijer, Tjerk. 2001. Semi-lexicality and underspecification in serial verb constructions. In Norbert Corver & Henk van Riemsdijk (eds.), *Semi-lexical categories*, 415-451. New York: Mouton de Gruyter.
- Hagemeijer, Tjerk. 2007. *Clause structure in Santome*. Doctoral Dissertation, Universidade de Lisboa.
- Hagemeijer, Tjerk. 2009a. Initial vowel agglutination in the Gulf of Guinea creoles. In Enoch Aboh & Norval Smith (eds.), *Complex processes in new languages*, 29-50. Amsterdam, Philadelphia: John Benjamins.
- Hagemeijer, Tjerk. 2009b. Aspects of discontinuous negation in Santome. In Norbert Cyffer, Erwin Ebermann, Georg Ziegelmeyer (eds.), *Negation patterns in West African languages and beyond*, 139-165. Amsterdam, Philadelphia: John Benjamins.
- Hagemeijer, Tjerk. 2011. The Gulf of Guinea creoles: genetic and typological relations. *Journal of Pidgin and Creole Languages* 26(1): 111-154.
- Hagemeijer, Tjerk; Génèreux, Michel; Hendrickx, Iris; Pereira, Amália; Tiny, Abigail; Zamora, Armando. 2014. The Gulf of Guinea creole corpora. *Proceedings of the 9th LREC conference*. [to be completed]
- Hagemeijer, Tjerk & Ogie, Ota. 2011. Edo influence on Santome: evidence from verb serialization and beyond. In Claire Lefebvre (ed.), *Creoles, their substrates, and language typology*, 37-60. Amsterdam, Philadelphia: John Benjamins.
- Harris, Alice C. 2008. Reconstruction in syntax. In Gisella Ferraresi & Maria Goldbach (eds.), *Principles of syntactic reconstruction*, 73-95. Amsterdam: John Benjamins.
- Harris, Alice C. & Campbell, Lyle. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Heine, Bernd. 2005. On reflexive forms in creoles. *Lingua* 115: 201-257.
- Jacobs, Bart. 2009. The Upper Guinea origins of Papiamentu. Linguistic and historical evidence. *Diachronica* 26(3): 319-379.
- Jacobs, Bart. 2011. The origins of Papiamentu: Linguistic and historical ties with Upper Guinea. Doctoral Dissertation, Ludwig-Maximilians-Universität.

- Lewis, M. Paul. 2009. *Ethnologue: Languages of the world, 16th edition*. Dallas: SIL International.
- Lightfoot, David W. 2002. Myths and the prehistory of grammars. *Journal of Linguistics* 38(1): 113-136.
- Longobardi, Giuseppe & Guardiano, Cristina. 2009. Evidence for syntax as a signal of historical relatedness. *Lingua* 119: 1679–1706.
- Lord, Carol. 1993. *Historical change in serial verb constructions*. Amsterdam, Philadelphia: John Benjamins.
- Lorenzino, Gerardo. 1998. The Angolar creole Portuguese of São Tomé: Its grammar and sociolinguistic history. City University of New York: Ph.D. Dissertation.
- Martinus, Efraim F. 1996. *The kiss of a slave: Papiamentu's West-African connections*. Ph.D. Dissertation, University of Amsterdam.
- Matos, Raimundo J. da Cunha. 1842. *Corographia histórica das ilhas de S. Thomé, Príncipe, Anno Bom e Fernando Pó*. Porto: Typographia da Revista.
- Maurer, Philippe. 1992. L'apport lexical bantou en Angolar. *Afrikanische Arbeitspapiere* 29. 163-174.
- Maurer, Philippe. 1995. L'Angolar: un créole afro-portugaise parlé à São Tomé. Hamburg: Helmut Buske Verlag.
- Maurer, Philippe. 1999. El verbo locativo *poner* en santomense, principense y angolar. In Klaus Zimmermann (ed.), *Lenguas criollas de base lexical española y portuguesa*, 89-100. Frankfurt/Madrid: Vervuert/ Iberoamericana.
- Maurer, Philippe. 2009. Principense — Grammar, texts, and vocabulary of the Afro-Portuguese creole of the island of Príncipe. London: Battlebridge Publications.
- Maurer, Philippe. 2013. Papiamentu. In Susanne Michaelis, Phillippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages, Vol. II. Portuguese-based, Spanish-based, and French-based languages*, 163-181. Oxford: Oxford University Press.
- Melzian, Hans Joachim. 1937. *A concise dictionary of the Bini language of southern Nigeria*. London: Kegan Paul, Trench Trubner & Co.

- Michaelis, Susanne; Maurer, Philippe; Haspelmath, Martin; Huber, Magnus (eds.), *The survey of pidgin and creole languages, Vol. II. Portuguese-based, Spanish-based, and French-based languages*. Oxford: Oxford University Press.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5(2-3): 125-166.
- McWhorter, John. 2005. *Defining creole*. Oxford: Oxford University Press.
- Mufwene, Salikoko. 2001. *The ecology of language evolution*. Cambridge. Cambridge University Press.
- Muysken, Pieter & Smith, Norval. 1994. Reflexives. In: Jacques Arends, Pieter Muysken & Norval Smith. (eds.), *Pidgins and creoles – An introduction*, 271-288. Amsterdam, Philadelphia: John Benjamins.
- Parkvall, Mikael. 2000. *Out of Africa*. London: Battlebridge Publications.
- Pires, Acrísio & Thomason, Sarah G. 2008. How much syntactic reconstruction is possible? In Gisella Ferraresi & Maria Goldbach (eds.), *Principles of syntactic reconstruction*, 27–72. Amsterdam: John Benjamins.
- Plag, Ingo. 2008. Creoles as interlanguages: Syntactic structures. *Journal of Pidgin and Creole Languages* 23: 307-328.
- Post, Marike. 1992. The serial verb construction in Fa d'Ambu. In Ernesto d'Andrade & Alain Kihm (eds.), *Actas do colóquio sobre crioulos de base lexical portuguesa*, 153-169. Lisbon: Colibri.
- Post, Marike. 1997. Negation in Fa d'Ambô. In Ruth Degenhardt, Thomas Stolz & Hella Ulferts (eds.), *Afrolusitanistik – eine vergessene Disziplin in Deutschland?*, 292-316. Bremen: Universität Bremen.
- Quint, Nicolas. 2000. Le Cap Verdien: Origines et devenir d'une langue metisse. Paris: L'Harmattan.
- Roberts, Sarah. 2000. Nativization and genesis of Hawaiian Creole. In John McWhorter (ed.), *Language change and language contact in pidgins and creoles*. Amsterdam, Philadelphia: John Benjamins.
- Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.

- Thomason, Sarah G. & Kaufman, Terrence. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thornton, John. 1992. *Africa and Africans in the making of the Atlantic world, 1400–1680*. Cambridge: Cambridge University Press.
- Veenstra, Tonjes. 2008. Pidgin/creole genesis: the impact of the language bioprogram hypothesis. In Silvia Kouwenberg & John Singler (eds.), *The handbook of pidgin and creole studies*, 219-241. London: Blackwell.
- Vogt, John. 1973. The early Sao Tome-Principe slave trade with Mina: 1500–1540. *International Journal of African Studies*, VI (3): 453–467.
- Walkden, George. 2013. The correspondence problem in syntactic reconstruction. *Diachronica* 30(1): 95-122.
- Zamora, Armando. 2010. *Gramática descriptiva del fa d'ambô*. Barcelona: CEIBA Ediciones.