Segmental Typology of African Creole Languages: Examining Uniformity, Simplification and Simplicity

Thomas B. Klein
Georgia Southern University

1. Introduction

Until very recently, phonology has been neglected as an area in the study of Creole languages (Bhatt & Plag 2006b, Plag 2003b). Thus, it is perhaps not surprising to note that typological work on Creole phonologies is exceedingly rare. Consequently, when phonological generalizations intended to hold across Creole languages are invoked in the literature, they tend not to be founded on a systematic sample, or the geographical focus centers either on Atlantic or Pacific creoles. African creoles in particular are rarely, if ever, considered as a group. Furthermore, only Creole languages with certain Indo-European lexifiers tend to be highlighted. This practice excludes a fair number of creoles on the African continent that have Arabic or certain Niger-Congo languages as their lexifiers. What is missing are phonological studies of the synchronic typology of Creole languages in general and of African creoles in particular in which all types of lexifiers are included and in which a systematic comparison to non-Creole languages is made.

Maddieson’s (1984) work stands out as a comprehensive and influential work on the typology of phonemic segments. It is based on the UCLA Phonological Segment Inventory Database (UPSID) of 317 languages, none of which are pidgins or creoles. This database lends itself nicely to a comparison with Creole languages, in part because it is designed and documented well and because it has been analyzed extensively from a typological perspective. Furthermore, this work offers a straightforward measure of structural complexity that may be applied to other types of languages including creoles.

There are a number of typological claims about Creole languages including that their linguistic structure is more alike or simpler than that of non-Creole languages. Influential concepts such as Bickerton’s bioprogram (see Bickerton 1981 et seq.) are based on the seeming surface similarity of Creole languages. Note, however, that the evidence in support of the bioprogram hypothesis was drawn from morphosyntax, not phonology. Substratist and superstratist thinking also relies to a good degree on the idea that a given set of structures is more or less iterated across Creole languages, thus making them appear more similar than the linguistic diversity in the formative contact situation would suggest. The idea that Creole languages are grammatically simplified in comparison to their source languages, especially the superstrates, is considered a truism by many, professional linguists and laypersons alike. Others have argued that the linguistic structure of Creole languages is simple in some absolute sense, not just as compared to the languages in the contact situation, but also in comparison to non-Creole languages in general. This idea of absolute simplicity has been advocated prominently in recent work. In particular, it has been claimed that the world’s simplest grammars are Creole grammars (McWhorter 2001a, b). According to this work, creoles display significantly less complexity as a group than the rest of the world's natural grammars. This entails that the linguistic structures of Creole languages should cluster with the structures of non-Creole languages that may be considered simple. Conversely, Creole languages should not cluster with middle range or complex patterns.

---

1This chapter has benefited from discussion with the participants of ACAL 36 in Savannah, in particular Mike Cahill, David Odden, and Philip Rudd, from the work with the editors Frank Arasanyin and Michael Pemberton, and from the comments of two anonymous reviewers. Responsibility for any errors lies solely with the author.
Creole and creoloid languages in Africa display a great diversity of geographic locations, substrate languages, and superstrate lexifiers. Hence, they offer an excellent testing ground for notions of structural uniformity, simplification and simplicity. They also vary widely in age and degree of creolization or restructuring. For example, the Gulf of Guinea creoles are considered prototypical, and they are among the oldest creoles on earth. Others such as Nigerian Pidgin have been creolized very recently. The status as full-fledged Creole languages has been debated for some varieties such as Réunionnais and Shaba Swahili. Such languages, however, are eligible for inclusion in the present work as creoloid languages.

The aims of the present chapter are to construct and analyze a representative typological sample of phoneme inventories from African creoles and creoloid languages and to compare this sample with the inventories of lexifiers and potential substrate languages. Using the measures in Maddieson (1984), the results of the analysis of phoneme inventory size, vowel qualities, and stop series are contextualized and interpreted in light of the themes of uniformity, simplification and simplicity.

2. A typological sample of African Creole languages

It is an explicit goal in constructing the present sample to represent the range of creoles in Africa, that is, insular and continental languages with European and non-European lexifiers. No more than two closely related languages are admitted to avoid over-representation of particular types of creoles. For instance, two Gulf of Guinea creoles are included in the sample to represent insular Portuguese-lexified Creole languages even though there are additional creoles with this lexifier in that particular geographical area. Acrolectal varieties are generally excluded from consideration because they are least distinct from the lexifier; consequently, the focus is on basilectal or mesolectal varieties. Marginal phonemes are generally excluded. For instance, whenever a given phoneme is described as occurring only in recent loans from the lexifier, it is not included in the inventory of that particular Creole language. Finally, there is a bias towards the quality of the available descriptions. Only languages for which sufficiently detailed information on the phonology is available have been considered for this sample.

According to the survey in Smith (1995), Africa is home to about twenty-two Creole languages. The current sample contains fourteen languages. Table 1 lays out how the creoles are classified according to geographic area and lexifier language and which sources have been used to gather the phonological information.

<table>
<thead>
<tr>
<th>IE-lexifier</th>
<th>Creoles</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nigerian Pidgin</td>
<td>Faraclas 1996</td>
</tr>
<tr>
<td>French</td>
<td>Mauritian</td>
<td>Baker 1972, Papen 1978</td>
</tr>
<tr>
<td></td>
<td>Seychellois</td>
<td>Bollée 1977, D’Offay &amp; Lionnet 1982, Papen 1978</td>
</tr>
<tr>
<td></td>
<td>Cape Verdean (Santiago)</td>
<td>Macedo 1989, Mendes et al. 2002, Quint 2000, Rougé 2004</td>
</tr>
<tr>
<td></td>
<td>Angolar</td>
<td>Lorenzino 1998, Maurer 1995</td>
</tr>
<tr>
<td></td>
<td>São Tomense (Foro)</td>
<td>Ferraz 1979, Schang 2003</td>
</tr>
<tr>
<td>Arabic</td>
<td>Juba</td>
<td>Kaye &amp; Tosco 2001, Owens 1997</td>
</tr>
<tr>
<td></td>
<td>Kinubi</td>
<td>Owens 1997, Pasch &amp; Thelwall 1987</td>
</tr>
</tbody>
</table>
English, French, and Portuguese are the central Indo-European lexifiers in the African context. Importantly, representatives of creoles with non-Indo-European lexifiers are also included in the sample.

3. Size and structure of phoneme inventories

The inventories of the Creole languages are measured against the inventories of the potential substrates and the non-Niger-Congo lexifiers to render the comparison meaningful. To this end, phoneme counts were extracted for the Niger-Congo (*Niger-Kordofanian* in Maddieson’s (1984) terminology; N=31), Nilo-Saharan (N=21), and Afro-Asiatic (N=18) languages represented via the phoneme charts in Maddieson (1984). Afro-Asiatic languages included in UPSID but not spoken on the African continent (Hebrew, Socotri, Neo-Aramaic) were not considered for this count. Khoisan languages were excluded because they are not potential substrates for any of the Creole languages in the present sample. Phoneme counts for the Indo-European lexifiers and Arabic were gathered from extant sources. This section displays the findings from the sample of African creoles as to total size of the phoneme inventory, the number of distinctive vowel qualities, and the number of stop series and contrasts these numbers with the results from the UPSID database and the counts for the lexifiers.

Maddieson’s (1984) conventions for phoneme counts were followed exactly. Marginal phonemes have been excluded throughout. Nasal vowels contribute towards the phoneme inventory size, but count for the vowel quality inventory only if the quality of a given nasal vowel is distinct from the oral vowels in the language. For example, the qualities of the Mauritian nasal vowels are added to the vowel quality inventory because they are distinct from the qualities of the oral vowels. The vowel qualities of nasal vowels are not distinct from those of the oral vowels in languages such as Angolar or Cape Verdean; consequently, they are not added to the vowel quality inventory count (see below). All phonemic nasal vowels, however, appear in the count for the phoneme inventory size. Diphthongs are not counted unless they feature vowel qualities that do not occur elsewhere. Long consonants or vowels are not counted unless the length feature appears as idiosyncratic to a segment or if certain vowel qualities appear only with long vowels. For example, geminates in Arabic are not counted for the inventory because this length is interpreted as a suprasegmental feature (Maddieson 1984: 162). On the other hand, long and short vowels are counted because the source on Arabic that Maddieson (1984) uses describes different qualities for long versus short vowels.

The smallest African inventory has twenty phonemes (Efik), whereas the largest has fifty-nine (Igbo). This range is considerable, but it is not as great as in the UPSID database as a whole. The smallest inventories there have eleven phonemes, whereas the largest has 141 phonemes. Maddieson (1984) employs a tripartite metric of simplicity/complexity. Simple inventories have fewer than twenty phonemes. Typical phoneme inventories contain between twenty and thirty-seven segments. Seventy percent of the total UPSID languages (c. 222 languages) fall within the typical range. Inventories with more than thirty-seven phonemes are classed as complex. No African languages in the current sample have simple inventories by this measure. Sixty African languages (86%) in this sample fall within the typical range, whereas ten (14%) have thirty-eight or more phonemes.

English, French, Portuguese, and Arabic are examined as lexifier languages for the present sample of Creole languages. Varieties of Arabic have been logged as having thirty-six phonemes (Maddieson 1984) and thirty-two phonemes (Thelwall & Akram Sa’adeddin 1999), English as thirty-six (Ladefoged 1999) and thirty-seven (Cruttenden 1994), French as thirty-seven (Maddieson 1984) and thirty-four (Fougeron 1999), and European Portuguese as thirty-three (Cruz-Ferreira 1999, Mateus &
Thus, the phoneme inventory sizes of these four lexifier languages fall within the typical range, albeit at the upper end of it.

Table 2 displays the number of phonemes found in the inventories of African creoles.

<table>
<thead>
<tr>
<th>Creole</th>
<th># of phonemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kituba</td>
<td>20</td>
</tr>
<tr>
<td>Juba Arabic</td>
<td>22</td>
</tr>
<tr>
<td>Seychellois</td>
<td>23</td>
</tr>
<tr>
<td>Réunionnais</td>
<td>24</td>
</tr>
<tr>
<td>Kinubi, Mauritian, Shaba Swahili</td>
<td>26</td>
</tr>
<tr>
<td>Krio, Nigerian Pidgin</td>
<td>30</td>
</tr>
<tr>
<td>Kriyol, São Tomense</td>
<td>31</td>
</tr>
<tr>
<td>Cape Verdean</td>
<td>34</td>
</tr>
<tr>
<td>Sango</td>
<td>35</td>
</tr>
<tr>
<td>Angolar</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 2. Creole phoneme inventory size

All creoles in the sample have between twenty and thirty-seven phonemic segments. This means that all Creole languages sampled have an inventory size that is typical for non-Creole languages. None have quantitatively complex inventories. None of the creoles in the sample exhibit an inventory size that has been classified as simple in the analysis of the non-Creole languages.

The inventories of distinctive vowel qualities in all the UPSID languages range in size from three to fifteen. Five-vowel systems are most frequent (31%), followed by six-vowel systems (19%) and seven-vowel systems (15%). Maddieson (1984) suggests a complexity metric that classifies systems of three to four distinctive vowel qualities as simple and eight to fifteen phonemic vowel qualities as complex. Typical non-Creole languages display five to seven distinctive vowel qualities. This group makes up almost two thirds of the UPSID languages. The vowel quality inventories in the African languages range from three (Shilha) to ten (Akan, Amo, Dan, Diola, Hamer, Sebei, Somali). Shilha is the only language with fewer than five vowels in the present sample. Eighteen languages have five-vowel systems (26%), nine have six-vowel systems (13%), and twenty have seven vowels (29%). The remaining twenty-two African languages (31%) have between eight and ten distinctive vowel qualities. Thus, fifty-three percent (37/70) of African languages in the sample fall within the range typical for UPSID languages overall. The vast majority of the African languages ranges between five and ten vowels. The range of the lexifiers appears to be greater. Dialects of Arabic may have three (Thelwall & Akram Sa’adeddin 1999), or eight (Maddieson 1984) distinctive vowel qualities, French has twelve (Maddieson 1984, Fougeron 1999), varieties of English have twelve (Ladefoged 1999) or thirteen (Cruttenden 1994) vowels, and European Portuguese has a nine-vowel system (Cruz-Ferreira 1999, Mateus & d’Andrade 2000).

The vowel quality inventories of the African Creole languages are displayed in (1) through (4). Following a notational convention in Maddieson (1984), the mid vowel phonemes in quotation marks may be tense or lax.

(1) Five-vowel inventory: Juba Arabic, Kinubi, Kituba, Kriyol, Seychellois, Shaba Swahili

\[ \text{i, } u, \ "e", \ "o", \ a ] \]
Six-vowel inventory: Réunionnais

\[
\begin{array}{c}
i \\
e \\
o \\
\alpha/\alpha \\
\end{array}
\]

Note that the literature on Réunionnais posits variable qualities for the low front or central vowel. Papen (1978) calls it a schwa, whereas Staudacher-Valliané (1992) posits a distinction among the low vowels. These descriptive reports are interpreted as the variable realization of a sixth vowel in this system.

Seven-vowel inventory: Angolar, Krio, Nigerian Pidgin, Sango, São Tomense

\[
\begin{array}{c}
i \\
e \\
o \\
\varepsilon/\varepsilon \\
a \\
\end{array}
\]

Eight-vowel inventories (a) Cape Verdean (b) Mauritian

\[
\begin{array}{c}
i \\
e \\
o \\
\varepsilon/\varepsilon \\
\ddot{a} \\
\ddot{a} \\
\end{array}
\]

Given the data in (1) through (4), it becomes clear that African creoles only exhibit between five and eight phonemic vowels. Five- and seven-vowel systems are most frequent (11/14 = 79%). Only fourteen percent (2/14) show a complex vowel inventory by Maddieson’s measure. Interestingly, no African Creole language in the sample boasts a simple vowel system of just three or four vowels.

In the UPSID database, the number of stop phoneme series by manner of articulation in a single language ranges from one to six. There is a total of fourteen different stop series attested overall. Languages with two stop series are most frequent (51%). Languages with only one stop series, which is almost invariably plain voiceless, are considered simple. Languages with two or three stop series, which generally exhibit plain voiceless and plain voiced stops plus one other series, are considered typical, whereas languages with four to six series are considered complex. Note that Maddieson (1984) groups labio-velar stops with place, not manner, so they do not figure in this count even though a number of African Creole and non-Creole languages possess them. African languages range from one (Sebei) to six (Igbo) stop series. The great majority (60/70; 86%) of African languages fall within the typical range of two or three stop series, however. Two stop series (26/70) and three-stop series (24/70) are almost equally distributed within this range. With the exception of Igbo, only languages with four series of stops (8/70) are found in the complex range. All lexifier languages in the present sample have two series of stops; note that the Arabic pharyngeals are grouped with place, not manner.

The number and the types of stop series by manner found in the African Creole languages are displayed in (5) through (7).

Two stop series

- Plain voiceless /p, t, k/
- Plain voiced /b, d, g/
(6) Three stop series

(a) Plain voiceless /p, t, k/ Sango, Shaba Swahili
Plain voiced /b, d, g/
Prenasalized voiced /mb, nd, ng/

(b) Plain voiceless /p, t, k/ São Tomense
Plain voiced /g/
Voiced implosive /b, d/

(7) Four stop series

(a) Plain voiceless /p, t, k/ Angolar
Plain voiced /g/
Voiced implosive /b/
Prenasalized voiced /mb, nd, ng/

(b) Plain voiceless /p, t, k/ Kriyol
Plain voiced /b, d, g/
Prenasalized voiceless /mp, mt, ngk/
Prenasalized voiced /mb, nd, ng/

Following Maddieson’s (1984: 26) convention, a single phoneme is sufficient to establish a distinct stop series, even though such a series may be considered defective on other counts. Only two to four stop series are found in a given creole, and there are just five types of stops attested. Series of two stops are most frequent (9/14), and nearly all African creoles (12/14 = 86%) show two or three series. Thus, the great majority of creoles show what is considered typical for the UPSID languages in terms of the number of stop series. Two African creoles in the sample may be considered complex, given that they boast four types of stops. Importantly, none of the African creoles display just a simple one-stop series. The occurring types of stops can show significant substrate influence, in particular through the presence of implosives and prenasalized stops.

Table 3 summarizes and displays the numbers for the range, the median, and the average of phoneme inventory size, vowel quality inventory, and number of stop series for the African creoles, the African non-Creole languages, the four lexifiers combined, and all the languages in Maddieson’s (1984) UPSID sample.

<table>
<thead>
<tr>
<th></th>
<th>Phoneme inventory</th>
<th>Vowel qualities</th>
<th>Stop series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creoles</td>
<td>20-37</td>
<td>28</td>
<td>28.2</td>
</tr>
<tr>
<td>African NCs</td>
<td>20-59</td>
<td>30.5</td>
<td>31.4</td>
</tr>
<tr>
<td>Lexifiers</td>
<td>32-37</td>
<td>36</td>
<td>35.0</td>
</tr>
<tr>
<td>UPSID</td>
<td>11-141</td>
<td>28.5</td>
<td>&gt;31</td>
</tr>
</tbody>
</table>

Table 3. Numerical distribution

Maddieson (1984: 7) reports that the mean number of segments per language in the UPSID sample is a little over thirty-one whereas the median falls between twenty-eight and twenty-nine. Means or medians for the other parameters in focus cannot be recovered from his work.

Table 3 shows that the lexifiers occupy the narrowest range in terms of the size of the phoneme inventory. The spread for the Creole languages is significantly wider, but does not extend beyond what is considered typical for the world-wide sample of UPSID languages. The African non-Creole languages display a greater range of phoneme inventory sizes, whereas the range of the UPSID languages combined is the largest. The medians and the averages are close for the great majority of the parameters under discussion, indicating fairly even distribution across the languages under investigation. Exceptions to this finding are the medians and averages for the phoneme inventories in...
the UPSID languages and the vowel quality inventories for the lexifiers. The latter mismatch reflects that fact that the spread of vowel qualities is quite significant in the lexifiers, but most of them cluster towards the upper end of the continuum. The numbers in Table 3 show that the Creole languages have on average the smallest phoneme inventories of the four language groups considered. The inventories of the African non-Creole languages and the UPSID languages are on average a full series of phonemes larger that the creoles. This difference is more pronounced in comparison with the lexifier languages. They average around seven phonemes more than the creoles; this is roughly the equivalent of two full series of phonemes. The Creole languages display the narrowest spread in terms of vowel quality inventories. The other three language groups have quality inventories ranging from simple to complex, whereas the creoles are within or just above the typological middle ground. The African creoles and non-creoles center on a seven-vowel system. The lexifiers, in distinct contrast, average three more vowels. African non-creoles and the UPSID languages exhibit the widest spread in terms of the number of observed stop series. Creole languages are narrower along this parameter, again centering on the typological middle ground or extending just above it. The four lexifier languages all show two series of stops; hence, they are the most uniform of the four language groups in terms of this parameter.

4. Discussion and conclusion

Greater uniformity in the African Creole and creoloid languages is observed consistently only in comparison to the potential substrates. The creoles appear more alike in terms of their phonological segments than the diverse set of potential substrates and the vast geographical area over which they are distributed might suggest. They are significantly narrower in the number and range of the observed phoneme inventory sizes, vowel qualities, and manner stop series than African non-Creoles. However, the superstrates examined are actually more uniform than the African creoles in terms of phoneme inventory size and number of stop series, whereas vowel quality inventories of the creoles cluster more strongly than those of the lexifier languages. It seems remarkable that the African creoles do not display any nine-vowel systems even though this is found in Portuguese and is not uncommon in the substrates.

Broad confirmation for the notion of simplification in Creole languages has been found only in terms of the size of the total inventory. The Creole phoneme inventories are significantly smaller than those of the potential substrates, and even more so in comparison to the superstrates. Simplification in comparison to the potential substrates seems minor in terms of the vowel quality inventories and the number of stop series. Notwithstanding a number of outliers among the African non-Creoles, the averages for these two parameters are quite comparable. Simplification appears more pronounced in comparison to the vowel quality inventories of the Indo-European lexifiers. The rich vowel sets of French and English, in particular, are not found in any of the Creole languages. No simplification of the stop series of the lexifiers takes place, however. Instead, certain African Creole languages display additional stop series such as implosives and pre-nasalized stops typical for their substrates and are, hence, more complex than their lexifiers on this count.

None of the creoles and creoloid languages in the sample may be classed as having simple segmental structures by any of the typological measures applied. This is noteworthy evidence against the common idea that creoles in general or African creoles in particular have simple grammars in some absolute sense. Nonetheless, the attested structures show a clear typological pattern overall. Nearly all types of structure investigated for African Creole languages in the present chapter are located within what is considered typical for non-Creole languages. Examples of typological complexity may be found as well, but they tend to cluster in the Portuguese-lexified West African creoles. The strong showing of the typological middle is by no means unique to African creoles or to segmental parameters. Klein (2006) has demonstrated that typical segmental inventories predominate in a worldwide sample of Creole languages as well. In addition, the syllable template of Creole languages has recently been shown to be more complex than is often claimed; it is certainly not just CVCV. Notwithstanding certain simplifications of the syllables of Indo-European lexifiers, (C)(C)V(C) emerges as a common creole syllable template upon close examination (Klein, ms.; Plag & Schramm 2006). Thus, there is significant evidence that African creoles and their kin around the world do not
display simple phonological structures in an absolute sense, in particular as far as phonemic inventory parameters and syllable templates are concerned, but instead occupy a typological middle.

This chapter has demonstrated that significant novel generalizations may be obtained from the scrutiny of the phonological typology of African Creole languages, an area that has been treated only cursorily in the literature to date. Popular conceptions of the alleged uniformity, simplification and simplicity in the structure of Creole languages need to be revisited in light of the evidence presented here. It will be interesting for future work to develop the phonological typology of African creoles and creoloids along additional parameters and to examine Niger-Congo languages as lexifiers. When fine-grained typologies of creoles, their substrates, and their superstrates are in place, we should be able to see more clearly which Creole structures are more uniform, simpler or more complex than those of the superstrate or substrate languages. The typological approach should also facilitate objective measurement of creoleness for disputed African creoloid languages such as Afrikaans. Future work should also find it interesting to figure out plausible reasons why the typological patterning of the phonology of Creole languages appears the way it does.

References


Klein, Thomas B. Ms. Diversity and complexity in the typology of syllables in Creole languages. Statesboro, GA: Georgia Southern University.


