The formation of the Gulf of Guinea creoles: between languages, genes, and history

Tjerk Hagemeijer
(Universidade de Lisboa)
Goals

• To show that the four Portuguese-related Gulf of Guinea Creoles (GGC) spread from a single proto-language

• To argue that this proto-language shows a founder effect from southern Nigerian populations and languages based on evidence from:
  – History
  – Linguistics
  – Genetics
The Gulf of Guinea Islands

• Settlement
  – São Tomé: 1493
  – Príncipe: ±1500 (from São Tomé)
  – Annobón: ±1550 (from São Tomé)

• Settlers
  – Portuguese men (convicts, clergy members, soldiers, etc.)
  – Jewish orphans under 8 years old
  – Large numbers of continental Africans (predominantly slaves)

Language contact results in the creation of a new language by the slave population through the standard **pidgin-creole cycle**
Settlement and slave trade areas

1. Homestead society (aprox. 1493-1515)
   → predominance of the Niger Delta (southern Nigeria) slave trade

2. Plantation society (aprox. 1515-1600)
   → predominance of Bantu slave trade areas, corresponding roughly to the Congo and Angola

(Caldeira 2008; Pereira 1506; Ryder 1969; Teixeira da Mota 1976; Thornton 1992; Vogt 1973, a.o.)
Sugar mills on São Tomé - 16\textsuperscript{th} century

<table>
<thead>
<tr>
<th>Year</th>
<th>1517</th>
<th>1522</th>
<th>1550</th>
<th>1567</th>
<th>1580</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nr.</td>
<td>2/3</td>
<td>6</td>
<td>60</td>
<td>120</td>
<td>70</td>
</tr>
</tbody>
</table>
Linguistic founder effect

• To what extent did the languages spoken in these two slave trade areas affect the formation of the creole language on São Tomé?

• The linguistic features of the GGCs show that the languages of southern Nigeria were crucial in the formation of the proto-language that spread from the island of São Tomé in the 16th century
Especlation of the proto-language (16th century)

Proto-language (proto-Gulf of Guinea Creole)

- Angolar
- Fa d’Ambô
- Lung’le
- Santome
Creoles and genetic classification

- Creole grammars are a compromise between grammars in contact and innovation.
- Creoles derive most lexicon from one source (lexifier or superstrate language).
- The Comparative Method
  - Cognates
  - Regular sound correspondences
- Creoles constitute young branches of their lexifier language.
Examples of cognates and regular sound changes

<table>
<thead>
<tr>
<th>Modern Portuguese</th>
<th>Santome</th>
<th>Sound change</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>chorar</em></td>
<td><em>sola</em></td>
<td>(t)ʃ &gt; s</td>
<td>to cry</td>
</tr>
<tr>
<td><em>chave</em></td>
<td><em>sabi</em></td>
<td>(t)ʃ &gt; s</td>
<td>key</td>
</tr>
<tr>
<td><em>chuva</em></td>
<td><em>suba</em></td>
<td>(t)ʃ &gt; s</td>
<td>rain</td>
</tr>
<tr>
<td><em>dinheiro</em></td>
<td><em>djêlu</em></td>
<td>r &gt; l</td>
<td>money</td>
</tr>
<tr>
<td><em>febre</em></td>
<td><em>feble</em></td>
<td>r &gt; l</td>
<td>fever</td>
</tr>
<tr>
<td><em>rir</em></td>
<td><em>li</em></td>
<td>r &gt; l</td>
<td>to laugh</td>
</tr>
<tr>
<td><em>pai</em></td>
<td><em>pe</em></td>
<td>aj &gt; ɛ</td>
<td>father</td>
</tr>
<tr>
<td><em>vai</em></td>
<td><em>be</em></td>
<td>aj &gt; ɛ</td>
<td>to go</td>
</tr>
<tr>
<td>*barriga &gt;<em>baiga</em></td>
<td><em>bega</em></td>
<td>ai &gt; ɛ</td>
<td>belly</td>
</tr>
</tbody>
</table>
Lexical layers in the GGCs

- Portuguese lexicon (approx. 90%):
  - shared by the four GGC → primary layer

- Nigerian lexicon (Edoid):
  - frequently shared by all or several GGC → primary layer

- Bantu lexicon (Kongo, Kimbundu):
  - not shared and almost absent from Lung’le → secondary layer
## Portuguese-related cognates

<table>
<thead>
<tr>
<th>ST</th>
<th>LU</th>
<th>FdA</th>
<th>ANG</th>
<th>Meaning</th>
<th>Portuguese</th>
</tr>
</thead>
<tbody>
<tr>
<td>sangi</td>
<td>isengi</td>
<td>sangi</td>
<td>thangi</td>
<td>blood</td>
<td>sangue</td>
</tr>
<tr>
<td>ūa</td>
<td>ūa</td>
<td>wan</td>
<td>ūa</td>
<td>one</td>
<td>uma</td>
</tr>
<tr>
<td>kôlê</td>
<td>kwê</td>
<td>khôlê</td>
<td>kôlê</td>
<td>to run</td>
<td>correr</td>
</tr>
<tr>
<td>alê</td>
<td>arê</td>
<td>alê</td>
<td>alê</td>
<td>king</td>
<td>el-rei (archaic)</td>
</tr>
<tr>
<td>bluku</td>
<td>buuku</td>
<td>buuku</td>
<td>buuku</td>
<td>mean</td>
<td>bronco(?)</td>
</tr>
<tr>
<td>saya</td>
<td>saa</td>
<td>saa</td>
<td>thaa</td>
<td>to pull</td>
<td>salhar (archaic)</td>
</tr>
</tbody>
</table>
## Edo-related cognates

<table>
<thead>
<tr>
<th>ST</th>
<th>LU</th>
<th>FdA</th>
<th>ANG</th>
<th>meaning</th>
<th>Edo etymon</th>
</tr>
</thead>
<tbody>
<tr>
<td>bôbô</td>
<td>bôbô</td>
<td>bôbô</td>
<td>bôbô</td>
<td>to carry a child on the back</td>
<td>vívovo</td>
</tr>
<tr>
<td>idu</td>
<td>idu</td>
<td>idu</td>
<td>iru</td>
<td>louse</td>
<td>iru</td>
</tr>
<tr>
<td>izè</td>
<td>izè</td>
<td>inze</td>
<td>idhè</td>
<td>crayfish</td>
<td>izè</td>
</tr>
<tr>
<td>lêlê</td>
<td>lêlê</td>
<td>lêê</td>
<td>lêlê</td>
<td>to follow, along</td>
<td>lele</td>
</tr>
<tr>
<td>ubwê</td>
<td>igbe</td>
<td>ôgê</td>
<td>ongê</td>
<td>body</td>
<td>egbe</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>impersonal pronoun</td>
<td>a</td>
</tr>
</tbody>
</table>
# Bantu lexicon in Angolar

<table>
<thead>
<tr>
<th>ST</th>
<th>LU</th>
<th>FdA</th>
<th>ANG</th>
<th>Mean</th>
<th>PORT.</th>
<th>KIMBUNDU</th>
</tr>
</thead>
<tbody>
<tr>
<td>labu</td>
<td>urabu</td>
<td>labu</td>
<td>nkila</td>
<td>tail</td>
<td>rabo</td>
<td>mukila</td>
</tr>
<tr>
<td>pixi</td>
<td>pexi</td>
<td>pixi</td>
<td>kikiê</td>
<td>fish</td>
<td>peixe</td>
<td>kikêle</td>
</tr>
<tr>
<td>nôtxi</td>
<td>unôtxi</td>
<td>nôtxi</td>
<td>n’thuku</td>
<td>night</td>
<td>noite</td>
<td>usuku</td>
</tr>
<tr>
<td>xinku</td>
<td>xinku</td>
<td>xinku</td>
<td>tano</td>
<td>five</td>
<td>cinco</td>
<td>tanu</td>
</tr>
<tr>
<td>vwa</td>
<td>vwa</td>
<td>vôa~va</td>
<td>pupuka</td>
<td>to fly</td>
<td>voar</td>
<td>pupuluka</td>
</tr>
</tbody>
</table>
Typological features

Macro-Sudan

Bantu

(Güldemann 2008)
## Labial velars /kp/ e /gb/

<table>
<thead>
<tr>
<th>ST</th>
<th>LU</th>
<th>FdA</th>
<th>ANG</th>
<th>Edo</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>kwaku</em></td>
<td><em>ukpaku</em></td>
<td>?</td>
<td>?</td>
<td><em>ukpakon</em></td>
<td>traditional tooth brush</td>
</tr>
<tr>
<td><em>kwali</em></td>
<td><em>ukperi</em></td>
<td>okwali</td>
<td>?</td>
<td>?</td>
<td>basket</td>
</tr>
<tr>
<td><em>ukwê</em></td>
<td><em>ikpe</em></td>
<td>?</td>
<td><em>ikwe~inkwe</em></td>
<td><em>ikpe</em></td>
<td>seed, grain</td>
</tr>
<tr>
<td><em>ubwami</em></td>
<td><em>ugbami</em></td>
<td>ogomu</td>
<td>?</td>
<td><em>agbanmwen</em></td>
<td>chin</td>
</tr>
<tr>
<td><em>ubwa</em></td>
<td><em>ugba</em></td>
<td>?</td>
<td>?</td>
<td><em>ogba</em></td>
<td>fence</td>
</tr>
<tr>
<td><em>ubwê</em></td>
<td><em>igbê</em></td>
<td>ôgê</td>
<td>ôngê</td>
<td><em>egbe</em></td>
<td>body</td>
</tr>
</tbody>
</table>
Labial-velars in Africa

(Maddieson 2005)
Body reflexives

(1) È mata ubwê dê  (Santome)
(2) È mata igbê sê  (Lung’ie)
(3) È mata ôgê dêli  (Fa d’Ambô)
(4) È mata òngê rê  (Angolar)
    3SG kill body POSS 
    ‘S/he committed suicide.’

(5) Ọ gbé-ègbé èré ruà  (Edo)
    3SG kill-body POSS PART 
    ‘S/he committed suicide.’
NP final interrogative (where is/are ...?)

(1) *Inen*  *bô?*  
3PL  INT
Where are they?’

(2) *Kasô*  *tê*  *ba?*  
dog  POSS  INT
‘Where is your dog’?

(3) *Bo*  *bô?*  
2SG  INT
‘Where are you?’

(4) *Lêlu*  *ô*  *bô?*  
money  POSS  INT
‘Where is your money?’

(5) *Rèn*  *vbòó?*  ([vòó])  
3SG  INT
‘Where is s/he?’

(Santome)

(Lung’le)

(Fa d’Ambô)

(Angolar)

(Edo)
Serial verb constructions

Directional

(1) Ê kolê ba ke (Santome)
(2) Ê kwê we kaxi (Lung’le)
(3) Ê khôlê ba khadji (Fa d’Ambô)
(4) Ê kolê ba kai (Angolar)
   3SG run go house
   ‘S/he ran home.’

(5) Íran rhûlé làá òwá (Edo)
   3PL ran enter house
   ‘They ran into the house.’
Serial verb constructions

Comitative

(1) Zon na lêlê Maya xê fa. (Santome)
Ìvié má lèlé Òdúwà kpàá. (Edo)
PN NEG follow PN leave NEG
‘Zon/Ìvié didn’t leave in Maya/Òdúwà’s company.’

(2) Inen zunta kume lôsô. (Santome)
Íràn kùgbé-rè rrí izè. (Edo)
3PL gather eat rice
‘They ate rice together.’
Verb serialization in Africa

(Dimmendaal 2001: 383)
Summary

• The numerous lexical, phonological, and syntactic resemblances between the four GGCs, either directly or through historical reconstruction, show their genetic relatedness

• The GGCs share many lexical and typological features with southern Nigerian languages, in particular with Edoid languages

• Correlation between linguistic features and the predominant slave population during the homestead society
Population genetics

- Predominantly African genetic contribution
- Surprisingly high of the Nigerian genotype on the population of São Tomé

Coelho et al. 2008: Comparison of the distribution of betaglobine S haplotypes
The puzzle of the Angolares

- Main hypothesis: The Angolares are the descendants of the numerous slaves that ran away from the plantation system from the 16th century on
- The Angolares are therefore expected to be genetically similar to the remaining population of São Tomé.
Structural correspondence: Angolar and Santome

(1) \( n \ ga \ fa \ ku \ ê, \ si \ n \ të \ tempu \) (Angolar)
\( n \ ga \ fla \ ku \ ê, \ xi \ n \ të \ tempu \) (Santome)
1SG  FUT  talk  with 3SG if  1SG  have  time
‘I will speak to him, if I have time’

(2) \( a \ na \ ka \ bendê \ kikiê \ nge \ wa \) (Angolar)
\( a \ na \ ka \ bendê \ pixi \ nai \ fa \) (Santome)
IMP  NEG  HAB  sell  fish  here  NEG
‘Fish is not sold here’
DNA of the Angolares

- The Angolares are genetically less diversified than the remaining population of São Tomé
  - Y chromosome: high level of differentiation from the rest of the island, showing evidence of ancestral lineages related to the Bantus of the forest (pygmy gene flow)
  - Mitochondrial DNA (maternal inheritance) contributes genetic diversification to the Angolares

(Coelho et al. 2008)
HAPLOGROUP Profiles
(Tomás et al. 2003)
Coelho *et al.* (2008): Genetic differentiation on São Tomé based on a random analysis
Female genetic differentiation: Angolares vs. Population of S. Tomé

Coelho et al. (2008): mtDNA sequence variation
Male genetic differentiation: Angolares vs. Population of S. Tomé

Coelho et al. (2008): Y-chromosome haplotype diversity
The Angolares: genes and language

1. The Angolares are not a pre-European population and do not speak a Bantu language (against some previous misinformed claims)

2. The Angolares are likely to be the result of a bottleneck event, i.e. a close-knit masculine runaway Bantu group that survived outside the plantation system

3. Historically documented abduction of (creole) women from the plantations may have contributed greater genetic diversification of this group
The Angolares: new questions

4. Hypotheses to be further tested wrt the linguistic and genetic differentiation of Angolar:

4.1. Did the original runaway group shift from a Bantu language to Creole and the Bantu features are a case of retention?

4.2. Did the original runaway group speak a Creole that borrowed Bantu features (in particular, lexicon and phonology)?

4.3. How much convergence occurred between Angolar and Santome over the centuries?
Conclusions

• The four GGC descend from a single proto-language that resulted from the contact between middle/classic Portuguese and southern Nigerian languages, during the homestead society (= fast creolization)

• The founder effect of this African layer is corroborated by evidence from different disciplines: history, linguistics, genetics

• The Bantu layer is mostly the result of secondary contact, but left a substantial lexical and genetic imprint, in particular on the Angolares
DÊSU-PAGA!
(Obrigado)
Selected references


