How to become a Macro-Sudan belt language: the Gulf-of-Guinea creole (GGC) case

Tom Güldemann1,2 and Tjerk Hagemeyer1
1Humboldt University Berlin, 2MPI-SHH Jena, 3Universidade de Lisboa

1 Introduction

1.1 Cross-African areal typology and the Macro-Sudan belt

<table>
<thead>
<tr>
<th>No.</th>
<th>Macro-area</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)</td>
<td>Sahara spread zone</td>
<td>genealogical offshoot of (II)</td>
</tr>
<tr>
<td>(II)</td>
<td>Chad-Ethiopia</td>
<td>contact area</td>
</tr>
<tr>
<td>(III)</td>
<td>Macro-Sudan belt</td>
<td>contact area</td>
</tr>
<tr>
<td>(IV)</td>
<td>Bantu spread zone</td>
<td>genealogical offshoot of (III)</td>
</tr>
<tr>
<td>(V)</td>
<td>Kalahari Basin</td>
<td>contact area</td>
</tr>
</tbody>
</table>

Table 1: Proposed recent linguistic macro-areas in Africa (Güldemann 2010)

Map 1: Linguistic macro-areas in Africa before recent large-scale colonizations

Macro-Sudan belt as a large contact-mediated area south of the Sahara, north of the rain forest, and west of the Ethiopian plateau, pre-figured by Greenberg (1959, 1983) but established in more detail by Güldemann (2003, 2008) and Clements and Rialland (2008)

<table>
<thead>
<tr>
<th>Family</th>
<th>Stock</th>
<th>Greenberg's supergroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mande</td>
<td>-</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Kru</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Gur</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Kwa</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Benue-Congo (except Narrow Bantu)</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Adamawa-Ubangi</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Bongo-Bagirmi</td>
<td>Central Sudanic</td>
<td>Nilo-Saharan</td>
</tr>
<tr>
<td>Moru-Mangbetu</td>
<td>Central Sudanic</td>
<td>Nilo-Saharan</td>
</tr>
<tr>
<td>Atlantic</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Dogon</td>
<td>-</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Songhai</td>
<td>-</td>
<td>Nilo-Saharan</td>
</tr>
<tr>
<td>Chadic</td>
<td>Afroasiatic</td>
<td>-</td>
</tr>
<tr>
<td>Ijoid</td>
<td>-</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Narrow Bantu (Benue-Congo)</td>
<td>Niger-Congo</td>
<td>Niger-Kordofanian</td>
</tr>
<tr>
<td>Nilotic</td>
<td>East Sudanic</td>
<td>Nilo-Saharan</td>
</tr>
</tbody>
</table>

Note: Bold = core family, italic = major GGC adstrate

Table 2: Families partaking in the Macro-Sudan belt (Güldemann 2008)

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Implosive consonants</td>
<td>Maddieson (2005a)</td>
</tr>
<tr>
<td>2</td>
<td>Labial-velar consonants</td>
<td>Maddieson (2005c)</td>
</tr>
<tr>
<td>3</td>
<td>Three+ level tones</td>
<td>Maddieson (2005d), Clements and Rialland (2008)</td>
</tr>
<tr>
<td>4</td>
<td>ATR vowel harmony</td>
<td>Hall et al. (1974), Dimmendaal (2001), Maddieson (2005b)</td>
</tr>
<tr>
<td>5</td>
<td>Nasalized vowels</td>
<td>Hajek (2005)</td>
</tr>
<tr>
<td>6</td>
<td>&quot;Lax&quot; question prosody</td>
<td>Clements and Rialland (2008)</td>
</tr>
<tr>
<td>7</td>
<td>SBJ-(AUX)-OBJ-V-X</td>
<td>Gensler and Güldemann (2003)</td>
</tr>
<tr>
<td>8</td>
<td>'Sur'pass' comparative</td>
<td>Greenberg (1983), Stassen (2005)</td>
</tr>
<tr>
<td>9</td>
<td>Logophoricity system</td>
<td>Güldemann (2003b)</td>
</tr>
<tr>
<td>10</td>
<td>Post-V/clause-final NEG</td>
<td>Dryer (2009), Idiatov (2010)</td>
</tr>
<tr>
<td>11</td>
<td>Serial verbs</td>
<td>Dimmendaal (2001)</td>
</tr>
<tr>
<td>12</td>
<td>Plural word</td>
<td>Dryer (2013)</td>
</tr>
<tr>
<td>13</td>
<td>Focus-verb doubling</td>
<td>Bond and Anderson (2014)</td>
</tr>
</tbody>
</table>

Table 3: Relevant linguistic features of the Macro-Sudan belt
1.2 The Gulf-of-Guinea creoles

+ Gulf-of-Guinea creoles:
- family of four languages spoken on three formerly uninhabited islands west of Gabon and Equatorial Guinea, emerged at the end of the 15th century lexically based on Portuguese
- national (minority) languages not used in education and threatened by local lingua franca

Table 4: The four Gulf-of-Guinea creoles (after Ethnologue)

+ different historical phases of GGCs - 2 crucial early phases (cf. Hagemeijer 2011):
(I) Portuguese “superstrate” + Niger Delta “substrate” > founder or proto-creole
(II) Slightly later but intensive contact with Bantu of zone H (Kongo, (K)Mbundu)
(III) yet later strates with various other languages, partly in different places: Portuguese, Kabuverdianu, Umbundu; Fa d’Ambu only: Spanish, Pichi

> relatively short time between “speciation” of proto-creole and additional contact with other African languages > non-trivial distinction between different linguistic influences

2 Macro-Sudan features and the Gulf-of-Guinea creoles

F1 Implosive (= glottalic ingressive) consonants
+ present in GGCs (Ferraz 1979; Maurer 1995, 2009; Zamora 2010)
+ present in Niger Delta including Edo (Eluge 1986, 1989)
+ absent in Bantu H, only present in restricted Bantu subareas (Güldemann 2011: 115)
+ absent in Portuguese

F2 Labial-velar consonants
+ today only present in Principense, labialization and other reflexes in the remaining creoles (Table 5) > feature can be reconstructed for Proto-GGC, as Principense is also in other respects structurally conservative
+ present in Niger Delta including Edoid (Elugbe 1986, 1989)
+ largely absent in Bantu including zone H (Güldemann 2011: 115)
+ absent in Portuguese

F3 Three and more level tones
+ absent in GGCs but controversial analysis of their prosody as tone or accent systems:
+ remarkable that tone is at all entertained: unexpected in “new” restructured languages
  (McWhorter 2005), very rare in creoles (cf. Good 2004 for similar Saramaccan exception)
+ present in Niger Delta including some Edoid languages (Elugbe 1986, 1989)
+ largely absent in Bantu including zone H (Güldemann 2011: 115)
+ absent in Portuguese

F4 ATR harmony
+ Sáotomense: vowel harmony “in the properties frontness, height, etc.” (Ferraz 1979: 111)
1. [i] [u]
2. [e] [o]
3. [a]

(2) Sáotomense
- [o/me] ‘middle’ vs. [’me] ‘man’
- [me/se] ‘to want’ [mese] ‘master’
- [mole] ‘to die’ [ləvɛ] ‘dew’
- [bo/lo] ‘to rub’ [pota] ‘door’

> Hagemeijer (2009) argues for an analysis in terms of mid-vowel harmony that can be viewed as a reduced ATR-system restricted to the open-mid and close-mid vowels (similar to the system in Saramaccan under Gbe influence (McWhorter and Good 2012, N. Smith p.c.)

+ present in Niger Delta including Edoid (Elugbe 1986, 1989)
+ largely absent in Bantu including zone H (Güldemann 2011: 115)
+ absent in Portuguese

F5 Vowel nasalization
+ present throughout GGCs (Ferraz 1979; Maurer 1995, 2009; Zamora 2010)
+ present in Niger Delta including Edoid (Elugbe 1986, 1989)
+ very rare in Bantu, absent in Bantu zone H (Güldemann 2011: 115)
+ present in Portuguese

+ abstract set of cross-linguistically unusual features of final question marking like open vowel, L tone, sentence-final falling intonation, lengthening, and combinations thereof
(3) Allomorphs of lax question prosody in Ncam (Gur)
  a. -a appears after a consonant-final root:
    ḏ ɔc̃ ʔɛ́ S/he walked ḏ ɔc̃ ʔɛ́ ṭə́? Did s/he walk?
  b. -a is added after other short vowels, where it undergoes a variety of assimilations
  c. -a replaces a short final i, which is most often epenthetic:
    ḏ ɡājki S/he repaired ḏ ɡājki ṭə́? Did s/he repair?
  d. a final long vowel has extra length, with no change in quality:
    aɾi sə́? It’s rotten aɾi sə́? Is it rotten?
  e. falling intonation, final lengthening and breathy termination are regularly present
  (Clements and Rialland 2008: 78)

+ final a or â in polar questions present across GGCs with low boundary tone in Principep
  (Maurer 2009: 146, cf. Figure 1), with falling intonation in Sáotomense and Fa d’Ambu
  (Hagemeijer &., Zamora p.c.), with unknown prosody in Angolar (Maurer, Araújo p.c.)

Figure 1: Prosody in declarative (1033), polar question without a (1034), and polar question with a (1035) in Principep (Maurer 2009: 148)

can be dropped with remaining low tone still marking interrogative function
(4) Òṣé ḏéè ṭé ṭai?
  PN bought book Q
  Did Ozo buy a book? (Agheyisi 1990: 98)

+ so far not widely attested in Bantu (Güldemann 2011: 115)
+ absent in Portuguese
F7 Word order S-(AUX)-O-V-X
+ conditioned or predominant preverbal position of object
(5) Koranko (Mande)
\[ \text{ń sí wò lè-ème yô rò} \]
\[ 1S \text{ PROSPECTIVE that .one CAUS-fall water in} \]
I'm going to throw her into the water (Kastenholz 1987: 117)

+ absent in GGCs, basic clause order is strictly S-AUX-V-O
+ present in some Niger Delta languages but apparently absent in Edo
+ present in Northwest but absent elsewhere including zone H (Güldemann 2007, 2011)
+ present in Portuguese with pronouns (similar to cases in Niger-Congo)

F8 Surpass comparatives
+ comparative construction with a “surpass” verb marking the standard
(7) Swahili (Bantu, Benue-Kwa)
\[ a-na-andika haraka ku-pita mimi \]
1-PRS-write fast INF-pass 1S
he writes faster than me

+ present in GGCs but synchronically a grammaticalized form without verbal features
(8) Sãotomense
\[ ê sa longô pasa mu \]
3S be tall surpass me
he is taller than I

- also with a variant for intensification ?and superlative
(9) Angolar
\[ ëe xa masi dhangaru patha ëto \]
one be more high surpass other
one is higher than the other (Maurer 1995: 52)

F9 Logophoricty
+ obligatory disambiguation of (non)coreference in reported discourse
(12) Kera (Chadic, Afroasiatic)
a. \[ wə mëntí tò këoré \]
3M.S. QUOT 3M.S.LOG go.away
b. \[ wə mëntí wò këoré \]
3M.QUOT 3M.S. go.away
Er sagte, daß er weggehe [he said he would go] (Ebert 1979: 260)

+ absent in GGCs
+ present in Niger Delta including Edoid (possibly not the intensifying variant)
(11) Edo (Edoid, Benue-Kwa)
\[ Özö mosè së ét Azărí \]
PN be.beautiful surpass PN
Ozo is more beautiful than Azari. (Hagemeijer and Ogie 2011: 50)

+ recurrent in Bantu (Güldemann 2011: 115-6, cf. Chatelain 1888/9: 114) for Mbundu
+ absent in Portuguese

F10 Post-V/clause-final NEG
+ originally treated as a subareal feature according to Dryer (2009), extended to entire Macro-Sudan belt by Idiatov (2010)

+ present across GGCs (cf. Hagemeijer 2011: 133-4)
(13) Sãotomense
\[ non na tê awa fa \]
1P NEG have water NEG
we don’t have water (Hagemeijer et al. 2014b)

+ present in Niger Delta including Edoid (Agheyesi 1990: 55-6, Hagemeijer f.n.)
+ largely absent in Bantu including zone H (Güldemann 2011: 116)
+ absent in Portuguese

9 Workshop “Areal phenomena in northern sub-Saharan Africa” at WOCAL8 Kyoto

(15) Principense

\[ \text{in ten posan fa} \]
1S go until town NEG
I didn’t go to town (Maurer 2009: 133)

+ present in Niger Delta languages like Igbo and parts of Edoid (Elugbe 1977) but not in Edo itself (Güldemann and Hagemeijer 2006)
+ recurring in Bantu (Güldemann 2011: 116–7), particularly salient in zone H (Güldemann and Hagemeijer 2006)
+ absent in Portuguese

F11 Serial verbs
+ distinct type of multi-verb construction defined by structural and semantic-functional characteristics (cf., e.g., Stahlke 1970, Ameka 2003)

(16) Yoruba (Yoruboid, Benue-Kwa)

\[ \text{màmá mi á sọ fún mi pé} \]
mother 1S IRR speak give 1S QUOT
My mother would tell me ‘...’ (Bamgbose 1966: 167)

+ salient in GGCs with a range of different types like directional, locative/goal, instrumental, comitative etc. (Maurer 1995, 1999, 2009; Post 1992; Hagemeijer 2000; Hagemeijer and Ogie 2011)

(17) Sãotomense

\[ \text{ê fe ubwa loja ke dê} \]
3S build fence encircle house 3S:POSS
s/he built a fence around his/her house (Hagemeijer 2011: 122)

(18) Principense

\[ \text{inen zunta we posan} \]
3P gather go town
they went to town together (Maurer 2009: 118)

(19) Angolar

\[ \text{kathô tambu n’kila rê pé kosí bega} \]
dog take tail 3S:POSS put under belly
the dog hid his tail (under his belly) (Hagemeijer 2011: 126)

Fa d’Ambu

\[ \text{amu sa ma xoe bi do-l} \]
1S be take thing come give-3S
I brought it to her (Post 1992: 163)

+ widely present in Niger Delta including Edoid (Hagemeijer and Ogie 2011)
+ largely absent in Bantu, including zone H (Güldemann 2011: 117)
+ absent in Portuguese

F12 Plural word
+ plural word as major means of marking nominal plurality

(21) Chalcatongo Mixtec (Oto-Manguean)

\[ \text{ni-xá =rí k’ažá lážá káni xinzá} \]
COMP-buy =1 many rope long P
I bought many long ropes (Macaulay 1996: 113, cited by Dryer 2013)

> Macro-Sudan feature according to world survey by Dryer (2013)
> recurrent origin in 3rd-person plural pronoun (cf., e.g., Westermann 1947)

+ present in GGCs with likely etymological source in Edo element(s) (Hagemeijer 2011: 129–30) > possibly internal grammaticalization of substrate-based 3P pronoun

(22) Sãotomense

a. \[ \text{inen kume káni} \]
3P eat meat
they ate meat

b. \[ \text{inen migu mu kume káni} \]
P friend 1S:POSS eat meat
my friends ate meat

(23) Fa d’Ambu

\[ \text{ta ku nan pe nen se sa kha bay, …} \]
when COMP P man P DEM be TAM go when these men left, … (Hagemeijer 2015: 307)
F13 Focus-verb doubling for predicate-centered focus
+ focused verb doublet for state-of-affairs focus or general predicate-centered focus, in two syntactic configurations, viz. in-situ or in fronted position
(24) Nweh (Grassfields, Benue-Kwa)
   Atem a ki? n’èqè akendéy èqù
   PN 3S PST1 ?boil plantains boil
   Atem BOILED plantains (Nkemnji 1995: 138)

(25) Aja (Gbe, Benue-Kwa)
   dò, dà (yò) è dà
   no cook FOC 3S cook
   [The woman ate the beans.] No, she COOKED (them) (Fiedler p.c.)

> wide-spread across Africa (Bond and Anderson 2014), but particularly recurrent in Macro-Sudan belt (?and Chad-Ethiopia) and only sporadic or even absent elsewhere (Güldemann ongoing research)
> both constructional variants particularly frequent in Gulf-of-Guinea coast subarea of Macro-Sudan belt (cf., e.g., Williamson 1986: 10-2)
3 Discussion

3.1 Linguistic and social history of the Gulf-of-Guinea creoles
+ GGCs have a typical Macro-Sudan belt profile in (partially) displaying 10 of 13 features
  > transparent differential contribution of three major linguistic population components:
  a) Edo(id): almost all relevant features in line with its membership in the Macro-Sudan belt
  b) Portuguese: just 1 of 10 relevant features despite its central lexical contribution
  c) Bantu zone H: just 3 of 10 relevant features in line with Bantu profile different from
     Macro-Sudan belt (cf. Güldemann 2011) > Bantu impact far less extensive than originally
     assumed by Ferraz (1979) (cf. Güldemann 2013)
     > Niger Delta languages and Edo in particular as the major source of the non-Portuguese
     features in GGCs

  + strong linguistic Macro-Sudan affiliation of GGCs fully in line with other evidence for
    Hagemeijer’s (2011) historical model
  > GGCs present a strong case for a specific substrate creole providing a good
    explanation for an individual linguistic profile (against anti-substrate creole theories)
  > macro-areal typology as a potentially useful heuristic for reconstructing historical
    trajectories of particular linguistic lineages (cf. Güldemann 2011)

  + broad sociolinguistic history of GGCs relatively well understood:
    - Niger Delta population quickly pidginized and nativized Portuguese to a proto-creole on
      São Tomé island
    - language shift of slaves accompanied by significant retention of Niger Delta substrate
      features in the contact language due to heavy imbalance of slave-white ratio and limited
      access to and acquisition of the target language Portuguese
    - creole (or extended pidgin) also became the target language for new slave arrivals (mainly
      Bantu speakers) and diversified to a family, notably by spreading to two more islands,
      Príncipe and Annobón
    > linguistic mechanism for contact-mediated feature transfer in crucial early phase was
      “shift-induced substrate interference” rather than “borrowing” in terms of Thomason and
      Kaufman (1988)

3.2 GGCs and the historical modelling of the Macro-Sudan belt
+ provided language contact scenarios can explain linguistic macro-areas, what is the
  relative role of the two major patterns by Thomason and Kaufman (1988)
  > most previous accounts of the Macro-Sudan belt tend to privilege borrowing over shift-
    induced substrate (or don’t consider the latter in the first place) by entertaining family
  > alternative approach: substrate is (A) at least as or (B) more important than borrowing

  + hypothesis (A) should be viewed as null hypothesis:
    (I) substrate as a more natural explanation for degree and kind of feature transfer
    (II) substrate also explains well other robust contact-induced areas in and outside Africa:
      Dravidian substrate in Indo-Aryan
      Cushitic substrate in Ethiosemitic
      Tuu + Kx’a substrate in Khoe (Güldemann and Fehn forthcoming)
      etc.
    (III) GGCs etc. provide concrete and relatively transparent cases for reconstructing the
      emergence and dynamics of a synchronic typological profile of a language (family)

  + indigenous lineages in macro-areas like the Macro-Sudan belt with a largely opaque
    history in terms of populations dynamics
  > GGCs potentially attractive as a partial analogue for a more abstract modelling of macro-
    areas, this in various ways:
    a) new linguistic entity in a certain geographical area ends up in a heavily restructured form
       with a typical local linguistic profile
    > sociolinguistically marked creole genesis representative for long-term precolonial history
       in Africa? - cf. strong substrate effects in non-creole cases
    b) language resulting from the original contact event expands to a larger family under
       retention of numerous features
    c) provided an appropriate environment, features can be stable despite heavy subsequent
       contact with languages of a different profile and demographic superiority

  + more extreme hypothesis (B) would tend to explain old contact-induced areas by a strong
    early founder effect, where the substrate “oozes up” into new linguistic population layers,
    implying a model according to which “linguistic features sit and populations move”
  > Is that a possible scenario for the deep history of macro-areas?
Abbreviations

CAUS Causative, COMP Complementizer, DEM Demonstrative, FOC Focus, INF Infinitive, IRR Irrealis, LOG Logophoric pronoun, M Masculine, NEG Negative, P Plural, PN Proper name, POSS Possessive, PQ Polar question, PROG Progressive, PRS Present, PST Past, QUOT Quotative, S Singular, VN Verbal noun

References

Agbeyesi, Rebecca. 1990. A grammar of Edò. UNESCO.


Clements, Nick and Annie Rialland. 2008. Africa as a phonological area. In Heine and Nurse (eds.).


Güldemann, Tom. 2008. The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In Heine and Nurse (eds.).


