A null theory of Creole formation

Michel DeGraff
degraff@MIT.EDU

Goals

My working hypotheses about Creoles go against the dogma in most textbooks and current research:

1. *Creole languages are NOT the outcome of a “Pidgin-to-Creole cycle”*

2. Creole formation is a normal instance of language change.
A debate about family values?

1. *Are Creoles genealogically related to their European superstrate languages?* (Yes: Baissac 1880, Hall 1958...)

2. **No**: Creoles are “genetically ‘orphans’ [with] two ‘foster-parents’: one that provides the basic morphological and/or syntactical pattern, and another from which the fundamental vocabulary is taken” (Taylor 1956:413; cf. Adam 1883, Lefebvre 1998).

3. **No**: Because they originate in Pidgins, Creoles are outside the scope of the Comparative Method.

4. *Creoles are “simplest” languages* due to “pidgin” origins (Baissac 1880, Adam 1883, Taylor 1956, Hall 1958, McWhorter 1998ff, Plag 2008ff, Bakker & al 2011, etc.)

Uncontroversial facts

1. **Early definition of “Creole languages”**: Varieties spoken by the “Creole” communities of the colonial Caribbean (17th–19th centuries).

2. Creole languages developed among Europeans and Africans in the colonial Caribbean.

3. Creole languages are linguistic side-effects of an early sort of ‘globalization.’
The geo-politics of the “classic Creoles”

1. Manufactured goods
2. Slaves
3. Raw materials i.e. sugar, cotton, tobacco, metals

The “classic Creole” Geography

Holm 1989 Pidgins and Creoles

ENGLISH-BASED
1. American CE
2. Oceanic CE
3. Amazonian CE
4. New Guinea CE
5. Melanesian CE
6. Polynesian CE
7. Micronesian CE
8. Philippine CE
9. East Asian CE
10. Malay-Indonesian CE
11. Andaman Islands CE
12. Batak, Manet, Sukhoi CE
13. West Papuan CE
14. Melanesian CE
15. Kula-Adaman, etc. CE
16. Austronesian CE

ENGLISH-ADDON
1. English
2. German
3. French
4. Spanish
5. Dutch
6. Portuguese
7. Russian
8. Italian
9. Arabic
10. Japanese
11. Korean
12. Chinese
The dogma on “Pidgin” and “Creole” (Holm 2000)

A *pidgin* is a reduced language that results from extended contact between groups of people with no language in common; it evolves when they need some means of verbal communication, perhaps for trade, but no group learns the native language of any other group for social reasons that may include lack of trust or close contact. Usually those with less power (speakers of *substrate* languages) are more accommodating and use words from the language of those with more power (the *superstrate*), although the meaning, form and use of these words may be influenced by the substrate languages.

1.2 Creoles

A *creole* has a jargon or a pidgin in its ancestry; it is spoken natively by an entire speech community, often one whose ancestors were displaced geographically so that their ties with their original language and sociocultural identity were partly broken. Such social conditions were often the result of slavery. For example, from the seventeenth to the nineteenth century, Africans of diverse ethnolinguistic groups were brought by Europeans to colonies in the New World to work together on sugar plantations.
The “Pidgin-to-Creole cycle” in textbooks

“Contact situations where speakers have restricted access to each other's language can sometimes lead to the formation of a PIDGIN—a rudimentary language with minimal grammatical rules and a small lexicon. By definition, a pidgin has no native speakers.”

(O’Grady et al’s Contemporary Linguistics p. 503)

“...[C]hildren [of pidgin speakers] may learn the pidgin as a first language, and it may be adopted as the native language of the new community. [Then] the pidgin becomes a full-fledged language known as a CREOLE. When a Pidgin becomes a Creole, its inventory of lexical items and grammatical rules expands dramatically, usually in only one or two generations.” (Idem, p. 504)
The Niger-Congo family tree:

Corollaries for **Human Language** evolution and complexity hierarchies

1. The Pidgin-to-Creole Cycle recapitulates the *transition from Homo Erectus’s pre-human protolanguage to Homo Sapiens’s first human language*.

2. “Creole languages are the missing linguistic fossils . . . the equivalent of the Galapagos to Darwin” *(Newsweek 1982)*

3. “[Creoles] most approximate some of the early languages... *Creoles . . . are the only languages which have started again*” *(New York Times, Science Section, 2001)*

4. “... of the well documented creoles, none equals the complexity [...] of a non-creole language.” *(Dixon 2010)*
**One benchmark: Haitian Creole**

1. One of the most studied Creole languages, and the one with the largest community of speakers, is Haitian Creole (“HC”).

2. A “most Creole of Creoles” (McWhorter 1998)

3. HC emerged in the 1650s from contact among:
   a) regional and colloquial varieties of French
   b) African languages from West and Central Africa (Niger-Congo)

---

**A map through the evidence:**

**Haitian Creole in a comparative perspective**

1. A “maracaronic” lexicon?
2. Familiar diachronic patterns in verb syntax
3. Creole Simplicity?
   a) Bit complexity?
   b) Creole simplicity as a reflex of early interlanguages?
4. Conclusion
### Lexicon of English as “genetic” language

**English morphemes with French etymology:**

1. beef (*bœuf*), veal (*veau*), pork (*porc*), . . . 
2. count (*comte*), noble (*noble*), parliament (*parlement*) . . . 
3. battle (*bataille*), ally (*allié*), alliance (*alliance*), admiral (*amiral*), march (*marche*), 
4. enemy (*ennemi*), peace (*paix*), . . . 
5. judge (*juge*), jury (*jury*); justice (*justice*), court (*court*), defendant (*défendant*), 
6. crime (*crime*), petty (*petit*), marriage (*mariage*), . . . 
7. clergy (*clergé*), altar (*autel*), miracle (*miracle*), pray (*prier*), sermon (*sermon*), 
8. virgin (*vierge*), saint (*saint*), friar (*frère*), . . . 
9. cuisine (*cuisine*), sauce (*sauce*), boil (*bouillir*), filet (*filet*), soup (*soup*), 
   pastry (*patisserie*), fry (*frire*), roast (*rôîr*), . . . 
10. -or (*-eur* as in *dictateur*), -tion (as in *célébration*), -ment (as in *déguisement*), 
11. -ee (as in *payée*), -able (as in *agréable*), -age (as in *espionnage*), . . .

### Lexicon of English as “genetic” language

**Morphemes with Latin etymology:**

1. kitchen (*coquina*), pan (*panna*), cup (*cuppa*), dish (*discas*), parsley (*petrosilium*) . . .
2. priest (*presbyter*), bishop (*episcopus*), nun (*nonna*), angel (*angelus*) . . .
3. ex- (ex as in *expatriare*), pre- (prae as in *præcursīor*), pro- (pro as in *proclāmare*), dis- (dis as in *disjunctio*), re-, (re as in *recreare*), inter- (inter as in *intermuralis*) . . .

*The lexicon of English is 65% non-Germanic, lots of it due to the Norman invasion in 1066.*
Lexicon of HC as “non-genetic” language

The bulk (> 90%) of the HC vocabulary, including its affixes, is derived from French:

1. -te (cf. Fr -té), as in bèlte ‘beauty’ (cf. bèl ‘beautiful’)
2. de- (cf. Fr de-), as in derespekte ‘to disrespect, to insult’ (cf. respekte ‘to respect’)
3. en- (cf. Fr in-), as in enkoutab adj. ‘foolhardy (cf. koute ‘to listen’)
4. ti- (cf. Fr petit ), as in Ti-Yèyèt
5. -ab (cf. Fr -able ), as in enkoutab
6. -ay (variant: -aj ), as in plasay (variant: plasaj)
7. -è (cf. Fr -eur ), as in djolè n. ‘boaster’ (cf. djòl n. ‘mouth’)
8. -en /-èn (cf. Fr. -en /-enne ), as in the pair Dominiken / Dominikèn n. (cf. Dominikani n. ‘Dominican Republic’)
9. -èt (cf. Fr -ette ), as in Boukinèt (cf. Bouki)
10. -man (cf. Fr -ment ), as in kozman n. ‘talk, gossip’ (cf. koze ‘to talk’)
11. -man (cf. Fr -ment ), as in kòrèkteman adv. (cf.kòrèk adj. ‘correct’)
12. -syon (cf. Fr -tion ), as in admirasyon (cf admire ‘to admire’)
13. -adò (cf. Spanish -dor ), as in babyadò n. ‘griper’ (cf. babye ‘to whine’)

Feb. 15, 2013
AAAS - Boston
Implications

1. English, a “genetic” language, has 65% of its lexicon with non-Germanic origins.

2. The etymology of most (over 90%) of HC lexicon is French.

3. The English lexicon is more “macaronic” than HC.

English has out-creoled Haitian Creole, even though Haitian Creole is called a “most Creole of Creoles”

(reductio ad absurdum—Q.E.D.)

More general implications

The “Gold Standard” of the Comparative Method applies whenever cognate sets can be reasonably established within a time depth of some 10,000 years (Dunn et al 2008; Wichmann & Saunders 2007)

1. Cardinal numbers from French: en ‘1’, de ‘2’, twa ‘3’, kat ‘4’ ... san ‘100’ ... mil ‘1,000’ ... from French un, deux, trois, quatre ... cent ... mille ...

2. Ordinal numbers, including the suffix /-jɛm/ and its morphophonology (sandhi, suppletion, etc.), from French: premye ‘1st’, dezyèm ‘2nd’, twazyèm ‘3rd’, katryèm ‘4th’, ... santyèm ‘100th’ ... milyèm ‘1,000th’ ... from French premeir, deuxème, troisième, quatrième, ... centième, ...

3. Kinship terms from French frè ‘brother’, sè ‘sister’, kouzen ‘cousin’, kouzin ‘cousin (feminine)’ ... from French frère, soeur, cousin, cousine ...

4. Color terms from French: blan ‘white’, nwa ‘black’, rouj ‘red’ ... from French blanc, noir, rouge ...

More “Gold standard” for HC

6. TMA markers from French: *te* ‘ANT’, *ap* ‘PROG, FUT’, *ava* ‘IRREALIS’, *fimi* ‘COMPLETIVE’ ... from French *étai*, *était*, *été* (imperfect and participle of ‘to be’), *après* ‘after’, *va(s)* ‘go +3sg/2sg+PRES’, *finir/fini(s)* ‘to finish’...

7. Prepositions from French: *nan* ‘in’, *pou* ‘for’, *apre* ‘after’, *anvan* ‘before’, *devan* ‘in front of’, ... from French *dans, pour, après, avant, devant, ...*

8. Determiners, demonstratives, etc., from French: *yon* ‘a’, *la* ‘the’, *sa* ‘this/that’ from French *un, la/là, ça*.

9. Pronouns from French: *m(wen)* ‘1sg’, *ou* ‘2sg’, *li* ‘3sg’, *nou* ‘1pl, 2pl, yo *3pl’ ... from French *moi, vous, lui, nous, eux* ...  

10. Complementizers from French: *ke* ‘that’, *si* ‘if’, *pou* ‘for’ ... from French *que, si, pour* ...

---

HC meets the Gold Standard!

11. Almost all derivational morphemes from French: for example, HC *de-* as in *deboutonen* ‘to unbutton’ and *dezose* ‘to debone’ from French *de-* which, like HC *de-*, has inversive and privative uses.

12. Morphophonological phenomena with French roots: liaison phenomena as in *an Bèljik* ‘in Belgium’ vs. *ann Ayiti* ‘in Haiti’; *de zan /de zā/ ‘two years’, *twa zan /twa zā/ ‘three years’, *san tan /sa tā/ ‘one hundred years’ ... (cf. the pronunciation of the HC and French ordinal and cardinal numbers above; Cadely (2002, 2003) provides further examples of HC-French correspondences in phonology)

These HC-French correspondences meet Nichols’ (2006) “individual-identifying” threshold: these correspondences contain a fair amount of “faits particuliers” (or “language-particular idiosyncratic properties” in Meillet’s terminology) that reliably rule out chance correspondences.
Verb placement in Haitian Creole

1. a) *Bouki *deja pale Kreyòl
   Bipiki "Bouki already speaks Creole"

b) Bouki *pa pale Kreyòl
   Bipiki NEG speak Creole ‘Bouki doesn’t speak Creole’

2. a) *Bouki pale deja Kreyòl
   b) *Bouki pale pa Kreyòl

3. ... Neg ... Adv ... [V V Obj ]

Verb placement in French

1. a) *Joujou parle déjà Créole
   *Joujou speak already Creole
   ‘Joujou already speaks Creole’

b) Joujou (ne) parle pas Créole
   *Joujou NEG speak NEG Creole
   ‘Joujou doesn’t speak Creole’

2. a) *Joujou déjà parle Créole
   b) *Joujou (ne) pas parle Créole

3. V fin ... Neg ... Adv ... [V V Obj ]
Key question:

Are the “discontinuities” in the history of Creole languages of a fundamentally distinct nature than the “discontinuities” that are found in the history of non-Creole languages?

Bouqui aimera Bouquinette = Bouqui [\(v\)aim + \(T^0\)er] + \(Agr^0\) Bouquinette
V-raising vs. V-in-situ across time and space

Key observation:

With respect to the syntax (and morphology) of verbs, there is no fundamental difference between “creolization” as in the history of HC vs. “language change” as in the history of English
How about complexity or lack thereof?

“Bit complexity” (à la McWhorter, Parkvall, Bakker et al, etc.): 

Counting of overt markings for a relatively small and arbitrary set of morphological and syntactic features

E.g.: gender, number, person, perfective, evidentiality ... number of vowels and consonants, number of genders ... suppletive ordinals, obligatory numeral classifiers ... passive, antipassive, applicative, alienability distinction, difference between nominal and verbal conjunction

Haitian Creole more “complex” than French?

HC:
Kouto sa a pa fèt pou koupe pen
knife this NEG make for cut bread
‘This knife is not made for cutting bread.’

Fr:
Ce couteau n’est pas fait pour couper le pain
this knife IS NOT made for cut the bread
‘This knife is not made for cutting bread.’
More “complex” CP layers in HC?

Kouto sa a pa fèt pou li koupe pen
knife THIS NEG make for 3SG cut bread
‘This knife is not made for cutting bread.’

Li te ale nan fèt la pou li te danse
3SG ANT go to party THE for 3SG ANT dance
‘S/he went to the party to dance’

More “complex” DPs in HC?

Jinyò achte yon (gwo) chwal (blan)
Junior bought a (big) horse (white)
Jinyò bought a (big) (white) horse

Jinyò achte (gwo) chwal (blan) an
Junior bought (big) horse (white) the
‘Jinyò bought the big white horse’

Jinyo rive a?
Junior arrive the?
‘Has Junior arrived, as expected?’
Bit complexity is *un*-interesting

“To prove the claim [that Creole grammars are overall simpler than non-Creole grammars], one would need to show that for every single subdomain of grammar (not just for an eclectic range of subdomain) all creoles score lower or equal to all non-creoles.” (Deutscher 2009: 250)

“Simplicity” as a feature of early interlanguages in L2A?

1. **The Early Interlanguage Hypothesis (EIH):** Creole formation mostly depends on processing procedures that are attested in the *earliest* stages of language acquisition by adults (Plag 2008).

2. **Corollary:** Creoles are exceptional languages, with origins from early L2A interlanguages in a fossilized state of ‘arrested development.’
Two processing procedures that are absent in *early* interlanguages *and* in Creoles (Plag 2008, 2009)

L2A sequence: $t_1 < t_2 < t_3 < t_4 < t_5$

1. Intra- and inter-phrasal information exchange: at $t_3$ and $t_5$, respectively (à la Pienemann 2008)
2. Clausal embedding: the “S’-procedure”, at $t_5$

**Our claim:** These “advanced” processing procedures exist in all natural languages, including Creoles—necessarily so (i.e., as part of $C_{HL}$).

### Pienemann’s (2008) exchange levels

<table>
<thead>
<tr>
<th>Stage</th>
<th>Loci of Exchange</th>
<th>Example</th>
<th>Illustration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence</td>
<td>Within sentence</td>
<td>Peter sees a dog</td>
<td><img src="image" alt="Sentence Example" /></td>
</tr>
<tr>
<td>Phrase</td>
<td>Within phrase only</td>
<td>Two kids</td>
<td><img src="image" alt="Phrase Example" /></td>
</tr>
<tr>
<td>Category</td>
<td>No exchange</td>
<td>Talk-ed</td>
<td><img src="image" alt="Category Example" /></td>
</tr>
</tbody>
</table>

Note: The illustrations show the syntactic structures corresponding to each stage of exchange.
Information exchange and inflectional morphology

1. We agree with the ELH assumption that information exchange is a necessary condition for the appearance of overt syntactically-determined inflectional morphology (= “contextual inflection”).

2. But information exchange is not a sufficient condition for inflectional morphology.

3. Booij 2005 includes cases of “agreement” (i.e., information exchange in syntax) without agreement morphology.

   “These examples [Dutch het as definite singular neuter determiner vs. de as default non-neuter determiner for definite NPs] illustrate that agreement is not always marked by means of morphology, but may also be marked through the choice of a specific lexical item”.

Agreement sans inflection in Haitian Creole (HC) and Saramaccan

1. senk liv yo */la (HC)
   five books DET.PL / DET.SG
   ‘the five books’

2. a) di man kodo womi (Saramaccan)
   the.SG one single man
   ‘the single/unique man’
   b) dee dii womi
   the.PL three man
   ‘the/these three men’

   (Rountree & Glock 1992; Aboh, to appear)
Agreement *plus* clausal embeddings in Saramaccan (Aboh, to appear)

1. **Di** fisí [\(CP_{di}\) mi tata kisi ] bigi
   the.SG fish that.SG my father caught big
   ‘the fish that my father caught is big’

2. **Dee** fisí [\(CP_{dee}\) mi tata kisi ] bigi
   the.PL fish that.PL my father caught big
   ‘the fish that my father caught are big’

   (Rountree & Glock 1992; Aboh, to appear)

Clausal embeddings in Creoles across time and across space

1671 Mermaid text from Martinique (Hazaël-Massieux 2008)

mouchié faire [\(TP_{yon autre negre}\) courir après li [\(CP_{pour prendre li avec ligne}\)]]
man make a other negro run after 3sg for take 3sg with line
‘The man made another negro run after it in order to catch it with a line’

Also consider: *Wh*-movement, focus constructions and predicate clefts across unbounded domains in Saramaccan (Aboh 2006), Jamaican Creole (Durrelman 2008), Capeverdean (Alexandre 2012) and various other Creoles (Byrne & Winford 1993, Holm & Patrick 2007, etc.)
Conclusion

The emergence of new languages and new language varieties sheds light on the interplay of first- and second-language acquisition as new grammars are built from complex and variable input.

THANKS TO:

• Enoch Aboh
• Trevor Bass
• Robert Berwick
• Noam Chomsky
• Morris Halle
• Salikoko Mufwene