Is phonology different?

Untangling a discrepancy in the “simplicity hypothesis”

Jeff Good
jcgood@buffalo.edu
University at Buffalo (The State University of New York)

1 Introduction

Background assumptions (not necessarily uncontroversial)

If not exactly “right”, work like McWhorter (2001a) touches upon something noteworthy.

Even if we only go so far as to say creoles are “locally simple” (Plag 2008:117), we’re still dealing with something pretty important.

Even if we weaken the claim further to limit it to “canonical” creoles, it’s still something to take note of.

The current state of the evidence suggests, in fact, we probably don’t need to weaken the claim (Parkvall 2008).

Framing the question

The useful question is not whether or not creoles are “simple” But, rather, how they are simple

Big-picture ideas

They instantiate some cognitively “default” grammar (Bickerton 1984)
They are “streamlined” in two ways (McWhorter 1998:792)

Product of a social environment not favoring grammatical exuberance
Lack of history (i.e., time) to acquire exuberance

Goals of (fairly programmatic) talk

Introduce problems raised for simplicity by “phonology”
Develop a typology of complexity
Apply it to selected cases of creole “simplicity”/“complexity”
Map out possible refinements to the “streamlining” research program

The “meta-goal”: Suggest a reorientation in debates about creole “simplicity”

Away from an unnecessarily reductionist debate centered around whether the claim is correct or incorrect...

Towards the more complex question: What does this sort of typologizing get us, and how can we make sure we get it right?

In short, the study of pidginization, by requiring us to study simplification, may lead to recognition of a sociolinguistic universal. Simplification may prove to be, not an isolated phenomenon, but one pole of a continuum applicable to outer form in all languages. (Hymes 1971:73).

2 Complex phonology?

From this perspective, it must be concluded that the Creole simplicity hypothesis makes the wrong predictions. The segmental inventories of Creole languages are not simple, notwithstanding a very small number of exceptions. On the other hand, they do not tend to be very complex either. Instead, Creole inventories have a strong affinity to the typological middle. (Klein 2006:18)

Perhaps the most common comment among these replies is that my address of phonology is superficial and preliminary. This is quite true. (McWhorter 2001b:391)

While I still believe that older languages are more complex overall than a crucial subset of creoles, if it is found that older languages can have less complex phonologies than creoles, then it will be interesting and fruitful to seek a principled reason why this should be so in this particular module of language but not others. (McWhorter 2001b:391)

Prosodic complexity


Good (2004, 2006) on Saramaccan tone and intonation

Less striking patterns, but still relevant studies: Gooden (2003), James (2003), Hualde and Schwegler (2008)
Is phonology different?

[13] Conclusion (to anticipate): Probably not. To find out, we need to make sure we’re not comparing complex apples with complex oranges.

[14] Parade example, segment inventories (McWhorter 2001a:138–139)


[r] Consonants

| p | t | c | ě | k | q | ň |
| p’ | t’ | ě’ | k’ | q’ |
| b | d | g |
| m | n |
| f | s | š | x | h | h |
| v | z | ž | y |

Plus: Phonemic length and pharyngealization

[v] Vowels

| i | u |
| e | o |
| ia | oe | uo |
| a |
| ae | aa | oaa |

Plus: Phonemic length and nasalization; various vowel combinations

[b] Saramaccan

[r] Consonants

| p | t | tj | k | kp~kw |
| b (ň) | d (õ) | dj | g | gb~gw |
| mb | nd | ndj | ng |
| m | n | nj |
| f | s | ň |
| v | z |
| l | j | w (hw) |

[v] Vowels

| i | u |
| e | o |
| r | œ |
| a |

Plus: Phonemic length and nasalization; various vowel combinations

[c] Rotokas (Firchow and Firchow 1969:273)

[r] Consonants

| p | t | k |
| ň | f | g |

[v] Vowels

| i | u |
| e | o |
| a |

Plus: Phonemic length and nasalization; various vowel combinations

[15] Parade morphology example, noun classes (McWhorter 2001a:139)

[a] Chechen

The Chechen noun class system (based on data from Johanna Nichols)

| SG | PL |
| D | D |
| B | B |
| J | J |
| V | |

Cickuo  ch’aara  bu’u.
cat.ERG fish.ABS B.eat.PRS
“The cat eats the fish.”

Ch’eeruo  cicig  du’u.
fish.ERG cat.ABS D.eat.PRS
“The fish eats the cat.”

Superficially looks somewhat “normal”—except that only a subset of verbs and adjectives actually show this alternation.
Is phonology different?

Bantu

The Proto-Bantu noun classes (simplified from Maho (1999:51)). Pairings indicate common singular/plural pairs.

<table>
<thead>
<tr>
<th>SINGULAR</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mü-</td>
<td>2 ṣa-</td>
</tr>
<tr>
<td>3 mü-</td>
<td>4 mi-</td>
</tr>
<tr>
<td>5 lì-</td>
<td>6 mà-</td>
</tr>
<tr>
<td>7 kì-</td>
<td>8 ṣi-</td>
</tr>
<tr>
<td>9 nì-</td>
<td>10 lì-nì-</td>
</tr>
<tr>
<td>11 lù-</td>
<td>13 tù-</td>
</tr>
<tr>
<td>14 ṣù-</td>
<td>15 kù-</td>
</tr>
<tr>
<td>16 pà-</td>
<td>17 kù-</td>
</tr>
<tr>
<td>18 mü-</td>
<td>19 pì-</td>
</tr>
</tbody>
</table>

Prefixal concord in Swahili (Katamba 2003:111)

- **M**-toto **m**-dogo *a-me-fika.*
  - 1-child 1-little 3s-TNS-arrive
  - “The little child arrived.”

- **Ki**-kapu **ki**-dogo *ki-me-fika.*
  - 7-basket 7-little 7-TNS-arrive
  - “The little basket arrived.”

Saramaccan

Saramaccan does not have noun classes in the usual sense.

The closest thing it has are words which could be analyzed as having vestigial noun class prefixes.

There are even a few words with “prefixed” and bare alternate forms, for example, *(a)kulî ‘Hindustani’, (a)masîni ‘machine’, (a)tengûengû ‘limping’, and (a)dikpôkpo ‘mushroom’.*

Are segment inventories and noun classes the same kind of thing?

3 The sociohistorical model

What do we mean by *pidgin*?

(i) Not mother tongue, (ii) medium for interethnic communication, (iii) normalized, (iv) not primary language of any speech community (see, e.g., Bakker (2003:4-5) for a brief overview)

For “streamlining”, the relevant stage cannot be a pidgin in this sense.

Rather it would be what has been termed the *jargon* stage, preceding a pidgin “in which people experiment with forms and structure, before any norms establish” (Bakker 2003:4).

“Pidginization”, then, refers not necessarily to the development of a pidgin used over a long period of time, but simply to a particularly extreme degree of reduction, unknown in language change unintermediated by a break in transmission. (McWhorter 2001b:406–4067).

See also Hymes (1971:70) on “imperfective” and “perfective” senses of pidginization and discussion of the related term pre-pidgin continuum.

So, we arrive at something like the following picture (with possible variations):
In pidgin genesis [=jargonization], all the participants are making guesses, and “right” guesses are those that are understood by everyone. (Thomason 1993).

Note that what we care about here is “language” as a social artifact, not a mental representation.

What’s important here is that the speech variety is not yet “normed”: There is not yet structure (or, at least, very much structure).

4 Typologizing complexities

Hypothetical scenario: Pidgin developing from interaction between two languages with clicks.

Language A word for ‘dog’: ʃa
Language B word for ‘dog’: ʃa
Pidgin word for ‘dog’: complex ʃa or simple ʃa?

Three kinds of complexities (based on Moravcsik and Wirth (1986:7))

Syntagmatic complexity: Complexity deriving from the structure of a given linguistic object.
Typological “markedness”: Not complexity, per se, but typological rarity or phenomena at the “top” of an implicational hierarchy (Haspelmath’s (2006:33–37) “markedness as abnormality” grouping).
Paradigmatic complexity: Complexity deriving from the range of subdistinctions available within a particular, grammaticalized (in a broad sense) linguistic category.

Some sample complications

Syntagmatic: A syntagmatically unmarked syntactic structure would hold a one-form–one-meaning pattern; discontinuous marking (e.g., French-style discontinuous negation) would constitute a syntagmatic complication.
Typological: A rara whose substance can be reasonably understood to be “difficult” in some sense, e.g., a labiodental flap, tonal TMA marking, German split word order.
Paradigmatic: The more members of the paradigm, the more complicated. Extreme examples are arbitrary declension/conjugation classes (as found in, for example, Latin).

Are all typological rara complicated?

It doesn’t seem possible to say given what we now know about areal skewing (see, e.g., Nichols (1992), Dryer (1989, 1992)).
Rara may be rare because they involve an improbable series of diachronic events, not because the end state is somehow “marked” (see Harris (2008)).

Complications of structure versus complications of substance

For syntagmatic and paradigmatic complexities, the substance of the relevant categories could be an extra, independent complication.
Discontinuous negation, for example, would probably be less substantively complex than, say, discontinuous marking of a discontinuous tense (e.g., the Dagaare hodiernal+remote past marker (Schwenter 1998)).
A voiced fricative will always be more complicated, in an articulatory sense than a voiceless fricative (see Ohala (1999)).

Structural/typological definition of jargon: The unstructured enumeration of linguemes that has been understood in a given jargon’s setting.

Lingueme: A unit of linguistic structure that can be inherited in replication (see Croft (2000:200–205))
Substance linguemes: A lingueme (e.g., a morpheme) with phonological substance (representable, for example, as a phonemic string)
Schematic linguemes: A lingueme without phonological substance into which substance linguemes can be placed (e.g., “SOV” or “CVCV”)

How complexities can be transferred

Syntagmatic: An input lingueme’s complex syntagmatic properties are introduced into the jargon and correctly interpreted by the jargon community.
Typological: An input lingueme’s rare typological properties are introduced into the jargon and correctly interpreted by the jargon community.
Paradigmatic: A set of linguemes evincing a paradigm’s (in a broad sense) structure is introduced into the jargon and the members of the set are correctly identified as belonging to the same abstract paradigmatic structure.

Another option for each case: misanalysis.
Number of linguemes necessary for a complexity to make it into a jargon

- Syntagmatic: One
- Typological: One
- Paradigmatic: The number of distinctions found in the paradigm

What’s a segment inventory

- A structured object representing the available segmental contrasts in a speech variety. (This what we’d find in a “language”)
- An enumeration of the segments that happen to be present in that variety. (This is what we’d find in a “jargon”)

How to get a non-simple segment inventory: Fill up the jargon with linguemes whose phonological forms, in total, contain more segments than a Rotokas.

How to get inherent inflectional morphology (see Booij (1996))

- Not required by syntactic context (paradigm example: nominal number)
- Heuristic lingueme pool in jargon: cat and cats
- Distinction: form1 vs. form2 linked to SINGULAR vs. PLURAL
- Minimum size of paradigmatic set: Two

How to get contextual inflectional morphology (see, again, Booij (1996))

- Appearance sensitive to syntactic context (paradigm example: case)
- Heuristic lingueme pool in jargon
  - Puer currit.
    - boy,NOM run,PRS.3S
    - “The boy runs.”
  - Puella puerum vidit.
    - girl,NOM boy,ACC see,PAST.3S
    - “The girl saw the boy.”
- Distinction: form1 vs. form2 linked to SUBJECT vs. OBJECT
  - form1 + SUBJECT = “nominative”; form2 + OBJECT = “accusative”

Compare Latin with Aghem (Hyman 1979:47–49)

- m mò bùnò
  - 1S.B DPST fall FOC
  - “I fell.”
- ò mò kòò mùè
  - 3S.B DPST see 1S.A
  - “He saw me.”
- à zìá mùè bèkò
  - DS eat,INC 1S.A fufu,B
  - “I am eating fufu.”

- Distinction: form1 vs. form2 linked to TOPIC vs. FOCUS.

Transfer of contextual inflection into a jargon requires

- Transfer of morphological paradigm
- Transfer of constructional paradigm

Acquiring singular/plural

- [ cat ]SINGULAR
- [ cats ]PLURAL

Acquiring nominative/accusative

- [ [ puer ]SUBJECT [ ]PREDICATE ]INTRANSITIVE
- [ [ ]SUBJECT [ puerum ]OBJECT [ ]PREDICATE ]TRANSITIVE

The easier route: acquiring the morphological paradigm, but no construction

- [ m ]TOPIC
- [ mò ]FOCUS

See also Manessy (1977:135–141) on formative invariability and paradigmatic and syntagmatic univocity in pidginization.

Under the interlanguage hypothesis put forward here the loss of inflectional categories and their exponent is the predictable consequence of universal constraints on language processing in early second language acquisition. (Plag 2008:131).
Is phonology different?

[47] Different predictions?

[a] Streamlining model: Paradigmatic complexities are lost. No predictions regarding presence/absence of other complexities.

[b] Interlanguage model: Paradigmatic and syntagmatic complexities are lost (?) (and possibly typological complexities, too, depending on one’s theory of grammar).

[48] The meaning of “⊕”

[a] Simplest possible meaning: “association” (see Gil (2008))

[i] ayam makan
    chicken eat
    (Riau Indonesian (Gil 2008:114–115))

[ui] “The chicken is eating.”

[ii] “Someone is eating the chicken.”

[iv] “Someone is eating for the chicken.”

[v] “the chicken that is eating”

[vi] “where the chicken is eating”

[vii] . . .

[viii] A(CHICKEN, EAT)

[b] More complex meanings

[i] [AGENT] ⊕ [PREDICATE] (e.g., John died.)

[ii] [MODIFIER] ⊕ [MODIFIED] (e.g., strong man)

[49] Compositional associational semantics would seem to be the “simplest” (i.e., least specified) logical possibility for concatenation (see Gil (2008:116)).


MOST ASSOCIATIONAL
Minangkabau
Sundanese
Yoruba
Vietnamese
Bislama
Cantonese
Twi
Papiamentu
Hebrew
English
MOST ARTICULATED

[51] Under the streamlining model, the meaning of concatenation in a jargon will presumably be subject to “compromise”—not simplification—effects

[52] And the data is suggestive of a compromise (but don’t take me too seriously)

[a] The Atlantic creole falls between the European and African languages

[b] The Pacific creole falls between the European and Austronesian languages

[53] See also Farasclas (2003) who documents significant syntagmatic complexity in Tok Pisin, attributing it to substrate languages.

[54] Is “tone” special?

[a] Of course, it depends on what you mean by “tone”

[b] A language with tone is one in which an indication of pitch enters into the lexical realisation of at least some morphemes. (Hyman 2006:229)

[c] Tone would seem to be “different” in the relevant way: The nature of its phonetic substance makes discovery of the full range of tonal contrasts in a language difficult without paradigmatic comparison.

[d] This also can be true of certain segmental contrasts, e.g., /b/ vs. /p/ vs. /p/. (Byman 2006:229)

[e] But it is pervasive for “true” tone (i.e., tone without substantial secondary accompaniments)

[f] Hypothetical examples: [tata /] vs. [tata ‘] vs. [tata ‘]

[g] Each involves a rise: So, when you encounter one of them on its own, is it tata, tata, or tata?

[h] A single lingueme can tell you some things about tone, but misclassification is almost a given unless the jargon contains a rich set of forms evincing the relevant tonal paradigms.

5 Conclusion

[55] So, is phonology different?

[a] Not yet.

[b] We need a better understanding of the structure of different complexities.

[56] Segment inventories ultimately derive from the importation of syntagmatic complexities into a jargon.
We need to look at paradigmatically complex phonological phenomena

(a) Morphophonological alternations
(b) Phonosyntactic alternations
(c) Unnatural phonological alternations

The kind of comparanda McWhorter should have used: The following data is from Nyakyusa. (See Hyman (2003:74) and Good (2007:214–215.).)

<table>
<thead>
<tr>
<th>ROOT TRANS</th>
<th>APPL</th>
<th>APPL-TRANS</th>
<th>ROOT GLOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sok-</td>
<td>-sos-y-</td>
<td>-sok-el-</td>
<td>-sok-es-y-</td>
</tr>
<tr>
<td>-lek-</td>
<td>-les-y-</td>
<td>-lek-el-</td>
<td>-lek-es-y-</td>
</tr>
<tr>
<td>-syt-</td>
<td>-syus-y-</td>
<td>-syt-el-</td>
<td>-syuk-es-y-</td>
</tr>
<tr>
<td>-kind-</td>
<td>-kus-y-</td>
<td>-kind-il-</td>
<td>-kik-is-y-</td>
</tr>
<tr>
<td>-jong-</td>
<td>-jos-y-</td>
<td>-jong-el-</td>
<td>-jok-es-y-</td>
</tr>
<tr>
<td>-ag-</td>
<td>-as-y-</td>
<td>-ag-il-</td>
<td>-ak-is-y-</td>
</tr>
<tr>
<td>-tup-</td>
<td>-tuf-y-</td>
<td>-tup-il-</td>
<td>-tuk-if-y-</td>
</tr>
<tr>
<td>-pub-</td>
<td>-puf-y-</td>
<td>-pub-il-</td>
<td>-puk-if-y-</td>
</tr>
<tr>
<td>-lim-</td>
<td>-lim-y-</td>
<td>-lim-il-</td>
<td>-lim-ik-is-y-</td>
</tr>
<tr>
<td>-lum-</td>
<td>-lum-y-</td>
<td>-lum-il-</td>
<td>-lum-ik-is-y-</td>
</tr>
</tbody>
</table>

And this is nothing compared to Tiene (Ellington 1997, Hyman and Inkelas 1997), which may have been in the Loango “catchment” area.

An important shift

(a) Bickerton’s (1984) simplicity hypothesis is grounded in the notion of “language as mental grammar”
(b) McWhorter’s simplicity hypothesis is grounded in the notion of “language as cultural artifact”

This is important since it already tells us that criticisms like Lefebvre (2001) and DeGraff (2001) misconstrue the relevant notion of simplicity.

Can a language’s sociohistorical profile be correlated with its grammar?

(a) Research agenda: Use languages whose history is well-known to discover correlations between sociohistorical profiles and grammatical types.

Antecedents can be found at least as far back as the 1930’s in work trying to explain divergences between Germanic and the rest of Indo-European (see Polomé (1996)).

The reason why pidgins and creoles are interesting here: We know more about their sociohistorical background than all but a handful of other languages.

Work like this is timely in light of a shift within the field of typology from the question “what’s possible” to “what’s where why” (Bickel 2007).

This is a potentially important line of research and is, therefore, worth getting right, which means:

(a) Properly circumscribing its explanatory possibilities
(b) Properly understanding the structure of a valid falsification

Cartesian-uniformitarian linguistics is anti-exceptionalist: it aims at understanding the speaking mind, and thus our very humanity, which includes the humanity of Creole speakers, notwithstanding the “accidents” of (post)colonial history. (DeGraff 2003:791)

But it is these “accidents” of history which make creole languages (or at least Saramaccan) exceptionally interesting:

Taking the long view, one might say that in 1690, Saramaka society and culture had not yet been born, but that by 1765, when the first window opened to the outside, it was already fully formed. The miracle of creolization, then, occurred over just several decades. (Price 2007:298)

Indeed, one might even suggest that societies born of creolization—creole societies—are not, as some would have it, unusually poor but unusually rich in cultural resources, in their cultural “building blocks” and “grammar”, and, especially, in the processes by which they play with, transform, and remodel these resources into something fresh. (Price 2007:298)
References


Is phonology different?

University of Toronto, 22 August, 2008


