

## Inflectional morphology through creolization: richness and predictability of verb inflection in Portuguese and Portuguese-related creoles

Alain Kihm (CNRS – Université Paris-Diderot)

### 1. Introduction

Creole languages are commonly considered to be morphologically much simpler than their lexifiers, verging on total absence of morphological processes.

Portuguese verb inflection: 3 finite modes (indicative, subjunctive, conditional), 4 tenses (present, simple past, perfect, future), 6 persons, “inflected” infinitive, and so forth. Apparently more complex than in creolized offsprings: all or most TMA and person-number values expressed by concatenating invariable verbs, “particles”, and clitic pronouns:

- |  |                                    |
|--|------------------------------------|
| (1) <i>Cantarei.</i><br>/kant-ar- e/       | PORTUGUESE                         |
| SING-STEM.FUT-1SG/<br>I will sing.         |                                    |
| (2) <i>N na kanta.</i>                     | KRIYOL (GUINEA-BISSAU / CASAMANCE) |
| 1SG PROG sing<br>I am singing / will sing. |                                    |
| (3) <i>N ta kanta.</i>                     | KRIYOL (GUINEA-BISSAU / CASAMANCE) |
| 1SG HAB sing<br>I sing.                    |                                    |

Morphological complexity confounds two dimensions: (i) RICHNESS, i.e. number of forms; (ii) PREDICTABILITY, i.e. possibility to predict form  $y$  knowing form  $x$  related to  $y$  (Blevins 2006). Cf. Latin and Wolof.

Latin verb-to-verb derivation limited: inchoatives with /-sk-/ suffix, e.g. *florescere* ‘to bloom’ (cf. *florere* ‘to be in bloom’); frequentatives with /-(i)t/ suffix, e.g. *clamitare* ‘to shout ceaselessly’ (cf. *clamare* ‘to shout’); desideratives suffixed with /-ur-/ , e.g. *esurire* ‘to be hungry’ (cf. *edere* ‘to eat’); diminutives with /-ill-/ , e.g. *sorbillare* ‘to drink a little, to sip’ (cf. *sorbere* ‘to drink’) (Leumann 1977: 536-555; Väänänen 1963:91). Wolof’s flourishing: 28 forms in Fal et al. (1990:27-29).

Wolof derivational morphology appears more complex than Latin’s in the sense of being richer (more forms).

Latin is the more complex in terms of predictability. Four derivations do not apply to all verb stems, and no formal or semantic rule for knowing to which stems they apply. Deverbal form never securely predictable: *iacere* ‘to throw’ > *iacitare* /jak-t-a-re/ ‘to throw repeatedly’, regular (predictable from the verb stem); *pellere* ‘to push’ > *pulsare* /pul-s-a-re/ ‘to hit’, involving alternative stem /pul-/ instead of /pell-/ and /-s-/ allomorph of the /-(i)t/ suffix.

Compare Wolof derivational network *dog* ‘to cut’, *dogat* ‘to cut up’, *dogaat* ‘to cut again’, *dogle* ‘to help cutting’, *dogloo* ‘to make cut’, etc. *Modulo* semantic compatibility, every verb base can be associated with every suffix. Derivation almost fully productive. Shape of derived forms always predictable. Wolof less complex than Latin as far as predictability is concerned.

Complexity and predictability are grammarian’s notions. Assume, however, that they entertain some relation to psycholinguistic, performance-based concepts such as complexity and efficiency in Hawkins’s (2004:8-10) sense.

“Complexity increases with the number of linguistic forms and the number of conventionally associated (syntactic and semantic) properties that are assigned to them when constructing syntactic and semantic representations for sentences” (Hawkins 2004:9). Number appears to be the key factor. Corrected, by the efficiency factor, “increased by selecting and arranging linguistic forms so as to provide the earliest possible access to as much of the ultimate syntactic and semantic representation as possible” (Hawkins 2004:9).

Predictable form such as Wolof *dogloo* fully satisfies “earliest possible access” requirement: if you already know what *dog* means and that *-loo* causativizes the verbs it attaches to – minimal demands on anyone who claims to know Wolof – *dogloo*’s meaning is immediately apparent. Not so with Latin *clamitare*: because derivation is not systematic, you first have to recognize that the *-it-* sequence is not part of the root, but has a specific meaning.

Wolof derivational paradigms are more efficient and more predictable than the Latin ones, while being numerically more complex.

When comparing creoles and lexifiers’ morphologies, authors are generally content with counting forms. Predictability must also be taken into account.

IDENTIFYING THE FORMS TO BE COUNTED IS NOT SELF-EVIDENT, BUT IT DEPENDS ON THE THEORETICAL FRAMEWORK ONE IS CONDUCTING THE STUDY IN.

## 2. What is a form ?

Cf. (1)-(3). *Cantarei*, a possibly complete utterance, is obviously a word. *N na kanta* is supposed to have to be a sentence. Forming *cantarei* is assumed to be a matter for morphology or word syntax (Halle & Marantz 1993; Ackema & Neeleman 2004); constructing *N na kanta* pertains to syntax *tout court*. Such a component interchange would constitute a characteristic feature of the creolization process : cf. Haitian *M tap chanté* vs. French *Je chantais*, Krio (Sierra Leone) *A bin sing* vs. English *I sang*, and so forth.

Such assumptions are actually grounded in a particular, CONSTRUCTIVIST view of the morphology of natural languages.

According to this view, *cantarei* results from successively adding /-r-/ and /-ei/ suffixes to the stem /kanta-/, obtained by adding /a/ theme vowel to root /kant/. Root and suffixes – not the stem: problem for constructivist approaches owing to meaningless or “morphomic” /a/ (Aronoff 1994) – are supposed to exist independently in the lexicon, out of which they are taken to be glued together according to the rules of word syntax, analogous, although not quite identical, to the rules of syntax proper.<sup>1</sup>

ABSTRACTIVE view – “Word-based Morphology” or “Word and Paradigm” (Blevins 2006) – is psychologically more realistic. It takes the WORD-FORM (Matthews 1974), e.g. *cantarei*, as its basic element. Word-forms only fully interpreted in contrast with other word-forms, with which they entertain paradigmatic relations, e.g. *cantarás* ‘you will sing’, *cantei* ‘I sang’, etc. PARADIGMS are the actual basic elements, i.e. sets or *séries associatives* (Saussure 1915/1982 :175) of forms sharing a root.<sup>2</sup>

Saussurean definition founded on mental association ⇒ paradigms can be inflectional (cf. *cantarei*, *cantarás*, etc.) or derivational (e.g. *sing*, *singer*, *song*). Only inflectional paradigms considered here.

	SINGULAR	PLURAL
1	<i>cantarei</i>	<i>cantaremos</i>
2	<i>cantarás</i>	<i>cantareis</i>
3	<i>cantará</i>	<i>cantarão</i>

Table 1 : Future indicative paradigm of Portuguese *cantar* (1<sup>st</sup> conjugation)

Table 1 = subparadigm within grand paradigm including all inflected forms of lexeme CANTAR (see Aronoff 1994 for a definition of the lexeme).

Abstractive approach: roots, stems and affixes have no independent reality anteceding the word-forms in which they appear; they are abstracted from the reunion of these forms like regularities over a backcloth of variations. Putting together *cantarei* ‘I will sing’ and *canto* ‘I sing’ brings to the fore both the regularity /kant/, minimal *signifiant* of lexeme-*signifié* CANTAR, and variations /are/ and /u/, exponents of the *signifiés* or feature sets {TENSE *future*, MODE *indicative* AGREEMENT *Isg*} and {TENSE *present*, MODE *indicative* AGREEMENT *Isg*}. Neither /kant/ nor /are/ nor /u/ exist outside of the word-forms that make them apparent.

<sup>1</sup> E.g., /-ei/, expressing the subject, follows the verb stem. Subject NP’s canonically precede verbs in Portuguese. Constructivist theories are also called “lexical incremental” (Stump 2001).

<sup>2</sup> Not a necessary condition : see suppletive forms within paradigms (e.g. *go* / *went*).

Psychological reality can be attributed to the abstractive approach from an acquisitional perspective. Children hear words, not roots and affixes. The existence of a specific language faculty remains an unproved hypothesis. The ability to perceive variations within repetition (the Art of the Fugue) pervades all our mental achievements beyond language (e.g. music, arithmetic, etc.), so we may be sure it is part of the human cognitive endowment.

Paradigms include not only “compact” words as in Table 1, but periphrastic constructs as well such as Portuguese *tenho cantado* ‘I have sung’, *hei-de cantar* ‘I will sing’, and so forth.

The root-feature associations that constitute paradigm cells can be realized according to two modes : synthetic as in *canto*; periphrastic as in *tenho cantado*.

“a lexeme’s compound tense constructions have the same morpholexical status as the synthetically realized members of its paradigm” (Ackerman, Stump & Webelhuth to appear).

“Where lexeme L belongs to category X and  $\sigma$  is a set of morphosyntactic properties associated with L, the realization  $w$  of the pairing  $\langle L, \sigma \rangle$  is synthetic and  $w$  may be inserted as the head of XP.” (Ackerman, Stump & Webelhuth to appear)

“Where lexeme L belongs to category X,  $\sigma$  is a set of morphosyntactic properties associated with L, and the realization  $w_1$  and  $w_2$  of the pairing  $\langle L, \sigma \rangle$  is periphrastic,  $w_1$  and  $w_2$  may be inserted as the heads of the respective phrases XP and YP. Language-specific stipulations may require XP and YP to form a nested constituent  $[_{XP} X YP]$  or may allow them to be independent and discontinuous from one another.” (Ackerman, Stump & Webelhuth to appear)

*Tenho cantado a cantiga* ‘I have sung the song’ can *a priori* be analysed as either (4) or (5) :

(4)  $[_{VP} [_{VP} [_{V} \textit{tenho}]]] [_{V'} [_{V} \textit{cantado}]] [_{NP} \textit{a cantiga}]]]$

(5)  $[_{VP} [_{V'} [_{V} \textit{tenho cantado}]]] [_{NP} \textit{a cantiga}]]]$

According to (4), *tenho* and *cantado* are two verb forms (one auxiliary, the other main) syntactically concatenated; according to (5), *tenho cantado* is one verb form. How do we tell periphrases from syntactic strings?

“(a) If an analytic combination C has a featurally intersective distribution, then C is a periphrase ; (b) if the morphosyntactic property set associated with an analytic combination C is not the composition of the property sets associated with its parts, then C is a periphrase; (c) if the morphosyntactic property set associated with an analytic combination C has its exponents distributed among C’s parts, then C is a periphrase.” (Ackerman & Stump 2004)

Fourth criterium: (d) What syntax may undo, syntax did it (see below).

Criteria apply simultaneously or separately. Criterion (a): C is a periphrase if the same feature set it expresses discontinuously is also expressed through compact forms in the language : cf. English *more important* vs. *bigger*.<sup>3</sup> Does not apply to *tenho cantado*: no simple verb form means the same. Applies to French *passé composé* (e.g. *j’ai chanté*), part of whose meaning intersects that of *passé simple* (e.g. *je chantai*), or to Catalan periphrastic past *vaig cantar* ‘I sang’, fully synonymous with synthetic past *cantí* ‘I sang’.

Criteria (b) and (c) apply to Portuguese *tenho cantado* : *tenho* = {TENSE *present*, MODE *indicative* AGREEMENT *1sg*} ; *cantado* refers to indeterminate past: cf. *a cantiga cantada o ano passado* ‘the song (that was) sung last year’ ; combination expresses perfect: refers to event that began to occur before speech time and is still taking place at speech time, or that took place during the same time period that also contains the speech referring to it.<sup>4</sup> Combination’s global meaning not reducible to composition {present + indeterminate past} (criterion a). Exponents of features {present} and {indeterminate past} associated each

<sup>3</sup> I am grateful to Olivier Bonami (p.c.) for clarifying this point.

<sup>4</sup> In Portuguese as in English, *Tenho cantado* ‘I have sung [during a natural time period where I still am, typically earlier today]’ thus contrasts with *cantei* ‘I sang’ [during a natural time period that I now consider fully over]. Cf. Comrie’s (1976:60) *Perfect of persistent situation*.



Contrast of Portuguese *cantarei* and Kriyol *N na kanta* is not a matter of swapping components, from morphology to syntax, but of changing the morphological expression mode, from the synthetic to the periphrastic.

Entails that Kriyol clitic subject pronouns, functional equivalents of Portuguese person suffixes, are part of the morphological periphrasis. Also includes Kriyol object enclitic pronouns (cf. *N na kanta-l* ‘I will sing it’) just as it hosts Portuguese enclitic and mesoclitic object pronouns (cf. *Canto-o* ‘I sing it’, *Canta-lo-ei* ‘I will sing it’).<sup>7</sup>

Syntactic structure of *N na kanta kantiga* ‘I’ll sing the/a song’ not as in (4), but as in (7), in which *N na kanta*, cell in the paradigm of the lexeme KANTA, heads the VP like *cantarei* does in (8):<sup>8</sup>

- |  |            |
|--|------------|
| (7) [S [VP [V' [V <i>N na kanta</i> ][NP <i>kantiga</i> ]]]]   | KRIYOL     |
| (8) [S [VP [V' [V <i>cantarei</i> ][NP <i>uma kantiga</i> ]]]] | PORTUGUESE |

### 3. A numerical comparison of the Portuguese and Kriyol verb inflections

#### 3.1. Portuguese

For the (European, Standard) Portuguese paradigms, see Williams (1944), Teyssier (1976), Mateus et al. (1989).

52 distinct synthetic forms for 1<sup>st</sup> conjugation active verbs.<sup>9</sup> 43 periphrastic compound past forms (e.g. *tenho cantado* ‘I have sung’); 6 periphrastic future *haver de*-infinitive forms (e.g. *hei-de cantar* ‘I’ll sing [some day]’); 160 synthetic future / conditional forms with infixed simple (accusative) or complex (dative + accusative) “mesoclitic” pronoun (e.g. *canta-lo-ei* ‘I’ll sing it’, *canta-lha-ia* ‘I would sing it to her/him’) = 261 forms.

Plus 3030 forms = 101 forms not in simple future / conditional × 30 enclitic simple and complex object pronouns (e.g. *canto-a* ‘I sing it’, *cantam-no-los* ‘they sing them to us’, *tem-na visto* ‘s/he has seen her’); 117 passive forms *ser* + past participle (e.g. *é cantado* ‘it is sung’, *tem sido cantado* ‘it has been sung’, *há-de ser cantado* ‘it will be sung’); 702 passive forms × 6 dative enclitic pronouns (e.g. *foi-me entregue* ‘it was handed to me’).<sup>10</sup>

**Total: 4110 distinct forms.**

#### 3.2. Kriyol

Kriyol TA marking: 3 primary forms = functional application to verbal lexemes of the feature sets {VFORM *finite*}, {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, ASPECT<sub>2</sub> *punctual*, TENSE *nonanterior*} and {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, ASPECT<sub>2</sub> *nonpunctual*, TENSE *nonanterior*} (Kihm 1994, chap. 3).

Punctuality and durativity (nonpunctuality) are independent from the perfective vs. imperfective contrast :

“[...] imperfectivity means viewing a situation with regard to its internal structure (duration, phasal sequences), and durativity simply refers to the fact that the given situation lasts for a certain period of time (or at least is conceived of as lasting for a certain period of time) [...] The opposite of durativity is punctuality, which thus means the quality of a situation that does not last in time (which is not conceived of as lasting in time), one that takes place momentarily.” Comrie’s (1976 :41-42)

Perfectivity “represents the action pure and simple, without any additional overtones” (Comrie 1976:21)., Perfectivity is always punctual in Kriyol. It is one possible reading of bare dynamic predicates, the default state of affairs for described actions or activities. Other interpretations in different contexts. Bare stative predicates denote present states, again the default reading. No expression for (*a priori* conceivable) nonpunctual perfectivity.

<sup>7</sup> I assume that Portuguese proclitic object pronouns (e.g. *Quer que o canto* ‘S/he wants me to sing it’) do not form a morphological unit with the verb. Kriyol only knows enclitic object pronouns.

<sup>8</sup> This makes Kriyol a pro-drop language like Portuguese.

<sup>9</sup> Only phonologically distinct forms are counted.

<sup>10</sup> I do not take into account the various gender and number forms of the past participle agreeing with the subject.

No incompatibility between imperfectivity and punctuality, if the latter also characterizes eventualities that last during a time that is (conceived of as) narrowly bounded, typically speech time (S) (Reichenbach 1980:287ff.) if a present eventuality.<sup>11</sup> Hence progressive reading of dynamic imperfective punctual predicates. With stative predicates, punctual imperfectivity is interpreted inchoatively or as a future. The latter reading is also possible for dynamic predicates. Pragmatic explanation: common sense and felicity conditions command it whenever the context, sentential (e.g. adverb such as *amañan* ‘tomorrow’) or discursive, reveals that the described eventuality is actually not taking place.<sup>12</sup>

Nonpunctual imperfective = habitual or iterative.

(9)  $V \sigma \{VFORM \textit{finite}\} \Rightarrow V$

(10)  $V \sigma \{VFORM \textit{finite}, ASPECT_1 \textit{imperfective}, ASPECT_2 \textit{punctual}, TENSE \textit{nonanterior}\} \Rightarrow na \oplus V$

(11)  $V \sigma \{VFORM \textit{finite}, ASPECT_1 \textit{imperfective}, ASPECT_2 \textit{non punctual}, TENSE \textit{nonanterior}\} \Rightarrow ta \oplus V$

Associating finite value of attribute VFORM with verbal lexeme V is realized as the phonological root or bare form of the given lexeme (V): e.g. *N kanta* ‘I sang’, *N sibi* ‘I know’, *I burmeju* ‘It is red’ (9). Associating finiteness, imperfective, punctual, and nonanterior with a verbal lexeme V is realized as the concatenation ( $\oplus$ ) of proclitic *na* and the bare form of the lexeme: e.g. *N na kanta* ‘I am singing’ or ‘I will sing’, *N na sibi* ‘I will know’, *I na burmeju* ‘It is getting red’ (10). Associating finiteness, imperfective, nonpunctual, and nonanterior with a verbal lexeme V is realized as the concatenation of proclitic *ta* and the bare form of the lexeme: e.g. *N ta kanta* ‘I sing’, *N ta sibi* ‘I know (habitually)’, *I ta burmeju* ‘Its is red (as an inherent property)’.

*V*, *na* $\oplus$ *V*, and *ta* $\oplus$ *V* are morphological periphrastic forms. They associate with the 6 proclitic subject pronouns to form 18 periphrastic forms:

	SINGULAR	PLURAL
1	<i>N na kanta</i>	<i>no na kanta</i>
2	<i>bu na kanta</i>	<i>bo na kanta</i>
3	<i>i na kanta</i>	<i>e na kanta</i>

Table 2 : punctual imperfective paradigm of *kanta*

By default, the intervals denoted by *na* $\oplus$ *V* and *ta* $\oplus$ *V* include speech time (S). Also by default, *V* refers to a time before S if the predicate denotes an action, to S itself if the predicate denotes a state.

Associating {TENSE *anterior*}, exponent  $\Theta ba$ , to the 3 primary forms cancels the default. Bare forms then refers to simple past with state verbs: e.g. *N sibi ba* ‘I knew’, to interval preceding some reference time (R) that itself precedes S with dynamic verbs: e.g. *N kanta ba* ‘I had sung’. Imperfective forms get translated to the past: e.g. *N na kanta ba* ‘I was singing’, *N ta kanta ba* ‘I used to sing’.

(12)  $V \{VFORM \textit{finite}, TENSE \textit{anterior}\} \Rightarrow V... \Theta ba$

(13)  $V \{VFORM \textit{finite}, ASPECT_1 \textit{imperfective}, ASPECT_2 \textit{punctual}, TENSE \textit{anterior}\} \Rightarrow na \oplus V... \Theta ba$

(14)  $V \{VFORM \textit{finite}, ASPECT_1 \textit{imperfective}, ASPECT_2 \textit{nonpunctual}, TENSE \textit{anterior}\} \Rightarrow ta \oplus V... \Theta ba$

18 more forms.

Dots in (12)-(14) show that *ba* may be separated from *V* by a complement of *V*. For many speakers, the complement may only be a clitic object pronoun as in *N kanta-l ba* ‘I had sung it’, not an NP.<sup>13</sup> These speakers analyse such forms as pronominally infixed periphrases, analogous to Portuguese *canta-lo-ei* ‘I will sing it’.

Object pronoun must not be featurally identical to subject pronoun– cf. *\*N oja-N* vs. *N oja ña kabesa* ‘I saw myself’ (lit. ‘I saw my head’).<sup>14</sup> Simply sharing person feature (not number feature) yields

<sup>11</sup> See below for past eventualities.

<sup>12</sup> The English progressive present also authorizes both readings depending on context.

<sup>13</sup> At least they never produce *V NP ba* constructions, although they accept them from other speakers (see below).

<sup>14</sup> Not the case in Portuguese: cf. *(Eu) vi-me* ‘I saw myself’.

grammatically dubious sentences – cf. ?*N oja-no* ?‘I saw us’, ?*Bu oja-bo* ?‘You saw you-all’. Hence a paradigm of 6 (subject proclitics) × 4 (object enclitics) × 3 (aspectual compositions) = 72 forms:

<i>N oja-u ba</i> ‘I had seen you’
<i>N oja-l ba</i> ‘I had seen her/him/it’
<i>N oja-bo ba</i> ‘I had seen you-all’
<i>N oja-elis ba</i> ‘I had seen them’

Table 3 : 1SG anterior perfective with infix object pronoun

Kriyol  $V\oplus$ pronoun sequences are periphrases like their Portuguese counterparts, hence 72 more forms (the same minus *ba*).

Other, less numerous speakers find it acceptable to insert object NP between *V* and *ba*:

- (15) *N tene bu foto ba* KRIYOL  
 1SG have 2SG.POSS picture ANT  
 I had your picture.<sup>15</sup>

Perhaps older stage of the language where *ba* is not an inflectional feature exponent, but a time adverb ⇒ 90 *ba*-forms (18 without infix object pronoun + 72 with) deducted from internalized grammars.<sup>16</sup>

However, (i) the same speakers also produce, in free variation,  $V\oplus ba$  NP constructions (e.g. *N tene ba bu foto*), the unmarked construction in present-day language; (ii) the size of the intrusive object NP is severely limited : bare noun or, at most, noun with one modifier (e.g. *bu foto* ‘your picture’); (iii) *ba*’s position is not free : immediately follows object NP.

Suggests that speakers do have the 90 *ba*-forms in their competences, but (a) they kept obsolescent adverbial *ba* in their lexicons next to modern inflectional (non-lexical) *ba* ; or (b) infixation is less constrained for them than for speakers who only accept clitic pronouns.<sup>17</sup>

Two periphrastic futures : punctual future  $na\oplus bin\oplus V$  and nonpunctual future  $na\oplus ba\oplus ta\oplus V$  (Kihm 1994:108-117).<sup>18</sup> With subject clitic pronouns = 12 forms × 4 (object enclitics) = 48 forms.<sup>19</sup>

Provisional total: 228 distinct forms.

Kriyol verbs inflect for 2 voices : passive with *-du* suffix; causative with *-nt/dV* suffix.<sup>20</sup>

- (16) *Kantiga kantadu.* KRIYOL  
 /kantiga kanta-du/  
 song sing- PASSIVE  
 The song was sung.
- (17) *E bibinti omi biñu.* KRIYOL  
 /e- bibi- nti omi biñu/  
 3PL drink-CAUS man wine  
 They made the man drink wine.

I assume causative voice to be as fully productive as passive, *modulo* semantic compatibility: shunned whenever predicate’s denotation does not fit well with notion of direct, physical causation morphological causativization implies in contradistinction with syntactic causativization, also available in Kriyol (Kihm 1994:246ff.; Biagui 2010).

Passive and causative are compatible each with the 6 persons, the 3 aspects, the anterior tense, and the 2 periphrastic futures, hence 72 forms : e.g. *Kantiga na kantadu ba* ‘The song was being sung’, *No na bin*

<sup>15</sup> I am grateful to Jean-Louis Rougé for this example.

<sup>16</sup> For *ba*’s origin, see Kihm (1994:102-103) and Rougé & Kihm (2008).

<sup>17</sup> NP infixation does not endanger the periphrastic status of  $V(\dots)ba$ , as periphrases can be discontinuous (see §2).

<sup>18</sup> As an autonomous verb, *bin* means ‘to come’. *Ba* is a reduced form of *bay* ‘to go’. The overall meaning is partially compositional, but not quite.

<sup>19</sup> These forms cannot be translated to the past through *ba*.

<sup>20</sup> Prenasalized coronal stop unvoiced if onset of final syllable of verbal base includes voiced C, voiced otherwise. Vowel of suffix copies last vowel of base.

*bibinti-l* ‘We’ll make her/him drink’, etc., × 4 (object enclitics) = 288 new forms. Causatives can be passivized = 72 forms:

- (18) *Omi bibintidu.* KRIYOL  
 /omi bibi- nti- du/  
 man drink-CAUS-PASSIVE  
 They made the man drink. (Lit. “The man was made to drink”.)

660 forms × 2, predicate negation *ka* being prefixal:

- (19) *N ka na kanta-l.* KRIYOL  
 1SG NEG PROG sing- 2SG.O  
 I’m not singing it / I won’t sing it.

Makes Kriyol typologically similar to Atlantic and Mande languages it has been in contact with for several centuries. All include affixal predicate negations syncretized with TA marking:

- (20) *Woyuma.* WOLOF  
 /woy-u- ma/  
 sing- NEG.PF-1SG  
 I didn’t sing.

**Total: 1320 distinct forms.**

About 1/3 of the 4110 Portuguese forms. Very far from the crushing imbalance of several thousands to one invoked by the supporters of creole extreme morphological poverty for whom only synthetic forms call for a morphological analysis.

Extinction of 2 forms out of 3 is not due to sheer reduction in the inventory of TA and person-number features of the lexifier: most of the synthetic expressions of the latter are replaced by equivalent periphrastic expressions: cf. *Cantarás* vs. *Bu na kanta* ‘You will sing’.

Primary responsible: decrease in number of TA feature sets with differentiated expressions: absence of subjunctive and conditional forms in Kriyol. About 30 TA feature sets of Portuguese; 8 in Kriyol, discounting voices:

- i. {VFORM *finite*} (*V*)
- ii. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*} (*ta V*)
- iii. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, ASPECT<sub>2</sub> *punctual*} (*na V*)
- iv. {VFORM *finite*, TENSE *anterior*} (*V... ba*)
- v. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, TENSE *anterior*} (*ta V... ba*)
- vi. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, ASPECT<sub>2</sub> *punctual*, TENSE *anterior*} (*na V... ba*)
- vii. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, TENSE *future*} (*na ba ta V*)
- viii. {VFORM *finite*, ASPECT<sub>1</sub> *imperfective*, ASPECT<sub>2</sub> *punctual*, TENSE *future*} (*na bin V*)

Other culprits: absence of complex enclitic pronouns such as *lho* ‘it to her/him’; accusative-dative non-distinction in 3<sup>rd</sup> persons pronouns : cf. Portuguese accusative *o(s)/a(s)* vs. dative *lhe(s)* where Kriyol only has *l* and *elis*;<sup>21</sup> loss of gender contrast splitting 3<sup>rd</sup> person object clitics in Portuguese.

Extinctions were not so catastrophic as is often claimed. Numerical comparison alone, i.e. merely counting forms, does not suffice to contrast Portuguese with Kriyol.

#### 4. Comparing Portuguese and Kriyol verb inflection for predictability

<sup>21</sup> Kriyol equivalent of *Dei-lho/a* ‘I gave it to her/him’ = *N da-l el*. Enclitic = addressee, autonomous pronoun = object. Singular *l* ‘her/him/it’ refers indiscriminately; plural *elis* ‘them’ tends to refer to persons only.

Predictability: ratio from the number of forms that have to be listed – i.e. stored in the lexicon and assumedly memorized as such – to the number of forms that can be deduced from the listed ones (Blevins 2006).<sup>22</sup> If few listed forms allow one to deduce many or all the other forms, predictability is strong. If many forms have to be listed and few are deducible, this makes for a relatively opaque system. Opacity is the opposite of predictability.

Numerical richness and predictability are independent dimensions. English verb inflection is scarce. Yet, strong verbs are opaque: no form in paradigms such as *sing / sang / sung* can be predicted from the other forms: compare *bring / brought*. Weak verbs are predictable provided one knows roots-stems (necessarily listed) and feature exponents (/ -s/, / -ed/, / -ing/, / will/, / would/, / have... -ed/, etc.).

#### 4.1. Portuguese

Portuguese verb inflection is exceedingly rich. Opacity has nothing to do with amount of forms. First cause: 3 inflection classes characterized by 3 theme vowels: /a/ for 1<sup>st</sup> conjugation, /e/ for 2<sup>nd</sup> conjugation, /i/ for 3<sup>rd</sup> conjugation. The shape of many TMA exponents depends on which conjugation a verb belongs to: cf. *cantei* ‘I sang’ vs. *parti* ‘I left’. Only knowing the forms which show the theme vowel allows one to predict the exponents. 1SG indic. pres. *parto* ‘I leave’ does not reveal the verb belongs to 3<sup>rd</sup> conjugation: cf. 1<sup>st</sup> conjugation *canto* ‘I sing’ and 2<sup>nd</sup> conjugation *temo* ‘I fear’; 2SG *partes* ‘you leave’ excludes 1<sup>st</sup> conjugation (cf. *cantas* ‘you sing’), but leaves open possibility of 2<sup>nd</sup> conjugation (cf. *temes* ‘you fear’); and so forth.

Nonpredictability also comes from having several themes even for “regular” verbs such as *cantar*: 1SG indic. pres. *canto* ‘I sing’ implies theme /kant/, similar to the root (exponent /-o/); 2SG *cantas* ‘you sing’ implies theme /kanta/ (exponent /-s/) – the same for the other persons. In simple past, theme /kant/ in 1SG and 3SG *cantei* ‘I sang’ and *cantou* ‘s/he/it sang’, /kanta/ in other persons: *cantaste* ‘you sang’, *cantámos* ‘we sang’, *cantaram* ‘they sang’. Future *cantarei*, *cantarás*, etc. built on theme /kantar/ identical to infinitive: cf. mesocliticized forms such as *olhar-vos-ei* ‘I will look at you’.

Things are even more complex with irregular verbs. *Ferir* ‘to wound’ shows theme /fir/ in the 1SG indic. pres. (*firo* ‘I wound’) and every person of subj. pres. (*fira* ‘that I wound’, *firas* ‘that you wound’, etc.), but /fer/ elsewhere: cf. *feres* ‘you wound’, *feriu* ‘s/he/it wounded’, etc. Many verbs (by no means all!) with the same morphophonological pattern (-ir infinitive and /e/ in penultimate syllable) behave like *ferir*. Still more complex is *trazer* ‘to bring’ and its 4 themes: /traz/, /trag/ (e.g. *trago* ‘I bring’), /tros/ (e.g. *trouxe* ‘I brought’) and /trar/ (e.g. *trarei* ‘I’ll bring’).

Opacity is brought about not only by the multiplicity of themes, but also by their being variably assigned to diverse subparts of the paradigm. These two factors plus the inflectional classes contribute to load Portuguese verb inflection with a significant measure of opacity, such that more than one form must be listed for every verb in any event.

#### 4.2. Kriyol

Kriyol stands below Portuguese for inflectional richness, but above English (lesser degree of periphrase grammaticalization in the latter). Kriyol verb inflection exhibits total predictability of English weak verbs. Thanks to the absence of inflectional classes and the unvarying identity of the verb theme with the root, knowing the latter and the exponents of the morphosyntactic feature sets is always enough to predict all cells of the global paradigm.<sup>23</sup> Major separation between Kriyol and Portuguese.

Such an absence of morphological opacity does not constitute a general property of creole languages, contrary to widespread assumptions (e.g. McWhorter 1998). Cf. Korlai Indo-Portuguese: verbs belong to 4 inflectional classes that determine the past tense form: cf. *yo kato* ‘I sung’ (*kata* ‘to sing’), *yo kumew* ‘I ate’ (*kume* ‘to eat’), *yo subiw* ‘I went up’ (*subi* ‘to go up’), *yo tapu* ‘I warmed’ (*tapu* ‘to warm’) (Clements 2007; Luís 2008).

Mauritian verbs show two syntactically conditioned forms: long form (LF, e.g. *chante*) and short form (SF, e.g. *chant*). SF cannot always be predicted from LF via deletion of final vowel as in *chante* vs. *chant*: cf. *rantré / rant* ‘to come in’, *tonbe / tom* ‘to fall’, *reste / res* ‘to stay’, etc. (Baker 1972; Bonami & Henri 2010).

<sup>22</sup> Being deducible does not entail that forms are not listed as well, especially if they are frequent (Baayen et al. 2003).

<sup>23</sup> The vowels ending Kriyol verbs (e.g. *kanta* ‘to sing’, *sibi* ‘to know’, *kume* ‘to eat’, etc.) do not signal distinct classes, they are simply part of the root.

Plural formation in Nubi, Arabic-related Creole spoken in Uganda and Kenya, mobilizes 7 suffixes, whose assignment to a given noun root is unpredictable (Wellens 2005 ; Kihm 2011, to appear).

In contrast, Afrikaans, commonly viewed as a strongly decreolized creole or “semi-creole” (Holm 2000), shows a verb inflection that looks as predictable as that of Kriyol and is much more streamlined: 4 forms at most: uninflected form meaning pres. and imperative, e.g. *ek sing* ‘I sing’, *sing!* ‘sing!’;<sup>24</sup> 3 possibly periphrastic forms: past (e.g. *ek het gesing* ‘I sang’), future (e.g. *ek sal sing* ‘I’ll sing’), conditional (e.g. *ek sou sing* ‘I would sing’) (van Schalkwyk 1988).<sup>25</sup> If the same doubts concerning the degree of grammaticalization can be raised against these forms as against their English counterparts, then the whole of Afrikaans verbal inflectional morphology consists in only one form, which makes it perfectly predictable indeed.

## 5. Conclusion : What does the present account imply for creole formation?

If periphrastic forms proceed from the morphological component just like synthetic forms do, creole languages can no longer be considered morphologically deprived. Significant differences do exist between creoles, some more minimalist (e.g. Saramaccan), others more luxuriant (e.g. Korlai). Never, however, do we descend to the lone or two or three forms conceded by scholarly tradition.

Complete predictability is not a general property of creole languages. Although comparison with the lexifiers always betrays greater global transparency of the creole paradigms, this may encapsulate zones of equal or even greater opacity: Mauritian short forms are opaque to a degree which finds no equivalent in the French verb paradigms (Bonami & Henri 2010).

Such observations throw serious doubt on theories according to which creoles realize some unmarked version of Universal Grammar (Bickerton 1984, 1988) or a primeval, unadulterated form of prototypical human grammar (McWhorter 1998). Both theories share the notion that morphological processes constitute an accretion, an outgrowth due to the outrages of time, of which younger languages ought to remain largely or entirely unencumbered. In that, they are faithful to the “syntactocentrist” outlook (Jackendoff 1997) which has been dominating language studies for the past fifty years. For the paradigm-centred framework applied in the present study, in contrast, morphological processes are an essential component of human language, as well as a distinct component with respect to syntax.

Given this, the true question raised by creole emergence is: why do creole languages favour periphrastic expression to the expense of the synthetic expression that is dominant or at least well entrenched in the lexifiers? In other words, why do they give preference to transparency without giving up paradigm richness?

A satisfactory answer to these questions is provided by the hypothesis that creole emergence results primarily (as well as rarely, due to special circumstances) from an unguided second language acquisition (SLA) process involving adults (Becker & Veenstra 2004; Plag 2008).

Adults have lost a capacity children acquiring their native language exercise prominently: the capacity to recognize and effortlessly assimilate vast synthetic paradigms in which relations between cells are either opaque to varying degrees (as in Portuguese *firo* vs. *feres*, or *trago* vs. *trouxe*, or *sou* ‘I am’ vs. *era* ‘I was’) or predictable, but based on small form variations (as in *cantas* vs. *canta*). It is therefore fully understandable that most or even all of these forms disappear in the course of the SLA process, especially when conditions are such that the process cannot go to its end (learning the L2 in a quasi-perfect fashion), but stops at an “interlanguage” that is sufficiently elaborated for the needs of everyday, practical communication, but not more (Selinker 1972 ; see the notion of “Basic Variety” in Klein & Perdue 1997). Yet, all the concepts these unacquired forms expressed remain indispensable for the development of the complex interactions and symbolic productions (narratives, tales, songs, etc.) a reformed linguistic community requires. Periphrases, taken over from the lexifier and/or created via grammaticalization (Bruyn 2008), are the means to overcome this challenge.

<sup>24</sup> Afrikaans personal pronouns cannot be analysed as clitics.

<sup>25</sup> The status of forms such as *wees besig om te sing* ‘to be singing’ or *gaan sing* ‘be going to sing’ is unclear. That the auxiliary and the main verb can be separated by complements (e.g. *Ek het ‘n lied gesing* ‘I sang a song’) does not run counter to the periphrastic character of the form (Ackerman et al. to appear).

## References

- Ackema, Peter & Ad Neeleman (2004). *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.
- Ackerman, Farrell & Gregory T. Stump (2004). Paradigms and periphrastic expression. In L. Sadler & A. Spencer, *Projecting Morphology*, 111-157. Stanford: CSLI Publications.
- – & Gert Webelhuth (à paraître). Lexicalism, periphrasis and implicative morphology. In B. Borsley & K. Börjars (eds), *Non-transformational Theories of Grammar*. Oxford: Wiley-Blackwell.
- Aronoff, Mark C. (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge (Mass.): MIT Press.
- Baayen, R.H., J.M. McQueen, T. Dijkstra & R. Schreuder (2003). Frequency effects in regular inflectional morphology: revisiting Dutch plurals. In R.H. Baayen & R. Schreuder (eds), *Morphological Structure in Language Processing*, 355-390. Berlin: Mouton de Gruyter.
- Baker, Peter (1972). *Kreol. A Description of Mauritian Creole*. London: C. Hurst & C°.
- Becker, Angelika. & Tonjes Veenstra (2004). The survival of inflectional morphology in French-related Creoles: The role of SLA processes, *Studies in Second Language Acquisition*, 25:283-306.
- Biagui, Noël Bernard (2010). L'expression du factitif en casamançais et dans les autres créoles apparentés. Paper to the 10th ACBLPE Conference, Villejuif (France), 1-3 July 2010.
- Bickerton, Derek (1984). The Language Bioprogram Hypothesis, *Behavioral and Brain Sciences* 7, 173-188.
- (1988). Creole languages and the bioprogram. In F. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, vol. 2, 267-284. Cambridge: Cambridge University Press.
- Blevins, James P. (2006). Word-based morphology, *Journal of Linguistics* 42, 531-573.
- Bonami, Olivier & Fabiola Henri (2010). How complex is creole inflectional morphology? The case of Mauritian. *14<sup>th</sup> International Morphology Meeting*, Budapest, May 2010.
- Bruyn, Adrienne (2008). Grammaticalization in pidgins and creoles. In S. Kouwenberg & J.V. Singler (eds), *Handbook of Pidgin and Creole Studies*, 385-410. Oxford : Wiley-Blackwell.
- Cantel, R. (1962). *Précis de grammaire portugaise*. Paris : Librairie Vuibert.
- Clements, J. Clancy (2007). Korlai (Creole Portuguese). In J. Holm & P.L. Patrick (eds), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars*, 153-173. London: Battlebridge.
- Comrie, Bernard (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge : Cambridge University Press.
- Enç, Mürvet (1987). Anchoring conditions for tense, *Linguistic Inquiry* 18, 633-657.
- Fal, Aram, Rosine Santos & Jean-Léonce Doneux (1990). *Dictionnaire wolof-français*. Paris: Karthala.
- Fradin, Bernard (2003). *Nouvelles approches en morphologie*. Paris : PUF.
- Gazdar, Gerald, E. Klein, Geoffrey Pullum & Ivan Sag (1985). *Generalized Phrase Structure Grammar*. London: Basil Blackwell.
- Halle, Morris & Alec Marantz (1993). Distributed morphology and the pieces of inflection. In K. Hale & S.J. Keyser (eds), *The View From Building 20*, 111-176. Cambridge (Mass.): MIT Press.
- Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Holm, John (2000). *An Introduction to Pidgins and Creoles*. Cambridge : Cambridge University Press.
- Jackendoff, Ray (1997). *The Architecture of the Language Faculty*. Cambridge (Mass.): MIT Press.
- Kihm, Alain (1994). *Kriyol Syntax: The Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam: Benjamins.
- (2011). Pidgin-creoles as a scattered sprachbund: Comparing Kriyol and Nubi. In P. Bhatt & T. Veenstra (eds), *Creoles and Typology, Journal of Pidgin and Creole Studies* 26:1 (special issue), 43-88.
- (to appear) Plural formation in Nubi and Arabic: A comparative study and a word-based approach, *Brill's Annual of Afroasiatic Languages and Linguistics* 3: 1-20.
- & Jean-Louis Rougé (2011). O verbo nu em Kriyol (Guiné-Bissau / Casamansa). Paper presented to the 11<sup>th</sup> ACBLPE Conference, Porto, 7-8/07/11.
- Klein, Wolfgang & Clive Perdue (1997). The Basic Variety, or couldn't natural languages be much simpler? *Second Language Acquisition Research* 13(4), 301-347.
- Leumann, M. (1977). *Lateinische Laut- und Formenlehre. Lateinische Grammatik von Leumann-Hofmann-Szantyr*, 1er Band. München: C.H. Beck'sche Verlagsbuchhandlung.

- Luís, Ana R. (2008). Tense marking and inflectional morphology in Indo-Portuguese creoles. In S. Michaelis (ed.), *Roots of Creole Structures: Weighing the contribution of substrates and superstrates*, 83-121. Amsterdam: Benjamins.
- Mateus, Maria Helena Mira, Ana Maria Brito, Inês Duarte & Isabel Hub Faria (1989). *Gramática da Língua Portuguesa*. Lisboa: Caminho.
- Matthews, P.H. (1974). *Morphology: An introduction to the theory of word-structure*. Cambridge: Cambridge University Press.
- McWhorter, John H. (1998). Identifying the creole prototype: vindicating a typological class, *Language* 74: 788-818.
- Plag, Ingo (2008). Creoles as interlanguages: Inflectional morphology, *Journal of Pidgin and Creole Languages* 23(1), 114-135.
- Pollard, Carl. & Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.
- Reichenbach, Hans (1980). *Elements of Symbolic Logic*. New York: Dover Publications.
- Rougé, Jean-Louis & Alain Kihm (2008). A origem da marcação da anterioridade *ba(n)* no crioulo da Guiné e Casamansa (Kriyol), paper presented to the ACBLPE Conference, Lisbon, July 2008.
- Saussure, Ferdinand de (1915/1982). *Cours de linguistique générale*. Paris: Payot.
- Selinker, L. (1972). Interlanguage, *International Review of Applied Linguistics* 10, 209-231.
- (1992). *Rediscovering Interlanguage*. London: Longman.
- Stump, Gregory T. (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Teyssier, Paul (1976). *Manuel de langue portugaise (Portugal – Brésil)*. Paris: Klincksieck.
- Väänänen, V. (1963). *Introduction au latin vulgaire*. Paris: Klincksieck.
- van Schalkwyk, H. (2008). *Teach Yourself Afrikaans*. Sevenoaks: Hodder & Stoughton.
- Wellens, Ineke (2005). *The Nubi Language of Uganda: An Arabic Creole in Africa*. Leiden: Brill.
- Williams, Edwin B. (1944). *An Introductory Portuguese Grammar*. New York: F.S. Crofts & C°.