On the origin of Creoles: A Cartesian critique of Neo-Darwinian linguistics

On the origin of creoles:
A Cartesian critique of Neo-Darwinian linguistics
by *Michel DeGraff*

1. Neo-Darwinian creolistics: Whence and whereto? A sketch

The main goal of this essay is to constructively deconstruct the age-old myth that "the world's simplest grammars are creole grammars" and to demystify the methodological (mis)practices that underlie this myth and its corollaries throughout creole studies and beyond.

I start with some notes on historiography and methodology, connecting certain trends in 20th- and 21st-century creolistics to outdated (quasi) Darwinian concepts in early 19th-century comparative-historical linguistics. Then I move to linguistics per se, inspecting the empirical and theoretical bases of creolists' foundational assumptions about creole diachrony and synchrony. This will (re-)establish the epistemological limits of certain key terms in creole studies, including "pidgin(ization)", "creole/creolization", "young" vs. "old", "simple(st)" vs. "(most) complex", etc. I will argue that these terms, although per-

haps useful as a-theoretical heuristics and as sociohistorical approximations, cannot serve as theoretically-grounded linguistic-structural taxa: Given Universal Grammar and its Cartesian-Uniformitarian foundations (Chomsky 1966, 1981, 1986, 1995, etc.), there cannot be any invariant and sui generis set of structures and processes that fall under the labels "creole" and "creolization". Assuming Universal Grammar, creolization reduces at the individual level to the same sort of cognitive processes that underlie idiolect formation through language change, and so are creole languages aprioristically undistinguishable from non-creole languages – that is, there is no synchronic creole typology that excludes non-creole languages. As I show below, such claims go against the grain of the most ancient and the (still) most prevalent dogma in creole studies.

1.1.   *Historiography and epistemology: From Schleicher to Popper*

My essay is best introduced by the following quotes from Schleicher (1863), Saint-Quentin (1872), and Adam (1883) on simple(st) grammars and from Osthoff & Brugman (1878), Foucault (1972), and Popper (1965) on methodology. The first two sets of quotes ((1)–(2)) are words from the past and the last three ((3)–(5)) are words of caution for the future, and all five are relevant to linguists' time-honored search for simplest grammars. These quotes speak for themselves, and eloquently so.

1.1.1.   *(Pre-)Darwinian linguistics: Schleicher on the "Tree of Language"*

(1)   a.   THE RULES NOW, WHICH DARWIN LAYS DOWN WITH REGARD TO THE SPECIES OF ANIMALS AND PLANTS, ARE EQUALLY APPLICABLE TO THE ORGANISM OF LANGUAGES, THAT IS TO SAY, AS FAR AS THE MAIN FEATURES [of Darwin's theory] ARE CONCERNED.   (Schleicher 1863 [1983: 30], emphasis in original; also see 1983: 16–17)

        b.   The construction of all languages points to this, that the eldest forms were in reality alike or similar; and those less complex forms are preserved in some idioms of the simplest kind, as, for example, Chinese. [...] In this remote stage of the life of speech, there is consequently no distinction [...] between verbs and nouns; there is neither declension nor conjugation.[1]
                                                (Schleicher 1863 [1983: 51])

---

1. Darwin himself was somewhat ambivalent about the use of morphosyntax as a measure of complexity cum perfection; see, e.g., Darwin (1871: Chapter 2, 61–62).

    c.   [L]anguage is of significance not only for the elaboration of a scientific [i.e., taxonomic] systematization of humanity, but also for the evolutionary history of man. [...] [T]he various stages of languages are to be considered as the perceptible, characteristic traits of various grades of man. [...] Now language has revealed itself to science as something that has evolved very gradually [...] The comparative anatomy of languages shows that the more highly organized languages evolved very gradually out of simpler language organisms, probably in the course of very long time spans. (Schleicher 1863 [1983: 79])

1.1.2. *Neo-Darwinian creolistics: Saint-Quentin and Adam on simple(st) languages*

(2)   a.   [Creole grammar] is, therefore, a spontaneous product of the human mind, freed from any kind of intellectual culture. [...] When one studies its structure, one is so very surprised, so very charmed by its rigor and simplicity that one wonders if the creative genius of the most knowledgeable linguists would have been able to give birth to anything that so completely reaches its goal, that imposes so little strain on memory and that calls for so little effort from those with limited intelligence. An in-depth analysis has convinced me of something that seems paradoxical. Namely: if one wanted to create *ab ovo* an all-purpose language that would allow, after only a few days of study, a clear and consistent exchange of simple ideas, one would not be able to adopt more logical and more productive structures than those found in creole grammar.
(Saint-Quentin 1872 [1989: 40–41]; my translation)

    b.   [Cayenne Creole] grammar [...] is nothing but the grammar that is common to the languages of Guinée. The latter we can call *langues naturelles* as opposed to *langues cultivées*. For the botanist, plants that are *naturelles* are superior to plants that are *cultivées* to the extent that the former are pristine products that are free of intentional adulteration. Likewise, for the linguist, the speech of peoples considered primitive has primacy over the speech of civilized peoples: the former is closer to the sort of grammatical instincts of which children's utterances reveal processes that are simple, logical and fast. [...] [Cayenne Creole] grammar is more *naturelle* than that of Sanskrit, Latin, and French. But this grammar did not spontaneously emerge in Guy-

ane; it was imported from Africa.

(Adam 1883: 4–5; my translation)[2]

In reality, the Malagasy slaves [who created Mauritian Creole] have brought along to Mauritius their native grammar, but not the forms that I have previously mentioned [inflectional gender marking, inflectional plural marking on verbs, the *avoir* ('to have') auxiliary, the verbal copula]. These forms are the product of an evolution that has not happened in the Polynesian and Melanesian languages. In dealing with such forms, Malagasy speakers kept their native grammar [...] In Mauritius, this native grammar reasserted its influence.

(Adam 1883: 7; my translation)

1.1.3.   *On Cartesian-Uniformitarian linguistics: The Neogrammarians*

(3)   a.   These [methodological] principles are based on a two-fold concept, whose truth is immediately obvious: first, that language is not a thing which leads a life of its own outside of and above human beings, but that it has its true existence only in the individual, and hence that all changes in the life of a language can only proceed from the individual speaker; and second, that the mental and physical activity of man must have been at all times essentially the same when he acquired a language inherited from his ancestors and reproduced and modified the speech forms which had been absorbed into his consciousness.

(Osthoff & Brugman 1878 [1967: 204])

b.   If someone could once and for all manage to get rid of these generally harmful expressions "youth" and "old age" of languages! These and many others in themselves quite innocent grammatical terms have so far been almost exclusively a curse, hardly a blessing. For the child who was born in Greece in the Homeric age, who became aware of the speech forms of his linguistic community by hearing them, and who then reproduced them in order to make himself understood by his fellow men – for that child were these speech forms ancient?

(Osthoff & Brugman 1878 [1967: 205–206])

---

2. It is worth noting that, in spite of his own language-as-organism comparisons, Adam (1882: 3) warns against the "false analogy" and the "poetry" and "mysticism" of botanical metaphors when applied to linguistic structure. More generally, Adam's 1882 book criticizes some of the structural and epistemological bases of Schleicher and others' morphological-cum-genealogical classifications. Yet, Adam himself, like Schleicher, relies on morphology as an index of evolutionary progress; see Adam (1882: 24–31, 62, etc.).

1.1.4. *Foucault on "chimera and reverie" in linguistics*

(4)   Suffice it to recall that the quest for primitive language, a perfectly acceptable theme up to the eighteenth century, was enough, in the second half of the nineteenth century, to throw any discourse into, I hesitate to say error, but into a world of chimera and reverie – into pure and simple linguistic monstrosity.

(Foucault 1971 [1972: 223])

1.1.5. *Popper on (criticisms of) myths as science*

(5)   Thus science must begin with myths, and with [. . .] the critical discussion of myths, and of magical techniques and practices. [. . .] The critical attitude, the tradition of free discussion of theories with the aim of discovering their weak spots so that they may be improved upon, is the attitude of reasonableness, of rationality. It makes far-reaching use of both verbal argument and observation – of observation in the interest of argument, however.   (Popper 1965: 50)

1.2. *Schleicherian roots of Language and route to progress: From isolating to agglutinative to inflectional/fusional*

Taken together, the quotes in (1)–(5) suggest that the search for "the world's simplest grammars" has been going on for quite a while, and so has the dogma that "the world's simplest grammars are creole grammars". This search for "simplest grammars" is part of a larger search for, and larger myths about, the origins and evolution of our species. This search seems driven by an apparently innate drive to (in Schleicher's phrase) "classify humanity".

Schleicher's *Glottik* is linguistics as natural history. By definition, its morphological taxonomy of languages qua organisms is regulated by universal natural laws and reducible to the genealogical mapping of their teleological development toward "idioms of higher organization" (Schleicher 1850, 1863, 1865, etc.).

In the intellectual climate of early 19th century, Schleicher's pre-Darwinian language-as-organism evolutionary approach belonged to the "normal science" of his period ("normal" in the Kuhnian sense). In retrospect, Schleicher's *Glottik* can be viewed as one of the central myths that gave vigor, status, and popularity to comparative-historical linguistics; see Hoenigswald & Wiener (eds.) (1987) and Alter (1999) for comprehensive overviews and, specially, for an array of positions similar to Schleicher's through much of the 19th century and beyond.

At the core of Schleicher's language-as-organism evolutionary hypothesis was a then-attractive congruence between, on the one hand, the simple-to-

complex Darwinian evolution of biological organisms from "one-celled organisms" to "higher living beings" and, on the other hand, the postulated historical progression of languages, through variation and subspeciation, from isolating to agglutinative to inflectional/fusional (Schleicher 1863 [1983: 50–60], also Schleicher 1850 [1852: 6–13]). In the Schleicherian organic school, linguistic evolution, on a par with biological evolution, was to be modeled by a "tree of life", a *Stammbaumtheorie* – a genealogical ("family" tree) diagram depicting phylogenetic relationships, with (at most) one parent for each daughter node.[3]

Schleicher tries to establish his morphology-as-biology congruence by, inter alia, metaphorically taking the "simple cell" and the "simple root" as "the common primitive forms" of biological and linguistic evolution, respectively (Schleicher 1863 [1983: 55]). More explicitly, Schleicher takes "the radical elements [i.e., isolating root morphemes] as the CELLS of [prototypical primitive] speech" (1863 [1983: 53], emphasis in original) and posits an isolating proto-language – one made up exclusively of affixless roots – at the evolutionary beginning of each language phylum. The proto-language's monomorphemic words are the linguistic analogues of the "one-celled organisms" at the roots of biological evolution. Thus, Chinese with its tendency toward isolating morphology is most primitive while Sanskrit with its inflectional/fusional morphology is most advanced since, according to Schleicher and others, inflectional/fusional morphology marks the highest degree of complexity and perfection (see (1b) and (2b); also see Schleicher 1850). It is thus that morphology – inflectional morphology, in particular – has long served as the chief measure of evolutionary progress and/or structural complexity (but see Note 1).[4] Related measures are found in 20th- and 21st-century linguistics, as in the works of, e.g., Jespersen (1922: 233–234), Whinnom (1971: 109–110), Samarin (1980: 221), Seuren & Wekker (1986), Bickerton (1988: 274–276), Comrie (1992: 208–209), Seuren (1998: 292–293), McWhorter (1998, 2000a, b, this volume),

---

3. For (re-)evaluations of *Stammbaumtheorie* principles, see, e.g., Osthoff & Brugman (1878), Paul (1890), Schuchardt (1885), Vennemann & Wilbur (eds.) (1972), Wilbur (ed.) (1977), Labov (1994), Dixon (1997), Mufwene (1998, 2000b, 2001).

4. This is an overly simplified synopsis of 19th-century (pre-, post-, and quasi-)"Darwinian" linguistics. For related, if somewhat distinct, definitions and elaborations of morphology as a genealogical and/or complexity index, see Schlegel (1808 [1849: 446–453]), Humboldt (1836 [1988: Chapters 14, 19, 21–25]), Bopp (1833 [1985: vii, 102–103]), Müller (1868: 7–22). For additional original texts, more subtle critiques, current debates and extensive bibliographies, see Lehmann (ed.) (1967), Koerner (ed.) (1983), Koerner (1983), Maher (1983), Hoenigswald & Wiener (eds.) (1987), Aarsleff (1988), Seuren (1998), Alter (1999), Labov (2001: Chapter 1). The 19th-century original texts and their 20th-century exegeses should be on the reading lists of creolists who (tacitly) promote neo-Schleicherian scenarios for creole genesis; I myself still have much to learn from the mistakes and insights of our intellectual forebears.

etc. (cf. Note 6; see Chaudenson 1994 and DeGraff 2001a, b for overviews and critiques).

Time (i.e., language age) is a critical factor in Schleicherian models of linguistic complexity. In such genealogical-cum-teleological approaches to the linguistic *systema naturae*, time is built-in as a prerequisite for the development of complexity in both biological and linguistic evolution: complexity qua "higher organization" takes "very long time spans" to evolve (see (1c)). Thus, the following three related propositions:

(6)   a.   Complex species (i.e., those more advanced in the evolutionary hierarchy, connoting greater perfection) must be relatively old.[5]

         b.   Young languages and the early ancestors of old languages are necessarily simple (i.e., more primitive in the evolutionary hierarchy).

         c.   The youngest/earliest languages are the simplest (i.e., most primitive).

In the domain of Language, the hypothetical original species – Schleicher's ultimate *Ursprachen* – have had no time to evolve any complexity; thus they

---

5. Beyond old age, languages in Schleicherian linguistics enter into a stage of "senil[ity]" in which their inflection (if any) becomes moribund. Schleicher calls this degenerative phase "historical life" as opposed to "language evolution", with the former being a "retrograde metamorphosis" of the latter. In other words, "history" follows, and undoes the effect of, "evolution". And Schleicher adds that "retrograde metamorphosis" in morphology (e.g., the reduction of morphological complexity) is proportional to the "historicity" of the corresponding people. As one case study, Schleicher compares English with Icelandic. English speakers have been more historically active than Icelandic speakers: the latter remain relatively immune from language contact in continental Europe. Therefore morphology is more simplified in English than in Icelandic. (See Schleicher 1850 [1852: 23–30]; also see Maher 1983: xxviii–xxix for discussion of Schleicher's views on "historicity" and Trudgill 1989 and references therein for sociolinguistic interpretations of "historicity" sans "senility".)

As of Humboldt (1836 [1988: 203–213]), he takes "the wearing-down of inflections [which] is an undeniable fact" as a consequence of "the mind's progress", alongside "vernacular[ization]" and language contact (e.g., "foreign immigrations"). Given Schleicher's view above, the interesting, if surprising, observation here is Humboldt's (1836 [1988: 205]) remark that morphological decay results from "the mind's progress": "The more mature the mind feels itself to be, the more boldly it works in combinations of its own, and the more confidently it casts away the bridges that language constructs for the understanding." For Humboldt, this "more mature" genius, whose maturity is due to inter alia the intellectual development made possible by inflection, can cleverly decide to replace synthetic structures (e.g., affixes for nominal case and verbal tense) with analytic structures (e.g., prepositions and preverbal auxiliaries), thus promoting semantic transparency and ease of articulation (Humboldt 1836 [1988: 206]).

Taking Schleicher's and Humboldt's teleological-genealogical programs to their logical consequences, one must then claim that Prototypical Creoles, with their (alleged) "lack of inflection" (see (11)), are either most "senile"/"retrograde" or most "mature" in revealing "the mind's [utmost] progress"(!).

are located at the prototypical primitive stage (i.e., the isolating/affixless stage) while "languages of a higher organization [e.g., Indo-Germanic] have arisen from simpler forms, through a process of gradual development" (Schleicher 1863 [1983: 50]). It is in this vein that Schleicher (1865 [1983: 79]) takes language to be "of significance not only for the elaboration of a scientific [i.e., taxonomic] systematization of humanity, but also for the evolutionary history of man" (cf. Note 4).

1.3. *A foundational myth: Creoles as contemporary* Ursprachen

"Systematization of humanity" via linguistic structures and the (alleged) lack thereof is also found at the very inception of creolistics and throughout its existence. Creolists' own "classic" systematization is based on the age-old orthodoxy that creole morphology is EXTRAORDINARILY simple or simplified from a diachronic and/or synchronic perspective.[6]

In the 17th to 19th centuries, such orthodoxy had explicitly "race"-based underpinnings. One taken-for-granted piece of "normal science" revolved around the notion that non-whites were inferior human beings, and so was non-whites' speech considered inferior to whites' speech. For candidate (perhaps ambiguous) illustrations of such beliefs, see the works of early creolists such as Pelleprat (1655), Saint-Quentin (1872) (see (2a)), Baissac (1880), Adam (1883) (see (2b)), along with 18th/19th-century dictionary and encyclopedia entries for the word "creole" (see, e.g., Pierre Larousse's 1869 *Grand dictionnaire universel du XIX siècle* and Vinson's entry in the 1889 *Dictionnaire des sciences anthropologiques*). In such works, the distinctive features of creoles were not due to socio-historical factors only, but they were also taken to reflect the inferiority of their (non-European) speakers. The latter were deemed to be cognitively unable to master the "complexities" of European languages. Per this orthodoxy, "primitive" people spoke "primitive" languages (see Section 3.1; cf. (2) and Note 35).

1.4. *Toward Cartesian-Uniformitarian creolistics*

At the turn of the 19th century the Neogrammarians, while adopting Schleicher's seminal insights in comparative-historical methodology (e.g., with re-

---

6. See Pelleprat (1655), Saint-Quentin (1872), Baissac (1880), Adam (1883), Schuchardt (1909), Jespersen (1922), Bloomfield (1933), Hjelmslev (1938), Hall (1953), D'Ans (1968), Samarin (1980), Whinnom (1971), Valdman (1978), Bickerton (1988), Seuren & Wekker (1986), Seuren (1998), McWhorter (1998, 2000a, b, this volume), etc. For further discussion of linguistics–ideology interactions in the terminologies, taxonomies, and theorizing of creole studies as well as in creole-related language policies and practices, see, e.g., Prudent (1980), Alleyne (1994), Y. Dejean (1999a, b, forthcoming), Mühleisen (2000), DeGraff (2001b, in preparation), Corcoran (2001).

spect to the reconstruction of unattested proto-forms), fought hard to get rid of his excessively organic and teleological metaphors. The Neogrammarians' stated goal was to understand Language and language change through the study of variation among (related) idiolects and through the study of universal laws (e.g., sound-change laws) that are ultimately rooted in the psychology and physiology of individual speakers. This was a shift of interest from the reconstruction of *Ursprachen* from archives to the analysis of contemporary idiolects in vivo (i.e., as manifested in individual speech). For Neogrammarians (see (3)), grammars live in speakers' minds, not in society; thus, the study of individual grammars as manifestations of a "psychical [i.e., psychological] organism" should take epistemological priority over "Historical Grammar" (i.e., "descriptive grammars of different periods [...] tacked together"): only the former is truly "scientific"; see Paul (1890 [1970: xxxvii, xliii, 1–19]). As Osthoff & Brugman (1878 [1967: 198]) put it, what needed correction is the methodology whereby "[l]anguages were indeed investigated most eagerly, but the man who speaks, much too little".

Continuing this "Cartesian" (i.e., universalist and mentalist) trend into the 20th century, Boas and his students, including Sapir, assumed both the existence of language universals, the "psychological reality" of idiolects, and the independence of phenotypes, language and culture across humanity; see, e.g., Boas (1911: 11), Sapir (1921: ix, 207–220). Both Boas and Sapir reacted strongly against the sort of Romantic ethnocentrism inherent in Schleicherian correlations between morphology, age, and complexity. About the (non-)rapport between linguistic morphological types and cultural evolution, Sapir (1921: 219) wrote:

(7)   [A]ll attempts to connect particular types of linguistic morphology with certain correlated stages of cultural development are vain. Rightly understood, such correlations are rubbish. [...] Both simple and complex types of language of an indefinite number of varieties may be found spoken at any desired level of cultural advance. When it comes to linguistic form, Plato walks with the Macedonian swineherd, Confucius with the head-hunting savage of Assam.

Sapir's view is robustly supported by crosslinguistic evidence both within and across genetic phyla, among languages of similar and dissimilar time depths (e.g., English, Icelandic, German, Chinese, Wampanoag, Kivunjo, and Nicaraguan Sign Language). In a related observation, Thomason (1980: 361) pointedly notes that inflectional morphology can vary greatly within single genetic groupings, as in Indo-European:

(8)   [T]he general pattern of development from flexional to isolating morphology is well known. But this progression is much more advanced

in some branches of the family than others, and the least elaborate inflectional systems in modern Indo-European languages – those of the western European languages, say – bear little resemblance to the most elaborate systems, such as noun declension in most Balto-Slavic languages. It is not just a matter of a reduction in the number of cases or gender distinctions; often, the categories themselves have changed.

Sapir's and Thomason's observations are worth keeping in mind throughout the discussion in this paper.

The Schleicherian genealogical classification of language by age cum complexity loses further grounding with the advent of Chomsky's "Cartesian Linguistics" in the second half of the 20th century (see, e.g., Chomsky 1966, 1981, 1986, 1995). As in the Neogrammarian dogma in (3), grammars are inherently parts of human biology, not autonomous living organisms that undergo birth, age, senility, and death independently of their speakers. Generative linguistics' objects of study are, in Cartesian mode, internal properties of individual minds (i.e., mental grammars qua I(NTERNAL)-LANGUAGES). Thus, generative linguistics is "internalist biolinguistic inquiry" (Chomsky 1995: 1–11, 2001: 41–42) and is intrinsically Uniformitarian (cf. Descartes's assumption that "reason is by nature equal in all men" and Descartes's notion of knowledge as mental representations; see Chomsky 1966 for relevant discussion).

Per current assumptions in biology, the basic morphology of the human brain is uniform across the species. It has thus become more difficult, if not impossible, to theoretically correlate biological evolution with crosslinguistic variation, unless one adopts a quasi-Lamarckian view of language change whereby crosslinguistic typology can be reduced to genetic variation across human groupings. In this quasi-Lamarckian scenario, language-specific structures (e.g., isolating vs. agglutinative vs. inflectional/fusional morphology) would be correlated with variations in the human genetic blueprint. This is not a likely scenario given current results in language acquisition and biology: linguistic structures do not evolve and are not transmitted like DNA.[7]

---

7. Any congruence between genetic and linguistic evolution is mediated via geographical isolation, population displacements, and the like, not by any biological causal relationship. See Hoenigswald & Wiener (eds.) (1987), Bateman et al. (1990), and Cavalli-Sforza (2000: Chapters 5, 6) for some comparison of (the mechanisms underlying) genetic and linguistic evolution. *Current Anthropology* (1990, vol. 31, numbers 1–4) offers diverging opinions on the methodology of such comparisons. But the point remains that *Stammbaumtheorie*, even if useful for approximating population displacements through time, is at the very best only in rough correlation with genetic transmission. (Also see Note 18.)

1.5.   *Back to Schleicher(ian creolistics)'s* Ursprachen

Notwithstanding current advances in Cartesian and Uniformitarian (bio)linguistics, there is a sense in which Schleicher's approach has survived the 19th century into the 20th and now 21st century, albeit under new theoretical guise. While modern linguistics is making steady advances in its exploration of our intrinsically human and species-uniform Universal Grammar (UG), "creolistics" has kept up, and even revived, early 19th-century notions of language evolution. Indeed, creolistics is perhaps the only field where the search for a genealogical and typological class of "simplest grammars" is still at the center of contemporary research. It is thus that certain trends in creolistics are reviving Schleicher's *Glottik* with creole languages as the new class of youngest, thus structurally simplest, linguistic species. In this modern *Glottik*, creole languages are living specimens of *Ursprachen*, i.e., contemporary proto-languages – "the world's only instantiation of spoken language having been 'born again' " in McWhorter's (Section 2.3) evangelical phrase.

However, an empirically and theoretically grounded implementation of the claim that the world's simplest and/or most optimal grammars are creole grammars still constitutes the holy grail of creolists through an unbroken lineage of research programs (see Note 6). In the 20th and 21st centuries, starting with, e.g., Jespersen (1922) and up to McWhorter's target article in this volume (henceforth WSG[8]) linguists have used morphosyntax to try and identify sufficient and necessary conditions that would make creoles deeply special in a structural and synchronic sense. Perhaps unsurprisingly, many of these contemporary observations on creole morphosyntax are quite reminiscent of those encountered in 17- to 19th-century texts; see DeGraff (2001b: 88–98) for further details. For example, consider the statement that "creoles are natural languages reborn from a radical reduction of their source languages into makeshift jargon" (WSG: 144). This statement finds direct antecedents throughout creole studies, from its very inception (see Note 6).

An explicit quantification for the alleged maximal simplicity of creole grammars is offered in WSG. Even though it makes no reference to Schleicher, McWhorter's recent work (1998, 2000a, b, WSG) is, of late, the most sustained and perhaps most widely read effort to articulate a structural basis for Schleicherian linguistics, with the aim of categorizing languages according to some explicit complexity hierarchy. McWhorter tacitly assumes the Schleicherian dogma whereby linguistic typology must, at all costs, include a genealogically and structurally well-defined class of "simplest grammars". In his

---

8. When making reference to (parts of) WSG, I will explicitly use "WSG", possibly followed by (sub)section or page number. Otherwise, (sub)section numbers without further indication refer to my own commentary on WSG.

revamped complexity scheme, Prototypical Creoles are the new *Ursprachen*, languages created *ab ovo* from virtually "ground zero" complexity (WSG: Section 4.4). This is somewhat reminiscent of Saint-Quentin's structural claims in (2a). Thus, McWhorter's hierarchy continues the Schleicherian tradition of putting certain languages (here Prototypical Creoles) in a deeply-special class of linguistic neonates – contemporary fossils of Language at its evolutionary incipience. In a nutshell, the argument is that "because so much of a grammar's complexity results from the operation of random accretion over time, creoles display less complexity than the rest of the world's natural grammars" (WSG: 133).

The central assumptions here are (i) that creole languages are markedly younger languages than non-creoles, and (ii) that this age difference is linguistically measurable and significant, contra Osthoff & Brugman's (1878) admonition in (3a). These assumptions are related to THE foundational claims in creole studies, namely the oft-repeated dualist statement that creoles are "non-genetic" languages that emerge via an abnormal "break in transmission" whereas non-creole languages gradually evolve "genetically" via "normal transmission" (the modern locus classicus for this claim is Thomason & Kaufman 1988: 8–12, 206, and passim; see Section 3.3 for discussion). In the classic creole-genesis scenarios, creole youth stems from the "pidgin-to-creole life cycle" as signaled by the concomitant morphological bottleneck (see references in Section 1.2 and in Note 6). In the most recent exponent of this dogma, creoles are "born as pidgins, and thus stripped of almost all features unnecessary to communication" (WSG: Abstract).

1.6. *Toward Cartesian creolistics (redux): A guide for "learning by debunking"*

As will become obvious through the development of the present critique, current research on creoles as contemporary *Ursprachen* (as, e.g., in WSG) presents us with a "modern" collage of pre- and neo-Darwinian claims about language evolution. The antecedents of such claims go back to the structural-cum-genealogical speculations of, e.g., Schleicher (1863) (see (1)) and Saint-Quentin (1872) (see (2a)). Thus, a full-fledged Popperian (see (5)) critique of (neo-)Schleicherian-cum-Quentinian creolistics can proceed via a close examination of the proposal in WSG about age-complexity correlations. This proposal conveniently provides us with an updated adaptation of linguistic-genealogical arguments that have run virtually uninterrupted through the past two centuries. A critique of this proposal will, I hope, clear up the scene for empirically-responsible and theoretically-grounded Cartesian-cum-Uniformitarian creolistics (i.e., the sort of creolistics that does not assume any a priori

fundamental structural distinction between creoles and non-creoles).[9]

As we will see below, modern creolists' measures for age-complexity correlations – like their early 19th-century intellectual antecedents – amount to an empirically, theoretically, and logically flawed view of creole formation and language change. Bearing in mind that "science must begin with the criticism of myths", these re-formulations of Schleicherian linguistics provide a valuable point of departure for learning from our mistakes and for advancing our knowledge (in Popperian mode; see (5)). My critique will thus exemplify "learning by debunking" (in Gould's 1996: 351–353 terminology). Indeed I share in Umberto Eco's optimism in his book *Serendipities* (1998: iix) where he so describes various brands of "lunatic" linguistics (cf. Foucault's allusion to "chimera and reverie" in (4)): "[E]ven the most lunatic experiments can produce strange side effects, stimulating research that proves perhaps less amusing but scientifically more serious".[10] Here are the two "serendipities" to be derived from the discussion of neo-Darwinian (or neo-Schleicherian) creolistics below: (i) establish the epistemological limits of the terms "pidgins", "creoles", "young" vs. "old", "simple(st)" vs. "(most) complex" as linguistic-structural and historical-phylogenetic taxa; and (ii) promote the study of language contact and its ubiquitous outcomes (including "creole genesis") in a Cartesian and Uniformitarian analytical framework.

Fundamental to neo-Schleicherian arguments about creoles' lack of complexity are the two notions "pidgins" and "features unnecessary to [basic] communication". These notions still remain ill-defined. I will argue in Section 4 that there can hardly be a core set of basic-communication features that defines the structural essence of ALL pidgins. As for the exclusively-creole combination of structural features claimed to be "predictable from the history of creoles in pidginization", this typology will be argued to be deeply problematic, on both empirical and theoretical grounds (see Section 2 and references in Note 12). Problematic as well, at least from a linguistic-theoretical standpoint, is the notion "age of languages": I will argue that thus far no well-defined and

9. See DeGraff (2001b: 98–99, in preparation) for some of the sociological costs that are associated with the dogma that creoles are structurally distinct from "regular"/"normal" languages (cf. Note 6). Such dogma has, inter alia, undermined the role that creole languages should play in the education of creole speakers and in the exercise of their human rights. With respect to Haiti, see critiques in P. Dejean (1989, 1993), Y. Dejean (1975, 1993, 1999b, forthcoming).

10. Darwin himself (1871: Chapter 21, 385) is worth quoting in that respect: "False facts are highly injurious to the progress of science, for they often long endure; but false views, if supported by some evidence, do little harm, for every one takes a salutary pleasure in proving their falseness; and when this is done, one path towards error is closed and the road to truth is often at the same time opened." Darwin may have been overly optimistic. For the sort of harm that can be caused by false views in the human sciences, see, e.g., Gould (1996); cf. Note 51.

independent (i.e., testable and non-circular) LINGUISTIC metric can objectively measure the age of languages.

Yet it is the age of languages that in creole studies is often taken (in a manner reminiscent of early 19th-century – Schleicherian – linguistics) as one crucial factor leading to complexity differentials between creoles and non-creoles. For neo-Schleicherians, creoles' lesser complexity is yet another "predictable result of their youth" (see, e.g., WSG: Section 1), notwithstanding the fact that there is still no reliable litmus test for young languages as a linguistic class (see (7)–(8); also see Sections 3 and 6).

What about "complexity" per se? Here too creolists enlist terminology and assumptions that will be shown to be ill-defined and empirically- and theoretically-controversial in fundamental ways. For examples, the "complexity metric" in Section 2.4 of WSG is theoretically peculiar: it is based on an arbitrary list of superficial linguistic features with no psychologically-relevant, theoretically-grounded, or independently-motivated unifying basis. In fact, the conceptual foundations of this metric are either left undefined (along with, e.g., "communicative necessity", "basic [human] communication", and language "age") and/or are made largely incompatible or irrelevant to what we (seem to) know about linguistic typology, historical linguistics, language acquisition, language processing, and theoretical linguistics (cf. WSG's interpretations of "complexity", "Universal Grammar", etc.). Furthermore, the choice and testing of creolists' complexity metrics is often circular, tendentious and empirically flawed (see Sections 5 and 6).

Chomsky once wrote that "linguistic theory must be constructed with explicit and precise definitions and operational tests" (1957: 233). In the absence of empirically- and theoretically-grounded definitions of "creoles", language "youth", "pidgins", "basic [human] communication", "complexity", and so on, what are we to make of the proposition that "the world's simplest grammars are creole grammars" because creoles, and only creoles, are "born again" languages? I take it that, in the absence of "explicit and precise definitions and operational tests" this most simplistic proposition can provide us with only a mismeasure of creole languages and, indeed, a mismeasure of creole SPEAKERS if one assumes a Cartesian (or Humboldtian) approach whereby languages are properties of minds.[11]

In what follows, I will document the theoretical, empirical, and bibliographical lapses that undermine the recent (and not-so-recent) claims that "the world's simplest grammars are creole grammars". In particular I will review the foun-

---

11. Cf. Humboldt's "inner linguistic sense" which is intimately related to the "genius" (e.g., the "mental capacities") of particular peoples at particular evolutionary stages; see Corcoran (2001) for one recent discussion in the context of creole-genesis scenarios; also see Humboldt's remarks on the evolution of inflection in Note 5.

dational assumptions upon which traditional and recent claims about creole complexity (or lack thereof) rest. These assumptions touch on the following questions, which I address in turn:

(i) How are "creoles" defined? (Section 2)

(ii) How does one measure the "youth" of languages? (Section 3)

(iii) Can we determine the features "necessary to basic human communication"? What is the relationship between "basic human communication" and "simplification" in "pidgins"? Does Universal Grammar define "basic communication" requirements? (Section 4)

(iv) What is the epistemological status of complexity metrics in creole genesis scenarios? (Section 5)

(v) Given (non-)answers to all of the above, (how) can we falsify the claim that "a subset of creole languages display less overall grammatical complexity than older languages"? (Section 6)

2. Defining "creole"

What's in a name? Long ago, Francis Bacon (1620 [1994: 64–65]) warned us that that there "are names for things that do not exist (for just as there are things without names because they have never been seen)" and that certain "obscure and deep seated [terminology] is derived from an incorrect and unskilled abstraction". Furthermore, he wrote (1620 [1994: 55]):

(9)   [W]ords are applied according to common understanding. And in consequence, wrong and inappropriate application of words obstructs the mind to a remarkable extent. [. . .] [W]ords plainly do violence to the understanding and throw everything into confusion, and lead man into innumerable empty controversies and fictions.

In 1878, the Neogrammarian manifesto also warned linguists against "[believing] that they have then fathomed the essence of the phenomena when they have devised a name for the thing" (Osthoff & Brugman 1878 [1967: 202]).

What exactly are creolists from Saint-Quentin to McWhorter comparing when they make claims to the effect that "the world's simplest grammars are creole grammars"? What is the exact scientific purpose of this comparison? The postulation of some structural essence common to all "creole" languages is central to the present discussion as it has been from the beginning of creole studies. That this structural essence is the outcome of "simplification" is another age-old tenet in creole studies. As Chaudenson (1994: 41) notes, "the idea that creoles are simplified versions of their European ancestors is as old as creole studies".

Yet it has often been argued in Uniformitarian (non-dualist) fashion that "creolization is a sociohistorical, not a structural, process" (Mufwene 2000a).

In this view, creoles cannot be distinguished a priori from non-creoles on strictly synchronic structural grounds. I myself have now adopted a strictly language-external definition whereby "creole" is a sociohistorical attribute that connotes the results of particular types of (abrupt) language contact marked by exacerbated social distance cum power imbalance; see DeGraff (1999a, b) and references therein for overviews. I also take it, as a working hypothesis within the generative framework I work in, that:[12]

(10)   [C]reoles [as mental entities, i.e., I(nternal)-languages] are no more and no less than the result of extraordinary external factors coupled with ordinary internal factors [...] [Within mentalistic approaches to language creation and language change,] THE NOTION OF "CRE-OLIZATION" AS A UNITARY AND DISTINCT LINGUISTIC PHENOM-ENON EVAPORATES.   (DeGraff 1999b: 477; emphases added)

The Cartesian-Uniformitarian position in creole studies has long been opposed by the dualist orthodoxy that assumes a strong form of "creole exceptionalism". Per this dualism/exceptionalism, creole languages – thus creole speakers – are deeply special, with genealogical and structural properties that are fundamentally distinct from their non-creole counterparts. This is grosso modo the position of neo-Schleicherian creolists who posit both a diachronic exceptionalism ("pidgin ancestry" and creole "youth") and a synchronic exceptionalism (a "Creole Prototype"). These distinctions are straightforwardly Schleicherian in the rapport they try to establish between complexity and time: "creoles are the world's only instantiation of spoken language having been 'born again', when speakers expanded pidgins" (WSG: Section 2.3), and creoles' exceptional youth (i.e., their exclusive "pidgin ancestry") entails that "the world's simplest grammars are creole grammars". As in Schleicher's evolutionist scenario, it is assumed that "much of a grammar's complexity results from the operations of random accretion over time" (WSG: Section 2.3). However, these genealogical and structural distinctions lack theoretical and empirical substance.

2.1.   *Haitian Creole: An anti-prototype Prototypical Creole*

On the synchronic structural front, the Creole Prototype is equated to the combination of the following traits (WSG: Note 1):

---

12. For related discussion, see, e.g., Van Name (1870), Meillet (1906), Schuchardt (1909), Hall (1966), Alleyne (1971), Hoenigswald (1971), Givón (1979), Marantz (1983), Muysken (1988), Mufwene (2001), DeGraff (1997, 1998, 2000, 2001a), etc. Also see the discussion, and the references cited, in DeGraff (1999b: 477, 519, etc.) and DeGraff (2001b: 108, Note 13).

(11)   a.   lack of "inflectional affixation";
           b.   lack of "tone distinguishing monosyllabic lexical items or encoding morphosyntactic distinctions";
           c.   lack of "opaque lexicalization of derivation–root combinations".

The assumption is that inflection, tone, and opaque derivation–root lexicalizations increase complexity. As of the lack of complexity-inducing morphological processes in the Creole Prototype, it is assumed as a consequence of the hypothetical pidgin-to-creole cycle, its concomitant morphological bottleneck, and the correlates thereof vis-à-vis the *ab ovo* growth of creole affixes. See the following quotes from WSG: Section 5.2:

(12)   a.   When language "begins anew" amidst pidginization, the linguistic vehicle consistently lacks affixation entirely or exhibits it only minimally, with affixes developing only slowly even when the pidgin is creolized.
           b.   Observed and documented processes of language change make it clear that the main source of affixes is erstwhile free morphemes.

There is already a plethora of data and observations that, taken together, invalidate the empirical and theoretical claims in (11)–(12) about creole morphogenesis.[13,14] Yet, in creolistics and in other human sciences, complexity rankings and genealogical scenarios are no simple matters, scientifically and sociologically (see, e.g., Gould 1996 for relevant caveats). Thus, the need to address the claims in (11)–(12) from ground-zero up, starting with the theoretically and empirically central aspects of the Creole Prototype.

To start with, let's take Haitian Creole (hereafter HC). This sociohistorically prototypical plantation creole has been argued in DeGraff (2001b: 69–88) to manifest both "inflectional affixation" and "opaque lexicalization of derivation–root combinations", contra the predictions in (11). Moreover virtually all HC affixes are etymologically related to French affixes (i.e., virtually

---

13. See Ducœurjoly (1802), Hymes (ed.) (1971), Alleyne (1988), Muysken (1988, in press), Thomason & Kaufman (1988), Keesing (1988, 1991), Syea (1992), Dijkhoff (1993), Mufwene (1993, 2000a, b, 2001), Chaudenson (1994), Kouwenberg (1994), Arends, Muysken, & Smith (eds.) (1995), Jahr & Broch (eds.) (1996), DeGraff (1997, 1999a, b, 2001a, b, forthcoming), Thomason (ed.) (1997), Freeman & Laguerre (1998), Migge (1998), Fattier (1998), Siegel (1999), Muysken & Law (2001), Valdman (2001), etc.

14. Creolists often pay lip service to the age-old dogma that "morphology [is] essentially alien to creole languages" (Seuren & Wekker 1986: 66) while their very data illustrate robust patterns of both inflectional and derivational morphology, including opaque lexicalizations. In fact, some of the relevant data in these works suggest that certain creoles may well have MORE affixes than certain non-creoles, including a subset of the creole's ancestor languages; e.g., many affixes in Haitian Creole have no counterparts in the Fongbe substrate; see DeGraff (2001b: 58–69) contra Lefebvre (1998: Chapter 10).

none of these affixes result from grammaticalization), thus shedding doubt on the claims in (12).

Well-documented facts of Haiti's sociohistory and demographics teach us that HC should count as a bona fide creole, even a "most creole of creoles", one whose historical conditions "have been perfect for the preservation of a basilectal creole" (McWhorter 1998: 809, 812; 2000b: 206). But, in creole studies, even established historical facts can be tinkered with to fit the scenario *du jour*: After robust HC data were advanced as counterexamples to pro-prototype claims (see DeGraff 2001b), HC's privileged status as "basilectal creole" and "most creole of creoles" got revoked. HC is now taken to "not exemplify the Creole Prototype in the purest possible form" because of alleged "contact over the centuries with French" (WSG: 143).

Historically, linguistic interaction in colonial Haiti between Europeans and Africans was by far the most intense at the onset of contact, with French structures having had the most influence in the formation of (proto-)HC quite early on, in late 17th through early 18th century. Thereafter, contact with French speakers was greatly reduced after the sugar boom in the middle of the 18th century: the labor needs of expanding sugar plantations led to a drastic increase in the arrival rates of Africans. Before the sugar boom, the colony was still made up of mostly small homesteads – the *société d'habitation* – many of which subsequently and gradually gave way to the brutal segregation of the plantation economy – the *société de plantation*. (See Baker & Corne 1982, Chaudenson 1992, Chaudenson & Mufwene 2001, and Singler 1996 for an overview of the historical and demographic details and for pointers to the relevant literature.)

After the independence battles of 1791–1803, the French presence was virtually eliminated. There have always been, and still exist, various degrees of contact between HC and French, specially at the higher echelons of society, with concomitant contact phenomena in both languages as naturally expected. However, Haiti's history has allowed post-genesis HC to evolve in relative isolation from its socially-remote, albeit prestigious, lexifier. Haiti is the only New World plantation society that eliminated most of its "lexifier" population through war. From the *société de plantation* onward, French has been especially remote from (most of) the monolingual peasantry, Haiti's numerical majority. More generally, the vast majority of contemporary Haitians in Haiti are monolingual creolophones with relatively little contact with other languages (one notable exception is the region alongside the Haitiano-Dominican border where HC speakers are regularly exposed to Spanish varieties; see DeGraff 2001b: 68 for one linguistic-structural consequence of this contact in the domain of morphology). Such a degree of linguistic isolation is quite unlike the situation in other creole-speaking communities in the Caribbean. These well-known sociohistorical and demographic facts explain why Haiti does not

offer the same sort of "creole continuum" that is found in the former English colonies of the Caribbean. Unlike (say) the English-lexicon Caribbean creoles, HC by and large (i.e., the majority of HC speakers) did not "remain in contact with [the] lexifier" (contra WSG: Note 17). It thus seems unlikely that the bulk of HC morphology would have been created post-genesis via late borrowings from French (e.g., after the elimination in 1803 of the majority of potential "lenders").[15]

Furthermore there is no evidence that there ever was an earlier stage of (proto-)HC where all creole varieties were uniformly devoid of all affixes. Such an affixless stage is even less likely considering that most of the HC lexicon is etymologically French and that HC speakers – like any other human speakers in our documentable past – are, in principle, able to extract morphological (e.g., affixal) information from stored patterns in their lexicon. In fact, whatever mental capacities would enable late-borrowing(-cum-restructuration) of French affixes in the course of HC's post-genesis diachrony would have

---

15. McWhorter himself seems well-aware of these sociohistorical facts, when needed for his theorizing elsewhere. For example, these facts play a key role in his Afrogenesis speculations (see Appendix B). Witness the following four assumptions about the genesis of plantation creoles (McWhorter 2000b: 200–207): (i) "the lexifier was available to all slaves not only during the *société d'habitation* phases but even later, during the plantation stage"; (ii) "adult slaves were capable of obtaining a viable second-language register of the lexifier"; (iii) "plantation-born children were even better situated to acquire the local standard than their parents"; (iv) "[t]he intimate conditions within which blacks and whites lived in *sociétés d'habitation* would have made acquisition of the lexifier even more compelling"; (v) "in a society like Haiti, where French speakers were ousted early in the colony's history, conditions have been perfect for the preservation of a basilectal creole". Assumption (i) seems controversial for well-documented sociohistorical and demographic reasons, specially at the plantation stage: given marronage, social segregation, low European-to-African ratios, etc., not ALL slaves could be exposed to, and learn, the lexifier; see Pelleprat (1655), Girod-Chantrans (1785), Moreau de Saint-Méry (1797), Descourtilz (1809), etc. But assumptions (ii)–(v) seem much less controversial, independently of the use to which they are put in McWhorter's "Afrogenesis Hypothesis"; see, e.g., Baker & Corne (1982), Chaudenson (1992), Chaudenson & Mufwene (2001), Singler (1996), Mufwene (2001) for related facts.

    In the case of Haiti, propositions (ii)–(iv) straightforwardly entail that many early HC speakers had a "diglossic competence between an L2 variety of the lexifier and a [...] pidgin" and that both the founding slaves and plantation-born children could acquire "the local standard" (or, at least, a variety thereof). If so, then one can reasonably argue that "diglossic" early HC speakers were in a position to adopt and adapt affixes and other patterns from their "L2 variety of the lexifier" in order to structurally expand the emerging "creole". This renders unlikely any scenario whereby incorporation of French-derived affixes into HC is necessarily a late post-genesis phenomenon (cf. Appendix A). The linguistic evidence points to the same conclusion: affixes in all varieties of HC – including any contemporary "preservation of a basilectal creole" (see (v) above) – have cognates in French affixes; see, e.g., Fattier (1998), DeGraff (2001b), contra Lefebvre (1998).

    (On the empirical and theoretical status of McWhorter's (2000b) "Afrogenesis Hypothesis" and the "imported pidgins" therein, see Appendix B.)

also enabled the adoption/adaptation of similar French affixes at the earliest stages of (proto-)HC. This is especially so given the numerical preponderance of French over Africans and the lesser racial segregation at the early stages of contact. In other words, the sociolinguistic context was more learner-friendly at the onset of contact – during the *société d'habitation* – than later on – during the *société de plantation* (see Note 15). No known (psycho- or socio-)linguistic principle could have forced all the speakers in the (early) contact situation to systematically ignore all morphological patterns in the available lexifier varieties (also see Appendix A).

Similar remarks apply to the diachrony of "opaque lexicalizations of derivation–root combinations" in HC – throughout I am assuming Uniformitarianism (e.g., that innate mental capacities for Language have remained uniform across the species in the past few millennia and across socioeconomic contexts; see (3)). Indeed, if HC speakers could create such morphology via "borrowing" through POST-genesis (reduced) contact with French speakers, then no known (psycho)linguistic constraint would prevent the creators of HC from creating this morphology at the ONSET of contact (e.g., during the *société d'habitation*, when the demographic and sociolinguistic factors were MORE favorable for the acquisition-cum-reanalysis of such forms via relatively HIGHER exposure to French speakers and, thus, more intimate familiarity with the lexifier). I must note again that, as far as we can tell, the bulk of the HC lexicon (both roots and affixes) has always been etymologically related to French (see Section 3.3). This means that creole creators from the get-go massively adopted (at the very least) French lexical forms, including opaque lexicalizations.

Thus, we can conclude that HC – a bona fide prototypical creole in the sociohistorical sense – with its inflectional and derivational morphology and its opaque lexicalizations represents a robust counterexample to classic neo-Schleicherian claims about creole morphogenesis via an affixless pidgin (also see Appendix A).

2.2. *The Creole Prototype is an anti-Saussurean artificial language*

On the theoretical front, given the very nature of the lexicon (in the Bloomfieldian sense) as the repository of Saussurean arbitrariness and given now-standard results in (psycho)linguistics, there seems to be no reason to expect "opaque lexicalization of derivation–root combinations" to take millennia to develop.[16] Indeed, it can be reasonably argued that "lexicalization entails that

---

16. The claim about the absence of creole "opaque lexicalizations" is not new. Seuren & Wekker (1986: 66–68) take "idiosyncratic exceptions", "highly specialized lexical items", and "richer expressive means" to constitute a class of lexical "luxuries" that are exclusive to non-creoles (i.e., to "older or more advanced" languages). See DeGraff (2001b: 88–98) for related orthodoxies in creole studies, some of which go back to the 17th century.

the form is no longer generatable via metaphorical inference and now requires storage as an independent form" (WSG: Note 10). In this statement, "lexicalization" is clearly defined, in Cartesian fashion, in terms of mental processes in the heads of individual speakers, namely "metaphorical inference" and lexical "storage". The latter is what is normally (and crucially) involved in the creation of "opaque lexicalizations" as Saussurean (thus, holistic) signs in the speaker's mental lexicon. To also claim that such a mental process necessarily requires "millennia" to unfurl (McWhorter 1998: 792–793, 798, 812, etc.) leads to a scenario in which individual speakers can live for millennia – not a likely scenario.

A more realistic and theoretically-grounded alternative is to assume that opaque lexicalizations as properties of I-languages (thus, of individual minds, in concert with communal conventionalizations and cultural transmissions) do not require millennia to develop: like other creole lexical items, opaque lexicalizations can be innovated in creole grammars or they can be "inherited" from (and via exposure to speakers of) the source languages; see Fattier (1998), De-Graff (2001b: 76–82). Regarding the "inheritance" of opaque lexicalizations, a pro-prototype creolist could try and argue that such "inherited" lexicalizations should not count for testing the prediction in (11c) because they were not developed over millennia by the creole speakers themselves. But this argument is circular since (Prototypical) Creoles are a priori defined as young languages by pro-prototype creolists. The argument is fallacious in yet another way: many bona fide opaque lexicalizations in, e.g., contemporary English were not innovated by the current generation of English speakers – these lexicalizations were "inherited" by modern English speakers.[17]

---

17. (Neo-)Schleicherian linguistics makes recurrent appeal to "very long time spans" as necessary conditions for the development of structural complexity beyond the requirements of basic/primitive communication. Witness Schleicher's pronouncement in (1c) and contemporary beliefs that complexity requires "millennia" to develop its "above and beyond UG" overspecifications (see, e.g., WSG: Sections 1, 2.3, 4.2, 5). It thus seems that (neo-)Schleicherian linguistics, unlike Cartesian-Uniformitarian linguistics (see, e.g., (3a)), gives epistemological priority to some sociohistorically-oriented reification of E(xternal)-languages, at the expense of mental grammars qua I(nternal)-languages. This, in turn, leads to a confusing conflation of mental processes and historical factors (the latter can be considered "historical accidents"; see Chomsky 1995: 6–7, 11). If human biology (UG, say) sets boundaries on the shapes of all natural languages (including the crosslinguistic inventory of what WSG: Section 2.3 calls "'ornamental' elaborations"), then no language-specific properties (and no language-specific "ornament") need "millennia" to develop.

This also applies to opaque lexicalizations as discussed in the main text and to the other linguistic "ornaments" in (18) that are considered "incidental to basic communication". Thus "'ornamental' elaborations" cannot be excluded a priori from creole languages. Here too the ill-defined linguistic distinction "old" vs. "young" becomes a methodological trompe-l'œuil (see (3b)). The main text elaborates on this E-/I-language distinction in resolving the old-versus-young issue; also see DeGraff (1999a: 8–9).

2.3.   *Empirical and methodological considerations*

On the strictly-empirical front, tone, inflection, and opaque lexicalizations have been documented in a variety of creole languages; see, e.g., the observations and references in Muysken & Law (2001) and Muysken (in press). Muysken & Law (2001: 49) make the important observation that:

(13)   There is no doubt that in many core creoles there is very little in-flectional affixation. This is also to be expected since in the second language acquisition of the European colonial languages, during the early stages of the process of creole formation, inflection is often lost. However, the contributing superstrate and substrate languages were not very rich in their inflection either, and in several cases (Berbice Dutch Creole, Papiamentu, Cape Verdean) we do get some inflection. If we take creolization in typologically very different languages [...] there is simplification and regularization of inflection, but not loss of inflection. The absence of inflection is the singlemost frequently noted supposedly typological feature of creoles, and indeed may be the way many people identify a language as a creole. However, it may be the accidental by-result, from a scholarly point of view, of the limited typological spread in the languages contributing to the proto-typical creoles.

In a similar vein, Givón (1979: 20–21) remarks that the "reduction of inflec-tions" in Caribbean creoles is as expected given the inflectional profiles of the substratum. From this perspective, the synchronic structural criteria in (11) and the concomitant diachronic assumptions in (12) are empirically and theo-retically flawed. In the following section, I address, inter alia, the assumptions in (12). These assumptions crucially enter in the theoretical basis of (neo-) Schleicherian genealogical creolistics.

3.   Creoles, how "old" are you?

3.1.   *Creoles as "born again" languages?*

The widespread consensus across time, across space, and across theories in creolistics and beyond is that creoles are young languages – linguistic neonates that embody an evolutionary prior stage in relation to non-creole languages; see DeGraff (2001b) for an overview. Girod-Chantrans (1785 [1980]) found HC to be "nothing but French back in infancy". Adam (1883: 3) reports that "[i]n Europe, creole speech is universally considered an infantile jargon" (also see (2b)). Jespersen (1922: 228) wrote that creole creators spoke "as if their minds were as innocent of grammar as those of very small babies [...] [thus, cre-oles'] inevitable naïveté and [...] childlike simplicity [...]". In Hall's (1962)

pidgin-to-creole life cycle, creole languages are the one exception to the principle of " 'normal' language [being] handed down from generation to generation". Bickerton's Language Biogrogram Hypothesis turns creole speakers into linguistic "Adams and Eves" (in Richard Price's terminology, as cited in Corcoran 2001). Seuren & Wekker (1986: 66, 68) and Seuren (1998: 292) contrast languages that are "older or more advanced" and "sophisticated" to creoles, which are "younger or less advanced" – "beginning" – languages (see Note 16). Thomason & Kaufman (1988: 8–12, 206, 211, etc.) consider creoles as non-"genetic" (i.e., "parentless") languages that, unlike "genetic" languages like English, have evolved via some kind of "abnormal [break in] transmission" (see Note 22). WSG: 131 follows suit: "[A]ll [of the world's natural languages'] grammars [...] trace back tens of millennia. [...] with one exception [...]: [...] creole languages have, by definition, existed only for several centuries at the most. The oldest known creoles today [...] trace back to the late fifteenth century".

In quasi-Schleicherian mode (see (1)) and against Neogrammarian warnings (see (3)), the notions "younger" vs. "older" language often form the cornerstones of dualist claims about complexity differentials between creoles and non-creoles (see Note 17). For example, the proposition that "the world's simplest grammars are creole grammars" is argued to be "a predictable and, in the end, rather unremarkable result of the recent origins of creole languages" (WSG: 162). More explicitly (WSG: 132):

(14) [T]ens of millennia of drift would leave all grammars existing during that timespan equal in terms of the amount of complexity accreted beyond the bounds of the genetic specification for language. This stipulation predicts, then, that one subset of the world's natural languages, creoles, would differ from the rest of the world's natural languages in displaying less of this kind of needless complexity.

3.2. *Why/how do we "age" languages?*

There is one fundamental question that is evaded throughout: How does one scientifically measure the "age" of languages for the (presumably) scientific purpose of correlating language age with language complexity?

Individual speakers (be they creole-speaking or not) do not live through millennia. Neither do their respective "I(nternal)-languages" qua Cartesian mental grammars. So the statement that non-creole languages – and only non-creole languages – have existed for "tens of millennia" surely cannot refer to I-languages. Such "tens of millennia" estimates must apply to some extensional notion of language – perhaps some (reified succession of) socially-determined "E(xternal)-languages" or communal languages (see (3) and Note 17 above on relevant methodological caveats, going back to the Neogrammarians).

And still we are faced with the fundamental question: When do E-languages (or some communal/sociohistorical reification thereof) commence? Under certain definitions and for certain well-defined purposes, language age can certainly serve as a useful heuristic in various disciplines, including archeology, ethnography, anthropology, historical linguistics, etc. For example, one can say that, in a sociohistorical sense, HC is a "young" language as it marks the identity of a newly-created community – a community that did not exist before the 17th/18th century (I thank Heliana Mello for an enlightening discussion of this point). In a somewhat related sense, varieties of (say) Indian English and the corresponding (speech) communities may also count as "young". But notice that this sort of youth is not necessarily correlated with increased structural simplicity: in some ways, "young" Indian English may even be more complex than the Queen's "old" English (consider, say, the phonology of retroflex stops in Indian English). Similar remarks could, in principle, apply to any "old"–"new" language pair, as with, e.g., European and Brazilian Portuguese.

Notions such as language birth, age, and death are also assumed implicitly and a-theoretically when we use terms such as "Proto-Indo-European", "Latin", "Old French", "Middle French", "Modern French", etc., as classificatory devices. But, notwithstanding the popularity and sophistication of *Stammbaumtheorie* qua "Tree of Language" (cf. Darwin's Tree of Life), old vs. new linguistic species cannot be discriminated by any measure that looks like biological genetic criteria (e.g., DNA, interfertility). There is no clear notion whereby E-languages can be taken to reproduce like living organisms. Neither do we have clear linguistic-structural analogues for the DNA sequences that have now become so handy in tracing biological phylogenesis (see Section 1.4 above about the (im)possibility of quasi-Lamarckian linguistics).[18]

There are a number of fascinating sociological factors vis-à-vis why, when, and how certain languages start being (perceived as) "new" languages with birth certificates that distinguish them from their relatives. As a facile illustration, one can compare the status of Danish, Norwegian, and Swedish (or Span-

---

18. On a par with the linguistic criteria, the biological criteria for phylogenetic descent are still debated (e.g., on structural-cum-philosophical grounds: for example, is genetic overlap more important than interfertility in determining species boundaries? cf. chimps vs. humans, and lions vs. tigers vs. tigons/ligers). Yet, the genetic criteria for the phylogeny and identity of biological species are much better understood than their counterparts in historical-comparative linguistics: the biological criteria are grounded in increasingly sophisticated knowledge about genetics and evolutionary biology. It is now possible to try and retrace phylogenetic lines as far back as hundreds of millennia (e.g., back to our ancestral Eve; see Cavalli-Sforza 2000). This is quite unlike genetic linguistics (see the references in Note 3 for an array of challenges to *Stammbaumtheorie*; also see Note 7). Most linguists still have no clue about Eve's native language, notwithstanding speculations that "the first language had no affixes", which would make Eve THE Prototypical (Ur-)Creole speaker! (WSG: Section 5.2; also see references in Note 4).

ish, Portuguese, and Italian; or Serbian and Croatian) as "distinct" languages vs. that of Chinese as "one" language. Suffice it to say that these distinctions are more relevant for issues of identity (and) politics than for strictly linguistic (typological) matters.

Another example will drive the point home (in a metaphorical and literal, if not statistically-representative, sense): In New York, I once saw a sign that advertised (in Haitian Creole!): *Isit nou pale Franse* 'Here we speak French' (I have updated the sign's orthography to fit the official HC phonemic-spelling norms; see Y. Dejean 1980). The HC-as-French sign had been proudly displayed by a Haitian employee next to a Spanish sign advertising *Aquí se habla Español*.

One may well chuckle at this story and argue that it is surely not representative of Haitians' meta-linguistic attitudes, but we still need to ask: What are the precise linguistic-structural criteria – the operational typological threshold – that would classify, e.g., Missouri French and Cajun French as BONA FIDE varieties of French while HC is usually not so classified? Perceived notions of (non-)distinctness (in this case, between French and historically-related varieties) may be quite useful (or harmful) for creating and promoting stereotypes, political identities, and community boundaries, but they seem to have little to do with linguistic typology per se. It is thus not so surprising that the perception of separateness between a creole and its source languages is not uniform across all creole speakers; see Winford (1994: 45–48) and Mühleisen (2000: 84–92) for recent discussions.

For typological and sociological reasons (e.g., regarding creole-based education), I myself consider my native HC to be "distinct" from its Romance and Niger-Congo ancestors. For linguistic-typological reasons, one may well consider some variety of modern continental French (as spoken by Jean-Yves Pollock, say) to be distinct from its Latin and Old and Middle French ancestors as well as to its American relatives (Québec French, Missouri French, Cajun French, etc.); and so does Modern English fall into distinct varieties that in turn are distinct from Proto-Germanic and Old and Middle English varieties. In a somewhat related vein, I can also envisage, for methodological reasons, that children's early grammars are "distinct" from the grammars of their models – the older peers and caretakers that provide children with Primary Linguistic Data (see, e.g., Rizzi 1999 and references therein for discussions of children's "grammatical invention").

If our discussion of age-related complexity differentials (and the concomitant claim about "the world's simplest grammars") is to advance in a scientifically viable manner, we need an independent, theoretically grounded measure for measuring the birth and age of (new vs. old) languages in some linguistically relevant fashion that is impervious to our (often tacit) preconceptions, otherwise creolistics may well become the world's most simplistic science. Recall

Chomsky's (1957: 233) statement that "linguistic theory must be constructed with explicit and precise definitions and operational tests". Such explicitness and precision is even more urgent when dealing with languages that have generally been stigmatized from the very moment that they were identified and "baptized".[19]

In the absence of such measure (and given the discussion in Section 2; see Note 19), can we rely on our intuition to discriminate newborn from multi-millenarian languages? The question is not so simple.

Let's get back to the case of Jean-Yves Pollock as a speaker of Modern French. What exactly does it mean to say that (Modern) French is an old language? Here we must be terminologically picky in order to try and make sense of this question whose presuppositions are infected with metaphors that contradict one another. Do the idiolects of Pollock and his peers underlie an E-language that has existed for tens of millennia? If so, his (E-)French would actually include, and be the continuation of, a very long ancestry, including at least Middle French, Old French, Latin, etc. and all the varieties in between. And so would Spanish, Portuguese, Italian, Romanian, etc., be continuations (i.e., Schleicherian growths) of Latin and its descendants along the *Stammbaum* lines that connect Latin to the corresponding Romance variety. In other words, if each Romance language qua old language has, by definition, existed for "tens of millennia", then Romance speakers all speak a selfsame continuation of their common ancestor (say, Latin), which, in turn, implies that all Romance languages constitute the selfsame multi-millenarian language, by transitivity. Not a satisfying result, at least not for Romance linguists who have done much work to isolate robust parametric differences within Romance.

3.3.   *Creoles as multi-millenarian morphosyntactically wrinkled neonates*

Let us ask again: Can linguistic typology help us decide whether a given language has existed for "tens of millennia" or is "born again" and "begins anew"? Can neo-Schleicherian (i.e., genealogical) creolistics constructively engage linguistic theory in any scientific way (e.g., in any way that resembles how evolutionary biology engage (phylo)genetic analysis)?

Take morphosyntax. Can relatively well-understood typological properties (e.g., lexical semantics, derivational morphology, underlying word-order, scrambling, nominal case morphology, presence of definite articles) be used as genetic tracers for writing up the "birth certificate" of languages? Let's assume so (only for the sake of argument), and let's take these typological properties as the analogues of, say, DNA sequences in the dating of biological

---

19. Recall from Section 2 that the structural measures in (11)–(12) misdiagnose creole languages. So we cannot take the criteria therein as a litmus test for a creole typology (if any).

species (but see Notes 18 and 20 and Appendix B). Then, along the aforementioned typological dimensions, HC can be argued to be more similar to Modern French than Modern French is to Old French or to Latin. Thus, on the counts of both perceived similarity (in the eyes of certain creole speakers like that HC speaker in the Brooklyn store mentioned in Section 3.2) and typological closeness (along certain typological variables), HC seems no younger (or no older) than Modern French.[20]

What about morphology? It can be straightforwardly argued, as in Section 2.1 (also see Fattier 1998 and DeGraff 2001b, contra Lefebvre 1998), that most HC affixes historically derive from French affixes. Here the lexifier-creole similarities greatly exceed, both in cardinality and in systematicity, the sparse correspondences that have been used to argue for the "Afrogenesis" of French-lexicon creoles such as HC and Mauritian Creole from a single 17th-century West African pidgin ancestor (see Appendix B). In the case of HC, the lexifier-creole morphological continuity is not at all surprising: most of the HC lexicon is etymologically related to French. In turn, the French lexicon itself is mostly derived from Latin – with French emerging through language contact as occasioned by Roman imperialist conquests (compare with creole genesis in the context of Europe's imperialist conquests in Africa and the Americas).

Thus, the very morphology and history of HC (as compared to the morphology and history of, say, French) challenges the exclusive "born again" or "recent origins" status bestowed on radical creoles as a class by neo-Schleicherian creolists. As far as I can understand neo-Schleicherian techniques for linguistic phylogenetic analysis and for identifying language birth (but see Notes 18 and 20 and Appendices A and B), HC's etymological longevity and unifor-

---

20. This subset of typological features – lexical semantics, derivational morphology, underlying word-order, scrambling, nominal case morphology, presence of definite articles – is tendentiously chosen to illustrate (superficial) morphosyntactic similarities across creoles and their respective lexifiers and (superficial) morphosyntactic DIS-similarities within accepted *Stammbaumtheorie* phyla. Given the complex nature of Language (in an information-theoretic sense, within our current best theories) and given the concomitant vastness of typological diversity (which is somewhat still uncharted), it should be possible to handpick arbitrary sets of (superficial) morphosyntactic features along which any two languages will appear similar or dissimilar (see Appendix B). As Chomsky (1986: 151) reminds us, "even languages that have separated only recently may differ in a cluster of properties, something that has been observed in comparative studies". One such comparative study is Meillet (1929) where it is observed that "neo-Latin languages fall into a typological class that is quite remote from the structural type represented by Latin" (1929 [1938: 80], my translation; also see Meillet 1912 [1926: 148]). (For additional caveats on the use of morphosyntactic comparisons in phylogenetic linguistics, see Note 18; also see Thomason's caveat in (8).)

mity and its morphosyntactic inheritances-cum-restructurings should be the exclusive province of multi-millenarian languages.[21]

Recall the classic "pidgin-to-creole life-cycle" scenarios whereby pidginization creates a radical bottleneck for lexical and morphological development, thus forcing creoles to emerge from affixless ancestors (see Note 6; also see (12) and discussion in Section 4.2 below). A related and more general assumption is that "affixation [...] emerges from the grammaticalization, reanalysis, or reinterpretation of material which was not originally inflectional" (WSG: Section 5.2). Taken together, these assumptions entail that affixes in allegedly young languages such as HC must, in general, emerge via the grammaticalization of erstwhile free morphemes (see (12)). These scenarios are robustly contradicted by HC where, from genesis onward, almost all affixes have had, and still have, cognates in French affixes – which in turn often have cognates in Latin morphology. Etymologically these cognates diagnose millennia of seemingly unbroken transmission, quite an ancient pedigree for the morphology of a "most creole of creoles". Indeed, there is no documented stage in HC diachrony where the language was affixless or with most affixes derived from "erstwhile free morphemes" or with most affixes derived from outside of French (also see Section 2 and, especially, Appendix A).

In this respect, Goodman (1964: 26–28, 122–124) gives a variety of HC examples in the nominal and verbal morphosyntactic domains that suggest that "the initial speakers were exposed to a French which was virtually as complex inflectionally [...] as is standard French". Putting aside the fact that "standard French" (specially back then) was an artificially constructed language with few, if any, native speakers, Goodman's contention seems somewhat extreme, for at least two reasons: (i) most whites in colonial Haiti, and even in France, were more likely to be illiterate speakers of rural "patois" – Langue d'Oc, Langue d'Oïl, Norman French, etc. – than fluent speakers of standard/literate French (see, e.g., Chaudenson 1995: 18, Chaudenson & Mufwene 2001: 151–153); and (ii) not all creole creators were exposed to the same (non-native approximations of) native French varieties: right from the onset of contact, there must have existed a continuum of contact varieties, which were subsequently modulated through sociolinguistic factors into later varieties, including those known to us today (see, e.g., Alleyne 1971, Mufwene 2001; also see Note 15 and Appendices A and B).

---

21. In this regard, it is instructive to contrast HC – a "born again" language? – with English – a so-called "tens of millennia"-old Germanic language. In the latter, the majority of affixes are etymologically non-Germanic; this apparently is also true for the rest of the English lexicon which has been estimated to be 65 % non-Germanic. Furthermore, English diachrony and the genesis of HC exhibit comparable morphosyntactic differentials; see, e.g., DeGraff (1997, 2000, 2001a, forthcoming).

Nonetheless Goodman's point about HC's morphological complexity vis-à-vis its lexifier is valid to the extent that the available archival and comparative evidence suggests a robust degree of etymological continuity, which in turn disconfirms the radical morphological bottleneck posited by the pidgin-to-creole scenarios. Alleyne (1971: 172–174) makes a similar point, subject to similar caveats, when he gives linguistic evidence from French- and English-lexicon creoles that their lexifiers "in their full morphological systems where used in the contact situation". In the same vein, Mufwene (2000b: 9) writes that "to the extent that English pidgins and creoles, as well as indigenized Englishes, can ultimately be traced back to Old English, they all have a long history". Thus, as Goodman, Alleyne, Mufwene, and many others have argued before, there seems to be little, if any, evidence that creole genesis must prototypically proceed via "a radical reduction of [the] source languages into makeshift jargons" (cf. WSG: 144).

To recapitulate: Language age has long been taken as the crucial factor that determines level of complexity – this is in keeping with Schleicher's intuition about the genealogy of morphology (see Section 1.2). Schleicherian linguistics takes for granted the existence of some independent, precise and operational "language dating" algorithm for genealogical/phylogenetic analysis. Yet the language-dating heuristics that have thus far been used to diagnose language youth (e.g., pidgin-to-creole symptoms such as development of new affixes via grammaticalization) simply fail to account for robust data in HC. The latter's morphology is incompatible with its postulated ancestry in some hypothetical affixless pidgin. HC, as a sociohistorically prototypical creole language, manifests multi-millenarian morphological wrinkles.

At this stage, this reader is left begging what sorts of criteria are tacitly applied in creolists' genealogical heuristics. Thus far, it looks like we are dealing with either some arbitrary (perhaps sociologically motivated but unstated) presuppositions and/or some circular argument. The circularity would go something like this: Creole languages are "new" because they are creole languages whereas non-creole languages are "old" because they are not creoles.[22]

---

22. Adopting Thomason & Kaufman's (1988) model, one could equate their "genetic" languages with WSG's "old" languages and their "non-genetic" languages with WSG's "young" languages. For Thomason & Kaufman creole languages emerge "non-genetically" through some abnormal "break in transmission" whereas non-creole languages gradually evolve "genetically" via "normal transmission". Here is Thomason & Kaufman's litmus test for distinguishing "genetic" from "non-genetic" languages: "[I]f transmission has been interrupted, then there should be [. . .] a lack of correspondence among the various subsystems of the language, most probably between the lexicon as a whole and the grammar as a whole" (Thomason & Kaufman 1988: 11; also see pp. 8–12, 206, 211, etc.). Thomason & Kaufman's litmus test is challenged by the same sort of epistemological-methodological and empirical-structural problems already mentioned in the main text. Their structural criterion is not given any operational

Can creolists' theoretical elaborations on the concept "pidgin(ization)" and its import in the "pidgin-to-creole life-cycle" get us out of this conundrum?

4.   On "pidgins", "simplification", and "basic communication"

4.1.   *Epistemological issues: Vagueness, circularity, falsifiability, etc.*

One time-honored tradition in creole studies views pidgins and the proto-creoles they gave birth to as paragons of "basic [human] communication" with near-zero complexity. The precursors of this view go back to, e.g., Saint-Quentin (1872) and Schuchardt (1914); see (2a) and (19). We find similar views in 21st-century neo-Schleicherian creolistics (WSG: 126):[23]

(15)   [C]reole creators, in creating the pidgin that later developed into a creole, strongly tended to eschew traits from their native languages which were incidental to basic communication, and that such traits were therefore absent in the natural languages that the pidgins were transformed into.

I myself don't trust my intuitions on the elusive notion "basic communication", and specially not so in the teleological-functional context of (15) and its congeners in creole genesis scenarios. Above all, I don't know what the evolutionary and structural correlates of basic communication are. Has UG specifically evolved to perfectly implement basic communication? Is UG a, or THE, perfect (post-)pidgin grammar? What are the morphosyntactic requirements of "basic communication"? Are syntactic categories like N(P)s, V(P)s, C(P)s, etc., "functionally central"? What about XP-movement and other structural transformations? Couldn't "basic communication" do without them on a par with, e.g., formal languages? Ditto with respect to (abstract) Case marking

---

and testable measure. In the particular case of HC, lexical and grammatical correspondences between the creole and its lexifier show no greater amount of discrepancies than their counterparts in French and English diachrony. In fact, in certain domains (e.g., lexicon, morphology, underlying word order, and nominal inflection), the discrepancies in French and English diachrony may even be greater than in HC and Jamaican Creole diachrony, respectively (abstracting away from the rate at which discrepancies qua innovations propagate through the speech community). Similar non-creole, yet (perhaps?) "significant", discrepancies are also found in Indo-European and elsewhere (see, e.g., Thomason 1980 as cited in (8)). Pending an objective and falsifiable measure of "significant discrepancy" (what is Thomason & Kaufman's "significant" threshold?), certain lexical-vs.-grammatical "discrepancies" in non-creole diachrony seem as "significant" as in creole genesis. See Mufwene (1998, 2000a, b, 2001) and DeGraff (2001a, b, forthcoming) for further details on the theoretical abnormality of abnormal transmission in creole genesis (also see Notes 20 and 21).

23.   The statement in (15) controversially equates "creole creators" with those that "creat[ed] the pidgin". Besides, such equation is anachronistic: if (as claimed in (15)) creoles are "the natural languages that the pidgins were transformed into", then pidgin creation necessarily precedes creole creation.

and the X^0-vs.-XP distinction. Is there a maximum number of thematic roles per verb in "basic communication"? What is the shape of the lexicon in "basic communication"? What is the minimal phonemic inventory required by "basic communication"? Etc., etc. (Cf. Note 35.)

From what I gather in the UG literature (e.g., from *Linguistic Inquiry* articles), there is not much there that can be straightforwardly related to "basic communication" needs. And it even seems that there are sociolinguists interested in language change who, like generativists, are quite skeptical about teleological functionalism. For example, Labov (1994: Chapter 19) bears the title "The overestimation of functionalism". In the next chapter on "The maintenance of meaning", Labov concludes (1994: 598):

(16) A good many theories of language put forward recently would explain language structure as the result of the intentions of the speaker to communicate meaning to the listener. There is a part of language behavior that is subject to conscious control, to deliberate choice, to purposeful and reflective behavior. But as far as I can see, it is not a major part of the language faculty, and it has relatively little influence on the long-range of language structure.

Given such widespread, plus my own, skepticism about teleological explanations for language structure, I had hoped to find a clear structural definition of "basic communication" in 21st-century Schleicherian writings. After all, "basic communication" is THE linchpin of creole genesis scenarios based on age-complexity correlations. Yet "basic communication" is thus far left without any explicit and operational criteria. No independent algorithm is provided to derive the denotation of this term and its linguistic profile. This is (methodo)logically debilitating. In absence of independent criteria for "basic communication", Schleicherian creolists' arguments risk circularity of the following sort:

(17) a. "Basic communication" comprises whatever structural properties make up (many? most? all?) creole languages.
      b. Conversely, out of the vast array of superficial crosslinguistic distinctions, let's (arbitrarily?) select a quite small inventory of features that happen to not exist in (many? most? all?) creoles and let's make these features "incidental to basic communication". In other words, features that are absent in (many? most? all?) creoles are not "basic" to communication.

(18) Features "incidental to basic communication" include "ergativity, grammaticalized evidential marking, inalienable possessive marking, switch-reference marking, inverse marking, obviative marking,

'dummy' verbs, syntactic asymmetries between matrix and subordinate clauses, grammaticalized subjunctive marking, verb-second, clitic movement, any pragmatically neutral word order but SVO, noun class or grammatical gender marking (analytic or affixal), [. . .] lexically contrastive or morphosyntactic tone beyond a few isolated cases", and "lexicalized derivation–root combinations"

(WSG: Section 6 and Note 20)

The reasoning in (17) is the theoretical essence of neo-Schleicherian creolistics while the list in (18) is explicitly offered as a negative litmus test for basic communication. Yet, no independent justification and theoretical argument is advanced to explain why the features in (18) are "incidental to basic communication". Given the vast array of superficial crosslinguistic distinctions, why the ad hoc list of 15 features in (18) and not others? For example, why should ergativity, but not accusativity, be dispensable in basic communication? Pending answers to these and other questions above, it seems to me that appealing to some arbitrary list of scattered features to derive creoles' simplicity via a "basic communication" pidgin that lacks such features runs immediately into theoretical trouble (e.g., circularity, theoretical vacuity, and unfalsifiability).

4.2.   *On the making of "pidgins"*

What are "pidgins" and how do they emerge? One thing that we seem to know, based on a variety of comparative evidence, is that pidgins (be they "early"/"reduced" or "extended"/"expanded") cannot be uniformly reduced to some sort of lowest-common-denominator "basic communication" natural language: in standard descriptions, "early" and "extended" pidgins fall at opposite ends of the structural and functional continua.

About "EARLY pidgins", one common observation is that they often emerge as reduced communication systems used in restricted and specialized contexts (e.g., for sporadic limited exchanges outside of one's speech community). The prototypical, if controversial, definition for (early) pidgins is that they arise as "makeshift adaptations, reduced in structure and use, no one's first language" (Hymes 1971a: 3; also see Schuchardt 1909, Jespersen 1922: Chapter 12, Bloomfield 1933: 472–473, Hall 1962: 151–153, etc.; see references in Section 1.2 and Note 6). Bickerton (1999: 49), for one, considers early pidgins to be "reduced well below the minimum required by natural languages". It may even be argued that early pidgins may be unlike native languages to the extent that early pidgins' (lack of) structure seems to fall outside the formal boundaries for natural languages as set by UG; see DeGraff (1996b, 1999b: 499–500) for some discussion. It seems then that the "drastically reduced linguistic structure and lexicon" of early pidgins with restricted functions is the result of "the very first stage of rudimentary language learning" (Hall 1962: 151–153). These

structural and functional restrictions immediately disqualify early pidgins as candidates for viable full-fledged systems for human communication, IF – and this is a big "IF" – "viable full-fledged [...] human communication" entails the ability to encode for transmission the expressive needs of normal human beings across a functional range of topics. Something along the lines of this assumption is adopted in WSG: Section 2.3, where it is claimed that pidgins in the pre-expansion stage are "universally agreed to be rudimentary codes not fulfilling the needs of full language".

As for "EXTENDED pidgins" (e.g., in Melanesia), they seem to function beyond and above "basic" communicative needs: these pidgins incorporate various structural properties that are as formal (i.e., not "functionally central") as that of any full-fledged natural language (see, e.g., Hall 1962: 154–155, Keesing 1998, 1991, Siegel 1999; below I illustrate non-"basic" features in pidgins).

It thus appears that, as in other cases of (I-)language creation by adults, the making of pidgins leads to distinct looking results depending on contingent sociohistorical specifics. In Hymes's (1971b: 69) words, "the characteristics found in development to, and of, a pidgin admit of degrees. [...] pidgins and pidginization are instances *par excellence* of variable adaptation of means to an audience and situation".

Cartesian-Uniformitarian methodology (see, e.g., (3)) invites us to sort out historical processes and the external entities they create ("E-pidgins", say – social entities) from psychological processes and their concomitant individual-level creations ("I(nternal)-pidgins", say – mental entities); see Paul (1890), Andersen (1983), and Siegel (1999) for useful overviews and methodological caveats (also see Note 17). My (null) working hypothesis is that the making of I-pidgins (I-pidginization, if you will) enlists cognitive processes that commonly unfurl, not just in situations of abrupt and/or limited language contact, but also in the various instances of second language acquisition in "ordinary" contexts of language contact.

In this Cartesian (i.e., mentalist and internal) perspective, I(NTERNAL)-pidgins are not sui generis: they are the internal linguistic states – (transient or crystallized) interlanguages, if you will – in which adult language learners (qua second-language creators in need of a lingua franca) routinely find themselves. When viewed EXTERNALLY a stabilized (E-)pidgin, on a par with a communal (E-)language with native speakers, is a reification and conventionalization of the creations of individual speakers interacting in specific sociohistorical contexts, with their particular linguistic ecologies and their particular communicative requirements. From that perspective, the fact noted above that pidgins in distinct sociohistorical matrices may widely differ from one another (i.e., that pidginization and pidgins "admit of degrees", in Hymes's words) can be naturally and constructively related to, inter alia, the observed variability in

the outputs of second-language acquisition (e.g., with respect to the structural profiles and functional characteristics of adult learners' interlanguages and fossilizations thereof); see DeGraff (1999b: 479–508) for one overview and some references.

This Cartesian "I-pidgin(ization)" working hypothesis seems heretical to the field at large (see Siegel 1999 for one recent overview of diverging perspectives). Indeed, one (now familiar) truism in creolistics is that pidginization is a sui generis process that eschews (virtually) all morphology; see Jespersen (1922), Hjelmslev (1938), Bickerton (1988), Seuren & Wekker (1986), Seuren (1998), McWhorter (1998, 2000a, b, WSG), etc.). Per this truism, pidgins are uniformly (and, for some, teleologically) designed *ab ovo* as simplest languages. The corollaries of this truism – the pidgin-to-creole life-cycle and its concomitant morphological bottleneck – constitute received wisdom in language-contact and historical-linguistics textbooks. By definition, the pidgin-to-creole cycle is exclusive to creole formation and is radically different from processes underlying the diachrony of non-creole languages. It is noteworthy that this scenario is still part of the communis opinio in creole studies notwithstanding the fact that the "classic" pidgin-to-creole litmus test fails on representative creoles, including HC (see Section 2 above and also Alleyne 1971). At this point, a bit of critical historiography is in order, before exploring modern exponents of the pidgin-to-creole cycle.

The view that pidginization entails a morphological bottleneck – a "stripping" of language-particular morphology – is already found in Schuchardt's description of "the creole before [it] become[s] the native language of the majority" (1914 [1980: 91]):

(19)   For the master and the slave it was simply a matter of mutual comprehension. The master stripped off from the European language everything that was peculiar to it, the slave suppressed everything in it that was distinctive. They met on the middle ground.

Schuchardt's hunch echoes through much of contemporary creolistics. We still find allusion to "stripp[ing]" as in Bickerton's (1988: 272–278) claims that "a sharp, and in some cases quite radical, reduction in the structural properties of the original target language was an essential prerequisite for new language formation" and that such reduction entails that, in the formation of radical creoles, "the target's bound morphology [is] stripped [. . .] thoroughly".

More recently, Schuchardt's "simpl[e] matter of mutual comprehension" has been linked (as in (20a)) to "basic communication" and its "functionally central" features (see, e.g., (15) and (20b)). Here, one ill-defined term – Schuchardt's "middle ground" – is replaced in WSG by another ill-defined term – "basic communication" – which, in turn, is based on some ill-defined

mechanics for "stripping away virtually all of a language's complexity (as defined in WSG), such that the complexity emerging in a creole is arising essentially from ground zero". And both the early-20th and the early-21st century scenarios involve "deliberate design" – somewhat reminiscent of "naive or teleological design" as criticized by Labov (1994: Chapters 19–20), see (16) and also Paul (1890 [1970: xliv–xlvi]):

(20)   a.   "[P]idgins [are] stripped of almost all features unnecessary to communication"   (WSG: Abstract)

        b.   "[Pidgins are] communication vehicles deliberately designed to eschew all but the functionally central"   (WSG: Section 2.3)

In (19), Schuchardt takes some intuitive impression of "basic variety" cum "baby/foreigner talk" (cf. Bloomfield 1933: 472, Ferguson 1971, 1975, 1981) to a structural extreme. He assumes that speakers in contact situations can systematically suppress structures that are "peculiar"/"distinctive" to their respective native languages in order to create an "emergency language" for "mutual comprehension". But this entails that, whenever speakers of (say) languages X and Y need an "emergency language" for "mutual communication", X speakers can correctly decide which of their native structures will "meet with [Y speakers'] total incomprehension", and vice versa. And this would be why, in the European-African "emergency language", the Europeans eliminated European affixes (e.g., plural-marking suffixes such as English *-s*) while the Africans, in analogous fashion, suppressed the expression of African affixes (e.g., plural-marking prefix such as Duala *ma-*); see Schuchardt (1914 [1980: 91–92]).

I find Schuchardt's claims in (19) and its modern implementations (see paragraphs below) theoretically and empirically challenging, even if seemingly common-sensical. In particular, I don't understand the psycholinguistics of finding the "middle ground". Ferguson (1971, 1975, 1981) seems right that every native speaker can resort to some recognizable and negotiable simplified register for speaking to linguistically handicapped foreigners. Yet, notwithstanding the broad tendencies identified by Ferguson and others toward universals of simplification (see, e.g., contributions to Clyne (ed.) 1981), foreigner talk doesn't seem to constitute a crosslinguistically well-behaved "simplest" structural type. Ferguson himself (1971: 146, 148; 1981) shies away from positing (simplification in) "foreigner talk" as a general, absolute, and sui generis process in pidginization. Instead, regarding simplification, Ferguson cautiously notes:

(21)   It is [...] clear that relative simplicities [e.g., smaller lexicon, less morphology and allomorphy] occur under a wide variety of circumstances, such as in pidginization, normal diachronic change, language

> acquisition, language pathology and register variation, although the
> details differ from one set of circumstances to another. It is also clear
> that simplification is rarely if ever the exclusive diagnostic charac-
> teristic of a particular language/variety/register in comparison with
> another [. . .].

It may well be the case that all contact situations entail simplification (how-
ever defined – e.g., as reduction in structural irregularities) to some noticeable
extent. It may also be the case that simplification can happen without (large-
scale) language contact (I return to this below).[24] However, speakers engaged
in language contact are neither telepaths, nor (psycho)linguists, nor fluent in
each other's languages. Therefore they cannot SYSTEMATICALLY decide what
in their native speech should unambiguously count as "peculiar"/"distinctive"
to the foreigner's ear. Furthermore, in deciding what's "peculiar"/"distinctive",
the "middle ground" creators must, strangely enough, abstract away from pho-
netics – the language-particular component that is most accessible to the for-
eigner's ear. This is much easier said than done. For now, the (psycho-)
linguistics of negotiating this "middle ground" strikes me as quite mysterious.
One (perhaps less mysterious) alternative then is to posit that speakers of any
pair of languages X and Y know in advance what the "middle ground" ought to
be, independently of any contrastive analysis of X vs. Y. A cognitive prerequi-
site for successfully establishing this "middle ground" is that speakers of X and
Y (and of all other languages) share a universal set of hardwired instructions for
finding this "middle ground" as "simply a matter of mutual comprehension".
(See Clyne (ed.) 1981 for bibliographies on foreigner talk; also see Section 4.4
on simplification in pidginization.)

One possible non-innatist, TELEOLOGICAL answer to the puzzle of Schu-
chardt's "middle ground" equates the latter to a pidgin qua (near-)perfect "ba-
sic communication" system (see (20)). This hypothetical pidgin is built al-
most exclusively on "functional central" features (i.e., it is "stripped off of
almost all features unnecessary to communication"). However, it is still not
clear how speakers (or linguists for that matter) can "deliberately" sort out
between "functionally central" vs. communicatively "unnecessary" linguistic
properties. Deciding what is "functionally central", and why, remains an ever-
elusive task that has long frustrated expert linguists who are deliberately tack-
ling this problem in the leisure of their research offices, thus the lively debates
in functional linguistics toward the discovery of deep-seated, (i.e., non ad hoc)

---

24. See, e.g., Meillet (1919 [1926: 201]), Weinreich (1953: Section 2.3), Givón (1979: 20–
    22), Thomason (1980: 361–362; see quote in (8)), Trudgill (1989: 228–229), Chambers
    (1995: 160), and also Note 5. DeGraff (1997, 2000, 1999d) and DeGraff (ed.) (1999) offer
    comparative case-studies across creole and non-creole diachrony.

correlations between function and structure. (Also see Mufwene 2000a: 72–76 for sociohistorical arguments against the teleological view of pidginization and creolization.)

There is one camp though where something like "basic communication" – qua universal set of INNATE instructions for finding the "middle ground" – has been proposed. Recently, Klein & Perdue (1997) have proposed that all second language learners go through a stable and universal "basic variety" (BV) stage. This BV, although not mentioned in WSG, somewhat looks like "basic communication" in WSG, at least in spirit: BV is a well-defined I-language, a predetermined state of the language faculty. The BV's prototypical features are: no inflectional morphology, no grammatical morphemes, (NP_1)–V–NP_2 order, tense-marking via adverbials, no movement, no complex hierarchical structure, etc. (Klein & Perdue 1997: 311–326, 332, 336). These "organizational principles" are genetically wired via some sort of "core UG", thus relatively independent of the native and target languages. Here is how Klein & Perdue locate BV vis-à-vis UG – compare (22a) with (20), (22b) with the "contingent accumulation of 'ornamental' elaboration that older grammars drag along with them" in WSG: Section 2.3, and (22c) with (29):

(22) a. [T]he human language capacity provides us with the potential to process very complex structures but does not force us to do so. [...] [BV] IS simple and still extremely functional.
                                                        (Klein & Perdue 1997: 302)
       b. [F]ully fledged natural languages are but elaborations of this BV. They add some specific devices, such as inflectional morphology or focus constructions; they also add some decoration, pleasant to the ear, hard to learn, but faithfully handed down from one generation to the next. But essentially, they build on the same organizational principles.   (Klein & Perdue 1997: 304)
       c. [T]he BV simply and directly reflects the necesary, rather than the more accidental, properties of the human language capacity.
                                                        (Klein & Perdue 1997: 304)

Not only is Klein & Perdue's BV controversial among second-language researchers, on both empirical and theoretical grounds (see articles in Jordens (ed.) 1997), but more to the point its postulated structure is quite unlike what we see across pidgins (see Section 4.3). That pidgins do not instantiate a uniform structural template is also recognized by Klein & Perdue (1997: 340). Furthermore Klein & Perdue's "basic variety" is expressively handicapped and lacks some of the structural characteristics (e.g., complex hierarchical structure) associated with full-fledged languages (Klein & Perdue 1997: 302, 333). Thus BV is not a good candidate for an expressively adequate "basic communication" system. Plus the ban on (complex) recursion in BV does make it

an unlikely candidate for a prototypical ("simplest") human language. In fact BV would require more, not fewer, constraints to prevent Merge from creating complex embeddings; this makes BV quite im-"perfect" as an I-language (pace Klein & Perdue's 1997: 337 claim that BV is a "perfect" I-language). If BV is as "perfect" as claimed by Klein & Perdue, then the ban on complex embeddings in early interlanguage is not a strictly I-language phenomenon rooted in some core parameter-setting; instead such ban must be a side-effect of rather superficial online production strategies that ease the processing burden of the non-native learner at the beginning of acquisition.

4.3. *Are pidgins designed* ab ovo *from "ground zero" complexity?*

The exclusively teleological-functional definition of pidgins – as speech that "eschew[s] all but the functionally central" (see (20)) – is incompatible with a variety of well-documented facts from psycholinguistics and contact linguistics.

On the empirical front, alongside robust evidence for various kinds of simplification in pidginization, there are well-documented pidgin structures that are "inherited" from (some of) the source languages (see below). In other words, these structures were not "eschewed" from the creole creators' native languages, whether or not these structures were "functionally central". The retention and reanalysis of source-language structures in pidginization (and creolization) is not surprising given what psycholinguistics and sociolinguistics have taught us about language transfer in second-language acquisition and about the dynamics of contact linguistics; see Weinreich (1953), Labov (1994, 2001), Mufwene (1990, 2000a, b, 2001), and DeGraff (1996b, 1999b, d) for some overview and further relevant comments. Besides, the fact that certain pidgin structures are absent in many "old" languages across phyla and across time suggests that not all pidgin structures are required for basic communication. This too is unsurprising. The sociolinguistic specifics of each instance of language contact are contingent on history. It is thus tautological that pidgins' source languages (and other relevant sociolinguistic factors in language contact) vary across time and space. Pidgins will thus "inherit" (and re-analyze) selected patterns from the languages in contact. Since it cannot be the case that every such "inheritance" exists in every ("old") language, these "inheritances" cannot all be taken as "functionally central" to "basic communication", lest many "old" languages are communicatively dysfunctional.

Counterexamples to (20) go as far back as Schuchardt (1914), if not earlier. Schuchardt documents widespread substrate influence on the developing "emergency language" at all levels of structure: syntax, lexicon, lexical semantics, proverbs, etc. For Schuchardt, such transfers are quite natural (1914

[1980: 93]):[25] "[T]he slaves spoke the creole not only with the Whites but also among themselves while their mother tongue was still in existence, the latter being moreover constantly revived to some extent by the continual immigration from Africa."

This rudimentary sketch of the socio- and psycho-linguistics of language contact in creole genesis has since been confirmed and refined across a wide variety of cases. It has now been painstakingly documented that pidgins are pregnant with (reanalyzed) structures from the languages originally in contact, alongside structural innovations. Such massive transfers, restructurations, and innovations give rise to an array of syntactic options for any given semantic function within and across pidgins. Unsurprisingly distinct pidgins select distinct functions for morphosyntactic marking and their morphosyntactic options often enter into competition for the expression of similar functions.[26]

Now, one could well try and argue that whatever source-language properties survive pidginization and make it into the creole must be "functionally central" properties that are "necessary to human [basic] communication". In other words, given evidence for admixture in creoles (see, e.g., references in Notes 26 and 31), if creoles begin anew at virtually ground zero, then it must be the case that language transfers do not introduce (substantial) complexity in creole genesis.[27] Witness the following quote from WSG: Section 4.4:

(23)   Creole languages are unique in having emerged under conditions which occasioned the especial circumstance of stripping away virtually all of a language's complexity (as defined in this paper), such that the complexity emerging in a creole is arising essentially from ground

---

25. For Schuchardt (1914 [1980: 91]), "the African languages exert a pace-setting influence [...] [n]ot after the creole had already become the native language of the majority; nor yet when it was created as an emergency language". In contemporary creolistics terminology, we can translate Schuchardt as saying that substrate influence via transfer is most felt sometime during the (more stable) pidgin phase, after the initial emergence of the "emergency language" (i.e., the jargon or early-pidgin), but before it acquires native speakers.

26. For some overviews and case studies of substrate influence and structural diversity within and across language-contact specimens, see, e.g., Hall (1944), Hymes (ed.) (1971), Bickerton (1977), Koopman (1986), Nichols (1986), Alleyne (1988), Faraclas (1988), Holm (1988), Thomason & Kaufman (1988: Chapter 7), Jahr & Broch (eds.) (1996), McWhorter (1997), Keesing (1988, 1991), Arends, Kouwenberg, & Smith (1995), Drechsel (1993, 1997), Singler (1996), Thomason (ed.) (1997), Migge (1998), DeGraff (ed.) (1999), Lumsden (1999), Siegel (1999), Mufwene (2001).

27. A somewhat related, though distinct, argument has been explicitly advanced by Adam (1883: 4–5); see (2b). For Adam it is the substratum that limits the complexity of the "hybrid" (read "relexified") creole. This is in opposition to Saint-Quentin (1872) where creole simplicity is "a spontaneous product of the human mind, freed from any kind of intellectual culture"; see (2a) and Note 35. See DeGraff (2001b: 90–98, 106: Note 7) for some discussion.

zero, rather than alongside the results of tens of thousands of years of other accretions.

This postulated "stripping away" entails that pidginization systematically filters out from the languages in contact "the results of tens of thousands of other accretions", allowing the retention of only features that are "functionally central" (cf. WSG: Notes 11, 13). However, given the right ecology, admixtures do carry along "incidental" features from the languages in contact. This carry-over of source-language features is favored by, e.g., relative homogeneity of (some of) the languages in contact, relative exposure to these source languages, the socio-psychological profiles of the speakers in contact, and other sociolinguistic incidental factors (see the references in Note 26, specially works by Alleyne, Mufwene, Siegel, Singler, and Thomason & Kaufman). Such factors may even include the whims of "one [pidgin] speaker"; see Nichols's (1986: 240) speculations about the grammaticalization of evidential marking in Chinese Pidgin Russian (see (24b)). No matter how they get carried over into the emergent contact language, these admixtures cannot all be subsumed under the (still elusive) category of "functionally central" properties. Indeed these admixtures and concomitant restructurings-cum-innovations do not seem required by basic communication; they even include some of the "incidental" features in (18). To wit, the very preliminary sample in (24) (also see the discussion in Section 6.4).

(24)   a.   Capeverdean Creole has object clitics and inflectional number marking (Baptista 1997: 262, 2001); Baptista (2001) also documents suffixal number marking in a variety of contact languages.
        b.   Chinese Pidgin Russian has one grammaticalized evidential marking (Nichols 1986; also see (25)).
        c.   Fanagalo Pidgin has noun classifiers (Heine 1978: 223).[28]
        d.   Kitúba exhibits a periphrastic aspectual marker preceding the verb and a bound tense marker suffixing onto the verb (Mufwene 1997: 179).
        e.   The Lingua Franca exhibits inflectional gender marking and agreement (Muusse & Arends forthcoming).

---

28. It is interesting to note that Fanagalo, which Heine (1978: 228) considers "an extreme case of pidginization", manifests six nominal class prefixes – more than Tsez (cf. WSG: Section 3.1.2) and more than (most) contemporary Germanic and Romance languages, all of which should count as much "older" than Fanagalo. This is as expected given the linguistic ecology surrounding Fanagalo's genesis and the accidental aspects of language creation. This also illustrates the theoretical futility and empirical vacuousness of neo-Schleicherian creolistics.

f. Mobilian Jargon has, inter alia, OsV word order where "s" denotes a pronominal subject, with "S" denoting a non-pronominal (full) NP subject (Drechsel 1993: 356–357, 1997: 128–135).

g. Ndjuka-Trio Pidgin exhibits OsV, in addition to sOV and SOV; and in WH-questions, the WH-object can occur either before or after the subject (Huttar & Velantie 1997: 105–108).[29]

h. Nubi Arabic has lexically and morphosyntactically contrasting tone (Heine 1982: 26, 41–43, etc.).

i. Palenquero has object clitics (Schwegler & Green forthcoming).

j. Palenquero has (at least) three strategies for sentential negation: preverbal negation, discontinous double negation (with simultaneous pre- and postverbal marking), and postverbal negation (Schwegler & Green forthcoming).

k. Papiamentu manifests both tone and stress (Rivera-Castillo 1998).

l. Solomons Pijin has a postverbal (and non-affixal) aspectual marker that exists alongside other tense/mood/aspect markers that are preverbal (Keesing 1991: 325–330).

m. Taimyr Peninsula Russian-based Pidgin has predominating SOV order (Wurm 1996: 86–87).

What this means from a superficial (E-language) perspective is that pidginization and creolization appear to drag, restructure, and (re-)create arbitrary amounts of so called "long-ago"/"baroque"/"random" accretions, thus perpetuating, and adding to, whatever complexity may have already existed in prior diachronic cycles. (Also see Note 26 and Muysken & Law's 2001: 49 caveat, cited above in (13).)

In a related vein, the sheer structural diversity of pidgins refutes the claim that pidgins are "communication vehicles deliberately designed to eschew all but the functionally central" (see (20)).[30] Pidgins' structural diversity includes

---

29. Notwithstanding McWhorter's (WSG: Section 5.2) disclaimers separating his own notion of complexity from other linguists' claims about (word-order) markedness, SVO is listed in WSG: Section 6 as one of the items that make up the structural profile of "the world's simplest grammars" (see the arbitrary list of "complex" features in (18)). Thus, along this particular word-order dimension and given the superficial metric in WSG, Ndjuka-Trio Pidgin (with, e.g., SOV, sOV and OsV) is both more marked and more complex than "old" languages that are straightforwardly SVO. Note that the Ndjuka-Trio word-order alternations are surely not functionally central to basic communication given their absence in many presumably-functional "old" languages.

30. Do the counterexamples in (24) really disconfirm neo-Schleicherian proposals? WSG: 151 offers a ready-made, but misleading, answer: "The identification of scattered exceptions in various creoles to the general tendency I have identified does not constitute a refutation of my argument". My counter-reply is simply based on the fact that this caveat is not heeded in WSG. For now, the features that are characterized as "incidental to basic communication"

patterns that are unfamiliar to speakers of "old" languages like English and French. Commenting on such diversity while sketching the "exotic" character of Chinese Pidgin Russian, Thomason & Kaufman (1988: 191) sum up this picture much better than I can, so I quote at length:

(25)   Like the other pidgins described in [the section on diversity in pidgin structures], Chinese Pidgin Russian has features that are unusual among the better-known pidgins and creoles with European lexical sources: SOV word order, postpositions as well as preposition(s), V NEG word order, and a few inflectional and derivational affixes. [...] None of these features could be predicted as the result of the operation of universal structural tendencies alone, because the suffixes represent marked constructions, and the word order features are different from the ones found in other contact languages. The presence of both preposition(s) and postpositions is itself rather highly marked in universal terms. [...]

We should emphasize, finally, that the examples given in this section do not by any means exhaust the instances of pidgin structures that are not promising candidates for simplified lexical source language features or features of universal grammar. [...] Our goal here has been to

---

(e.g., in (18)) seem even more "scattered" than my counterexamples. And so are the features of the Creole Prototype in (11). For example, what (if any) theoretical principles motivate the postulation that the Creole Prototype in (11) lacks "tone distinguishing monosyllabic lexical items or encoding morphosyntactic distinctions" (as in (11b))? "Monosyllabic lexical items" and "morphosyntactic distinctions" do not look like a natural class to me. The ad hoc disjunction in (11b) could well be replaced with the equally ad hoc statement "lack of tone distinguishing words with exactly two vowels or encoding distinctions in contrastive stress". While many of the pidgin structures above can, in principle, be related in a natural fashion to well-known (psycho)linguistic facts of language acquisition/creation and language contact (e.g., language transfer, restructuration, grammaticalization, regularization, simplification, innovation, etc.), the list in (18) hardly forms a theoretically justifiable natural class for computing complexity (see Section 5 below). Compare, say, the counterexamples and counterobservations referenced in the main text with the arbitrary list of scattered features in (18). The complexity metric defined (negatively) through (18) is based, not on a theoretically-cogent "general tendency", but on features that seem (relatively) rare crosslinguistically – rare, at least, among the "old" languages that Western linguists are most familiar with (including the lexifiers of Caribbean creoles). WSG even quotes the quite telling remark that some of its non-creole "test" languages (Kabardian and other languages of the Caucasus) are "extraordinarily complex by any linguistic standard" – if so, the "test" languages in WSG will surely make many other languages (creole or not) look extraordinarily simple by the very metric in WSG.

In any case, given the very complexity of Language and the vast space for potential distinctions at all levels of grammar (phonology, morphology, syntax, semantics, pragmatics, discourse, etc.), the ad hoc and scattered list of "incidental" features in (18) is truly "scattered": such a list cannot reliably estimate crosslinguistic rankings of overall complexity (see Section 6.4 below).

demonstrate that origin theories based solely on evidence from well-known, well-documented mainstream pidgins and creoles are inadequate to the extent that they fail to predict the kinds of features we have illustrated.

The last sentence in (25) applies straightforwardly to the definition of pidgins in (20) and to much else in the dogma that "the world's simplest grammars are creole grammars". And, alongside Muysken & Law's *mots justes* in (13), there is another caveat, from Thomason (1997: 6–7), that demystifies this dogma:

(26)   [T]he structural descriptions [of the "exotic" pidgins and creoles in Thomason (ed.) (1997)] provide a strong antidote to the still common view that all pidgins and creoles have similar and simple structures. Features like systematic OSV and SOV word order patterns of Hiri Motu, the noun class system of Kitúba, and the /kp/ and /gb/ phonemes of Sango will surely help to eradicate the idea that pidgins and creoles have maximally simple and more or less identical grammatical structures.

WSG systematically skirts all empirically "strong antidote" against the view that creole grammars are simplest. WSG offers no reference whatsoever to the "exotic" pidgins and the crosslinguistically rare distinctions that have been documented in Chinook Jargon, Chinese Pidgin English, Chinese Pidgin Russian, Fanagalo Pidgin, Hiri Motu, Kitúba, Mobilian Jargon, Ndyuka-Trio Pidgin, Nubi Arabic, Pidgin Delaware, Russenorsk, Sango, Taimyr Peninsula Russian-based Pidgin, etc. Yet there exist valuable treatises on many of these "exotic" pidgins. Mobilian Jargon is one case in point: it is closely examined by Drechsel (1993, 1997) who observes that "pidgins need not reflect universal patterns as thought earlier [...] but may actually exhibit highly marked features of syntax" (1993: 344).

Pidgin structures should then count as extremely diverse, with structures that often have no counterpart in many "old" languages. There is no space here to illustrate the complete range of such diversity: the possibilities may not be endless, but they do seem to go beyond what can reasonably be imagined as "basic communication"/"functionally central" requirements. This is not surprising given the history of many pidgins in conditions of intensive language contact. I encourage the reader to consult the relevant references in order to appreciate the extent to which the definition in (20) is empirically untenable; also see DeGraff (1999b: 479–508) for additional observations and references on the ontology of pidgins.[31]

---

31. Surprisingly, WSG's definitions in (20) and the creole genesis scenario therein contradict em-

With Schuchardt's and others' detailed evidence of admixtures, restructura-
tions, and innovations in creole genesis, any sui-generis definition of "creole"
as "language" [that] 'begins anew' amidst pidginization" is a fallacy. The avail-
able evidence thus far does not fare well for the teleological proposal in (20)
that pidgins are "communication vehicles deliberately designed to eschew all
but the functionally central". The semantics of "functionally central" should
not be up for grabs: the "functionally central" in basic human communication
– if definable – should be universal and should not be left to be determined on
a case-by-case basis by some ecological roll of dice, lest we run into empirical
and theoretical incoherence.

4.4.   *"Simplification":* Terminus a quo *and* terminus ad quem

We have now tasted our "strong antidote to the still common view that all
pidgins and creoles have similar and simple structures". Pidgin structures are
typologically diverse and they are definitely not a subset of "old" language
structures – many pidgin structures are nowhere to be found in functional "old"
languages such as English or French.

 This said, one must reckon that the products of (large-scale) language con-
tact do give the impression that they are, to a certain degree and in certain
domains, simpler than their corresponding source languages. For example,
overt morphological paradigms (e.g., phonetically-realized inflectional affixes
on nouns and verbs) tend to decrease in size, morphological irregularities tend
to be filtered out, various sorts of semantic transparency tend to increase, etc.

---

pirical and (anti-Bickertonian) theoretical observations in McWhorter's own work, such as:
"The presence of serial verbs in creoles, then, is not the result of a functional necessity. Their
presence is the result of their being A GRAMMATICALLY CENTRAL information-encoding
strategy of UNIFORM AND WIDESPREAD AREAL DISTRIBUTION in West Africa, such that
there is no reason to suppose that they would not be transferred into an emerging contact lan-
guage by West Africans, and then retained as the language developed through time and space
in the mouths of adults and children" (McWhorter 1997: 155, emphases in original). Here
McWhorter is arguing against Bickerton's view of "Saramaccan as UG *in vitro*" (McWhorter
1997: 12). Elsewhere we read, still in the pro-substratum anti-Bickerton vein: "A great many
structures that Bickerton designates innate are in fact much more likely to have been transfers
from the languages spoken by the slaves first brought to the Caribbean. Bickerton's claim
has been that any such similarities between Creole and African structures are mere accidents.
While it is hardly impossible that such accidents could have occurred – especially given the
'unmarked' nature of many of the structures in question – comparative analysis makes it rela-
tively unequivocal that many of the Caribbean-African correspondences are indeed transfers,
not spontaneous creations" (Kegl & McWhorter 1997: 20). While these two quotes capital-
ize on substratal influence to challenge Bickerton's Bioprogram Hypothesis, WSG (see, e.g.,
Note 11) systematically downplays similar substratum-influenced data in favor a catastrophic
Bickertonian scenario. Nonetheless the same sort of arguments levelled against Bickerton's
Biogrogram – including those in McWhorter (1997) and McWhorter & Kegl (1997) – also
applies against McWhorter's WSG proposal.

But it must be stressed that such simplification is not absolute. As documented through much of the creolistics literature, what we are dealing with is gradient simplification with respect to the languages in contact and their respective complexity in PARTICULAR domains of grammar. For each specific *terminus a quo* whose composition is determined by contingent sociohistorical factors, simplification leads to a necessarily distinct *terminus ad quem*. Given incidental variations in the particular linguistic ecology and in the relevant sociopsychological and demographic factors, the *terminus ad quem* in certain cases will be more complex in certain grammatical domains than the *terminus a quo* in other cases (for case studies, see references in Note 26 and also Romaine 1992: 217 for a similar observation and Muysken & Law's 2001 important caveat in (13)).

What are the sources of simplification in language contact as in the creolegenesis cases? This is a complex question. DeGraff (1999b: 491–499, 517–518) gives one, admittedly incomplete, answer, which is rooted in a Cartesian perspective that views pidginization and creolization as reducible to individual-level mental processes that are shared across the species (cf. (3)). There I argue that simplification stems from the cognitive limitations of adult language learners and the concomitant mechanics of second-language acquisition "under duress, in the initial stages of language acquisition in the context of language contact – contact that may be massive and abrupt, and that may involve considerable social and psychological distance between speakers in different language groups" (1999b: 491). This Cartesian (i.e., mentalist) view of simplification is Uniformitarian: the underlying (psycholinguistic) causes of simplification in creole genesis are not, and could not be, exclusively "creole" (see Section 4.2). Simplification patterns are due to the idiolect-formation mechanisms that are necessarily employed in the creation of both creole and noncreole languages. In fact, we find similar simplification patterns in well-studied cases of language change via language contact and, at the individual level, in the creation of interlanguages, and in child-language acquisition (see, e.g., contributions to DeGraff (ed.) 1999 and references therein).[32]

With this in mind, let us re-examine the conceptual basis of the frequent claim that simplification in pidginization, unlike simplification elsewhere, creates linguistic neonates that, as a class, start life from virtually "ground zero"

---

32. WSG's view of (second) language acquisition and its effects in creole genesis appears incoherent to me. On the one hand, WSG seems to accept the sensible view that (degrees of) pidginization can be reduced to adult acquisition, as in the history of the Riau dialect of Indonesian whose "unspecified nature is almost certainly due to a degree of pidginization in its life cycle, due to extensive acquisition by adults, having 'shaved away' a large degree of accreted complexity" (WSG: Section 4.4; also see WSG: Note 3). At the same time, WSG argues against the view that "creoles are born via the gradual 'streamlining' of a lexifier language via succeeding waves of second-language acquisition" (WSG: Note 3).

of complexity (see (23)). Let's put aside the grotesque claim that the *terminus a quo* of pidginization (i.e., the pidginizers' native languages) is invariably at "ground zero" complexity (cf. Note 27). How can it, then, be guaranteed that simplification of, and transfer-cum-restructuration from, the source languages, alongside structural innovations, uniformly creates pidgins at ground zero of complexity?

Let's consider the four relatively uncontroversial propositions in (27):

(27) a. "[Effects of] pidginization [can be] due to extensive acquisition by adults" (WSG: Section 4.4); "creolization is a cline phenomenon" (WSG: Note 3).

b. Second-language acquisition in the context of language (or dialect) contact is a common phenomenon across the history of many (creole and non-creole) languages; see Meillet (1926, 1938), Weinreich (1953), Trudgill (1989), Labov (1994, 2001), Chambers (1995), Mufwene (1998, 2000a,b, 2001), DeGraff (ed.) (1999), etc.

c. The structural results of second-language acquisition are characterized by various degrees of transfer from native languages, alongside various degrees of restructuration, simplification, and innovations – all based on species-uniform cognitive processes and the necessarily contingent nature of the particular linguistic ecology, its functional demands and socio-psychological profile; see Siegel (1999), DeGraff (1996b, 1999d), De Graff (ed.) (1999), Mufwene (1990, 2001) for overviews and bibliographies.

d. The source languages in creole genesis cannot (all) be at "ground zero" of complexity. This is as expected if these source languages are themselves "older language[s] [that] retain at all times a degree of complexity alongside [. . .] simplifications" (WSG: Section 4.4) and if "[o]ne would find a great many of [communicatively/functionally non-basic] features in the lexifier and substrate languages that were spoken by the creators of [the simplest] creoles" (WSG: Section 6).

If all four of these propositions hold, then the claim that "the complexity emerging in a creole is arising essentially from ground zero" in (23) is a non sequitur. The conjunction of the propositions in (27) entails, a priori, that the *terminus ad quem* in creole genesis vs. language change cannot be segregated at opposing poles of some non-arbitrary global complexity metric that takes into account the entire grammars of these languages. This is because, alongside simplification, restructuration, and innovation, language contact also entails language transfer through second-language acquisition, which will inevitably carry over into the emerging contact language some of the complexity from the

languages in contact (this is sketchily illustrated and discussed in (24) and surrounding comments). Thus, language contact cannot induce a natural class that represents "the world's simplest grammars". The latter, it seems to me, fall in what Foucault would call a "pure and simple linguistic monstrosity" – an unavoidable result of "the quest for primitive language [, which quest induces] a world of chimera and reverie" (see (4)).

All the aforementioned facts and observations about transfer, innovation, simplification, and typology in creole genesis and beyond should wake us up from any "chimera and reverie" whereby pidgins become systems for "basic communication" that eschew "all but the functionally central" (cf. (20)).[33]

4.5.   *"Basic communication": What are the basics?*

Popular scenarios for the emergence of "the world's simplest grammars" are fraught with epistemological problems: their theoretical foundations have long been undermined by notions left critically ill-defined – including "creoles", "pidgins", "young" vs. "old" languages, "basic communication", "functionally central" features. Regarding the latter two notions, in the absence of any independently justified theory for "functionally central" properties and their crosslinguistic realization, we still lack a coherent, non ad hoc notion of "basic communication" as a linguistic-theoretical concept. For now, "basic communication" remains vague and elusive.

What would be needed to adequately define "basic communication" is a theoretical framework (some universals of "basic communication", say) that would independently motivate the "functionally central" ingredients of "basic communication" and spell out how they are minimally realized crosslinguistically at all levels of grammar. Such a framework would, for example, predict the exclusion of the features in (18) from basic communication and explain why these features alongside other "incidental" features go beyond the requirements of "basic communication" (but see Labov's caveat about "naive or teleological design" in (16)). In contradistinction, the list of features that are "incidental to basic communication" (see (18)) is scattered through the space of typological variation and is constructed outside any independent theory of "basic communication".

---

33. Creolists who firmly believe that there exist synchronic structural diagnostics that exclusively define pidgins and creoles may well decide that the "exotic" pidgins mentioned here are not really "prototypical" pidgins. If so, it will be incumbent upon these pro-prototype creolists to provide explicit and operational structural definitions of "pidgins" and "creoles" (and of "young" vs. "old", and "simple(st)" vs. "(most) complex"). In order to reach a minimal level of scientific adequacy, such definitions must be neither ad hoc, nor circular, nor vague. Only then will they have theoretical and empirical bite (see Sections 5 and 6).

At this point, some pro-prototype creolist may optimistically respond that it is Universal Grammar (UG) itself, or some version of the Language Bioprogram à la Bickerton, that tells us about (the structure of) "basic communication". Or perhaps "basic communication" in creole formation is a community-wide fossilized instantiation of the "basic variety" seen in Section 4.2 (cf. (22)). And recall that the "basic variety" itself is postulated as the product of some "minimal/core UG". So we should ask: Does UG define the "functionally central" requirements of "basic communication"? Before answering this question, I first need to spell out some working assumptions about UG.

It is usually assumed that UG, by its very nature, does underspecify all idiolects, whether creole or non-creole. As far as I can tell, there is no sense (yet?) in which UG defines a scale whereby languages can be ranked as being more or less overspecified across all domains of grammar simultaneously – or more or less removed from some innate system for basic communication. Underspecification is the very essence of UG qua biological template for Human Language. UG only defines the set of PERMISSIBLE languages; no ACTUAL language is defined by UG alone; see, e.g., Chomsky (1986: 145–152, 1995: 6). This is akin to the way in which genotypes underspecify phenotypes. Language (with capital "L", in the singular) is innate, but languages (with small "l", in the plural) are not. That is, humans are hardwired for Language whereas the individual expression of this capacity as idiolects – (I-)languages with their particular phonetics, lexicon, morphology, syntax, semantics, etc. – is not biologically programmed, even though it is biologically-constrained.[34]

Generativists aim at one abstract implementation of this underspecified template for all human languages; see, e.g., Chomsky (1986: 145–152, 1995: 6–7, etc.). In this research program, this abstract Human Language template (aka UG) consists of "principles" and "parameters". Principles are presumably universal, ultimately hardwired in human biology. They exist alongside an array of underspecified parameter settings and/or an array of open slots for a language-particular lexicon with its concomitant phonemic inventory, morphology, lexical semantics, etc. The parameter-settings and lexical slots become (over)specified only after exposure to Primary Linguistic Data on an idiolect-specific basis (compare with the emergence of phenotypes via the interaction of innate genotypes with incidental environmental variables). In other words, UG specifies no actual parameter settings and no actual lexicon: it is the inevitably con-

---

34. This assumption is orthogonal to specific issues about how Language is represented in the mind/brain (e.g., questions of modularity – with respect to possible interactions between linguistic and non-linguistic cognitive processes). For my argument to go through here, it is only necessary to assume (uncontroversially, I suppose) that whatever properties ultimately enable language acquisition (i.e., the creation of human idiolects) are inscribed in human biology one way or another.

tingent linguistic experience that fills in – that "specifies" – the idiolect-specific information. This framework makes it axiomatic that every actually-occurring idiolect (including creole idiolects) will be "overspecified" with respect to UG. The latter only defines the space of – the boundary conditions on – possible human languages; it does not specify any one particular language or any one particular class of languages (pace Bickerton 1988; see (28)). Neither does UG specify a global hierarchy for classifying languages in terms of overspecification at all levels of grammar taken simultaneously.

Let's contrast this view with the proposition that creole languages, because of their alleged youth, represent the minimal – "simplest" – instantiation of some universal set of structural requirements as dictated by UG. This is the essence of Bickerton's Language Bioprogram hypothesis.[35] Here UG is taken as a sort of lowest-common-denominator grammar with respect to which specific languages are more or less overspecified. Such overspecification is claimed to go hand-in-hand with complexity: "old" languages are the most overspecified and they are the most complex, see the following quote from Bickerton (1988: 274):

(28) The present viewpoint accounts very naturally for this "simplicity" [of creole grammars]. In older languages, the universally shared set of syntactic principles is added to, and complicated by, a wide range of lexical and morphological properties as a result of millennia of diachronic change.

It thus seems that it is overspecifications vis-à-vis UG that, over millennia, take "old" languages away from the structural ideal of "basic communication" and away from "ground zero" complexity. Witness the following Bickertonian passages in WSG:

---

35. In 1872 already, a somewhat similar proposal was advanced by Saint-Quentin (see (2a)). A century or so later, we hear from Seuren & Wekker (1986: 64) that, as the main factor in creole genesis, "[Semantic Transparency] enabl[es] listeners to carry out semantic interpretation with the least possible machinery and with the least possible requirements on language learning" and from Seuren (1998: 292) that "[creole grammars] lack the more sophisticated features of languages backed by a rich and extended cultural past and a large, well-organized literate society". The later statement seems to imply that "sophisticated" morphosyntax must be backed by old and literate ("sophisticated"?) culture and that culturally un-"sophisticated" people (whatever that means) speak morphosyntactically un-"sophisticated" languages. But must "primitive" people speak "primitive" languages and "sophisticated" people "sophisticated" languages? The answer is no, given Sapir's well-founded observations in (7) on the "rubbish"-ness of morphology–culture correlations and my discussion in Section 6.4 of complexity–age correlations. (Also see Note 27 and DeGraff 2001b for recurring parallels between Saint-Quentin's and later proposals on creoles as "*ab ovo*" creations.)

(29) a. [I]n older grammars, millennia of grammaticalization and reanalysis have given overt expression to often quite arbitrary slices of semantic space, the result being a great deal of baroque accretion which, while compatible with [UG], is incidental to it, as well as to even nuanced human expression. In having not existed for long enough a time for drift to encrust them in this manner to any great extent, creoles are unique in reflecting the innate component of the human language capacity more closely than older languages do.   (WSG: 126)

      b. The observations I have made are couched in a view of older natural language grammars as vastly OVERSPECIFIED systems in comparison to the requirements of [UG].   (WSG: 160)

      c. [Are] older grammars' structures [. . .] completely, or even mostly, specified by, as opposed to merely compatible with, [UG?] [A]s I have argued, this is not the case.   (WSG: 132)

      d. [Are] creoles [. . .] closer to [UG] than other languages [?] The present paper is an attempt to explore and support [the] provocative idea [that they indeed are].   (WSG: Note 5)

      e. [C]reoles represent a fundamental layer of natural language, unobscured by the results of millennia of phonological, syntactic, and semantic drift which make [UG] such a challenge to glean in older languages [. . .].   (WSG: 155)

      f. In the realm of syntax, the hypothesis that creoles are closer to an ontogenetic foundation than many other languages appears promising.   (WSG: 157)

The quotes in (28)–(29) are incompatible with the conception of UG that I sketched above; see Marantz (1983) for a related critique in the context of Bickerton's Language Bioprogram Hypothesis.

As I have already pointed out, UG is not a "basic communication", or lowest-common-denominator, grammar in the sense of a minimal set of "functionally central" requirements. This is perhaps made clearer by taking the lexicon as example. The lexicon is yet another area of grammar where one language can "[give] overt and grammaticalized expression to more fine-grained semantic and/or pragmatic distinctions than another" (cf. WSG: Section 136). Individual lexica and the distinctions therein (e.g., vis-à-vis semantic distinctions as for, say, kinship terms – along with distinctions in phonemic inventory, contents of functional heads, affixal inventory, etc.) become fully specified only upon exposure to contingent Primary Linguistic Data in sociohistorically-determined environments. Independently of (say) phonemic inventory and complexity thereof, the Primary Linguistic Data give rise to arbitrary *signifiant–signifié* semantic oppositions of arbitrary complexity, including word-level ("opaque

lexicalizations") and phrase-level non-compositional semantics (e.g., idiom chunks) – much of this goes back to Saussure. The point here is that the lexica of natural languages are in no way fully specified sensu stricto by UG, no matter the eventual number of distinctions therein. Furthermore, lexical(ized) distinctions and the arbitrary semantic partitions they establish are, a priori, orthogonal to, say, phonemic inventories and/or their complexity: there is no reason to expect complexity qua number of distinctions (as in WSG) to increase in lockstep across all levels of grammar.

In any case, the actual (phonetic and semantic) make-up of lexical items IS constrained by universal laws of phonetics, by universal constraints on semantic interpretation, constraints on argument structure, and its linking to surface representations, etc. Take Haitian Creole as an example. At the phonemic level, HC may look simpler than French or English. But, at the lexical and morphosyntactic levels, HC uses operations like reduplication and predicate-clefting for semantic stress. The PRECISE morphological, syntactic, and semantic details of these operations are not necessarily "simpler" (i.e., with fewer overt distinctions) than the reduplication and cleft patterns that exist in "old" families like Romance and Germanic. Yet reduplication and predicate-clefting in HC must also obey UG strictures on a par with, say, affixation and verb-movement in Romance and Germanic.

As I have already mentioned, it is axiomatic that any given idiolect, though biologically bound by UG, will be overspecified with respect to UG's initial parametric and lexical slots and so on. These slots are necessarily left underspecified (i.e., open to parametric choices) in the initial state defined by UG. The crux of the matter is that, notwithstanding apparent crosslinguistic quirks, all such quirks and myriad others (including those found in creole languages) will, by assumption, fall within the boundaries defined by UG and indeed will help us discover the make-up of UG. No amount of complexity-building via diachronic drift can take languages "beyond the bounds of the genetic specification for language" (contra WSG: Section 6.3; see (14) above). Indeed, it is tautological that our "genetic specification for language" (i.e., the genetic encoding of UG) enables us to learn ANY "overspecification" – ANY " 'ornamental' elaboration" – in ANY human language; that is, the mind/brain is genetically pre-wired to acquire, store, produce, and process any and all the "overspecifications" that exist across the world's languages (see Note 17). Any linguistic feature that could not be so acquired, stored, processed, etc., would just not exist in any natural language, assuming with Descartes, Humboldt, Osthoff & Brugman, Paul, Chomsky, and others that natural languages are mental properties of *homo sapiens*. Thus, the necessity for crosslinguistic research: it is linguistic diversity that will help us elucidate the boundary conditions imposed by UG. It is by apprehending the diversity of specific languages that we will elucidate the unity of Human Language.

This Cartesian methodology puts an ironic epistemological twist on the (neo-)Schleicherian claims in (29). Is it really "such a challenge to glean [UG] in older languages" while "creoles represent a fundamental layer of natural language [that is] unobscured"? If creoles were really "the world's simplest grammars" with the fewest distinctions possible, then this would actually make it HARDER for pro-prototype creolists to "glean UG" and the diversity it affords. Prototypical Creoles make the "prototypical" creolist's job most trivial: "the world's simplest grammars" require no more than the world's simplest analyses. In this perspective, Prototypical Creoles, as defined in WSG, would have little, if anything, to contribute to theoretical progress in linguistics (be it in phonology, morphology, syntax, semantics, etc.). As Chomsky (1986: 149) writes, "[q]uite often, the study of exotic phenomena that are difficult to discover and identify is much more revealing, as is true in the sciences generally".

Let's imagine a would-be analogue of (29e) in the hard sciences. Imagine, say, some hypothetical claim to the effect that chemists have their best shot at discovering the molecular make-up of nature by examining the world's simplest molecules. If the latter were the only data that chemists had to experiment with, then they certainly would have no clue about the diverse complexity of nature's "Universal Chemistry". In particular, chemists working with "prototypically simplest" molecules (chemistry's equivalent of "Prototypical Creoles") would have no opportunity to glean the complex structures of proteins and DNA – the very molecules that make our existence possible. Similarly, if our field notes and intuitions were exclusively about Prototypical Creoles as defined in WSG, then we would have no clue about inflection, tone, Saussurean lexicalizations of root–affix combinations, ergativity, grammaticalized evidential marking, inalienable possessive marking, etc. The list in (18) hints at other features that UG makes available, but that linguists could never glean from the hypothetical "simplest grammars" defined by that list (cf. Note 36). In DeGraff (2001b: 76–78, 86–88), I argue that the Creole Prototype defined by (11) – by (11c) in particular – even lacks some of the basic Saussurean properties that are usually associated with natural languages and their lexicon (see Section 2).

If UG both defines the language learner's innate initial state and imposes boundaries on the outcome of acquisition and, thus, on each and every I-language, then it cannot be the case that UG is "a challenge to glean in older languages" because of their "incidental" features. To the contrary, UG is best studied through our exploration of the diversity of languages, and this exploration is best carried out when guided by our theoretical results about the unity of Human Language. In this view, Prototypical Creoles as theoretical constructs with simplest, unexpected, and ad hoc properties do constitute a typological-ontological (and epistemological) challenge to (the study of) our *faculté de langage* and the crosslinguistic structural possibilities it affords.

From the perspective of UG as sketched in this section, Prototypical Creoles become "linguistic monstrosit[ies]" (cf. (4)).

Be that as it may, we still lack any operational criteria for "basic communication" and its "functionally central" properties. Yet at this point we must raise the central question: How is "complexity" defined?

5. A most simplistic "complexity" metric

First, some words of caution: Any descriptively and scientifically adequate complexity metric requires an independent theory of complexity (that explains what is to be counted, why, and how) along with exhaustive descriptions for the languages to be compared (so we can list all that is to be counted), lest our complexity metric have exclusive scope on arbitrary bits of grammar with no consequence whatsoever for linguistic theory and global complexity across languages. No general claim about crosslinguistic levels of complexity is reliable if it focuses solely on a small set of disparate superficial patterns that are not unified in any kind of linguistic theory or psycholinguistics. More generally, simplistic and highly selective measures of complexity whose benchmarks focus on arbitrary and isolated aspects of surface strings in some handpicked sample of languages seem largely orthogonal to the theoretical and/or psychological foundations, and to the descriptive goals, of linguistic typology and theoretical linguistics (see Notes 37 and 40).

I thus agree with Muysken (1988: 288) that "the idea that creole languages are not grammatically complex in general only makes sense if one has a theory of grammatical complexity to fall back on". Chaudenson (1994) makes similar points, as he notes the absence of any coherent evaluation metric in past and current allegations of extraordinary creole simplicity. Without any independent theory and formal criterion for complexity, we cannot even begin to determine how particular properties (or absence thereof) contribute to global complexity. With this in mind, let's proceed to evaluate the most recent complexity metric in neo-Schleicherian creolistics.

5.1. *Defining complexity via description length (= number of information bits)*

The complexity metric in (30) is simply a count of "overt distinctions and/or rules" in (30a) (WSG: Section 2.4.3), which in turn is related to "length [of] descriptions" in (30b) (WSG: Section 2.4.2).

(30)   a.   The guiding intuition is that an area of grammar is more complex than the same area in another grammar to the extent that it encompasses more overt distinctions and/or rules than another grammar.

b. Our object of inquiry is differentials between grammars in degree of overspecification (as we will see, all grammars including creoles can be argued to be overspecified to some degree), to the extent that some grammars might be seen to require lengthier descriptions in order to characterize even the basics of their grammar than others.

Let's call this view of complexity BIT-COMPLEXITY.

Bit-complexity immediately faces a number of unresolved methodological and theoretical problems that render it scientifically unusable at best and tendentious at worst.

Let's paraphrase (30) in a transparent information-theoretic way. As sine qua non for a rigorous and objective application of the sort of complexity metric sketched in (30), we must at the very least get straightforward answers to the following questions:

(31)   a. The sort of complexity that is hinted at in (30) is proportional to "degree of overspecification", thus to "length [of] descriptions". For any given language *L*, what is the number (*n*) of information bits needed to describe the (entire set of) "overt distinctions and/or rules" therein? (*L*'s complexity increases with *n* which is proportional to the length of *L*'s description.)

b. The count *n* in (31a) presupposes a theory of grammar that would enable – or provide an algorithm for – the identification, then the counting, of language-specific "overt distinctions and/or rules". What is the (implicit) theory of grammar that identifies the items to be counted by *n*?

In Sections 5.2 and 5.3, I address these two questions in turn.

5.2. *Not all "bits" of grammar have theoretical bite*

In order to adequately answer the "*n*" question in (31a) for any given language, no less is needed than (an approximation of) the description of this language at all linguistic levels (i.e., for all "area[s] of grammar [with] overt distinctions and rules", including phonology, lexicon, morphology, syntax, semantics, discourse, etc.; cf. (30a)). Much progress has been made in typological linguistics, yet such exhaustive descriptions are not generally available for all of the world's languages. In the absence of such descriptions, no *n* can be reliably estimated toward an unbiased global comparison of all natural languages (e.g.,

in order to discover what "the world's simplest grammars" are).[36,37]

Pending such exhaustive descriptions for all areas of creole and non-creole grammars, what one should expect from an analysis that partitions natural language into simplest and most complex classes is a general theory whereby one can safely extrapolate from the (apparent) complexity of isolated and arbitrary linguistic properties (see, e.g., (18)) to global complexity. No such theory is hinted at: the very features in (18) from a very small and selective set of (apparently) "far out" languages seem to have been picked exactly so that the few creoles chosen as "test cases" show less bit-complexity in the corresponding domains than the few non-creoles chosen as "control cases" (see Note 37). What we have had thus far in many searches for the world's simplest languages are formulas for "rigged" experiments – experiments that are designed to guarantee the desired "results".[38]

---

36. Exhaustive and reliable grammars are notoriously lacking for creole languages, which still lack strong communities of native-speaker linguists (compare, for example, the study of Saramaccan syntax with that of Dutch syntax). And, to this day, the empirical basis of creole studies is weakened by the long-standing prejudices of linguists like Seuren who consider creoles to "lack the more sophisticated features of languages backed by a rich and extended cultural past and a large, well-organized literate society" (1998: 292) and by the too-common practice in creolistics to base one's arguments solely on skewed and sparse samplings of unanalyzed utterances as if creoles wore their entire grammars on few superficial strings (see Appendix B). It thus seems that any arbitrarily stipulated metric that ranks creoles as "the world's simplest grammars" is a self-fulfilling prophecy (see Note 38).

37. WSG hunts for the most unfamiliar non-creole languages with the most "exotic" linguistic properties (e.g., Tsez, Maori, Lahu, etc. with a combined total of less than one million speakers). Yet WSG systematically avoids the anti-Schleicherian "antidote" of less-familiar contact languages and their less-familiar – thus, certainly un-"basic" – features (cf. Sections 2, 4, and 6, especially the quotes in (13), (25), and (26) and the sort of data illustrated in (24) and Section 6.4). In effect, WSG: Section 1 ultimately pays only lip service to the stated "intention [for] a sustained investigation of creoles from the perspective of crosslinguistic configurational possibilities, beyond the Western European lexifier languages that have served as the primary focus of creolists' attempts to define the term creole". Such investigation also requires the thorough comparison, at all levels of grammar, of "old" Western European languages (e.g., English and French) with "born again" contact languages whose ancestors exclude European languages and include, say, "fearsomely elaborated" languages such as Tsez, Lahu, Maori – the non-creole benchmarks for "old language" complexity in WSG. Such systematic comparison is sorely missing, which makes hypothesis-"testing" in WSG look like a rigged experiment (see Section 6.4).

38. The coarse and a-theoretical complexity metric in WSG: Section 2.4.2 is justified as follows: "I believe that the difference in degree of complexity between older grammars and a subset of creole grammars is distinct enough that a complexity metric so fine-grained as to, for example, allow us to rank Romanian, Hausa, and Korean in terms of some general complexity quotient would be unnecessary to our project." This a priori belief that "the world's simplest grammars are creole grammars" becomes self-justification for replacing the theoretical necessity of a "general complexity quotient" in favor of a stipulated "metric" based on a scattered list of features based on a skewed sample of languages (see (18)). The metric itself is devised with the expressed goal of separating creoles from non-creoles; it ignores a large body of relevant

5.3. *How many "rules"? The looks of languages vs. the essence of grammar*

For the sake of the argument, I will now abstract away from the methodological issues in Section 5.2. Instead let's ask this: Is complexity only, or primarily, a matter of counting information bits – "overt distinctions and/or rules" as in (30a) – no matter the source of these bits (i.e., no matter the theory underlying these distinctions and rules)? If a complexity metric involves counting, then we better make sure we know what we are counting. Linguistic distinctions and rules are not pretheoretical objects that we can gather and count without prior analysis.[39]

In fact, bit-complexity bears no relation to any theory where grammatical phenomena are independently identified and analyzed. WSG (Sections 2.4.1–2.4.3 and Note 6) explicitly cuts off its complexity metric from the better understood areas of (psycho)linguistics, including grammatical theory, acquisition, production, and processing. In other words, bit-complexity may well have no basis in (what we know about) Language in the mind/brain – our *faculté de langage*. Bit-complexity, as defined in (30)–(31) is strictly a-theoretical: this is literally bit counting with no concern for psychological-plausibility and theoretical insights.[40]

---

evidence – including data in McWhorter's other work (see Note 31) – and it makes no recourse to, and no prediction vis-à-vis, (psycho)linguistic theory. It thus seems that the main purpose of the proposed metric is to justify the author's latest "belief" about creole genesis, against all the available counter-evidence and against existing theories (also see Notes 37 and 40).

39. In biology, where discussion of complexity is grounded in more solid empirical and theoretical results, it is not at all clear that naive counting (e.g., the counting of genes) could lead to any scientifically satisfying notion of complexity. For example, Szathmáry, Jordán, & Pál (2001) write: "Is the number of genes in an organism's genome an appropriate measure of biological complexity? [...] The recent flurry of completed genome sequences, including our own, suggests that this is not necessarily the case [...] Rather surprisingly, it turns out that the worm Caenorhabditis Elegans has 18,424 genes in its genome, the fruit fly Drosophila Melanogaster 13,601, the plant Arabidopsis about 25,498, and humans about 35,000. This suggests that there must be other, more sensible genomic measures of complexity than the mere number of genes." Szathmáry et al.'s proposal is to use "networks of transcription factors and the genes they regulate, rather than [...] simply counting the number of genes or the number of interactions among genes". Thus, biologists go beyond simple counting of overt items and they enlist inter alia abstract computational theories about "the connectivity of gene-regulation networks". These theories are at the core of our understanding of how genes work (e.g., the mechanics whereby certain genes are switched on and off in order to represent information and compute over these representations). No need to say that these theories go beyond my competence. Yet such a development strikes me as normal for the sciences. No matter the field, complexity measures – if they are to be scientifically constructive – must be related to broad empirical and theoretical concerns (cf. Darwin 1871: Chapter 2, 61–62 for relevant remarks on "how easily [our biological and linguistic complexity metrics] may err").

40. The apparent divorce in WSG between bit-complexity and complexity as understood by psycholinguists (e.g., language acquisition and language processing researchers) is all the more unexpected given the often discussed relations between creole genesis and language acqui-

At the onset, let's note one of the (bizarre) logical consequences of bit-complexity: the languages with the biggest lexica would be the most complex – independently of, say, their phonology and their syntax. Indeed, each new lexical item further partitions the speaker's semantic space (recall Saussure's view of the lexicon as a system of oppositions). For any given degree d of complexity, a large enough lexicon automatically carries enough "overt distinctions" to make the corresponding (I-)languages complex to the d degree. In this view, the proposed complexity metric (in particular, (30a)) applies even within the "same" (E-)language of a given speech community: within that community, the (I-)languages with the biggest lexica will unavoidably require the "length[iest] descriptions" (cf. (30b)) and will thus be the most complex languages ceteris paribus. This strikes me as a rather naive view of language complexity. Number of (superficial) "overt distinctions and/or rules" without regard for linguistic theory (assuming for the sake of argumentation that such sets can be made available) seems, to me at least, a rather crude and uninteresting way to approach linguistic typology.

A simplistic bit-complexity creolist could well try to save the "overt distinction" metric and argue that the lexicon altogether lies outside the scope of his metric. So let's now move from bit-complexity in the lexicon to bit-complexity in the syntax. For the syntax too, bit-complexity simplistically implies counting – here, counting of "rules" (WSG: 136):

(32)   A syntax is more complex than another to the extent that it requires the processing of more rules, such as asymmetries between matrix and subordinate clauses (e.g., Germanic verb-second rules), or containing

---

sition cum processing (compare with approaches to complexity in biological evolution; see Note 39). Separating complexity from psycholinguistics leads to incoherence within the very assumptions in WSG. If "creoles represent a fundamental layer of UG" and if "[i]n the realm of syntax, creoles are closer to an ontogenetic foundation than many other languages", then creoles should be the easiest languages to acquire (à la Bickerton). Furthermore, the bit-complexity of morphology is explicitly linked to "the development of morphophonological processes, which constitute an added component of a grammar to be LEARNED" (WSG: Section 137; emphasis added). Acquisition researchers have related complexity issues (say, phonological and semantic transparency in morphology) with ease of acquisition; see, e.g., Clark (1993). Here, complexity is linked directly to acquisition and storage. Thus, WSG does more than "take as a given that all languages are acquired with ease by native learners" (cf. WSG: Section 2.4.2). WSG's metric also implies that creoles should be acquired and/or processed with the most ease – as in Saint-Quentin (1872), Adam (1883), Seuren & Wekker (1986), Bickerton (1988), etc.; see (2) and Note 35). This contradicts WSG's attempts (Section 2.4.2) at isolating bit-complexity from acquisition and processing. (DeGraff (ed.) 1999 provides an overview of the constructive theoretical linkages among creole studies, historical linguistics, and language acquisition research, in a Cartesian framework that does not postulate fundamental structural differences between creoles and non-creoles.)

two kinds of alignment rather than one (i.e., ergative/absolutive and nominative/accusative) [...].

But (how) do we know when the "syntax [...] requires the processing of more rules" without a theory of syntax and without a theory of processing? Here, one cannot simply rely on the "looks" of language; instead one must evoke the "essence" of grammar – one must enlist a theory of grammar lest any complexity metric becomes fatally ill-defined. Deciding whether and where a particular syntax "process[es] more rules" presupposes an independent theory of grammar which the syntactician can use to discover the language-particular "rules" to be counted. Syntax is not directly readable from strings; see, e.g., Hawkins (1988) for similar points within a comparison of generative vs. typological approaches to grammar. Even creole languages, which in neo-Schleicherian creolistics "represent a fundamental layer of natural language [that is] unobscured" (see (28) and (29e–f)), do not bear their syntax on their strings. Actually, theoretical creolistics, like much else in linguistic theory, is the theater of vivid debates about the nature of creole "rules"; see, e.g., DeGraff (1999c, 2001b) and Y. Dejean (1999a); also see Section 6.4 and Appendix B for some anti-Prototype samples. In any case, to-date we have no exhaustive list of syntactic rules for the relevant languages, and comprehensive descriptions are even more sorely lacking for "exotic" contact languages, specially those created outside the Caribbean basin (see Section 4.2 and Note 36). Be that as it may, there is an inherent methodological cum conceptual fallacy in a complexity metric that relies on the counting of (language-specific) syntactic rules without an explicit theory of syntax for identifying and classifying said rules.

One example will straightforwardly illustrate the extent of this fallacy. Let's consider the statement that "asymmetries between matrix and subordinate clauses (e.g., Germanic verb-second rules)" entail an increase in complexity via the "processing of more rules" (see (32)). This is presumably because such asymmetries involve distinct "rules" for root vs. embedded clauses, thus an increase in the number n of "overt distinctions and/or rules" (cf. (31a)). No analysis is presented for the (added) rules that underlie this added complexity – the latter is taken for granted, and mistakenly so.

Since at least den Besten (1981) it has been argued, and it now seems quite likely, that there isn't any "asymmetric [...] Germanic V2 rule" per se. Typically, V2 in German(ic) results from the application of X^0- and XP-movement rules, both of which are made universally available by UG. The finite verb moves quite high outside of VP (e.g., to the C(omplementizer) position) while a maximal projection (e.g., some topic or operator) moves to the left of the verb (e.g., in Spec(CP)). The root–embedded "asymmetry" itself is understood by many Germanicists to be epiphenomenal, emerging as a surface side-effect of the interaction between abstract syntactic ingredients (e.g., head and phrasal

movement plus the features and contents of V, C, Spec(CP), etc., and their associated functional projections in root vs. embedded contexts).

Let's consider the German(ic) V2 asymmetry as an example – grossly oversimplifying the available analyses. German matrix clauses have a C(omplementizer) head that is usually empty, thus available as a possible landing site for V(erb)-movement, with the finite verb surfacing right-adjacent to a moved XP in Spec(CP), thus matrix V2. In German, the embedded C head is usually filled by an overt complementizer, which blocks V-to-C movement; the embedded finite V is thus stuck in the IP, in clause-final position; thus, the root–embedded asymmetry with respect to V2. Yet, in embedded clauses that allow an EMPTY C head, the "V2 asymmetry" disappears: V-to-C and XP-to-Spec(CP) take place, giving rise to an embedded V2 pattern that is "symmetric" with the root V2 pattern. Similar symmetry is robustly displayed in Germanic languages such as Icelandic and Yiddish. V2 in these languages can be argued to result from movement of the finite verb, not to C, but to a head lower than C, thus the lack of asymmetry since V2 does not depend on the (lack of) contents of C. Crosslinguistically, observed V2 (a)symmetries (as in German, Dutch, Yiddish, etc.) reduce to the complex interaction between X^0 and XP movement (and the triggers and/or semantics thereof) and language-specific properties of clause structure and functional heads, etc. (See Vikner 1995 for an overview.)

Notwithstanding current debates about the exact mechanics of V2 in Germanic and beyond, the lesson from syntax about the alleged complexity of "asymmetric" rules is clear. The above (simplified) analyses for V2 teach us that "V2 asymmetry" does not necessarily entail "the processing of more rules" than (say) "V2 symmetry". Even "symmetric" languages such as Yiddish and Icelandic instantiate similar sorts of movements, albeit within a different clausal topology. At the right level of analysis, the so called Germanic "movement rule asymmetries between matrix and subordinate clauses" become a rather superficial side-effect of a single uniform operation – head-movement of V into C and XP-movement to Spec(CP) – which applies whenever possible. The blocking of V-to-C due to overt C arises via independent morphosyntactic requirements (e.g., selectional requirements and the morphology and semantics of the CP layer). In this view, root-vs.-embedded (non-)V2 patterns in Germanic (and elsewhere) are not the result of distinct (root vs. embedded) transformations; instead such patterns result from deeper universal principles of the computational system of our *faculté de language* (e.g., structure building, selectional requirements, movement transformations, etc.) interacting with language-specific morphology and/of functional heads.

From this theoretical perspective, root–embedded asymmetries do not necessarily increase complexity in ways that are alien to creole languages. In fact, creole languages too manifest (superficial) root–embedded asymmetries,

contra the erroneous generalization in WSG: Section 6. Haitian Creole, for
one, has such asymmetries with respect to WH-phrases in direct questions:
the moved WH-phrase, no matter its underlying position, overtly moves to the
root Spec(CP). So in HC direct WH-questions, the WH-phrase is pronounced
in the root clause, never in an embedded clause; thus the appearance of a root–
embedded asymmetry. This is unsurprising in any theory where (direct) ques-
tions need to be typed as such at the root level and/or where semantic operators
need to take scope over their quantification domains. (Also see Syea 1997 for
another possible instance of creole root–embedded asymmetry, with respect to
copula distributions and V-to-C movement; cf. the Mauritian Creole data in
Appendix B.) In creoles (as in non-creoles), apparent asymmetries result from
deeper universals of syntax.[41]

   More generally, syntactic theory in the generative framework has witnessed
a fundamental move away from lists of language-specific and/or construction-
specific "rules". Current generative syntacticians have adopted the Principles-
and-Parameters/Minimalist hunch that "constructions" arise via the complex
interaction between, on the one hand, operations and constraints that are uni-
versal and, on the other hand, language-specific properties that reside mostly in
the lexicon – in particular, in (the morphology of) functional heads. The num-
ber of operations made available by UG may well be few (e.g., Merge, Move,
Agree, etc.). Yet these few universal operations interact in complex ways with
numerous language-particular properties, thus the vast and intricate array of
superficially distinct crosslinguistic phenomena (e.g., the "V2 asymmetry" in
German, the WH-movement asymmetries in Haitian Creole, the copula asym-
metries in Mauritian Creole, the Chinese in-situ WH-phrases, etc.). In such a
framework, complexity does not reside in the number of different "overt dis-
tinctions and/or rules" (e.g., distinct/asymmetric rules for root vs. embedded

---

41. With respect to WH-movement, it is instructive to test the complexity claim in (30a) by a
comparison of Haitian Creole (a "born again" language) and Chinese (an "old" language).
HC, but not Chinese, has overt WH-movement. Thus Chinese shows fewer "overt distinctions
and/or rules": both WH-questions and their declarative counterpart manifest the same surface
word order. There is no root–embedded asymmetry in the surface distribution of Chinese WH-
phrases. This is unlike the overt asymmetry in the positioning of WH-phrases vs. non-WH XPs
in HC. Yet, at some abstract level ("Logical Form"), it can be argued that Chinese – a WH-
in-situ language – does have WH-movement of the sort found in HC. Indeed the distribution
of Chinese WH-phrases and their structural relations to abstract scope positions obey some of
the same constraints that regulate the distribution of overtly moved WH-phrases as in HC; see,
e.g., Huang (1982) for the basic Chinese facts, Chomsky (1986: 152–155) provides a handy
summary. Here too, as with the V2 case, the view in (30) is theoretically naive; it is too tied to
the surperficial "looks" of languages to offer any deep insights into the evolution of grammar
and grammatical complexity. Indeed the metric in (30) ignores much of what syntacticians
have taught us about the abstract "essence" of grammar. (See Appendix B for similar flaws in
McWhorter's 2000b Afrogenesis Hypothesis.)

German clauses). Such construction-particular and language-particular "rules" may not even exist, although their labels are often retained for taxonomic descriptive purposes (Chomsky 1995: 170).

The moral of the story is clear: there can't be any "counting" in syntax without an explicit theory of syntax that independently tells us what needs to be counted, and how. Any comparative approach that gives the "looks" of languages priority over the "essence" of grammar runs the risk of becoming a most simplistic and misleading linguistic measure.

5.4. *Complexity is no simple matter*

Bit-complexity in (30) is quite ambitious: it is meant to rank the entire grammars of the entire set of creole languages against the entire grammars of the entire set of non-creole languages, with perhaps the few exceptions noted in WSG: Section 6. Recall the central claim that creole languages and non-creole languages tend to fall at opposite ends of the bit-complexity cline. This claim is not only about Lahu vs. Saramaccan, or Tsez vs. Saramaccan, or Maori vs. Saramaccan. Nor is this claim to be evaluated with respect to only a handful of linguistic features such as those in (18). Bit-complexity has universal scope: it is a claim about all languages across all areas of grammar (phonology, lexicon, syntax, semantics, pragmatics, etc.).

With this in mind and given the arguments above, it is worth stressing again that bit-complexity as "tested" in WSG enlists only an arbitrary set of linguistic properties (such as those in (18)). These properties are picked from a handful of exotic non-creole languages (e.g., those mentioned in paragraph above) without recourse to any independently-motivated theory of grammar, processing and/or acquisition. Thus, this metric has no principled implications for Language in the mind/brain (but see Note 40). Given our current state of knowledge and the complex nature of Language, we can't yet afford a global complexity metric with global crosslinguistic scope. In the meantime, the handpicking of languages and linguistic features in implementing and testing the metric in WSG belies the purpose stated therein to elaborate "a direct comparison of certain creole grammars with older language grammars, with a view towards making more precise my grounds for the claim that creole grammars constitute a synchronically identifiable class" (WSG: Section 1). If the few "test" languages and the few "test" properties are both prejudicially chosen without regard to any independently-established criteria, then whatever we may learn from this comparison is not enough to equate creoles to the (natural?) class with the label "world's simplest grammars".

Empirically it has already been argued above that a slightly larger sample of crosslinguistic data and typological/diachronic observations undermine neo-Schleicherian creolistics. The rest of this critique brings additional method-

ological and empirical observations that further undermine the bases of the new Schleicherian linguistics.

6.   "Learning by debunking": The empirical (non-)basis of age–complexity correlations

Here I focus on some of the specific empirical problems that undermine the "testing" of age–complexity correlations. In a nutshell, what we are dealing with is a set of "rigged" experiments where the "test" cases and "control" cases seem carefully handpicked to provide support for a neo-Schleicherian creolistics hypothesis. But this empirical support will be shown to be illusory: the empirical claims in neo-Schleicherian creolistics ultimately lead to theoretical incoherence, specially vis-à-vis purported complexity–age correlations.

6.1.   *Age before complexity? Complexity before age? Schleicherian circularity*

Pre- and neo-Darwinian linguistics from, say, Schleicher (1863) to WSG rests on the following premise: Complexity increases with age – as a language gets older, it gets more complex (but see Notes 5 and 24). It is further postulated that, after tens of millennia, old languages "all come to rest at a certain 'surplus complexity quotient' " – an evolutionary plateau of maximum complexity that excludes creole languages (WSG: Section 2.3), even though a few old languages like the Riau dialect of Indonesian and Southeast Asian languages may have slipped from the maximal-complexity plateau and acquired "pidgin-level syntax" due to "extensive adult acquisition" (WSG: Section 4.4). Temporarily putting aside the theoretical elusiveness of the notions "language birth" and "language age" qua linguistic constructs sensu stricto (see Sections 2 and 3 above), I will argue that (neo-)Schleicherian complexity–age correlations are robustly disconfirmed by the available diachronic and typological data.

Consider inflection, for example, which is taken as a marker of complexity-cum-age: "[I]nflectional morphology renders a grammar more complex than another one in most cases" (WSG: 137); "inflection almost always complexifies a grammar" (WSG: 138); "this language [Riau Indonesian] reveals its age in having [*inter alia*] three inflections [. . .]" (WSG: 155). How many inflections does a language need to "reveal its age"? As is noted in, e.g., DeGraff (2001b: 71–76) and Muysken & Law (2001) (see (13)), it is not true that creole languages lack inflectional morphology. In fact, even a(n early) pidgin can show inflectional morphology given the "right" language-contact ecology (as in, say, the case of overt transitive marking in Melanesian Pidgin in the mid mid-1880s described in Keesing 1991: 318–320; also see some of the contribu-

tions in Jahr & Broch (eds.) 1996 and Thomason (ed.) 1997).[42] If three inflections plus "some opaque derivation-root combinations, and optional numeral classifiers" are enough to "reveal [old] age", then creoles would have to be considered "old" languages. Considering their affixal inventories and their non-pidgin syntax, some creoles (e.g., HC, Mauritian Creole, Capeverdean) must be at least as old as the Riau dialect of Indonesian per the very description and criterion in WSG: Section 4.4. This is not expected within neo-Schleicherian proposals for age-morphology correlations.

Be that as it may, does inflectional morphology always increase with age? It has long been observed in the grammaticalization and historical-linguistics literature (see, e.g., Meillet 1912) that, from a diachronic perspective, grammatical systems (e.g., Case and Tense/Mood/Aspect – TMA – marking) often evolve along analysis–synthesis cycles whereby overt markers go through the ebbs and flows of syntax (analysis/periphrasis) and morphology (synthesis/word-level processes). Free-standing auxiliaries can become verbal inflectional affixes, pre-/postpositions can become nominal case affixes, and both verbal and nominal affixes can fuse to their stems and erode over time. Given such morphology–syntax cycles, hope springs eternal for any affixless language with affix envy: "Weep not, my children, for today's syntax is tomorrow's morphology" (Givón 1971: 413: 1; also see Hodge 1970, Giacalone Ramat & Hopper (eds.) 1998, Heath 1998, Haspelmath 2000, and Janda 2001 for some discussion of (de)grammaticalization phenomena and their theoretical bases; also see Note 46).

Schleicher himself was well aware of the erosion of morphology in "old" languages and clever enough to try and incorporate inflectional decay in his evolutionist scenario, distinguishing between "evolution" and "history" while strenuously holding on to his complexity–age correlations (see Note 5). Classic instantiations of the rise-then-decline of morphological marking are common in the history of Romance and Germanic and, more generally, throughout Indo-European and elsewhere. Compare, say, verbal inflection and nominal case in Old English vs. Modern English; ditto regarding, say, the evolution of nominal case and verbal inflection from Latin to Romance.

Here, we get the exact opposite of the Schleicherian complexity–age correlation. Vis-à-vis case and verbal morphology in the relevant stages of Germanic and Romance diachrony, "older" implies "simpler", assuming bit-complexity as in (30). Modern English, for example, has substantially fewer overt distinctions in, and fewer combinations of, verbal inflection than Old English. Ditto in the domain of overt case morphology on non-pronominal noun phrases: Mod-

---

42. Keesing qualifies Melanesian Pidgin inflection as "anomalous in the spectrum of pidgins/creoles" (but see Note 14 and the quote in (13)).

ern English has none of the overt case affixes that were once productive in Old English. WSG: 138 does note that "English expression of case is simpler overall than Latin's", but it is not noted that the Modern English expression of case is also much simpler than that of Old English, and so is the Modern French expression of case simpler than that of its ancestors – in terms of bit-complexity.

The English and French cases are counterexamples to the claim that "diachronic drift [...] encrusts older grammars [with complexities]": diachronic drift has reduced the overt inflectional paradigms of both English and French, thus decreasing their inflectional bit-complexity. Therefore, it is not at all clear that inflectional paradigms can be used as reliable indicators of language "age" (also see the case of Riau Indonesian above). In fact, within single language families we do find languages with drastically different degrees of overt inflectional morphology. Compare, e.g., Malayalam to other Dravidian languages, English to Icelandic, French to Italian, or Western European languages to Balto-Slavic languages as in (8) (also see Notes 5 and 24 and the comparative data in Hodge 1970). It has long been established that much morphological variation is expected within any single linguistic phylum. Assuming that all languages within a particular family "go back" to a single ancestor proto-language (abstracting away from the theoretical difficulties in language "dating"), then such variation within single families goes against the notion that language age can be correlated with bit-complexity. Of course, this is reminiscent of Edward Sapir's memorable quote in (7); also see (8).[43]

It must also be noted that acquisition processes (either in ordinary situations or in situations of abrupt language contact) exert an inevitable pressure toward regularization and/or morphological "simplification" in certain domains, even if this pressure is counteracted elsewhere in the grammar by other factors such as certain types of language transfer, grammaticalization, innovation, and the like. It is not only in the diachrony of Riau Indonesian that we find "a [certain] degree of pidginization [...] due to extensive acquisition by adults, having 'shaved away' a large degree of accreted complexity" (cf. WSG: Section 4.4). Acquisition by adults with its potential for morphological erosion is a widespread phenomenon in language contact and language change (see Notes 5 and 24). Furthermore, the standard fare in historical and

---

43. This observation also undermines the claim that "all natural languages [except creoles] would be equally complex by virtue of having all come to rest at a certain 'surplus complexity quotient' ". It is hard to conceive of an individual mental algorithm (as part of language acquisition, say) that would ensure that bit-complexity remains constant across "all of the grammars [that] trace back tens of millennia" (cf. WSG: Section 2.3). This would entail that speakers of (say) individual Indo-European languages must be able to check each other's bit-complexities in, for instance, inflectional morphology (cf. (8)) and compensate any discrepancies therein by adjusting bit-complexity in some other parts of their respective grammars. Such an algorithm is as improbable as the one critiqued in WSG: Section 2.2.

contact linguistics teaches us, for example, that: (i) sound change often leads to loss/assimilation of phonemic distinctions, (ii) morphological change often leads to regularization and analogical leveling (i.e., to loss of morphological distinctions), (iii) syntactic reanalysis often leads to structural simplification (or structural transparency), etc.; see Campbell (1999) for an overview.[44]

In fact, functionalists have often argued that ease of articulation, regularization, and rule extension are at the roots of language change, which is driven by functional factors – principles of least effort, economy, optimization; see Labov (1994: 547–568, 2001: 16–28) for a critical overview in the context of "maladaptive" language change. Some of these principles of economy entail a reduction of bit-complexity: as Labov notices, there is a long series of arguments that language change – sound change, in particular – may be "dysfunctional", and in the 19th century some of these arguments were advanced within language-as-evolving-organism approaches (but see Note 5 for a sample of concurring and diverging opinions). Labov (1994: 586–599, 2001: 10–14) surveys a number of areas where sound change does reduce overt distinctions, thus bit-complexity. Labov (2001: 10) remarks: "The almost universal view of linguists is [...] that the major agent of linguistic change – sound change – is actually maladaptive, in that it leads to the loss of the information that the original forms were designed to carry. Though there is a wide range of divergent opinions on the nature of sound change [...] there is general agreement on the negative character of this fundamental process".

In effect, this means that the "older" a language, the more opportunity these complexity-reducing "negative" processes would have had to reduce its bit-complexity, thus producing "loss of information" (in Labov's terminology). Here again, we get the equation "older = simpler" at least in certain domains, with "old age" entailing less, not more, overt distinctions (as with English and French case and verbal morphology). It is also expected that processes like grammaticalization (e.g., of free morphemes into bound morphemes) may ultimately offset some of the results of regularization, leveling, morphological erosion, etc. (cf. Givón's aforementioned quip "yesterday's syntax is today's morphology", which goes back to, e.g., Bopp's, Humboldt's, and Meillet's

---

44. Regarding syntax, the following is presented in WSG as an example of increased bit-complexity: an Irish pattern taken from Henry & Tangney (1999) which exemplifies "two kinds of alignment rather than one (i.e. ergative/absolutive and nominative/accusative)". However Henry & Tangney's discussion of these Case patterns suggests that in syntax too bit-complexity can decrease simply as a result of language acquisition sans creolization, as in the L2 acquisition of Irish by English-speaking children in immersion programs with non-native models. These children replace the "two kinds of alignment" in copula constructions with a uniform nominative/accusative pattern. Furthermore, Henry & Tangney (1999: 251) insist that the simplification they document "is not unique to acquisition in these circumstances, merely an acceleration of a normal process".

insights – Humboldt (1836 [1988: 205]; see Note 5) talks about "the wearing-down of inflection [as] undeniable fact"; also see Note 24).

6.2.   *Cycles of complexity in diachrony*

What these cycles suggest is a picture in which absence of (overt) distinctions in one domain of grammar (e.g., morphology with its affixes and other word-internal processes) can well be compensated by distinctions in another domain (e.g., syntax with its periphrases, word-order distinctions, selectional requirements, etc.).

For example, loss of overt inflectional morphology has often been correlated with increased rigidity in word-order. Vis-à-vis case morphology and word-order, it has long been noted, starting perhaps with Meillet 1912 [1926: 147–148], that languages seem to balance out overt case marking with grammatical marking in the syntax (via, e.g., pre-/postposition and word order). The exact correlation is hard to pin down, but the general, if overly simplistic, impression is that richness of Case morphology tends to be proportional with the scrambling of noun phrases away from their base positions (to wit: the history of English and of Romance). If so, measuring only overt morphological distinctions at the expense of abstract syntactic information leads to an incomplete and misleading metric. Similar caveats apply to the interface between any two (abstract) linguistic modules and the trade-offs therein in terms of grammatical information encoding. It seems to me that computing global complexity requires a theory of how grammatical information is encoded within and across the various linguistic modules (cf. Note 39).[45]

---

45. In this respect, WSG: 144 misinterprets my views on the relationship between inflectional morphology and the syntactic differences between a creole and its sources: "DeGraff [...] (1997, 1999b) argues that the differences between a creole grammar and that of its source languages are due to certain syntactic results following from loss of inflection during second language acquisition (such as lack of verb movement to I), with subsidiary results due to the filtering out of low-frequency features, and the ellipsis of certain functional categories, with the qualification that the effect of the latter two was no more marked than that upon other languages with heavy contact in their histories (DeGraff 2000)."

Although I myself have not (yet) explored the fascinating and probably enlightening history of Yiddish, what I did argue is that, like in other cases of language creation, CERTAIN (morpho)syntactic properties seem correlated to properties of (overt) inflectional morphology. This is a rather commonplace, if difficult to formalize, guiding intuition in much theoretical and historical work (see, e.g., van Kemenade & Vincent (eds.) 1997 and van Kemenade (ed.) 1999 for two recent anthologies on this topic). For example, the abstract linking of (degrees of) verbal inflection and (degrees of) verb raising, although somewhat controversial, is standard fare in current research – see the references in DeGraff (1997, 1999b: 501–502, 518–521, 2001a, forthcoming), Roberts (1999). Chomsky (1995: 6) attributes to Jespersen the hunch that crosslinguistic syntactic variations can in large part be reduced to variations in morphology. In this particular respect, certain (erstwhile) contact languages such as Capeverdean Creole, Chinook Jargon, Haitian Creole, and Saramaccan (where certain kinds of overt inflectional

Hawkins (1986) is one such attempt toward formalizing the trade-off between different sources of "complexity". Hawkins argues that, along the dimension of form–meaning opacity, English is more "complex" than German because of increased case syncretism in English. Independently of the merits of this proposal, the point here is that it can be logically argued that reduction of overt morphological distinctions (i.e., reduced bit-complexity in morphology) increases some other kind of complexity, namely form–meaning opacity which in turn increases semantic ambiguity. Hawkins places English and German morphology at opposite ends of the semantic-opacity continuum: German morphology allows for "a 'tighter fit'" between surface form and semantic representation" (1986: 122). Then again, English's increased "form–meaning opacity" due to case syncretism is somewhat compensated by its word-order which is more rigid than in German. Hawkins (1986: 216) speculates that:

(33)   [T]here is an inherent tension [...] between the rules generating linguistic forms on the one hand, and those assigning meanings to these forms on the other. Simplicity for the one means complexity for the

---

morphology seem more "economical" than in some of the creole's source languages) offer valuable databases – natural "test-tubes", if you will – to evaluate and refine current hypotheses vis-à-vis UG's constraints on language creation and language change – in particular, structural constraints on the interaction between inflectional morphology and word-order parameters; see, e.g., DeGraff (1992a, 1997, 1999b, d, 2000, 2001b, forthcoming), Veenstra (1996), Baptista (1997), Vrzić (1997).

This said, nowhere have I (or any other creolist, as far as I can tell) claimed that ALL "the differences between a creole and its source languages" are parasitic on (overt) inflectional morphology. In fact, the very work of mine cited in WSG insists on the following caveats: "In [morphology-driven approaches to syntax], crosslinguistic syntactic differences [...] are due to (AMONG OTHER THINGS) distinct inventories of inflectional paradigms" (DeGraff 1999b: 501; emphasis added), and "[O]ne should not expect creoles to reflect only unmarked parameters [...] There are certainly aspects of creole grammars that were influenced by structures in the source languages [...] Furthermore, even granting that creoles' tendency toward unmarkedness is rooted in their isolating [inflectional] morphology [...] , NOT ALL SYNTACTIC PROPERTIES NEED TO BE DIRECTLY TIED TO THE MORPHEMES THAT TEND TO BE LOST IN PIDGINIZATION. It is not clear, for example, whether the properties of (long-distance) wh-movement are parasitic on inflectional morphology" (DeGraff 1999b: 519; emphasis added). These quotes make it clear that, in my view, (loss of) inflectional morphology is not, and could not be, the exclusive factor that determines the overall syntactic shape of creole languages. DeGraff (1999b) goes on to discuss other factors in creolization (e.g., ecological and stochastic factors, substrate and superstrate influences, markedness, processing, learnability). Given the state-of-the-art in linguistic theory, it is not at all clear to me how OVERT morphology could ever be argued to be the SOLE determinant in syntactic change: indeed many aspects of syntax seem orthogonal to properties of overt inflection. Lastly, DeGraff (1999b) explicitly argues against "creole" as a structural type (see, e.g., the quote in (10)). The arguments in WSG against the view that "the differences between a creole grammar and that of its source languages are due to [...] loss of inflection [etc.]" address a (theoretically puzzling) straw-man of its own making.

other, and vice-versa, and the successful resolution of this tension de-fines a continuum along which the actually attested languages occur, from the least to the most complex set of formal rules, and corre-spondingly from the most to the least complex set of mappings be-tween form and meaning. Over time languages can drift from one part of this continuum to another, in response to independent (e.g., phonological) changes which have clear consequences for the map-ping between form and meaning.

No matter the fate of Hawkins' arguments, his speculations shed further doubts on the theoretical validity of the simplistic sort of (bit-)complexity met-ric advocated in (30): it can indeed be argued on empirical and theoretical grounds that, at least in some cases, ABSENCE of "overt distinctions" (com-pare, e.g., case affixes in English vs. German) can itself be a distinction that should enter into some global complexity metric. This is reminiscent of Meil-let's (1912) argument on how rigid (grammaticalized) word-order emerges to replace case morphology. Here again, we see why bit-complexity is of lit-tle use in the absence of a theory of grammar that motivates the items to be counted and relates them to larger linguistic concerns – with respect to mental representations, language acquisition, language processing, language change, etc.[46]

6.3. *Historical linguists' "endless cycles" vs. pro-prototype creolists' "ground zero"*

Keeping these opposing trends in the picture, it must then be concluded that, in creole genesis as in other cases of language contact, distinct grammatical do-mains in the languages in contact will belong to distinct points on, e.g., their re-spective morphology–syntax cycles (à la Meillet/Hodge/Givón) and semantic-opacity cycles (à la Hawkins). With this in mind, one is forced to conclude that creolization etc. (on a par with language change via language contact) will start, not "essentially from ground zero" (contra WSG), but from a contingent (i.e., sociohistorically determined) array of non-"ground zero" structural *termini ad quo*. The latter fit distinct points on diachronic cycles (e.g., with respect to

---

46. If "morphology" and "syntax" all belong to syntax – that is, if the structure-building op-eration "Merge" applies at both the word level and the phrase level merge, as in, e.g., Marantz, Miyashita, and O'Neil (2000: 3–4), – then it may not make much of a difference in terms of computational complexity whether Merge takes place within words or across words; see Haspelmath (2000) and DeGraff (2001b: 72–73: Note 19) for further comments. In other words, periphrasis may be cognitively as "cheap" or "expensive" as lexical storage or derivation-root combination (see Chafe 1970: 36–37). But these are all empirical questions that deserve further study.

their respective analysis–synthesis and semantic-opacity clines). These non-"ground zero" structural *termini ad quo* determine the ecology of the Primary Linguistic Data available in the formation of creole idiolects. In other words, various components of the (pre-)creole grammar will be extrapolated (via language transfer cum restructuration, innovation, grammaticalization, simplification, etc.) from patterns that are located across the analysis–synthesis and morphological-syncretism continua, as a synchronic reflection of prior cycles in the respective diachrony of the "old" languages in contact. In a similar manner, the diachronic *termini ad quem* will unavoidably be scattered across the relevant cycles in various grammatical domains. Ultimately, the trace of this scattering (i.e., the resulting I-language grammar(s) whose structures are necessarily bounded by UG) will depend both on the Primary Linguistic Data and the sociodemographic conditions – all of which result from a very complex set of historical contingent factors.

So, here too, the creolist-cum-complexity-theorist must employ much caution in dealing with the competing pressures on language complexity that arise from the sociohistorical cum typological specifics of the contact situation. As Hymes (1971b: 70) reminds us, we need "to recognize pidginization as a complex process, comprising the concurrence of several component process". At this stage, we are far away from the Schleicherian "older = more complex" perspective on language evolution. We are also quite far away from the *ab ovo* genesis of "born again" languages from "ground zero" complexity via "a radical reduction of [the] source languages into makeshift jargon" (cf. WSG: 149).

The empirically and theoretically responsible scenario is much more complex, thus much more fascinating, even if it removes creole languages from the category of contemporary fossils of Language evolution, and even if it deprives creolists of a most simplistic account for some of the most complex cases of diachronic development.

6.4. *The empirical test: To be "old"* AND *"born again"?*

How many non-creole languages could pass the WSG-type (neo-)Schleicherian structural litmus test for old-cum-complex languages?

WSG "carefully" picks its benchmarks for "old language" complexity, not from, e.g., Romance and Germanic (the well-documented source languages – the *terminus a quo* – of its Prototypical Creoles), but from Caucasian, Tibeto-Burman, and Polynesian languages (which have nothing to do with the Prototypical Creoles being surveyed). The complexity "control" cases – Tsez, Lahu, and Maori – would strike many linguists, not just creolists and creole speakers, as "fearsomely elaborated" or "extraordinarily complex by any linguistic standard" (cf. WSG: Section 3.2; also see Trudgill 1989: 237, who

can imagine "mainstream Europeans and North Americans find[ing Caucasian languages] 'exorbitant' or 'incredible' "). Plus these three "control" languages are all "esoteric" languages, languages spoken in small communities in relatively low-contact situations where, for social-network reasons, speakers are most likely to maintain (and perhaps even promote) features that may APPEAR complex and/or linguistically unusual (cf. Hymes 1971b: 73, Trudgill 1989: 236). Trudgill specifically cites the languages of the Caucasus, to which Tsez belongs, as prime examples of "low contact" languages. According to Grimes (ed.) (1996), Tsez has approximately 7,000 speakers, Lahu (Shi) 600,000, and Maori 110,000 (these are approximate totals across the relevant dialects). This sociolinguistic factor alone is a delibitating confound in these complexity experiments comparing Saramaccan with Tsez, Lahu, and Maori. What should be asked is whether a hypothetical creole derived from contact among, say, Caucasian languages of the Tsez-type (or Tibeto-Burman languages of the Lahu-type or Polynesian languages of the Maori-type) would end up looking like Saramaccan. (See Notes 30 and 37 above for related methodological remarks.)

As it turns out, the neo-Schleicherian "deck-stacking" methodology, to the extent that I understand how it can be applied without bias, unsurprisingly ranks the complexity of English alongside that of Saramaccan. Yet Saramaccan, but not English, is usually regarded as a most "radical" creole. In fact, WSG considers English to be an "old" language – much "older" than "born again" Saramaccan.

This is how we can try and check whether English is "born again" or "old" by applying the criteria and methodology exemplified in WSG:

English, like a language with "pidgin ancestry", lacks many of the traits that are claimed to be "incidental to basic communication". Indeed English lacks ergativity, grammaticalized evidential marking, inalienable possessive marking, switch-reference marking, inverse marking, obviative marking, noun class or grammatical gender marking, lexically contrastive or morphosyntactic tone, etc. (cf. (18)).

English, like (the radical creole) Saramaccan, is also much less "complex" than Tsez. Indeed, given the criteria in WSG, "Tsez's grammar is indisputably a more complex one" than English's. In fact, the complexity of Tsez, as measured in WSG, seems much higher than that of French, Spanish, Chinese, and many other "old" languages. Thus, shouldn't we also ask whether English, French, Chinese, etc., like Saramaccan, "have not existed for long enough a time for there to have arisen the sheer weight and depth of such features as in older languages like Tsez" (cf. WSG: Section 3.2)? Indeed, English and many other non-creole languages do resemble Saramaccan in lacking the following Tsez properties (cf. WSG: Sections 3.1–3.2): "pharyngealized uvulars"; "stops and affricates [with] phonemic ejective alternants"; "[noun] classes determined by the final segment of the stem"; "nouns [with] alternate forms for differ-

ent suffixes"; "markers for evaluative names"; in-situ adjunct wh-phrases vs. fronting of argument wh-phrases; "grammaticalized evidential markers"; "a lative marker"; "overt delineation of experiencer verbs"; etc.

English and many other non-creole languages, like (the radical creole) Saramaccan, is also much less "complex" than Maori, given the observations in WSG: Sections 3.2 and 4.3. Indeed, English and many other non-creole languages resemble Saramaccan in lacking the following Maori properties (taken from WSG: Section 4.3): "several different interrogative constructions according to the grammatical status of the constituent questioned" (furthermore one can also say of many Germanic and Romance languages, on a par with Saramaccan, that they do not exhibit "interrogation strategies [that] vary to anything approaching this extent according to grammatical relation"); "subtle possessive distinction [...] reminiscent of an alienable/inalienable distinction, but contingent basically upon dominance of possessor over possessee"; "subjects of intransitive verbs [that] are marked as possessives, while the verb itself is nominalized".

English, like the radical creole Saramaccan, is also much less "complex" than Lahu, given the observations in WSG: Sections 3.2 and 4.1–4.4. Indeed, English (along with many other non-creole languages) resembles Saramaccan in lacking the following Lahu properties (taken from WSG: Section 4.1): "bilabial, alveolar and velar, with both aspirated and unaspirated phonemic alternants in all five places of articulation"; "seven lexically contrastive tones"; "agentive markers distinguished by sex"; "[agentive markers] distinguished by sex to denote ownership or mastership, which can also be used to nominalize clauses"; "numeral classifier, for people, animals, shapes, and more general purposes"; nominal reduplication "as a classifier"; "an accusative marker [...] used with patients only to encode certain shades of emphasis"; "modal and pragmatic particles central to basic expression which are conventionalized into highly particular and idiosyncratic subdivisions of semantic and pragmatic space"; "verbs [that] occur only in concatenation with a specific other verb to convey completion"; "SOV"; "elaborate tonal system".

In a coda to the Lahu-vs.-Saramaccan discussion, it is noted that "English is also less complex than Lahu in all but one of the features cited (derivation)" (WSG: 149). Yet, in the same paragraph, an ad hoc list of English properties is pulled out in order to argue that "English is more complex, according to our metric, than Lahu in a great many aspects". This is yet another vacuous argument since similar ad hoc lists of non-Lahu "incidental" properties could be produced for any language. Indeed it remains possible in principle to produce a similar ad hoc list of Saramaccan features whereby Saramaccan too would look "more complex, according to [the WSG] metric, than Lahu in a great many aspects". In fact, given the structure of UG (see Section 4.5 above) and given the contingent aspect of parameter-setting and of the lexicon, it is always possible

(in principle) for any pair of languages *X* and *Y* to produce an ad hoc list of grammatical distinctions that exist in *X* but not in *Y*. Such lists exist for any pair of languages no matter whether *X* and/or *Y* are creole or non-creole. The handpicking of scattered "incidental" features does not constitute explanation or result. The contents of such lists are truly "incidental", that is, language- and construction-particular in a relatively superficial way (i.e., with no explanatory power).

As a (facile) exercise, I will now overtly rig a neo-Schleicherian experiment and produce a list of "incidental" Haitian Creole features of the Lahu/Maori/Tsez type above whereby HC "is more complex, according to [the WSG] metric, than [certain varieties of French and English] in a great many aspects":[47]

(34)   a.   Nasalized and non-nasalized vowels that, in certain contexts, occur in free variation, even when followed by a nasal consonant; nasal vowels can be analyzed as "a combination [of] oral vowel [plus] floating nasal consonant" (Cadely forthcoming).

    b.   "[A] regressive nasalization rule that [in certain contexts] applies within stems or underived morphemes and a progressive nasalization rule which [in certain contexts] takes place across a morpheme boundary" (Cadely forthcoming).

    c.   A set of personal pronouns with morphophonologically and syntactically conditioned clitic variants; the latter can behave as either proclitics or enclitics, according to complex morphophonological and syntactic conditions (DeGraff 1992a; Cadely 1994, 1995, 1997; Hilton 2000).

    d.   Morphophonologically conditioned allomorphy for the postnominal determiner with (at least) four allomorphs. The distribution of these allomorphs implicates the complex interaction of nasalization, vowel hiatus and glide-insertion phenomena and their respective underlying constraints (Y. Dejean 1980, Joseph 1988, Cadely 1995, Fattier 1998, Klein 2001). Klein (2001) argues that the morphophonology of the HC determiner leads to marked structures – "anti-markedness effects" – and that "[determiner] allomorphy defies views of creoles as 'simplified languages'".

    e.   Morphophonologically conditioned allomorphy for the prenominal determiner (at least four allomorphs) (Y. Dejean 1980, Joseph 1988, Fattier 1998).

---

47. Cadely's (1988) analysis of the HC syllable argues that HC has a much more extensive set of diphthongs than French: HC with its alleged "simplest" grammar has about twice as many diphthongs as its "old" French ancestor.

f. Demonstrative, definite, and plural-marking markers that are head-final in a language that's otherwise robustly head-initial (Joseph 1988, DeGraff 1992a).

g. Head-final definite articles alongside head-initial indefinite articles (Joseph 1988, DeGraff 1992a).

h. Two predicate-clefting strategies for focus purposes: (i) one with predicate-head doubling (with verbs, adjectives, and certain bare nominals); (ii) the other with a non-verbal pro-predicate left in the in-situ position of the moved predicate XP (the latter must be [−V]). (DeGraff 1992a, 1995, 1998; also see Appendix B for further discussion with a bit of data.)

i. Certain bare nominals can undergo predicate clefting with either the predicate-doubling strategy ((i) above) or with the (non-doubling) pro-predicate strategy ((ii) above). The predicate-doubling strategy allows both a permanent/essential (individual-level) interpretation and a temporary/provisional (stage-level) interpretation of the clefted predicate while the pro-predicate strategy forces the clefted predicate to be interpreted as individual-level (Damoiseau & Saint-Louis 1986).

j. Predicate clefting-cum-doubling for the formation of various adjunct clauses (Lefebvre & Ritter 1989).

k. Three strategies for the formation of causal clauses, each of which uses some distinct CP-related position – one of these strategies also employs the predicate clefting-cum-doubling pattern (Lefebvre & Ritter 1991).

l. A resumptive pronoun that surfaces in the small-clause subject position of DP(-like) predicates in affirmative clauses with no tense/mood/aspect marker (DeGraff 1992a, b, 1993, 1995, 1998).

m. A rule of apocope that applies to a subset of verbs with short and long variants; this rule is morphophonologically and syntactically conditioned (DeGraff 2001b: 74–75).

The above ad-hoc "incidental" list of HC non-English properties would thus lead to the conclusion that HC is more complex than English, if one assumes the style of argumentation in WSG. In addition, HC also manifests some (quasi) English-like features that constitute "overspecification [that] goes beyond the needs of a human grammar" (cf. WSG: Section 4.2). Such (quasi) English-like "ornamental"/"incidental" properties include:

(35) a. "[T]he overt and categorical marking of definiteness" on HC nouns (Joseph 1988).

b. An "overt marker of definiteness and indefiniteness, whose occurrence is determined by referentiality as well". In HC, we

say: *Mwen wè yon fim yè swa* 'I saw a film last night' where the marker *yon* marks the NP *yon fim* on a par with English *a* in *a film*) "despite [HC *fim*] being presupposed to the speaker, because it is not yet known to the hearer". (Here, HC, like English, "goes beyond the needs of a human grammar as far as encoding definiteness is concerned", which makes HC look older than "the thousands of grammars without such overt marking [...] (e.g., Chinese and Russian)"; cf. WSG: 161.)

c.   A "subtle distinction [that is] maintained" in HC between the *ap* future and the *pral* future. Note that, like English *will* and *going to*, HC *ap* and *pral* are other "feature[s] which give fine-grained and grammaticalized manifestation to a distinction lacking in Lahu as well as a great many other languages". So in this particular respect, HC – this "born again" language – is at the very least as "complex" than two "old" languages: English and the "mega-complex" Lahu.

d.   An intricate of set of (semi-)auxiliaries that can be used as (quasi) tense/mood/aspect markers (see e.g. HC *pral* above and its English analogue *going to*); also see HC *konn*(*en*) (habituality marker), *fin*(*i*) (completive marker), *dwe* (deontic/epistemic marker), etc.

One could argue that all the complex features of HC above – which, as in English, seem "incidental to basic communication" – have arisen "due to contact over the centuries with French" (cf. WSG: 143). But this would not work for a simple reason. Many of the features noted above have no direct counterpart in French. Furthermore, the French-like HC features (e.g., the auxiliaries *ap* and *pral*) seem to have been part of the language from very early on, resulting from the grammaticalization of French periphrastic constructions with *après* and *après aller*, respectively. Furthermore cognates of these preverbal tense/mood/aspect markers are regularly found in French varieties that are not labelled "creole" (cf. Appendix B).[48]

Such arbitrary and superficial comparison whereby "born again" languages may APPEAR more complex than "old" languages given some ad hoc list of language-specific features can be extended ad libitum. Indeed, for any choice

---

48.   For further discussion of preverbal tense/mood/aspect markers in the diachrony and synchrony of French and French-lexicon creoles, see (the data and references in) Van Name (1870: 143–146), Sylvain (1936: 138–139), Goodman (1964: 78–90), Gougenheim (1971), Chaudenson (1979: 80–82, 1992: 162–167), Baker & Corne (1982: 31–48), Chaudenson, Mougeon, & Beniak (1993), Chaudenson & Mufwene (2001: 178–192), Michaelis (1993), DeGraff (1996b, 1997, 2000, 2001a, forthcoming). Also see Section 2.1 above and especially Appendix A below for further refutation of late-borrowing scenarios in the diachrony of HC morphology.

of n (an integer), X (a so-called "born again" language) and Y (a so-called "old" language), X can be the source of an arbitrary list of n features that are absent in Y. Take, say, dual marking, whose presence in any language is considered (in WSG: Section 5.3) to increase complexity/markedness in the Greenbergian sense. Dual marking, which is absent in many "old" languages, does exist in a number of "born again" languages. In Cape York Creole, as described in Crowley & Rigsby (1979), the pronominal system offers both dual and plural in the non-singular forms (for 2nd and 3rd person), plus there is a grammaticalized distinction between inclusive vs. exclusive pronouns in the 1st person. If dual is dispensable from a communicative viewpoint (which it must be given its absence in many "old" languages), then Cape York Creole is surely more complex than, say, English in that respect. A similar example is provided by Chaudenson (1994: 50) who argues that dual marking in French-lexifier creoles in the Indian Ocean (e.g., in Réunion) makes these creoles more complex then their lexifier vis-à-vis the grammaticalization of number marking. Dual marking is also found in Taimyr Peninsula Russian-based Pidgin (Wurm 1996: 83). Such "incidental"/"ornamental" systems of pronominal reference and number marking are not found in "old" languages such as many English and French varieties. Here too, we have "born again" languages with arbitrary distinctions that are not found in "old" languages (also see Section 4 for further examples of pidgin/creole features that are "incidental to basic communications").

It is worth repeating that, given the partially contingent nature of parameter-setting and lexicon formation, any number n of arbitrary (and so called "incidental"/"ornamental") distinctions can, in principle, be found in ANY language, whether "old" or "born again". Indeed, the list of "incidental" features provided by creolists-cum-complexity-theorists in their creole-vs.-non-creole comparisons (see, e.g., (18)) is truly incidental.

Lastly, one could argue that the reason why English's bit-complexity looks low when compared to Tsez, Maori, and Lahu is that English is, after all, a "creole" as can be perhaps adduced from its history. Such argument would go on to claim that English's history of contact is the reason why so many creoles (such as HC and Cape York Creole, say) have features that are at least as "complex" as their English counterparts. Well, I suspect that many other languages besides English would also rate low given the complexity metric in WSG and its exotic benchmarks and "incidental features" list. In any case, the English-as-creole argument would drastically weaken the theoretical bite that WSG would like to assign to the term "creole", especially in the absence of an unambiguous STRUCTURAL definition for "creole".[49] In any case, English

---

49. Calling Middle English (ME) a "creole" – as, e.g., in Bailey & Maroldt (1977) – raises more questions (contradictions, really) than it resolves. Indeed, English (even as another "cre-

IS claimed as one of these multi-millenarian languages that are generally more complex than young languages (see, e.g., WSG: Section 4.2).

In fact, it can be argued that language contact in the history of English did contribute to reducing bit-complexity in restricted domains, as in, e.g., inflectional morphology. Meillet, Weinreich, Trudgill, and Chambers, among many others (see Notes 5 and 24) have argued that language learning in contact situations induces "erosion" of morphology and/or regularization of overt morphological distinctions. And it has been argued many times before that language acquisition itself can induce various degrees of "simplification" in the concomitant new idiolects.[50] Thus, any language (whether "old" or "born again") can undergo certain sorts of "simplification", independently of its creole status or origins.

What the above remarks suggest is that the complexity-reducing effects of "pidginization" in certain domains (e.g., in overt inflectional morphology) are quite widespread – even when there are no recognizable pidgins(-to-be) in sight (recall that (I-)pidginization as an individual-level process is in principle distinct from the creation of stable (E-)pidgins – the latter crystallize through the focusing of norms and other group-level sociolinguistic processes). This, of course, is not a novel observation. Schleicher himself considered that language contact (e.g., in the history of English) was a degenerative factor while Humboldt took the "undeniable fact" of inflectional erosion as a sign of intellectual maturity (see Note 5)! It has been commented over and over again in the sociolinguistics and historical linguistics literature that language contact, across space and time, often entails structural simplification in various domains (see Note 24). As Hymes (1971b: 73) wrote, "simplification may prove to be, not an isolated phenomenon, but one pole of a continuum applicable to [...] all languages" (see Note 50). Others have proposed that one key factor driving structural simplification in various domains is vehicularization, i.e., the sociohistorical process by which certain languages become lingua francas, as in the case of English; see Mufwene (2000b, 2001). At the individual level, the sim-

---

ole") still resembles a "genetic" Germanic language according to standard *Stammbaumtheorie* methodology (see, e.g., Campbell 1999 for an introduction). Furthermore, the sociohistorical specifics of English diachrony do NOT resemble those of ("radical") creole genesis as in the emergence of HC. For example, it is widely accepted that (varieties of) English remained the native idiolects for the majority of speakers in England throughout the Norman Conquest. Thus, I somewhat agree with Domingue (1977: 90) that: "[I]f ME [...] can be described as another creole, there must be many more creoles around than we think. On the contrary, if ME does not fit in the creole category, it will be because stringent characteristics of that category will have been established, a task yet to be done". Vis-à-vis "stringent [structural] characteristics of [the creole] category", the "task yet to be done" remains elusive still and seems to me impossible – thus, this paper.

50. See, e.g., Henry & Tangney (1999), Newport (1999), Roberts (1999), and references therein.

plification processes often noted in language contact and vehicularization must be ultimately rooted in the cognitive capacities and constraints that underlie second language acquisition (see, e.g., DeGraff 1996b, 1999b, c, d for discussion and references).

7. Envoi: The descent of the creole speaker

The various lists elaborated by (neo-)Schleicherian linguists in the past three centuries in order to isolate "the world's simplest grammars" still seem irreducible to any fundamental principle(s) of linguistic theory. Thus far, these lists cannot count as scientific explanation for any robust set of linguistic phenomena. At best, they identify scattered "historical accidents" in scattered domains of grammar in scattered samples of languages.

In any UG-based framework (along the lines sketched in Section 4.5), language-specific "incidental overspecifications" (such as pro-predicate morphemes, dummy verbs, dual marking, ergativity, grammaticalized evidential marking, pharyngealized uvulars, clicks, etc.) ARE "historical accidents" – the sociohistorically contingent choices made by particular idiolects cum societal conventions within the boundaries set by the biological necessities of our *faculté de langage*. Given the rich linguistic ecology and the (socio)linguistics of language contact, it is no surprise that "historical accidents" of various sorts also happen in creole genesis, as documented above.

One key question facing modern linguists is: What is the structure of UG such that, across the species, language learners faced with incidental and relatively shallow Primary Linguistic Data (PLD) unfailingly (re-)create idiolects with all sorts of abstract complex properties that are not evident from the PLD (cf. Chomsky's 1986 "Plato's Problem").

The central assumption here is that every idiolect is somewhat created anew at every instance of acquisition. A related, but distinct, assumption is also found in the grammaticalization camp. There too, notions like "old" and "new" languages seem to make little theoretical sense: "Students of grammatization realize that worrying about where one grammar ends and the next grammar begins is a totally meaningless and futile pursuit. For the 'new gramar is constantly being created on top of the wiling and yielding ruins of the old' [...]" (Matisoff 1991: 447)

From the Cartesian-Uniformitarian perspective espoused here, it can be reasonably argued that ALL (I-)languages evolve via an initial "break in transmission": grammars are not inherited, but (re-)created (Paul 1890; Meillet 1929; Halle 1962; Chomsky 1981, 1986, 1995; Lightfoot 1999; etc.). In each of its individual instantiations, language acquisition sensu stricto is not language transmission, but UG-guided language (re-)creation with contingent, sparse, and heterogeneous PLD drawn from idiolects (i.e., speakers) in contact (Chom-

sky 1981). The exact and particular nature of this "contact" is sociohistorically determined and so are the tempo, amplitude, and group-level effects of the individual "breaks". Yet, whether in "creole genesis" or through "language change", the PLD always underdetermine the attained grammar: there always exist structural "breaks" (however subtle) between "old" and "new" idiolects. Thus, the ineluctability of language change/creation. Whatever general tendencies may exist across instances of creole genesis partake of the same mental processes that underlie all other cases of (I-)language creation.

As of "(I-)pidginization", to the extent that it reduces to (aspects of) second-language learning (L2A) in adulthood, the "pidgin-to-creole life-cycle" will have congeners far and wide, in all cases of language contact where non-native utterances contribute to the PLD that are used in the formation of native idiolects (L1A). In a Cartesian-Uniformitarian perspective, the "pidgin-to-creole cycle" is naturally mirrored by an L2A–L1A cycle – or L2A–L1A "cascade relationship" in the terminology of DeGraff (1999b: 497, 504). In this vein, the discrepancies between "old" and "new" idiolects may seem more dramatic in the creole cases than in the non-creole cases, but it can reasonably be argued that the difference (if any) is a matter of degree, rate of spread, and/or subjective perception, not of quality.

If so, the sui generis structural category "world's simplest grammars" with "born again" genealogical status is only "chimera and reverie" and, worse yet, "linguistic monstrosity", as hinted at by Foucault in (4). Thus, the fallacy of dualist Neo-Darwinian scenarios for the origin of creoles. Given UG and given the sociolinguistics and ubiquity of language contact, there is not, and there could not be, a constant and exclusive set of creole structures that are FUNDAMENTALLY special, across time and across space, independently of the specific linguistic ecology (see references in Note 12). I thus agree with, *inter alios*, Mufwene and Muysken, as per the following quotations (compare with (10)):

(36)   a.   The very notion of a "creole" language from the linguistic point of view tends to disappear if one looks closely; what we have is just a language. (Muysken 1988: 300)

        b.   [N]o language-development processes were involved [in creolization] that were unique to [creole languages,] just the same ones usually assumed in historical linguistics except for the emphasis on language contact. (Mufwene 1996: 107)
What we have everywhere seems to be simple evolution of languages from one state to another in different ecological conditions. (Mufwene 1998: 324)

In my native Haiti and elsewhere in the Caribbean and wherever else we find creole speakers, arbitrary (pseudo-)linguistic measures are still employed

to "classify humanity" and de-humanize (monolingual) creole speakers (see references in Note 9). This "linguistic apartheid" is undermined (theoretically, at least) by observations like those in (36); also see the quote in (10) and the references in Note 12. As a creolophone creolist, I find that there is grandeur in this (Cartesian-Uniformitarian) view of Language: Creoles, on a par with ALL other languages (irrespective of genealogy), are reflections of our (species-uniform and species-specific) human biology, which is among the "most beautiful and most wonderful [forms that] have been, and are being, evolved" (cf. Darwin 1859 [1979: 459–460], Chomsky 1995: 1–4, 2001: 2).[51]

Appendix A: Were Haitian affixes "borrowed late"?

As we saw throughout the main text (see, e.g., Sections 2.1, 3.3, 6.1, 6.4), there is robust, inescapable evidence that HC – the creole formerly known as "most creole of creoles", " 'pure' [Creole Prototype] case", and "basilectal creole" (McWhorter 1998: 809, 812; 2000b: 206) – is far removed from a structurally "simplest" Creole Prototype with ancestry in a structurally "simplest" affixless pidgin. For instance, HC affixes (with cognates in French affixes) straightforwardly disconfirm the catastrophic pidgin-to-creole scenario whereby HC affixes would have emerged via grammaticalization of erstwhile free morphemes. How would a neo-Schleicherian creolist reconcile the postulation of a pidgin-to-creole catastrophic cycle (and its radical morphological bottleneck) with the well-documented French-based affixes of HC qua "most creole of creoles", "basilectal creole", etc.?

---

51. Reading this last sentence, some linguists may ask, with due caution, whether its contents jeopardizes scientific objectivity. Since this paper is ultimately about the mismeasure of creole speakers, my response is appropriately taken from Gould's *The Mismeasure of Man* (1996: 36–37):

> Scholars are often wary of citing [...] commitments [to social justice], for, in the stereotype, an ice-cold impartiality acts as the *sine qua non* of proper and dispassionate objectivity. I regard this argument as one of the most fallacious, even harmful, claims commonly made in my profession. Impartiality (even if desirable) is unattainable by human beings with inevitable backgrounds, needs, beliefs, and desires. It is dangerous for a scholar even to imagine that he might attain complete neutrality, for then one stops being vigilant about personal preferences and their influences – and then one truly falls victim to the dictates of prejudice. Objectivity must be operationally defined as fair treatment of data, not absence of preference. Moreover, one needs to understand and acknowledge inevitable preferences in order to know their influence – so that fair treatment of data and arguments can be attained! No conceit could be worse than a belief in one's own intrinsic objectivity, no prescription more suited to the exposure of fools. [...] The best form of objectivity lies in explicitly identifying preferences so that their influence can be recognized and countermanded.

> Gould then proceeds to debunk a number of rankings of human cognition across the "races", in, e.g., the practice of 19th-century craniometry and 20th-century psychometrics.

McWhorter (2000a: 107) claims that "[HC] fits quite neatly into the Creole Proto-type model as a case [...] in which contact with the lexifier over the centuries pulled the creole away from the Prototype to which it honed at its genesis". This claim is somewhat echoed in WSG (Section 3.2) where HC, "due to contact over the centuries with French", is considered to "have borrowed many French lexicalized derivation–root combinations and thus does not exemplify the Creole Prototype in the purest possi-ble form". Such a scenario is ahistorical given established socio-demographic facts of Haitian history as sketched in 2.1 above (also see Note 15 and Appendix B).

On the empirical front, McWhorter (2000a: 106) mistakenly relies on what he calls "[Goyette's (2000) demonstration] through painstaking historical analysis that the deri-vation markers in modern [Haitian Creole] [...] cannot have been incorporated into the creole at its birth, and in fact were borrowed from French in later periods."

Not only is Goyette's scenario as ahistorical as McWhorter's (see the sociohistorical sketch in Section 2.1), Goyette's empirical generalizations about (the diachrony of) HC and French are flawed, and so are his deductions.

Take, for example, Goyette's discussion of the timecourse of the HC prefix *re-* /re/ (cf. French *re-*). From Fattier (1998), he quotes the lone HC form *ekile* /ekile/ 'to move back' (a variant of HC *rekile* /rekile/). Goyette claims *ekile* as a cognate of some (di-alectal) metathesized French variant *erculer* (cf. Standard French *reculer* /røkyle/). In Goyette's scenario, *erculer*'s counterpart in HC is *ekile* and not *erkile* since most HC dialects forbid syllable-final /r/. Why is the single HC form *ekile*, taken from Fattier's 6-volume dissertation (1998), of such significance? Goyette's argument is based on the premise that "[i]n 17th-century French, [the prefix *re-*] was CONSISTENTLY metathe-sized" (emphasis added) while "where /re/ is a prefix [in HC] such metathesis is wholly unknown". Thus, Goyette argues, the first-syllable /e/ in HC *ekile* is the UN-productive remnant of the metathesized French prefix *er-*; crucially the productive prefix *re-* in contemporary HC was not part of proto-HC morphology. Goyette's conclusion: *re-* in contemporary HC was borrowed late, not inherited early.

To the extent that I can understand Goyette's argument, it seems to contain at least one empirical flaw, even if one abstracts away from Goyette's overly simplistic claims about the phonology, distribution, and diachrony of the *er-*/*re-* alternation across French dialects (the complex diachrony of French dialects sheds doubt on the categorical claim that "[i]n 17th-century French, [the prefix *re-*] was consistently metathesized"). Given (*inter alia*) that *ekile* and *rekile* have the same initial vowel /e/, Fattier (1998) reasonably takes *ekile* as a case of apheresis, not metathesis. Fattier also documents a third variant, which she writes ᴿ*ekile*, where the first segment (the superscript /r/) is a phonetically weakened variant of the /r/ in *rekile* (see Fattier 1998: Volume 1, 230–232, Volume 2, 449 for details). Furthermore, throughout her 6-volume thesis, Fattier documents ro-bust cases of apheresis in a variety of environments. Similar apheresis is documented in Ducœurjoly's (1802) creole-teaching manual. The latter also documents robust affix-ation in early HC, virtually all of it derived from French affixes (including apparently UN-metathesized *re-*). Thus vanishes Goyette's single data point in arguing for the late borrowing of HC *re-*.

The other HC suffix discussed by Goyette (2000) is agentive *-è* /ɛ/ as in *mantè* /mãtɛ/. For Goyette, "[HC *-è*] is the normal reflex of final French *-eur* [pronounced /œr/] [...]

IN THE SEVENTEENTH-CENTURY, [this French] SUFFIX LACKED A FINAL /r/ [...] (*menteux* instead of *menteur* 'liar'). [...] We should expect the modern [HC] form of the suffix to be [/e/ since French /ø/ is mapped into HC /e/]" (emphasis added). Here too, Goyette's scenario is overly simplistic and empirically inaccurate. Although exact dates are uncertain, it seems that the reduction and dropping of final /r/ in French *-eur* (as part of a larger pattern of final-consonant reduction) was most robust in the Middle French period up until the 16th century. The opposing trend (pronunciation of final /r/) gained strength in the second half of the 17th century, becoming re-established in many 18th-century dialects and their descendants up to the present.

So could it be true that "[i]n the seventeenth-century, [the *-eur*] suffix lacked a final /r/"? The crucial observation here is that, even at its apex, final-/r/ reduction in French was not categorical across all dialects; see, e.g., Fouché (1966: 669) and Morin (1986: 173–175). Instead many dialects manifested (phonologically, morphologically, semantically, and sociolinguistically conditioned) variation between dropping and retention of final /r/. In some cases, final /r/ would enter into external sandhi phenomena, being pronounced before a vowel or a pause. In other cases, final /r/ was pronounced for emphasis or semantic nuance or for ill-understood sociolinguistic reasons (e.g., hypercorrection). A very telling case is the 1547 example *Ajoustes si tu veux les Perfumeux, les Balleurs* 'Add if you want the *Perfumeux* and the *Balleurs*' where *-eux* and *-eur* alternate in the very same sentence (Brunot 1906: 290); also see Brunot (1913: 211–212, 1924: 671) for other instances of variation with semantic and/or sociolinguistic nuances and for further remarks on the diachronic course of final-/r/ reduction.

In addition to the variable rule of final-/r/ reduction, there are other factors affecting *-eur* in French diachrony. Discussing the passage from /o/ to /œ/ or /ø/ through Old and Middle French (cf. *flor > fleur*, *dolor > douleur*, etc.), Nyrop (1899: 163) mentions 15th- and 16th-century dialects where *œ* and *ø* do not exist and where *sœur* and *sur* enter into "near rhymes that are said to be either 'Provencal' or 'Gascon', either 'Normand' or 'from Chartres'". Nyrop adds that such near rhymes are widespread in the 15th and 16th centuries. Also relevant here is the following synchronic alternation: *chœur/choral*, *docteur/doctoral*, *fleur/floral*, *mœurs/moral*, *pasteur/pastoral*, etc. (I thank Dominique Fattier for bringing this alternation to my attention.)

Pending further research, one can reasonably speculate that the above alternations in French diachrony and synchrony are related to another case of variation in HC's synchronic morphophonology. As Fattier (1998) repeatedly notes (also see Freeman & Laguerre's 1998 dictionary), *è/ò* alternation in HC word-final syllables is quite widespread, although not generalized: *flatè/flatò* 'flatterer' (cf. *flate* 'to flatter'), *gadè/gadò* 'watchman' (cf. *gade* 'to watch'), *vòlè/vòlò* 'thief' (cf. *vòlè* 'to steal'), etc. (also see *fiyèl/fiyòl* 'godchild', *lè/lò* 'time', *sè/sò* 'sister', etc.); see Fattier (1998: Volume 1, 132, 269, Volume 2, 501, 740, 755). This *è/ò* alternation does not seem to have been borrowed late from French as spoken in Haiti: as far as I can tell, contemporary Haitian French manifests no analogous alternation or remnants thereof; see Pompilus (1961) for a sketch of Haitian French. (Although Goyette mentions HC *mantè* /mãtɛ/ from Fattier's thesis, he fails to mention the variant *mantò* /mãtɔ/ and the corresponding alternation vis-à-vis the *-è* suffix.)

On the conceptual front, it must be noted that, unlike, say, Québec French (per

Goyette), many contemporary (non-creole) French dialects (e.g., Standard French and Haitian French), on a par with HC, do not productively use *er-* for *re-* and neither do they productively drop their final /r/ in *-eur* /œr/ (contrast with Québec French *erculer* and *menteux* as cited by Goyette). Yet the absence of such alternations surely does not constitute evidence that the contemporary affixes *re-* and *-eur* in these French varieties sans *er-* and sans final-/r/ dropping "were borrowed [. . .] in later periods". Where would these "later period" French speakers "borrow" the (UN-metathesized) *re-* prefix and the /-œr/ suffix (WITH phonetic realization of final /r/) if ALL prior dialects consistently lacked such affixes?

Now, let's assume (purely for the sake of argument) that (some? most?) French dialects in 17th-century Haiti did metathesize *re-* into *er-* and did drop final /r/ in, e.g., *-eur*. Then, whatever (socio)linguistic processes led to contemporary absence of *er-* and the contemporary retention of final /r/ in (the "later periods" of) Standard French, Haitian French, etc., could, in principle, also account for the corresponding facts in HC without evoking an (unattested) affixless-pidgin stage. In any case, we must reckon with the complex French and HC variations noted in this appendix. Pending detailed work on the (socio)linguistics of Caribbean French colonists and their entourage, it is quite unlikely that all the relevant French varieties would "consistently metathesize [-*er*]" and that all "lacked a final /r/ [in their pronunciation of *-eur*]". (See, e.g., Chaudenson & Mufwene 2001: 145–153 for preliminary remarks and caveats on the intricate mix of Romance lects – standard, "patois", non-native varieties, koines, etc. – in the colonial French Caribbean and elsewhere in the New World.)

To sum up, the data and observations that disconfirm Goyette's (and McWhorter's) late-borrowing scenario are taken from Goyette's own bibliographical source (i.e., Fattier 1998) and from well-known facts about French diachrony as can be found in, e.g., Nyrop (1899, 1903), Brunot (1906, 1924), Zink (1986), Morin (1986), Pierret (1994). Lastly, Goyette's (2000) categorical claim about the postulated (but undocumented) existence of a totally affixless proto-HC is based both on faulty logic and on erroneous empirical generalizations about (the diachrony of) only two HC affixes – not a representative sample, by any measure.

Appendix B: On the "afrogenesis" of Haitian (and Mauritian) Creole

All available evidence, coupled with theoretical considerations, suggests that the ancestors of HC as spoken in colonial Haiti (then known as Saint-Domingue) were never affixless, contra the claims of the "classic" pidgin-to-creole scenario (see the data, argumentation and references in Sections 2.1, 3.3, and Appendix A). But what if the "real" ancestor of HC goes back even further, to some UNDOCUMENTED French-based proto-pidgin spoken somewhere in West Africa, say in Senegal, around a slave fort? Is it, then, this hypothetical pidgin qua HC's proto-ancestor that would have most closely honed to the ("simplest") Creole Prototype?

McWhorter's (2000b) "Afrogenesis Hypothesis" (hereafter AH) pushes the origin of Caribbean and Indian Ocean creoles back to a small number of hypothetical pidgins that would have been created around West African slave castles (cf. Goodman 1964: 129–132). The AH considers Caribbean and Indian Ocean French-lexicon creoles as expan-

sions of a SINGLE French-based pidgin ancestor that was created around a slave-trading fort on the Senegalese coast in mid-17th century. The AH is empirically, methodologically, and theoretically flawed. Only a small sample of these flaws can be discussed in this appendix, abstracting away from the lack of historical evidence (see, e.g., Bickerton 1998).

Empirically, the AH's linguistic evidence is based on a misleading and skewed survey based on superficial comparisons. The latter are string-based with no attempt whatsoever at structural, distributional, and semantic analyses. Independently of robust and documented morphosyntactic and interpretive differences, any superficial similarity between, say, tense/mood/aspect markers across Caribbean and Indian Ocean French-lexicon creoles is judged "too close to be attributed to chance" and taken to "suggest common ancestry" in the hypothesized SINGLE Senegalese French-based pidgin (McWhorter 2000b: 149). Given the granularity of such comparison, a host of fundamental differences among French-lexicon creoles (vis-à-vis their syntax and semantics) are either ignored altogether or dismissed as "minor paradigmatic variation [that] is not counterevidence to a common [Senegalese pidgin] parent".

Take tense/mood/aspect markers. On one hand, robust TMA-related differences across French-lexicon creoles are amply documented. On the other hand, non-creole French varieties (e.g., Québec French, Missouri French, Cajun French, and 17th-/18th-century French) exhibit cognates of the same preverbal TMA FORMS that are enlisted from HC and Mauritian Creole as support for AH. (The references in Note 48 document a variety of (dis)similarities in TMA and clause structure across creole and non-creole French-related varieties.) Do the "creole-like" preverbal TMA markers in regional and diachronic French varieties trace back to a single pidgin spoken somewhere in Africa? As many have noted, the origin of these markers is, in all likelihood, not from a Senegalese French-based pidgin, but from (the grammaticalization and, in some cases L1-influenced, restructuration of) verbal periphrases in earlier French varieties.

More generally, do loose and superficial similarities in the phonetics, distribution, and interpretation of preverbal TMA markers, IN ADDITION TO SYSTEMATIC MORPHOLOGICAL AND LEXICAL CORRESPONDENCES, constitute "conclusive evidence of a common origin" (cf. McWhorter 2000b: 148–151, 178–179, etc.)? If so, then HC and Mauritian Creole readily join Québec French, Cajun French, Missouri French, etc., as bona fide co-descendants with common origins in FULL-FLEDGED varieties of French. At this point, the discussion of (non)common PIDGIN origins for French-lexicon creoles becomes moot. (Also note that Baker's (1995: 14) survey of pidgin and creole characteristics tentatively takes the combination of preverbal markers as one of the few "potential candidates for linguistic features which might distinguish creoles from pidgins". If Baker is right, then the "rudimentary" would-be pidgin spoken around 17th-century West African slave castles could not have provided any stable structural model for the complex structures of HC and Mauritian Creole's TMA systems, contra the premises of the AH.)

Related methodological remarks apply to McWhorter's (2000b: 151–155) use of HC's *ye* and Mauritian Creole's *ete* as evidence for the AH. (See, e.g., HC *ye* in *Se yon lengwis Bouki ye* 'It's a linguist that Bouki is' and Mauritian *ete* in *En voler Malis ete* 'A thief Malis is'.) HC *ye* and Mauritian *ete* are classified, without structural analysis,

as (so called) "exposed copulas" and taken as tell-tales of HC and Mauritian Creole's common ancestry in a single Senegalese pidgin.

One problem here is that McWhorter skirts around central empirical and theoretical details of HC and/vs. Mauritian predication patterns and their (language-particular and universal/theoretical) implications. Here are some tidbits of ongoing research and debate on the nature of HC *ye* vs. Mauritian Creole *ete*, just enough to illustrate the fragility of the facile comparisons that underlie the AH.

On the Haitian side, I myself have analyzed *ye* as a morpheme that spells out certain traces of non-verbal (i.e., $[-V]$) predicates when the latter move outside the clause (i.e., outside IP into some operator position) for contrastive stress or WH-formation; see De-Graff (1992a, b, 1995, 1998, 1999c); cf. (34h)–(34j) above. In *Se yon lengwis Bouki ye*, *ye* would spell out the trace of the displaced nominal predicate *yon lengwis* 'a linguist', even when the (phonetically realized) trace is not in sentence-final position as in *Se yon lengwis Bouki ye vre* 'It's a linguist that he really is' (like HC *ye*, Mauritian *ete* does surface even when not "exposed" in sentence-final position; Syea 1997: 34). This disconfirms McWhorter's (2000b: 152) generalization that "copular overtness is sharply restricted to sentence-final position"; "exposed copula" is somewhat a misnomer. In reality, spell-out of $[-V]$ predicate traces by *ye* is subject to subtle syntactic and semantic constraints, having to do with the licensing of movement and traces and with the syntax-semantics of quantification. These constraints are rooted in UG, even if they result in distributional and interpretative details that appear specific to HC.

On the Mauritian side, Baker & Corne (1982: 46, 103) argue explicitly against the copula status of *ete* while Baker & Syea (1991: 172) take *ete* to result from Case-assignment requirements. As of Syea's (1997) analysis, it takes *ete* to result from the need to strongly head-govern a predicate trace; in root questions such as *Kot Malis (ete)?*, strong head-government of *kote*'s trace is ensured either via *ete* or via (the co-indexed trace of) a null copula that has moved from V to C and agrees abstractly with the moved predicate in Spec(CP). Though elegant, this analysis does not account for the (apparent?) restriction of *ete* to WH-, nominal, and prepositional predicates (do we get *ete* with movement of verbal and adjectival ($[+V]$) projections?). Neither does it account for the facts noted by Baker & Syea (1991: 167–170) whereby *ete* in contemporary Mauritian Creole is OBLIGATORY in matrix non-negated present-tense questions with *ki*, *ki kote*, and other WH-phrases distinct from *kot* (but see Syea 1997: 28–29 for crucially different – dialectal? – judgements). Lastly, if the null-vs.-overt alternation is driven by economy considerations vis-à-vis ECP satisfaction ("null except when it can't be"; Syea 1997: 52), then the null-vs.-overt alternation in *Kot Malis (ete)?* is incorrectly ruled out. Of course, these are all delicate theoretical problems that McWhorter's superficial "description" has nothing to say about.

Most relevant to the discussion here, four observations are in order on the cross-creole/crosslinguistic syntax of (so called) "exposed copulas": (i) HC *ye* and Mauritian *ete* seem to have evolved via distinct diachronic routes, judging from their respective etyma and from Baker & Syea's (1991) historical analysis; (ii) predication and predicate-movement patterns are not isomorphic across HC and Mauritian Creole: differences obtain, e.g., in root clauses and in comparative clauses; see Baker & Corne (1981: 31–48), Syea (1997: 30, Note 11); (iii) the HC and Mauritian predicate-movement patterns find

rough analogues in English varieties, which motivates Syea (1997) to extend his analy-sis to *John is/'s a teacher* vs. *I wonder what John is/*'s* (*now*) vs. *What is/'s John?*; in a similar comparative mode, DeGraff (1998) analyzes certain parallels between HC and Irish predication patterns and extends his analysis from HC to Irish; (iv) some of the Haitian and Mauritian non-verbal predication patterns find parallels in Hebrew, Arabic, Russian, etc. (DeGraff 1992a, b, 1998; Syea 1997).

Observations (i) and (ii) disconfirm McWhorter's (2000b: 152) claim that HC and Mauritian show "the same occurrence pattern", which would originate from a single common pidgin. Here, again, detailed distributional and structural analysis is of the essence: linguistics, after all, is about structure, not strings, since what you see is often NOT what the structure gets (see Section 5.3 above for related remarks). Observations (iii) and (iv) illustrate the methodological perils of scant and skewed comparisons as "linguistic evidence" in phylogenetic speculations: the structural resemblances in (iii) and (iv) surely do not suggest that HC, Mauritian Creole, English, Irish, Hebrew, Ara-bic, and Russian all descend from a Senegalese French-based pidgin.

In any event, as analyzed thus far, the HC predicate-movement strategies (see (34h)–(34j)) and their Mauritian analogues are built on the intricate interaction of delicate morphosyntactic and/or semantic constraints, and so are their TMA systems. As such, they could hardly qualify as pidgin features, specially given the definition of pidgins in WSG (Section 2.3) as youngest/simplest "rudimentary codes" that eschew all but the "functionally central" (also see Baker's 1995: 14 comment, cited above, on pidgins' apparent lack of TMA combinations). The predicate-movement and TMA strategies in HC and Mauritian do not exist in many (functional) "old" languages. Therefore, the syntax and semantics of TMA and predication in HC and Mauritian – alleged telltales of a common pidgin ancestor – could not have been part of any pidgin that was created on the Guinea Coast as "rudimentary code [that is unlike] full language".

Another conceptual flaw in the AH argumentation concerns its idiosyncratic use of the comparative method. Take, say, Romance languages and the uncontroversial fact that they, like French-lexicon creoles, exhibit structural correspondences aplenty at var-ious levels of grammar (including across their Latin-derived lexica). Pan-Romance cor-respondences are more reliably documented and more numerous than McWhorter's few HC–Mauritian correspondences (the latter number a dozen or so). The logic of the AH comparison and argumentation, when applied to Romance, would have us erroneously conclude that all Romance languages originated in a single locale via a single contact language created in a single encounter. However pan-Romance similarities are not due to a single encounter with Latin; they are due to historically and geographically sep-arate encounters between related varieties of Latin and diverse "substrate" languages. The Romance case teaches us that it is an error to claim monogenesis in a single pidgin in order to explain grammatical correspondences among certain French-lexicon creoles (e.g., HC and Mauritian Creole), lest we throw away our usual comparative-historical heuristics when dealing with creole genesis.

A logical flaw in AH concerns the (non-)evidence for a common French-based PID-GIN ancestor vs. related evidence for FULL-FLEDGED French ancestors. If pidgins are structurally reduced lowest-common-denominator compromises among the source lan-guages – "rudimentary codes not fulfilling the needs of full language [. . . that] eschew

all but the functionally central" (WSG: Section 2.3; also see discussion in Section 4.2 above) – then the few bits of superficial comparative data in the AH also count, in principle, as evidence for common ancestry in (native and non-native) FULL-FLEDGED French varieties, as spoken in the relevant contact situations. Given assumptions in the AH (and in WSG), full-fledged French varieties (or, rather, simplifications thereof) contributed to the reduced patterns in the hypothesized Senegalese (proto-)pidgin – recall that pidginization " 'shaves away' a large degree of accreted complexity" from the source languages (WSG: Section 4.4; also see McWhorter 2000b: 4). If so, there logically is no way to demonstrate single ancestry for, e.g., HC and Mauritian Creole in a common French-based PIDGIN while excluding separate ancestry in independent situations of contact between the corresponding FULL-FLEDGED varieties of French and (some of) the corresponding FULL-FLEDGED substrate(s), unless it can be documented that speakers of the relevant varieties entered into contact once and only once, in one and only one locale.

As den Besten, Muysken, & Smith (1995: 88–89) write, any sort of theory of monogenesis from a single pidgin is "fundamentally flawed" and "completely irrational" because "[a] UNIQUE example of any TYPE of phenomenon connected with human conceptual and cultural activity is just inconceivable – anything that can happen once can also happen more frequently". If (some) adult plantation slaves in Haiti, Jamaica, Barbados, Cuba, Columbia, Mauritius, Seychelles, Mauritius, Réunion, New Caledonia, etc., did approximate some (ANY) variety of some European language (cf. (ii) and (iv) in Note 15), then they, like language learners everywhere, could not have acquired that variety overnight. This is specially so in the psycho-social context of colonial plantations from the perspective of the African-born who were taken to the colony as adults (see the contemporary reports in Pelleprat 1655, Girod-Chantrans 1785, Moreau de Saint-Méry 1797, Descourtilz 1809, etc.). These African-born slaves (the *bossales*) must have passed through a "pidgin(ized)" (QUA EARLY INTERLANGUAGE) stage with some structural features similar to those of the corresponding "pidgins" created by their in-situ compatriots who dwelt around African slave forts. It is not accidental that the speech of the Bossales – the numerical majority on 18th-century Haitian plantations – was often ranked as markedly "inferior" and "unintelligible" as compared to the speech of the locally-born ("Creole") slaves. Sociolinguistic factors, some of which remain to be elucidated, would determine the eventual fate of these early pidgins/interlanguages in the Old and New Worlds. However Cartesian-Uniformitarian assumptions about language acquisition/creation (see Sections 4.2 and 4.4) guarantee the existence of these "pidgin(ized)" varieties, at least as transitory individual-level lects, across all instances of language contact in the Old and New Worlds and beyond.

As for full-fledged and stable French-lexicon creoles, systematic correspondences between them and the corresponding FULL-FLEDGED varieties of French as spoken in, say, the then-colonized Caribbean and Indian Ocean islands (e.g., systematic correspondences at the level of morphology and lexicon; see Section 3.3) are many times more robust and numerous than the few (about a dozen) superficial correspondences claimed by the AH as genetic tracers linking Caribbean and Indian Ocean French-lexicon creoles. Yet these systematic lexical-morphological correspondences across creoles and their respective lexifiers, while they belie exclusive ancestry in a structurally reduced

pidgin, are not taken into account by the AH. In effect, the AH enlists scant and skewed correspondences among Caribbean and Indian-Ocean French-lexicon creoles in order to argue for the un-broken transmission of some hypothetical French-based Senegalese pidgin while it discards massive and robust correspondences between these creoles and their French lexifier in order to argue for a radical break in the transmission of French. This methodological paradox is unlike standard practices in comparative-historical linguistics.

Keeping the latter in mind, we can conclude that CERTAIN commonalities among, say, Mauritian and HC need not be due to "the SAME encounter with French" and need not "trace back to the same pidgin" (contra McWhorter 2000b: 147, 150). Like in the better understood Romance case, commonalities can independently arise from SEPARATE encounters (in the plural) among overlapping sets of languages and from universal strategies of language acquisition/creation. Common ancestry and common patterns do not necessarily entail common birthplace. In this perspective, differences among French-lexicon creoles (SOMEWHAT on a par with differences across Romance) are due to, *inter alia*, ecological variations across contact situations – variations that are now being documented (which varieties were spoken and learnt where, how, by whom, to what ends, by how many, for how long, etc.?).

On a structural UG-related note, let's ask a question that goes beyond the specific concerns of the AH (this question is related to the general methodology in WSG; see Section 5.3 above): With respect to modern historical-comparative syntax, what is the status and import of string-based comparisons of few isolated and superficial patterns? With clever handpicking, any degree of ad hoc superficial (dis)similarity can be established between any pair of languages, whether they are historically related or not (see Note 20). Yet twelve or so superficial similarities seem enough for the AH to consider HC and Mauritian Creole "too alike not to have had a common [pidgin] ancestor", in spite of the lack of sociohistorical evidence for such a pidgin. One can't help but notice that there do exist, across a wide range of creoles and non-creoles across distinct genetic phyla, deep STRUCTURAL similarities in clausal structure, including TMA structure (see, e.g., Cinque 1999). In AH's parlance, these similarities become "too close to be attributed to chance", "suggest[ing] common ancestry". Thus, the widespread structural similarities in Cinque's large-scale and theoretically grounded crosslinguistic comparisons would erroneously suggest a monogenesis scenario for ALL languages – perhaps from a single (affixless?) Ur-creole spoken by Eve in prehistorical Africa! (See Note 18.) Humor put aside, Cinque's own, and more reasonable, conclusion is that certain crosslinguistically common patterns, even if they superficially look idiosyncratic, can – actually, MUST – arise independently across languages, simply due to UG. In this view, common patterns do not necessarily entail common ancestry in one single E-language. Common patterns are often due to the common biological ancestry of the species *homo sapiens* and universal constraints on I-languages. (See Marantz 1983: 16 for similar arguments, as part of a critique of Bickerton's Bioprogram Hypothesis.)

Correspondence address: Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, U.S.A.; e-mail: degraff@mit.edu

Abbreviations: AH Afrogenesis Hypothesis (see Appendix B), BV Basic Variety (Klein & Perdue 1997), HC Haitian Creole, PLD Primary Linguistic Data, TMA tense/mood/aspect, UG Universal Grammar, WSG McWhorter's target article in this issue of *Linguistic Typology*.

References

Aarsleff, Hans (1988). Introduction. In Humboldt 1988, vii–lxv.

Adam, Lucien (1882). *Les classifications, l'objet, la méthode, les conclusions de la linguistique*. Paris: Maisonneuve.

— (1883). *Les idiomes négro-aryen et maléo-aryen: Essai d'hybridologie linguistique*. Paris: Maisonneuve.

Alleyne, Mervyn (1971). The cultural matrix of creolization. In Hymes (ed.) 1971, 169–186.

— (1988). *Roots of Jamaican Culture*. London: Pluto.

— (1994). Problems of standardization of creole languages. In Morgan (ed.) 1994, 7–18.

Alter, Stephen G. (1999). *Darwinism and the Linguistic Image: Language, Race, and Natural Theology in the Nineteenth Century*. Baltimore: Johns Hopkins University Press.

Andersen, Roger (1983). Introduction: A language acquisition interpretation of pidginization and creolization. In Roger Andersen (ed.), *Pidginization and Creolization as Language Acquisition*, 1–56. Rowley, Mass.: Newbury House.

Arends, Jacques, Silvia Kouwenberg, & Norval Smith (1995). Theories focusing on the non-European input. In Arends, Muysken, & Smith (eds.) 1995, 99-109.

Arends, Jacques, Pieter Muysken, & Norval Smith (eds.) (1995). *Pidgins and Creoles: An Introduction*. Amsterdam: Benjamins.

Bacon, Francis (1620 [1994]). *Novum Organum*. [Translated and edited by Peter Urbach & John Gibson, Chicago: Open Court.]

Bailey, Charles-James & Karl Maroldt (1977). The French lineage of English. In Jürgen Meisel (ed.), *Langues en contact: Pidgins, créoles,* 21–53. Tübingen: Narr.

Baissac, Charles (1880). *Etude sur le patois créole mauricien*. Nancy: Berger-Levrault.

Baker, Philip (1995). Some developmental inferences from historical studies of pidgins and creoles. In Jacques Arends (ed.), *The Early Stages of Creolization*, 1–24. Amsterdam: Benjamins.

Baker, Philip & Chris Corne (1982). *Isle de France Créole*. Ann Arbor, Mich.: Karoma.

Baker, Philip & Anand Syea (1991). On the copula in Mauritian Creole, past and present. In Francis Byrne & Thom Huebner (eds.), *Development and Structures of Creole Languages: Essays in Honor of Derek Bickerton*, 159–175. Amsterdam: Benjamins.

Baptista, Marlyse (1997). The morpho-syntax of nominal and verbal categories in Capeverdean Creole. Doctoral dissertation, Harvard University.

— (2001). Number inflection in creole languages. Paper read at the meeting of the Society of Pidgin and Creole Linguistics, Coimbra, Portugal, June 2001. To appear in Anthony Grant (ed.), *Interface* (Bradford Studies in Language, Culture and Society). Bradford: University of Bradford.

Bateman, Richard, Ives Goddard, Richard O'Grady, V. A. Funk, Rich Mooi, W. John Kress, & Peter Cannell (1990). Speaking of forked tongues: The feasibility of reconciling human phylogeny and the history of language. *Current Anthropology* 31: 1–24.

Bickerton, Derek (1977). Pidginization and creolization: Language acquisition and language universals. In Albert Valdman (ed.), *Pidgin and Creole Linguistics*, 49–69. Bloomington: Indiana University Press.

— (1988). Creole languages and the bioprogram. In Newmeyer (ed.) 1988, 268–284.

— (1998). A sociohistoric examination of Afrogenesis. *Journal of Pidgin and Creole Languages* 13: 63–92.

— (1999). How to acquire language without positive evidence: What acquisitionists can learn from creoles. In DeGraff (ed.) 1999, 49–74.

Bloomfield, Leonard (1933). *Language*. New York: Holt.

Boas, Franz (1911). Introduction. In Franz Boas (ed.), *Handbook of American Indian Languages, Part 1*, 1–83. Washington: Government Printing Office.

Bopp, Franz (1833–1852 [1985]). *Vergleichende Grammatik des Sanskrit, Send, Armenischen, Griechischen, Lateinischen, Litthauischen, Gothischen, Deutschen und Altslawischen*. Berlin: Dümmler. [Translated by Edward B. Eastwick as *A Comparative Grammar of the Sanskrit, Zend, Greek, Latin, Lithuanian, Gothic, German, and Slavonic Languages*. London: Madden & Malcolm (1845–1853). Reprint: Hildesheim: Olms.]

Brunot, Ferdinand (1906). *Histoire de la langue française des origines à 1900. Tome II: Le seizième siècle*. Paris: Colin.

— (1913). *Histoire de la langue française des origines à 1900. Tome IV: La langue classique (1660–1715). Première partie.* Paris: Colin.

— (1924). *Histoire de la langue française des origines à 1900. Tome IV: La langue classique (1660–1715). Deuxième partie*. Paris: Colin.

Cadely, Jean-Robert (1988). Représentations syllabiques et distribution des diphtongues en créole haïtien. *Etudes Créoles* 11: 9–40.

— (1994). Aspects de la phonologie du créole haïtien. Doctoral dissertation, Université du Québec à Montréal.

— (1995). Elision et agglutination en créole haïtien: Le cas des pronoms personnels. *Etudes Créoles* 18: 9–38.

— (1997). Prosodie et cliticisation en créole haïtien. *Etudes Créoles* 20: 77–88.

— (forthcoming). Nasality in Haitian Creole. In Dany Adone (ed.), *Recent Developments in Creole Studies*. Tübingen: Niemeyer.

Campbell, Lyle (1999). *Historical Linguistics: An Introduction*. Cambridge, Mass.: MIT Press.

Cavalli-Sforza, Luigi Luca (2000). *Genes, Peoples, and Languages*. New York: North Point.

Chafe, Wallace (1970). *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

Chambers, J. K. (1995). *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell.

Chaudenson, Robert (1979). *Les créoles français*. Paris: Nathan.

— (1992). *Des îles, des hommes, des langues: Essai sur la créolisation linguistique et culturelle*. Paris: L'Harmattan.

— (1994). Les créoles français sont-ils des formes simplifiées du français? *Cahiers du Français Contemporain* 1: 41–54.

— (1995). *Les créoles*. Paris: Presses Universitaires de France.

Chaudenson, Robert, Raymond Mougeon, & Edouard Beniak (1993). *Vers une approche panlectale de la variation du français*. Aix-en-Provence: Institut d'Etudes Créoles et Francophones.

Chaudenson, Robert & Salikoko Mufwene (2001). *Creolization of Language and Culture.* London: Routledge. [English translation and revision of Chaudenson (1992).]

Chomsky, Noam (1957). Review of Charles F. Hockett, *A Manual of Phonology* (1955). *International Journal of American Linguistics* 23: 223–234.

— (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought.* New York: Harper & Row.

— (1981). Principles and parameters in syntactic theory. In Norbert Hornstein & David Lightfoot (eds.), *Explanation in Linguistics*, 32–75. London: Longman.

— (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

— (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.

— (2001). Derivation by phase. In Kenstowicz (ed.) 2001, 1–52.

Cinque, Guglielmo (1999). *Adverbs and Functional Heads : A Cross-Linguistic Perspective*. New York: Oxford University Press.

Clark, Eve (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

Clyne, Michael (ed.) (1981). Foreigner talk. *International Journal of the Sociology of Language* 28.

Comrie, Bernard (1992). Before complexity. In Hawkins & Gell-Mann (eds.) 1992, 193–211.

Corcoran, Chris (2001). Creoles and the creation myth: A report on some problems with the linguistic use of "Creole". *Chicago Linguistic Society* 36.

Crowley, Terry & Bruce Rigsby (1979). Cape York Creole. In Timothy Shopen (ed.), *Languages and their Status*, 153–207. Cambridge, Mass.: Winthrop.

Damoiseau, Robert & Gérald Saint-Louis (1986). Les verbo-adjectivaux en créole haïtien. *Modèles Linguistiques* 8: 103–175.

D'Ans, André-Marcel (1968). *Le créole français d'Haïti*. The Hague: Mouton.

Darwin, Charles (1859 [1979]). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* London: Murray. [Reprint: New York: Avenel.]

— (1871). *The Descent of Man, and Selection in Relation to Sex*. London: Murray.

DeGraff, Michel (1992a). Creole grammars and acquisition of syntax: The case of Haitian Creole. Doctoral dissertation, University of Pennsylvania.

— (1992b). The syntax of predication in Haitian. *North-Eastern Linguistics Society* 22: 103–117.

— (1993). A riddle on negation in Haitian. *Probus* 5: 63–93.

— (1995). Predicate-movement, quantification, events and properties. In Victor Manfredi & Karl Reynolds (eds.), *Niger-Congo Syntax and Semantics 6*, 69–82. Boston: Boston University African Studies Center.

— (1996a). Review article on Michaelis 1993. *Journal of Pidgin and Creole Languages* 11: 121–137.

— (1996b). UG and acquisition in pidginization and creolization. *Behavioral and Brain Sciences* 19: 723–724.

— (1997). Verb syntax in creolization (and beyond). In Liliane Haegeman (ed.), *The New Comparative Syntax*, 64–94. London: Longman.

— (1998). Nominal Predication in Haitian and in Irish. *West Coast Conference on Formal Linguistics* 16: 113–128.

— (1999a). Creolization, language change and language acquisition: A prolegomenon. In De-Graff (ed.) 1999: 1–46.

— (1999b). Creolization, language change and language acquisition: An epilogue. In DeGraff (ed.) 1999: 473–543.

— (1999c). Empirical quicksand: Probing two recent articles on Haitian Creole. *Journal of Pidgin and Creole Languages* 14: 1–12.

— (1999d). Parameter-setting in acquisition and through creolization and language change. In Rob Wilson & Frank Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, 624–627. Cambridge, Mass.: MIT Press.

— (2000). A propos des pronoms objets dans le créole d'Haïti: Regards croisés de la morphologie et de la diachronie. *Langages* 138: 89–113.

— (2001a). Against Creole genesis as "abnormal transmission". Paper presented at the 75th Annual Meeting of the Linguistics Society of America, Washington, D.C., January 2001.

— (2001b). Morphology in Creole genesis: Linguistics and ideology. In Kenstowicz (ed.) 2001: 53–121.

— (forthcoming). Morphology and word order in "creolization" and beyond. In Richard Kayne & Guglielmo Cinque (eds.), *Handbook of Comparative Syntax*. New York: Oxford University Press.

— (in preparation). *Postcolonial Linguistics: The Politics of Creole Studies* [working title]. London: Blackwell.

DeGraff, Michel (ed.) (1999). *Language Creation and Language Change: Creolization, Diachrony and Development.* Cambridge, Mass.: MIT Press.

Dejean, Paul (1989). *Survol des tentatives d'alphabétisation en Haïti par les services gouvernementaux, 1943–1988.* Port-au-Prince: Groupe d'Action et de Recherche pour l'Education.

— (1993). *Haïti: Alerte, on tue!* Montréal: CIDIHCA.

Dejean, Yves (1975). *Dilemme en Haïti: Français en péril ou péril français?* New York: Connaissance d'Haïti.

— (1980). *Comment écrire le créole d'Haïti.* Outremont, Québec: Collectif Paroles.

— (1993). An overview of the language situation in Haiti. *International Journal of the Sociology of Language* 102: 73–83.

— (1999a). L'archipel nous revient: Créolistes et données. *Etudes Créoles* 22: 85–114.

— (1999b). The native language as a medium of instruction: An issue revisited. In Arthur Spears (ed.), *Race and Ideology*, 93–103. Detroit: Wayne State University Press.

— (forthcoming). Créole, école, rationalité. *Chemins Critiques*.

den Besten, Hans (1981). On the interaction of root transformations and lexical deletive rules. *Groninger Arbeiten zur Germanistischen Linguistik* 20: 1–78.

den Besten, Hans, Pieter Muysken, & Norval Smith (1995). Theories focusing on the European input. In Arends, Muysken, & Smith (eds.) 1995: 87–109.

Descourtilz, Michel Etienne (1809). *Voyages d'un naturaliste et ses observations faites sur les trois règnes de la nature, dans plusieurs ports de mer français, en Espagne, au continent de l'Amérique Septentrionale, à Saint-Yago de Cuba, et à St.-Domingue.* 3 volumes. Paris: Dufart.

Dijkhoff, Marta (1993). Papiamentu word formation: A case study of complex nouns and their relation to phrases and clauses. Doctoral dissertation, Universiteit van Amsterdam.

Dixon, R. M. W. (1997). *The Rise and Fall of Languages.* Cambridge: Cambridge University Press.

Domingue, Nicole Z. (1977). Middle English: Another creole? *Journal of Creole Studies* 1: 89–100.

Drechsel, Emanuel J. (1993). Basic word order in Mobilian Jargon: Underlying SOV or OSV? In Anthony Mattina & Timothy Montler (eds.), *American Indian Linguistics and Ethnography:*

In Honor of Laurence C. Thompson (University of Montana Occasional Papers in Linguistics, 10), 343–367, Missoula: University of Montana.

— (1997). *Mobilian Jargon: Linguistic and Sociohistorical Aspects of a Native American Pidgin*. Oxford: Oxford University Press.

Ducœurjoly, S.-J. (1802). *Manuel des habitans de Saint-Domingue*. 2 volumes. Paris: Lenoir.

Eco, Umberto (1998). *Serendipities: Language and Lunacy*. New York: Columbia University Press.

Faraclas, Nicholas (1988). Nigerian Pidgin and the languages of southern Nigeria. *Journal of Pidgin and Creole Languages* 3: 177–197.

Fattier, Dominique (1998 [2000]). Contribution à l'étude de la genèse d'un créole: L'atlas linguistique d'Haïti, cartes et commentaires. 6 volumes. Doctoral dissertation, Université de Provence. [Reprint: Villeneuve d'Ascq: Presses Universitaires du Septentrion.]

Ferguson, Charles (1971). Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk, and pidgins. In Hymes (ed.) 1971, 141–150.

— (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17: 1–14.

— (1981). "Foreigner Talk" as the name of a simplified register. In Clyne (ed.) 1981: 9–18.

Foucault, Michel (1971 [1972]). *L'ordre du discours*. Paris: Gallimard. [Translation by Rupert Swyer reprinted in Michel Foucault, *The Archaeology of Knowledge*, 215–237. New York: Pantheon.]

Fouché, Pierre (1966). *Phonétique historique du français*. Paris: Klincksieck.

Freeman, Bryant & Jowel Laguerre (1998). *Haitian–English Dictionary*. 2nd edition. Lawrence, Kansas: Institute of Haitian Studies, University of Kansas.

Giacalone Ramat, Anna & Paul Hopper (eds.) (1998). *The Limits of Grammaticalization*. Amsterdam: Benjamins.

Gilbert, Glenn (ed.) (1980). *Pidgin and Creole Languages: Selected Essays by Hugo Schuchardt*. Cambridge: Cambridge University Press.

Girod-Chantrans, Justin (1785 [1980]). *Voyage d'un suisse dans différentes colonies d'Amérique pendant la dernière guerre*. Neuchâtel: Imprimerie de la Société Typographique. [1980 edition: Paris: Tallandier.]

Givón, Talmy (1971). Historical syntax and synchronic morphology: An archeologist's field trip. *Chicago Linguistic Society* 7: 394–415.

— (1979). Prolegomena to any sane creology. In Ian Hancock et al. (eds.), *Readings in Creole Studies*, 3–35. Gent: Story-Scientia.

Goodman, Morris (1964). *A Comparative Study of Creole French Dialects*. The Hague: Mouton.

Gougenheim, Georges (1971). *Etude sur les périphrases verbales de la langue française*. Paris: Nizet.

Gould, Stephen Jay (1996). *The Mismeasure of Man*. New York: Norton.

Goyette, Stéphane (2000). Creole wars—Episode I: The prototype menace. Paper read at the meeting of the Society of Pidgin and Creole Linguistics, Chicago, January 2000.

Grimes, Barbara (ed.) (1996). *Ethnologue*. 13th edition. Dallas: Summer Institute of Linguistics.

Hall, Robert A. Jr. (1944). Chinese Pidgin English grammar and texts. *Journal of the American Oriental Society* 64: 95-113.

— (1953). *Haitian Creole: Grammar, Texts, Vocabulary*. (The American Anthropologist: Memoir, 74; Memoirs of the American Folklore Society, 43.) Washington: American Anthropological Association.

— (1962). The life cycle of pidgin languages. *Lingua* 11: 151–166.

— (1966). *Pidgin and Creole Languages*. Ithaca: Cornell University Press.

Halle, Morris (1962). Phonology in generative grammar. *Word* 18: 54–72.

Haspelmath, Martin (2000). Periphrasis. In Geert Booj, Christian Lehmann, & Joachim Mugdan (eds.), *Morphology: An International Handbook on Inflection and Word-Formation*, Volume 1, 654–664. Berlin: Mouton de Gruyter.

Hawkins, John (1986). *A Comparative Typology of English and German: Unifying the Contrasts.* London: Croom Helm.

— (1988). On generative and typological approaches to Universal Grammar. *Lingua* 74: 85–100.

Hawkins, John & Murray Gell-Mann (eds.) (1992). *The Evolution of Human Languages*. Reading, Mass.: Addison-Wesley.

Heath, Jeffrey (1998). Hermit crabs: Formal renewal of morphology by phonologically mediated affix substitution. *Language* 74: 728–759.

Heine, Bernd (1978). Some generalizations on African-based pidgin languages. *Afrika und Übersee* 61: 219–229.

— (1982). *The Nubi Language of Kibera: An Arabic Creole.* Berlin: Reimer.

Henry, Alison & Denise Tangney (1999). Functional categories and parameter setting in the second language acquisition of Irish in early childhood. In DeGraff (ed.) 1999, 239–253.

Hilton, Dimitri (2000). Pronominal clitics in Haitian Creole: Phonological or syntactic? In Vicki Carstens & Frederick Parkinson (eds.), *Advances in African Linguistics*, 51–59. Trenton, N.J.: Africa World Press.

Hjelmslev, Louis (1938). Relation de parenté des langues créoles. *Revue des Etudes Indo-Européennes* 1: 271–286.

Hodge, Carleton (1970). The linguistic cycle. *Language Sciences* 13: 1–7.

Hoenigswald, Henry (1971). Language history and creole studies. In Hymes (ed.) 1971, 473–480.

Hoenigswald, Henry & Linda Wiener (eds.) (1987). *Biological Metaphor and Cladistic Classification: An Interdisciplinary Perspective*. Philadelphia: University of Pennsylvania Press.

Holm, John (1988). *Pidgins and Creoles*. Cambridge: Cambridge University Press.

Huang, James (1982). Logical relations in Chinese and the theory of grammar. Doctoral dissertation, Massachusetts Institute of Technology.

Humboldt, Wilhelm von (1836 [1988]). *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwickelung des Menschengeschlechts*. [Translated by Peter Heath as *On Language: The Diversity of Human Language-Structure and its Influence on The Mental Development of Mankind*. Cambridge: Cambridge University Press.]

Huttar, Georges & Frank Velantie (1997). Ndyuka-Trio Pidgin. In Thomason (ed.) 1997, 99–124.

Hymes, Dell (1971a). Preface. In Hymes (ed.) 1971: 1–11.

— (1971b). General conceptions of process: Introduction. In Hymes (ed.) 1971: 65–90.

Hymes, Dell (ed.) (1971). *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press.

Jahr, Ernst Håkon & Ingvild Broch (eds.) (1996). *Language Contact in the Arctic: Northern Pidgins and Contact Languages*. Berlin: Mouton de Gruyter.

Janda, Richard (2001). Beyond "pathways" and "unidirectionality": On the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* 23: 265–340.

Jespersen, Otto (1922). *Language: Its Nature, Development and Origin*. London: Allen & Unwin.

Jordens, Peter (ed.) (1997). Introducing the Basic Variety. *Second Language Research* 13–4 (special issue).

Joseph, Frantz (1988). La détermination nominale en créole haïtien. Doctoral dissertation, Université de Paris VII.

Keesing, Roger (1988). *Melanesian Pidgin and the Oceanic Substrate*. Stanford, California: Stanford University Press.

— (1991). Substrates, calquing and grammaticalization in Melanesian Pidgin. In Traugott & Heine (eds.) 1991, Volume 1, 315–342.

Kegl, Judy & John McWhorter (1997). Perspectives on an emerging language. In Eve Clark (ed.), *Proceedings of the Twenty-Eighth Annual Child Language Research Forum*, 15–38. Stanford, Cal.: Center for the Study of Language and Information.

Kenstowicz, Michael (ed.) (2001). *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press.

Klein, Thomas (2001). Conflicting trends in Creole syllable structure: Evidence from Haitian morphophonology. Paper read at the meeting of the Society of Pidgin and Creole Linguistics, Coimbra, Portugal, June 2001.

Klein, Wolfgang & Clive Perdue (1997). The basic variety (or: Couldn't natural languages be much simpler?) In Jordens (ed.) 1997, 301–347.

Koerner, Konrad (1983). Editor's foreword. In Koerner (ed.) 1983, ix–xvi.

Koerner, Konrad (ed.) (1983). *Linguistics and Evolutionary Theory: Three Essays by August Schleicher, Ernst Haeckel and Wilhelm Bleek*. Amsterdam: Benjamins.

Koopman, Hilda (1986). The genesis of Haitian: Implications of a comparison of some features of the syntax of Haitian, French, and West African languages. In Muysken & Smith (eds.) 1986, 231–258.

Kouwenberg, Silvia (1994). *A Grammar of Berbice Dutch Creole*. Berlin: Mouton de Gruyter.

Labov, William (1994). *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

— (2001). *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Larousse, Pierre (ed.) (1869). *Grand dictionnaire universel du XIX siècle*. Paris: Larousse.

Lefebvre, Claire (1998). *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge: Cambridge University Press.

Lefebvre, Claire & Elizabeth Ritter (1989). A note on verbal reduplication in Haitian Creole. *Revue Québecoise de Linguistique* 18: 173–182.

— (1991). A note on three types of causal clauses in Haitian Creole. *Journal of Pidgin and Creole Languages* 6: 279–283.

Lehmann, Winfred (ed.) (1967). *A Reader in Nineteenth-Century Historical Indo-European Linguistics*. Bloomington: Indiana University Press.

Lightfoot, David (1999). *The Development of Language: Acquisition, Change, and Evolution*. Malden, Mass.: Blackwell.

Lumsden, John (1999). Language acquisition and creolization. In DeGraff (ed.) 1999, 129–157.

Maher, J. Peter (1983). Introduction. In Koerner (ed.) 1983, xvii–xxxii.

Marantz, Alec (1983). Before Babel: The misguided quest for the origins of language. *The Sciences* 23–3: 16–20.

Marantz, Alec, Yasushi Miyashita, & Wayne O'Neil (2000). Introduction: Mind articulation. In Alec Marantz, Yasushi Miyashita, & Wayne O'Neil (eds.), *Image, Language, Brain: Papers from the First Mind-Articulation Project Symposium*, 1–9. Cambridge, Mass.: MIT Press.

Matisoff, James (1991). Grammatization in Lahu. In Traugott & Heine (eds.) 1991, Volume 2, 383–453.

McWhorter, John (1997). *Towards a New Model of Creole Genesis*. New York: Lang.

— (1998). Identifying the creole prototype: Vindicating a typological class. *Language* 74: 788–818.

— (2000a). Defining "creole" as a synchronic term. In Neumann-Holzschuh & Schneider (eds.) 2000, 85–124.

— (2000b). *The Missing Spanish Creoles: Recovering the Birth of Plantation Contact Languages*. Berkeley: University of California Press.

Meillet, Antoine (1906 [1926]). L'état actuel des études de linguistique générale. Opening lecture for the course on comparative grammar at the Collège de France. [Reprinted in Meillet 1926: 1–18.]

— (1912 [1926]). L'évolution des formes grammaticales. *Scientia (Rivista di Scienza)* 12 [26]: 384–400. [Reprinted in Meillet 1926: 130–148.]

— (1919 [1926]). Le genre grammatical et l'élimination de la flexion. *Scientia (Rivista di Scienza)* 25 [86]: 460–470. [Reprinted in Meillet 1926: 199–210.]

— (1926). *Linguistique historique et linguistique générale*. [Volume 1.] 2nd edition. Paris: Champion.

— (1929 [1938]). Le développement des langues. In *Continu et discontinu* (Cahiers de la Nouvelle Journée, 15), 119–131. Paris: Bloud & Gay. [Reprinted in Meillet 1938: 70–83.]

— (1938). *Linguistique historique et linguistique générale*. Volume 2. Paris: Klincksieck.

Michaelis, Susanne (1993). *Temps et aspects en créole seychellois: Valeurs et interférences.* Hamburg: Buske.

Migge, Bettina (1998). Substrate influence in the formation of the Surinamese Plantation Creole: A consideration of the sociohistorical data and linguistic data from Ndyuka and Gbe. Doctoral dissertation, Ohio State University.

Moreau de Saint-Méry, M. L. E. (1797 [1958]). *Description topographique, physique, civile, politique et historique de la partie française de l'Isle Saint Domingue*. 2 volumes. Philadelphia: Author. [New, revised edition: Paris: Société de l'Histoire des Colonies Françaises.]

Morgan, Marcyliena (ed.) (1994). *Language & the Social Construction of Identity in Creole Situations*. Los Angeles: Center for Afro-American Studies, University of California.

Morin, Yves-Charles (1986). On the morphologization of word-final consonant deletion in French. In Henning Andersen (ed.), *Sandhi Phenomena in the Languages of Europe*, 167–210. Berlin: Mouton de Gruyter.

Mufwene, Salikoko (1986). Can creole languages be defined without reference to their history? *Etudes Creoles* 9: 135–150.

— (1990). Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition* 12: 1–23.

— (1993). *Africanisms in Afro-American Language Varieties*. Athens, Ga.: University of Georgia Press.

— (1996). The founder principle in creole genesis. *Diachronica* 13: 83–134.

— (1997). Kitúba. In Thomason (ed.) 1997, 173–208.

— (1998). What research on creole genesis can contribute to historical linguistics. In Monika Schmid, Jennifer Austin, & Dieter Stein (eds.), *Historical Linguistics 1997*, 315–338. Amsterdam: Benjamins.

— (2000a). Creolization is a social, not a structural, process. In Neumann-Holzschuh & Schneider (eds.) 2000: 65–84.

— (2000b). Population contact and the evolution of English. *The European English Messenger* 9: 9–15.

— (2001). *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Mühleisen, Susanne (2000). Creole discourse: Exploring prestige formation and change across Caribbean English-lexicon creoles. Doctoral dissertation, Johann Wolfgang Goethe-Universität zu Frankfurt am Main.

Müller, Max (1868). *On the Stratification of Language*. London: Longmans, Green, Reader, & Dyer.

Muusse, Esther & Jacques Arends (forthcoming). Inflectional morphology in the Lingua Franca noun phrase. In Hans den Besten & Mikael Parkvall (eds.), *Pidgins: Their Nature and Significance. Papers from the 2001 Westminster Pidgin Workshop*. London: Battlebridge.

Muysken, Pieter (1988). Are creoles a special type of language? In Newmeyer (ed.) 1988, 285–301.

— (in press). Pidginization, creolization and language death. In Geert Booij, Christian Lehmann, & Joachim Mugdan (eds.), *Morphology: An International Handbook on Inflection and Word-Formation*, Volume 2. Berlin: Mouton de Gruyter.

Muysken, Pieter & Paul Law (2001). *Creole studies: A theoretical linguist's field guide*. Glot International 5/2: 47–57.

Muysken, Pieter & Norval Smith (eds.) (1986). *Substrata versus Universals in Creole Genesis*. Amsterdam: Benjamins.

Neumann-Holzschuh, Ingrid & Edgar Schneider (eds.) (2000). *Degrees of Restructuring in Creole Languages*. Amsterdam: Benjamins.

Newmeyer, Frederick (ed.) (1988). *Linguistics: The Cambridge Survey*, Volume 2: *Linguistic Theory: Extensions and Implications*. Cambridge: Cambridge University Press.

Newport, Elissa (1999). Reduced input in the acquisition of Signed Languages: Contributions to the study of creolization. In DeGraff (ed.) 1999, 161–178.

Nichols, Johanna (1986). The bottom line: Chinese Pidgin Russian. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The Linguistic Encoding of Epistemology*, 239–257. Norwood, N.J.: Ablex.

Nyrop, Kristoffer (1899). *Grammaire historique de la langue française, tome premier*. København: Nordisk Forlag.

— (1903). *Grammaire historique de la langue française, tome deuxième*. København: Nordisk Forlag.

Osthoff, Hermann & Karl Brugman (1878 [1967]). Vorwort. In Hermann Osthoff & Karl Brugman, *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, Volume 1, iii–xx. Leipzig: Hirzel. [Translated by Winfred Lehmann in Lehmann (ed.) 1967: 197–209.]

Paul, Hermann (1890 [1970]). *Principien der Sprachgeschichte*. 2nd edition. Halle/Saale: Niemeyer. [Translated by H. A. Strong as *Principles of the History of Language*. College Park, Md.: McGrath.]

Pelleprat, Pierre (1655). *Relation des missions des PP. de la Compagnie de Jésus dans les îles et dans la terre ferme de l'Amérique Méridionale*. Paris: Cramoisy & Cramoisy.

Pierret, Jean-Marie (1994). *Phonétique historique du français et notions de phonétique générale*. Leuven: Peeters.

Pompilus, Pradel (1961). *La langue française en Haïti*. Paris: Institut des Hautes Etudes de l'Amérique Latine.

Popper, Karl (1965). *Conjectures and Refutations: The Growth of Scientific Knowledge*. 2nd edition. New York: Basic Books.

Prudent, Lambert-Félix (1980). *Des baragouins à la langue antillaise: Analyse historique et sociolinguistique du discours sur le créole*. Paris: Editions Caribéennes.

Rivera-Castillo, Yolanda (1998). Tone and stress in Papiamentu: The contribution of a constraint-based analysis to the problem of creole genesis. *Journal of Pidgin and Creole Languages* 13: 297–334.

Rizzi, Luigi (1999). Broadening the empirical basis of Universal Grammar models: A commentary. In DeGraff (ed.) 1999, 453–472.

Roberts, Ian (1999). Verb movement and markedness. In DeGraff (ed.) 1999, 287–327.

Romaine, Suzanne (1992). The evolution of linguistic complexity in pidgin and creole languages. In Hawkins & Gell-Mann (eds.) 1992: 213–238.

Saint-Quentin, Alfred de (1872 [1989]). *Introduction à l'histoire de Cayenne* [. . .]. *Etude sur la grammaire créole par Auguste de Saint-Quentin*. Antibes: Marchand. [Reprint: Cayenne: Comité de la Culture, de l'Éducation et de l'Environnement, Région Guyane.]

Samarin, William (1980). Standardization and instrumentalization of creole languages. In Albert Valdman & Arnold Highfield (eds.), *Theoretical Orientations in Creole Studies*, 213–236. New York: Academic Press.

Sapir, Edward (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & Co.

Schlegel, Friedrich (1808 [1849]). *Ueber die Sprache und Weisheit der Indier*. Heidelberg: Mohr & Zimmer. [Translated by E. J. Millington in *The Æsthetic and Miscellaneous Works of Frederick von Schlegel*, 425–526. London: Bohn.]

Schleicher, August (1850 [1852]). *Linguistische Untersuchungen*, II: *Die Sprachen Europas in systematischer Uebersicht*. Bonn: König. [Translated by Hermann Ewerbeck as *Les langues de l'Europe moderne*. Paris: Ladrange.]

— (1863 [1983]). *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: Böhlau. [Translated by Alex Bikkers as *Darwinism Tested by the Science of Language*. London: Hotten (1869). Reprinted in Koerner (ed.) 1983, 1–73.]

— (1865 [1983]). *Über die Bedeutung der Sprache für die Naturgeschichte des Menschen.* Weimar: Böhlau. [Translated by Peter Maher in Koerner (ed.) 1983, 75–82.]

Schuchardt, Hugo (1885 [1972]). *Über die Lautgesetze: Gegen die Junggrammatiker.* Berlin: Oppenheim. [Translated by Theo Vennemann & Terence Wilbur in Vennemann & Wilbur (eds.) 1972, 39–72. Original reprinted in Vennemann & Wilbur (eds.) 1972, 1–37.]

— (1909 [1980]). Die Lingua franca. *Zeitschrift für romanische Philologie* 33: 441–461. [Translated by Glenn Gilbert in Gilbert (ed.) 1980, 65–88.]

— (1914 [1980]). *Die Sprache der Saramakkaneger in Surinam.* (Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Afdeeling Letterkunde, Nieuwe Reeks, 14–6.) Amsterdam: Müller. [Introduction translated by Glenn Gilbert in Gilbert (ed.) 1980, 91–126.]

Schwegler, Armin & Kate Green (forthcoming). Palenquero (Creole Spanish). In John Holm & Peter Patrick (eds.), *Comparative Creole Syntax.* London: Battlebridge.

Seuren, Pieter (1998). *Western Linguistics: An Historical Introduction.* Oxford: Blackwell.

Seuren, Pieter & Herman Wekker (1986). Semantic transparency as a factor in creole genesis. In Muysken & Smith (eds.) 1986, 57–70.

Siegel, Jeff (1999). Transfer constraints and substrate influence in Melanesian Pidgin. *Journal of Pidgin and Creole Languages* 14: 1–44.

Singler, John (1996). Theories of creole genesis, sociohistorical considerations, and the evaluation of the evidence: The case of Haitian Creole and the Relexification Hypothesis. *Journal of Pidgin and Creole Languages* 11: 185–230.

Syea, Anand (1992). The short and long form of verbs in Mauritian Creole: Functionalism versus formalism. *Theoretical Linguistics* 18: 61–97.

— (1997). Copula, wh-trace, and the ECP in Mauritian Creole. *Linguistics* 35: 25–56.

Sylvain, Suzanne (1936). *Le créole haïtien: Morphologie et syntaxe.* Wetteren, Belgium: de Meester.

Szathmáry, Eörs, Ferenc Jordán, & Csaba Pál (2001). Can genes explain biological complexity? *Science* 292 [5520]: 1315–1316.

Thomason, Sarah (1980). Morphological instability, with and without language contact. In Jacek Fisiak (ed.), *Historical Morphology*, 359–372. The Hague: Mouton.

— (1997). Introduction. In Thomason (ed.) 1997, 1–7.

Thomason, Sarah (ed.) (1997). *Contact Languages: A Wider Perspective.* Amsterdam: Benjamins.

Thomason, Sarah & Terrence Kaufman (1988). *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

Traugott, Elizabeth Closs & Bernd Heine (eds.) (1991). *Approaches to Grammaticalization.* 2 volumes. Amsterdam: Benjamins.

Trudgill, Peter (1989). Contact and isolation in linguistic change. In Leiv Egil Breivik & Ernst Håkon Jahr (eds.), *Language Change: Contributions to the Study of its Causes*, 227–237. Berlin: Mouton de Gruyter.

Valdman, Albert (1978). *Le créole: Structure, statut et origine.* Paris: Klincksieck.

— (2001). Sur les processus lexicogénétiques du créole haïtien. In Hartmut Schröder, Petra Kumschlies, & María González (eds.), *Linguistik als Kulturwissenschaft: Festschrift für Bernd Spillner zum 60. Geburtstag*, 43–53. Frankfurt am Main: Lang.

van Kemenade, Ans (ed.) (1999). Functional properties of morphosyntactic change. *Linguistics* 37–6 (special issue).

van Kemenade, Ans & Nigel Vincent (eds.) (1997). *Parameters of Morphosyntactic Change.* Cambridge: Cambridge University Press.

Van Name, Addison (1870). Contributions to creole grammar. *Transactions of the American Philological Association* 1: 123–167.

Veenstra, Tonjes (1996). *Serial Verbs in Saramaccan: Predication and Creole Genesis.* (HIL Dissertations, 17.) The Hague: Holland Academic Graphics.

Vennemann, Theo & Terence Wilbur (eds.) (1972). *Schuchardt, the Neogrammarians, and the Transformational Theory of Phonological Change*. Frankfurt am Main: Athenäum.

Vikner, Sten (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

Vinson, Julien (1889). Creole. In Louis Adolphe Bertillon et al. (eds.), *Dictionnaire des sciences anthropologiques*, 345–347. Paris: Doin.

Vrzić, Zvjezdana (1997). A minimalist approach to word order in Chinook Jargon and the theory of creole genesis. *MIT Working Papers in Linguistics* 31: 467–478.

Weinreich, Uriel (1953). *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York.

Whinnom, Keith (1971). Linguistic hybridization and the "special case" of pidgins and creoles. In Hymes (ed.) 1971, 91–115.

Wilbur, Terence (ed.) (1977). *The Lautgesetz-Controversy: A Documentation (1885–86)*. Amsterdam: Benjamins.

Winford, Don (1994). Sociolinguistic approaches to language use in the Anglophone Caribbean. In Morgan (ed.) 1994, 43–62.

Wurm, Stephen (1996). The Taimyr Peninsula Russian-based pidgin. In Jahr & Broch (eds.) 1996, 79–90.

Zink, Gaston (1986). *Phonétique historique du français: Manuel pratique*. Paris: Presses Universitaires de France.