The world's simplest grammars are creole grammars

JOHN H. MCWHORTER

Abstract

It is often stated that all languages are equal in terms of complexity. This paper introduces a metric of complexity, determined by degree of overt signalling of various phonetic, morphological, syntactic, and semantic distinctions beyond communicative necessity. By this metric, a subset of creole languages display less overall grammatical complexity than older languages, by virtue of the fact that they were born as pidgins, and thus stripped of almost all features unnecessary to communication, and since then have not existed as natural languages for a long enough time for diachronic drift to create the weight of "ornament" that encrusts older languages. It is demonstrated that this complexity differential remains robust even when creoles are compared with older languages lacking inflection, contra claims by theoretical syntacticians that the typology of creoles is largely a manifestation of parameter settings resulting from low inflection. The overall aim is to bolster a general paradigm arguing that creole languages are delineable synchronically as well as sociohistorically.

Keywords: *analytic–synthetic, complexity, creole, derivation, diachrony, gender, grammaticalization, inflection, Lahu, Maori, markedness, noun class, phoneme inventory, pidgin, Saramaccan, semantic transparency, tone, Tsez*

1. Introduction

It has long been a truism in creole studies that creoles are formally distinguishable from other languages only on the basis of their sociohistory, and that there is no logically possible synchronic distinction between creole grammars and older grammars. Elsewhere (McWhorter 1998, 2000) I have argued that a large subset of creole languages indeed display a particular confluence of traits, all predictable from the history of creoles in pidginization, which are unknown

Linguistic Typology 5 (2001), 125–166 1430–0532/2001/005-0125

to occur in combination in any older language grammar and thus constitute a synchronically identifiable Creole Prototype.[1]

Subsequently (McWhorter 2001) I presented further results of pidgin ancestry in creole grammars which reinforce my thesis that creoles, although natural languages, are qualitatively distinguishable from older grammars as a predictable result of their youth. This paper demonstrated that creole 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. Examples included, among others, evidential marking, ergativity, inalienable possessive marking, and inherent reflexive marking. The general conclusion was that in 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 Universal Grammar, 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.

This research program dovetails with Bickerton's Language Bioprogram Hypothesis (1981 and other works), which also proposes that creoles represent an underlying "layer" of language resulting from their roots in pidgins. Bickerton's main focus, however, was on the implications of his hypothesis for generative syntactic and acquisition theory, and as such he had little occasion to examine creoles from a crosslinguistic or typological perspective. My intention is 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".

This paper, a fourth and final installment in my preliminary exploration of the Creole Prototype theme, will continue in the vein of McWhorter (2001) in 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.

1. The traits absent in the prototype creole are inflectional affixation, tone distinguishing monosyllabic lexical items or encoding morphosyntactic distinctions, and opaque lexicalization of derivation–root combinations.

2. Are all natural grammars equally complex?

Just as it is a truism in creole studies that "creole" is strictly a sociohistorical term, it is a truism in linguistics in general that all languages are equally complex (e.g., Edwards 1994: 90, Bickerton 1995: 67, O'Grady et al. 1997: 6). The claim is generally made in reference to varying conceptions of the meaning of complexity, all broader than the constrained definition to be outlined later in this paper. However, there is a strong implication underlying such statements that anything "simple" in a given language will be "compensated for" by a "complex" feature, a typical example being Crystal's (1987: 6) provisional observation that "[a]ll languages have a complex grammar: there may be relative simplicity in one respect (e.g., no word-endings), but there seems always to be relative complexity in another (e.g., word-position)".

This must be designated a truism because, like the creole studies truism, it has long been asserted without having been subjected to systematic verification.

2.1. *Complexity and subsets of grammar*

Our first indication that the question is a richer one than generally assumed begins with something all linguists presumably agree upon: that one language can be more complex than another in terms of a particular area of grammar. For example, Kikongo distinguishes between four kinds of past tense including a completive (Welmers 1973: 350) while Japanese has only one overt marker of past tense, and has no grammaticalized indicator of completiveness exclusively. Thus, a single Japanese expression (2) corresponds to the four Kikongo sentences in (1).

(1) a. *nsuumbidingí nkóombo.*
'I bought a goat (today).'
b. *yásuumbidi nkóombo.*
'I bought a goat (yesterday)'
c. *yasáumba nkóombo.*
'I bought a goat (earlier).'
d. *nsuumbidi nkóombo.*
'I have bought a goat.'

(2) *yagi* *o* *katta.*
goat ACC bought
'I (have) bought a goat.'

It would certainly be mistaken to characterize the Japanese expression of past tense as "primordial" or unnuanced itself. For example, Japanese uses the past to describe events which have just come into perception where European languages (and probably most other languages) use the present: upon seeing a bus

coming into view, the Japanese person says *Basu ga kimashita* 'The bus came' rather than *Basu ga kimasu* 'The bus is coming'. It would also be difficult to overemphasize that complexity is a difficult notion, and we will shortly outline a principled characterization of complexity upon which the heart of the paper's argument will be founded. Nevertheless, under the assumption that despite its difficulties, complexity is not a concept of no epistemological validity whatsoever, we can most likely agree that Kikongo, in happening to have evolved as fine-grained an overt subdivision of pastness as in (1), has a more complex past-marking system than Japanese.[2]

Similarly, with verbs of motion, Russian and most other Slavic languages distinguish not only between imperfective and perfective as they do with almost all verbs, but make a further subdivision within the imperfective of what grammarians have called "determinate" versus "indeterminate". The indeterminate is expressed with a separate verb root. In the case of 'go', a separate pair of verb roots is used for 'going somewhere in a vehicle (or on a horse)', such that where English uses a single verb *go*, Russian uses no fewer than four (viz., *xodit'*, *ezdit'*, *idti*, *exat'*):

(3) Indeterminate imperfective
   a. *ja xožu v kino.*
      'I go to the cinema (often).'
   b. *ja ezžu v kino.*
      'I go to the cinema (often) (in a car).'

(4) Determinate imperfective
   a. *ja idu v kino.*
      'I am going to the cinema (now).'
   b. *ja edu v kino.*
      'I am going to the cinema (now) (in a car).'

(5) Perfective
   a. *pojdëm v kino.*
      'Let's go to the cinema.'
   b. *poedem v kino.*
      'Let's go to the cinema (in a car).'

2. In light of the argument in McWhorter (2001), and *contra* arguments that creolization merely eschews source languages' inflection and shaves away exceptional "wrinkles" in their grammars, it also bears mentioning that even in creoles in which Kikongo is claimed to have played a significant substratal role, such as Sranan, Saramaccan, and Ndjuka, there is no such fine-grained overt subdivision of pastness. This is particularly significant in the case of Palenquero Creole Spanish, in which Kikongo was essentially the only significant substratal influence (Schwegler 1998).

Again, it is obvious that in the particular case of subdividing the semantic space of movement, Slavic languages are more complex than English and many other languages.

While I could be accused of belaboring the obvious in showing that languages can differ in terms of complexity in particular areas, it is often assumed that overall, languages "balance out" in terms of complexity (cf. Crystal 1987 cited above), such that if Slavic has its aspect-determined motion verb pairs, English has a subtle system of articles overtly marking determination and referentiality that bedevil Slavic learners, or that where Kikongo has its past markers, Japanese has its paradigms of numeral classifiers, etc. Yet in the strict sense, there is no a priori reason to assume that all languages ultimately "tally" in terms of areas of typologically unusual complexity.

2.2. *Complexity and contingency*

On the contrary, there is a suggestion speaking specifically against such an assumption; namely, the very source of much of a grammar's complexity. There are many features commonly found in grammars which are the product of a gradual evolution of a sort which proceeded quite independently of communicative necessity, and must be adjudged happenstance accretion.

A particularly useful example is affixes marking grammatical gender. Obviously, grammatical gender marking is unnecessary to human communication, given how very many human languages lack it. Grammatical gender affixes, beyond the extent to which they distinguish natural (biological) gender, do not mark any real-world entity or category or serve any communicative need (see Trudgill 1999 for a useful discussion of this point). To be sure, the free classifier morphemes (Greenberg 1978) or former pronouns that gender markers usually arise from did mark real-world categories. However, as affixes, they no longer do, or at best correspond to perceivable categories only very approximately, their fundamental superfluity to communication being illuminated further by the fact that they sometimes even arise from happenstance phonetic correspondences (Nichols 1992: 141–142). The result, in requiring the classification of nouns according to two and often more classes (and as often as not requiring the control of associated morphophonemic rules and exceptions), gender marking is inherently more complex than its absence. Crucially, this added complexity emerges via chance, not necessity.

Grammatical gender affixes are so starkly devoid of semantic substance or syntactic function that their fundamentally happenstance essence is particularly clear, and they are useful in illuminating the similarly contingent nature of other features often found in grammars. The four Kikongo past markers, in finely subdividing the semantic space of pastness, clearly serve a function within the grammar, as do Slavic motion verbs in theirs. Yet like gender af-

fixes, both of these features emerged from grammars which lacked this particular complexity: Kikongo's Bantu relative LoNkundo has just two past markers (Welmers 1973: 348) as do many other Bantu languages; the suppletively distinct verbs of motion of Slavic languages have no analogue in their close Indo-European relatives in the Germanic branch. As such, despite their functional usefulness once having evolved, the emergence of these features was at root a chance elaboration.

Lass (1997: 367–368) is particularly illustrative on this underacknowledged but crucial point about human grammars:

> Not only does a language have by definition a sufficient set of structures and categories available for doing anything that a speaker "needs" to do; it will also have a lot of material that simply makes speakers do things, whether or not there is any (functional, discourse, pragmatic) "need" to do them. (English forces a speaker to mark durative aspect every time he utters a sentence in the present tense; German, Afrikaans, French and Swedish don't, but they have machinery for doing it if necessary.) We live perpetually with "decisions" of past generations. Somebody, somewhere (as it were) decided in the eighteenth century or thereabouts that the expression of progressive aspect should be obligatory in English, and as an English speaker I'm simply stuck with it.

If we acknowledge the contingent nature of such developments, then a question arises: Why, precisely, would such chance developments have occurred to an equal extent in all 6,000 or so natural languages on earth? If Kikongo has four past tenses and LoNkundo has just two, then according to the conventional wisdom that all languages are equally complex, we are forced to assume that LoNkundo not just may, but must have developed some quirky degree of complexity in some other area of grammar where Kikongo has honed to a more crosslinguistically unremarkable pattern (perhaps evolving two more noun classes, or a paradigm of evidential markers). However, what sustained linguistic investigation has been devoted to ascertaining that this is true in LoNkundo, or indeed among any conceivable subset of the world's languages? And more to the point, precisely what mechanism would we hypothesize as the source of this purported equalizer of complexity among the world's languages?

Surely the mechanism could not entail that grammars somehow calibrate themselves according to comparison with one another, which thus requires that this equalization would respond to some imperative of the internalized language competence (or I-language, in theoretical syntactic terms). Yet this would require in turn that the evolution of needless linguistic complexity was somehow advantageous to the evolution of our species, or, to wit, the passing on of genetic material. The growing current of linguistic thought comparing language change to natural selection might at this point suggest an analogy to cases such as the male peacock's tail, heavy and awkward to an extent

which hinders mobility. Yet the disadvantages of this burden are outweighed by the advantage conferred by its status as a sexual advertisement; meanwhile, however, it is unclear how elaborated verbs of motion or fine shades of pastness could have acquired men or women extra sexual partners. Furthermore, it would be extremely problematic to conceptualize a model of innate grammar which could specifically generate all of the innumerable possible random complexifications of particular areas of natural language grammar while also suppressing most of them in any one grammar, etc.

If we accept that Kikongo's past-marker paradigm and Slavic's unusually elaborated verbs of motion are traceable ultimately to chance elaboration, then we must also accept that we can conceive logically of neither a genetically selected neural linguistic blueprint, nor a cognitively or functionally determined mechanism, which would ensure that chance occurrences operate to a precisely equal degree upon every natural language grammar on earth. In other words, nothing in modern linguistic theory entails that all natural languages be equally complex, and as such, any assertion or even underlying assumption to this effect qualifies as a truism yet to be tested rather than as an established fact.

2.3. *Complexity and creoles*

There is, however, one conception under which it could be coherently proposed that all natural language grammars are equally complex. One might stipulate that after countless millennia of usage and drift, we might expect that by a certain point all grammars had, by the sheer dictates of chance, developed various random complexities in parts of their grammars. This might follow from the mounting evidence of the inherent tendency of natural systems to complexify with the passage of time according to apparently universal principles of self-organization (cf. Lightfoot 1999: 250). We might propose that the volume of such excrescence in each grammar eventually reached the limit of human propensity to process it (e.g., there are no grammars with a different suppletive form of every basic verb for every person/number combination, etc.). Under this scenario all natural languages would be equally complex by virtue of having all come to rest at a certain "surplus complexity quotient".

One crucial aspect of this scenario would be that all of the grammars in question trace back tens of millennia. This is unproblematic regarding all of the world's natural languages with one exception, which brings us back to the focus of this paper: creole languages. Creole languages have, by definition, existed only for several centuries at the most. The oldest known creoles today are the Portuguese-based creoles of Cape Verde and Guinea-Bissau, which trace back to the late fifteenth century; the Caribbean creoles mostly date to the late seventeenth and early eighteenth centuries; the English creoles of the Pacific trace to late eighteenth century interactions at the very earliest; Hawaiian

Creole English and the leading creoles and expanded pidgins based on African languages emerged in the decades surrounding the turn of the twentieth century.

This leads us to a hypothesis. Let us assume for these purposes that tens 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. Specifically, creoles are the world's only instantiation of spoken language having been "born again", when speakers expanded pidgins – universally agreed to be rudimentary codes not fulfilling the needs of full language – into natural language grammars.[3]

It is perhaps possible to assume at first glance that even a language "born again" would, by virtue of being generated by the same innate linguistic competence as all other languages, be immediately indistinguishable from older ones in any qualititative aspect.[4] This, however, presupposes that the structures older grammars are completely, or even mostly, specified by, as opposed to merely compatible with, Universal Grammar, and as I have argued, this is not the case. Creoles, in being recently borne of communication vehicles deliberately designed to eschew all but the functionally central (pidgins), are unique examples of natural languages with much less contingent accumulation of "ornamental" elaboration than older grammars drag along with them. As I will argue later, it would be empirically inaccurate to claim that creoles represent anything approaching the "ground zero" of human language, i.e., a perfect or "optimal" matching of irreducible semantic atoms with the simplest conceivable morphosyntactic realizations (cf. Kihm 2000 in agreement). Nevertheless, on the other hand, Kihm's (1984) early claim that "Creoles can display any de-

3. Some creolists argue that creoles are born via the gradual "streamlining" of a lexifier language via succeeding waves of second-language acquisition, with a general implication that the end result is more a "vehicularized" version of the lexifier than a new language entirely, and with a terminological and ontological dissociation of creoles from pidgins. I have argued comprehensively against this frame of reference in McWhorter (1998); cf. also McWhorter & Parkvall (forthcoming).
   For the purposes of clarity, I am oversimplifying here in a crucial respect: creolization is a cline phenomenon, and many natural languages were born via a language contact process that resulted in simplification of a degree less radical than pidginization: e.g., Afrikaans, Popular Brazilian Portuguese, Réunionnais French Creole, the English Caribbean mesolects, etc. Furthermore, there are languages in the world in which widespread second-language acquisition has simplified their grammars to a relatively minor but perceptible extent in comparison to their close relatives, e.g., Swahili, Cree.
4. This assumption is, for example, the fundamental one underlying DeGraff (1999) and related writings.

gree of complexity compatible with UG" does not correspond with the data set that the world's creoles actually constitute: on the contrary, because so much of a grammar's complexity results from the operations of random accretion over time, creoles display less complexity than the rest of the world's natural grammars.[5]

2.4. *Measuring complexity*

There is no conventionally agreed-upon metric for measuring complexity in grammars. This is partly because of the reign of the truism that all languages are equally complex, partly because the construction of a comprehensive diagnostic for precisely ranking any human language upon a scale of complexity is a daunting task, and perhaps most of all because the construction of such a diagnostic would be of little relevance or usefulness within most frames of reference in modern linguistic inquiry. Yet the issue of relative complexity is an important one in investigating the nature of creole languages and creolization.

2.4.1. *Complexity: Difficult but not epistemologically vacuous.* Complexity is certainly an ambiguous and malleable concept. For example, which plural marking strategy is more "complex": English, which marks plural only on the noun but with a marker that has three allomorphic variants; Swahili, which marks plurality redundantly on adjectives and nouns with a marker that varies according to several noun classes (*vitabu vizuri* 'beautiful books', *miti mizuri* 'beautiful trees'); or French, which (allowing for simplification of the full range of the facts) marks plural via a plural allomorph of the determiner and redundantly on adjective and noun via a proclitic which is, however, expressed only when the root is vowel-initial (*les arbres verts* [le zarbʁ vɛʁ] 'the green trees')? Rankings of complexity here will obviously differ according to acquisitional or structural framework, and whether one approaches the data from a first-language as opposed to second-language perspective; it is unlikely that any single ranking could be constructed that would stand as incontestable.

Yet areas of inherent ambiguity of this kind exist and do not, in any logical sense, render the entire notion of linguistic complexity a vacuous one. Indeed, the truism that languages cannot be compared in terms of complexity conflicts with the fact that linguists readily rank grammars in terms of complexity in terms of phonology and, to a lesser but considerable extent, morphology (e.g., Nichols 1992: 64–69). For example, in their work on phonology and morphology in isolated communities, scholars such as Andersen (1988) and Trudgill

5. Kihm (1984) wrote this more as a passing speculation than as a central assertion, and Kihm (2000) argues provisionally that creoles are indeed closer to Universal Grammar than other languages. The present paper is an attempt to explore and support that provocative idea.

(1989, 1996) have designated dialects in these communities as having more complex phonologies, allophonies, allomorphies, and derivational pathways from the phonemic to the phonetic than those of varieties spoken in a wider context. These authors have made these observations with no criticism of their assumption that complexity is gradient. I am unaware of any precise arguments as to why syntax and semantics would be, in contrast, inherently unamenable to complexity rankings. It is much less likely that any such unamenability exists than that the issue, its difficulties acknowledged, simply has not had occasion to receive much attention from modern linguists (cf. Comrie 1992).

In this light, there are ample cases in which complexity differentials are stark enough that most or all linguists' judgments could reasonably be assumed to concur regardless of frame of reference. It is complexity differentials of this degree which I will be concerned with in this paper, and from which my hypothesis will derive its support (and not subtler issues such as whether redundant marking is more or less complex than marking once, etc., which require paper-length analyses in themselves).

2.4.2. *Aims of this presentation.* Because linguistic investigation so rarely occasions sustained grappling with the complexity issue, there are various misconceptions that my treatment may foster in the reader.

First, my aim in this paper is not to construct a metric for ranking the world's languages on a scale of complexity, a goal whose ultimate intellectual benefit would be unclear in any case. My aim is a more constrained one: to provide a metric of complexity which will serve the purposes of elucidating and rendering falsifiable my specific claim that creole languages in general tend strongly to be less complex than older languages. 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.

Second, this metric does not stipulate that complexity is indexed with relative difficulty of production or processing. To wit, the model is not intended to imply that languages more complex according to its dictates are more difficult for the speaker to produce, nor that such languages are more difficult for the hearer to process; moreover, this metric takes as a given that all languages are acquired with ease by native learners. Our assumption is that human cognition is capable of processing great degrees of overspecification in language, and that possible differentials in ease of production or processing would be of import more to applied linguistic concerns than to academic and theoretical; as such, this issue will not be explored here. My 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.[6] (However, it bears mentioning that a highly elaborated grammar could be argued to be easier rather than harder to process, in making distinctions more clearly than a less elaborated grammar, and thus leaving less to context.)

Third, the metric as presented in this paper will not suggest what the most "optimal" language would be. This question, which will occasion different answers according to whether we are concerned with speaker or hearer, is obviously a rich one, which echoes throughout a great deal of phonological, syntactic, and acquisitional work. Obviously, however, the topic is much too large to be addressed gracefully in this paper, and would be tangential to its core argument.

2.4.3. *A metric of complexity.* With the aims of my project thus delineated, I propose the following four diagnostics of grammatical complexity, chosen to arouse the least possible controversy from as wide a spectrum as possible of linguists. 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 (this is not meant to subsume redundancy, which involves manipulation of the same form, not contrasting forms).

First, a phonemic inventory is more complex to the extent that it has more marked members. Markedness is here intended strictly in reference to crosslinguistic distribution: marked phonemes are those encountered less frequently in the world's languages than others conventionally deemed unmarked; e.g., ejectives, clicks, and labialized consonants vs. stops, rounded back vowels, and glides.

The motivation for treating such sounds as metrics of complexity is not a claim that these are more articulatorily complex in the sense explored by Trubetzkoy (1931); determining the precise measurement of articulatory complexity has proven too controversial to serve adequately as a foundation for the project in this paper. Rather, the motivation for using these sounds as a metric is implicational, in the Greenbergian (1966a, 1966b) sense: an inventory with a great many marked sounds (e.g., a click language) is more complex than one with all or almost all unmarked sounds (e.g., a Polynesian language) because the former type of inventory has marked members in addition to unmarked ones – the marked sounds IMPLY the concurrent existence of unmarked

6. Whether or not a language deemed more complex by our metric would be harder to acquire as a second language is a question of limited use, given that no matter how complex a language is, speakers of a closely related one will find it easy to acquire (Polish is not hard for Russians, Ojibwe is not hard for a Cree, etc.).

ones (there exist no phonemic inventories with only crosslinguistically marked phonemes). A larger phonemic inventory requires maintaining finer intersegmental distinctions within the bounds of the human vocal apparatus. There is obviously no claim that marked phonemic inventories are more difficult to acquire natively; our reference is solely to the number of distinctions entailed by such inventories in comparison to others.[7]

Similarly, a tonal system is more complex than another one when it has more tones, because this phonology requires mastery and processing of a larger set of contrasts and requires maintaining finer intertonal distinctions within the bounds of the human vocal apparatus.

Second, 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 two kinds of alignment rather than one (i.e., ergative/absolutive and nominative/accusative) (cf. Henry 1995, Henry & Tangney 1999).

Third, a grammar is more complex than another to the extent that it gives overt and grammaticalized expression to more fine-grained semantic and/or pragmatic distinctions than another. For example, the Muskogean Native American language Koasati uses different existential verbs for five types of object depending on their shape (with each verb varying suppletively for plurality):

(6) a. *nofó-k* *mat-haccáːl.*
birch-SUBJ afar-stand
'There's a birch over there.' (Kimball 1991: 453)
b. *óːla-k* *talláːk.*
town-SUBJ lie
'There's a town.' (Kimball 1991: 458)

In subdividing the semantic space of existentiality more finely than English does, Koasati marks existentiality in a more complex fashion than English does; see (7).

(7) *kom-mikkó-k* *cikkíːli-t* *kom-tohnóːci-fóːk-on* *mán*
our-chief-SUBJ keep-CONN to.us-send.off-when-SR/FOC also
toknáːwa *st-im-ako-yókpa-k* *im-máːya-k.*
money INST-3SOBJ-1PLSTAT-love-SS to.it-be.more-SS
'When our chief took care of it and sent off for things for us, we then grew to love money more.' (Kimball 1991: 491)

7. I leave unexplored the possibility that a relatively compact phonemic inventory may entail so many ALLOPHONIC distinctions overall as to encompass a larger number of possible segments than a relatively large phonemic inventory where overall allophonic variation is relatively minimal. However, by our metric, the former language's sound system would be more complex than the latter's.

In requiring overt specification of reference tracking (e.g., *-on*, *-k*) where English requires none, and in having particular pronominal affixes used only with stative verbs (e.g., *-ako-*) where English has no separate set of pronouns used according to verbal Aktionsart, according to our metric Koasati has a more complex grammar (in these areas) than English. In requiring the storage of a class of verbs with semantically opaque instrumental government (in this case, 'to love'), Koasati is akin to Germanic and Romance which also have rather arbitrary cases of verb government (German's dative-marked *Ich glaube ihm* 'I believe him', French's accusative-marked *je l'écoute* versus English's dative-marked *I'm listening to him*), but more complex than Japanese, which does not require storing a subset of verbs with SEMANTICALLY ARBITRARY case government specifications.

Fourth, inflectional morphology renders a grammar more complex than another one in most cases.

This last point requires discussion. The reader may justifiably object that there is no a priori reason to assume that inflectional encoding is inherently more complex than encoding the same feature with a free morpheme. This is, in itself, true. 'I wanted' is *ni-li-taka* (I-PAST-want) in Swahili but *wǒ yào le* (I want PERF) in Mandarin Chinese. There are no empirical grounds for assuming that Swahili is more complex in encoding 'I' and pastness with morphemes more phonetically incorporated with the head of the phrase than those in Mandarin Chinese are.

However, inflection more often than not has wider repercussions in a grammar than this, which are complexifying factors in terms of exerting a load upon processing.

Morphophonemics and suppletion. While some languages are as neatly agglutinative as Swahili, as often as not inflection leads to the development of morphophonological processes, which constitute an added component of a grammar to be learned. It can be argued that these processes have perceptible phonetic motivations and as such can be seen as a relatively minor addition to the processing component. Yet morphophonology also often leads to less phonetically predictable processes, such as Celtic consonant mutation (Welsh *ei gath* 'his cat' versus *ei chath* 'her cat') whose rules fall outside of the realm of phonology, or Germanic umlaut (Proto-Germanic [*foːt] 'foot' / [foːt-i] 'feet', with the plural form > [føːt-i] > English [føːt] > [fiːt]); in German, umlaut rules also occur variably in certain contexts, such as before the plural morpheme *-e* (*Kuh*/*Kühe* 'cow/cows' but *Schuh*/*Schuhe* 'shoe/shoes', etc.).

Meanwhile, suppletion also complexifies an area of grammar according to our metric. The various suppletive forms of *be* in English (*am*, *are*, *is*, *was*, *were*, *been*, and similarly in most Indo-European languages) render these languages more complex in this area than languages where the copula is invariable

across person and number in the present (i.e., Modern Swahili's *ni*) or is conjugated relatively predictably in the present and the past (i.e., Finnish *olla*).

Declensional and arbitrary allomorphy. Inflection also complexifies a grammar when encoding distinctions between noun classes or verb classes. English is not less complex than Latin in having *to the boy* rather than *puero*. However, the English expression of case is simpler overall than Latin's because Latin nouns could belong to any of five classes, each differing significantly in their case paradigms (on top of inflecting for singular and plural).

Similar are cases where a given alternation is expressed via a number of inflectional strategies, each of which must be learned and stored with the root. Russian imperfective *pisat'* 'to write' is rendered perfective via the addition of a prefix (*napisat'*), while *pit'* 'to drink' is rendered perfective via the addition of a different prefix for no synchronically perceptible reason (*vypit'*); meanwhile, imperfective *načinat'* 'to begin' is rendered perfective via the subtraction of a suffix (*načat'*) while the imperfective *klast'* 'to put' can only be rendered perfective via a different root altogether, *položit'*.

Agreement: Useful versus useless. As we have seen, a grammar with agreement-marking inflection is also inherently more complex than a grammar like English in that, unlike in the case of Swahili *nilitaka* versus Chinese *wǒ yào le*, agreement marking is not a different way of expressing a category marked in another fashion in English, but instead, marks a category not expressed in any fashion in English.

Thus my claim that inflection almost always complexifies a grammar is based not on an Anglophone's romanticization of Latin declensional paradigms, but upon (i) the effects that inflection typically has upon a grammar over time, and (ii) the fact that some inflection, such as gender marking and declensional noun classes, does not correspond to concepts expressed by all grammars, but is instead purely supplementary to a grammar's machinery.

3. Comparing complexity: An older grammar and a creole grammar

Based on the above definitions of complexity, we will now compare the grammar of the Nakh-Daghestanian (Northeast Caucasian) language Tsez (data from Comrie et al. 2000 unless otherwise noted) and the creole language Saramaccan, which emerged in the late seventeenth century and is thus approximately three centuries old (data collected by the author unless otherwise cited).

3.1. *Tsez versus Saramaccan*

3.1.1. *Phoneme system.* The Tsez phonemic inventory of some 42 segments includes uvulars, pharyngeals, and a separate series of pharygealized uvulars;

most stops and affricates have phonemic ejective alternants. Most consonants have labialized phonemic alternants. A long *a* is phonemic.

The only marked sounds in the Saramaccan phonemic inventory of 25 are three prenasalized stops and two prevelarized ones, entering into phonemic contrasts only occasionally. There are no uvulars, ejectives, or labialized consonants. There are some minimal pair distinctions for tense and lax mid vowels ([e]/[ɛ], [o]/[ɔ]).

3.1.2. *Noun class agreement.* Tsez has four noun classes, marked by agreement prefixes on most vowel-initial adjectives and verbs as well as elsewhere in the grammar (although vowel-initial adjectives and verbs are a minority in the lexicon). The classes are only about as semantically predictable as those determining gender in German. Tsez verbs fall into four classes determined by the final segment of the stem. Agreement is marked on verbs as well as nouns, but with the stipulation that it is marked only on (most) vowel-initial stems. Matrix verbs are normally marked for agreement with a complement clause, with the gender marker for "other", i.e., class 4 (8a), while focus upon an absolutive nominal inside the complement clause is marked via agreement with that nominal (8b). (Data are from Polinsky & Potsdam 2001: 584.)

(8) a. *eni-r* *už-ā* *magalu* *b-ā̄c-ru-λi*
mother-DAT [boy-ERG bread.CL3.ABS CL3-eat-PP-NOMZ]
r-iy-xo.
CL4-know-PRES
'The mother knows that the boy ate the bread.'

b. *eni-r* *už-ā* *magalu* *b-ā̄c-ru-λi*
mother-DAT [boy-ERG bread.CL3.abs CL3-eat-PP-NOMZ]
b-iy-xo.
CL3-know-PRES
'The mother knows that the boy ate the BREAD.'

Saramaccan has neither inflectional morphology nor free equivalents such as noun classifiers.

3.1.3. *Alternate stems.* A large number of Tsez nouns occur in an alternate stem form before certain inflectional suffixes. The changes are partially determined morphophonologically, but nouns differ as to exactly which suffixes the alternate occurs before, and some occur in different alternate forms for different suffixes. The generation of the alternate stem from the basic one is not a regular process, but instead follows various rules of thumb about as irregular as those determining plural marking in German: *giri* 'pole', *girimo-s* 'of the pole'

vs. *mec* 'tongue', *mecr-ebi* 'tongues', illustrating two segments (*-mo-*, *-r-*) used frequently to form alternate stems.

Saramaccan noun roots do not change according to case.

3.1.4. *Derivational morphology.* Tsez has a large number of derivational affixes for deriving nouns, verbs, and adjectives, including markers of evaluative names ('coward'), both denominal and deverbal abstract nominalizers, a marker of enveloping objects, a marker designating residents of a place, a marker meaning 'possessing X' (*kot'u* 'beard', *kot'u-r-yo* 'bearded man'), two denominal adjectival markers with an alienable/inalienable semantic distinction, a caritive suffix, and suffixes to derive transitive and intransitive verbs (from adjectives and adverbs) (14–15).

Saramaccan has two derivational suffixes; reduplication derives attributive adjectives and resultatives from transitive verbs (*mi lái dí gbóto* 'I loaded the boat', *dí láilái gbóto* 'the loaded boat', *dí gbóto dɛ́ láilái* 'the boat is loaded'). Reduplication also nonproductively generates nouns from verbs (*síbi* 'to sweep', *sísíbi* 'broom') (Bakker 1987: 21). One might reasonably surmise that Saramaccan might have analytic derivational strategies for all of the functions covered by the Tsez markers, in which case there would be no grounds for designating Tsez derivation more "complex". Saramaccan indeed marks, for example, causativity with a serial verb construction. However, in most cases – such as all of those listed in the previous paragraph – Saramaccan has no overt derivational strategy corresponding to the Tsez ones.

3.1.5. *Word order in questions.* Word order behavior of wh-words in Tsez depends on the adjunct/argument distinction:

(9) a. *neti obiy kidir-ā-ɣor Ø-ik'i-x?*
when father.ABS Kidero-in-to AGR-go-PRES
'When is father leaving for Kidero?'

b. *kidb-a šebi t'et'erxo?*
girl-ERG what.ABS read-PRES
'What is the girl reading?' (Comrie et al. 2000: 22)

Wh-movement in Saramaccan is straightforward: wh-words generated after the verb are regularly fronted.

3.1.6. *Evidentiality.* Tsez has an evidential distinction in past tense marking. Saramaccan has no grammaticalized evidential markers.

3.1.7. *Suppletion.* Tsez has some suppletive plurals and some suppletive transitive/intransitive pairs.

There is but one instance of suppletion in Saramaccan: before the verb *gó* 'to go', the imperfective marker *tá* is expressed as *nan*: *mi tá wáka* 'I am walking', but *mi nangó* 'I am going'. In entailing more suppletive plurals and lexically distinct transitive/intransitive pairs than Saramaccan, Tsez marks certain semantic distinctions overtly where Saramaccan does not.

3.2. *Implications of the comparison*

On the basis of the complexity evaluation guidelines proposed in Section 2.4, the grammar of Tsez is indisputably a more complex one than that of Saramaccan. The reader may justifiably wonder whether I have "stacked the deck" by highlighting aspects in which Tsez happens to be more complex than Saramaccan. In this light, for one, there is certainly no intention to imply that Saramaccan does not have its share of morphophonological processes: the 3rd person pronoun causes the assimilation of a preceding [a] such that *tapáɛn* 'cover it' is phonetically [tapɛ̃ɛ̃]); an epenthetic [m] surfaces morphophonologically between a few verbs and a following 3rd person singular pronoun: *fṍ* 'to beat', *a fṍ ɛn* ([a fomɛ̃ɛ̃]) 'he beat him'; there are some portmanteau morphemes in Saramaccan created by morphophonemic processes: e.g., before a pronoun beginning with a vowel, the vowel in *fu* becomes homorganic with the following one, and over the history of the creole the preposition has fused with the following word: *fu i* > *fí* 'your' (lit., 'for you'), *fu ɛ̃* > *fɛ̃ɛ́* 'his, her, its' (lit., 'for him/her/it'), etc. In addition, there are indeed some features in Saramaccan which would qualify as elaborated quirks. Saramaccan has overt and categorical marking of determination (via articles) whereas Tsez marks definiteness only on nouns with an attributive adjective (Bernard Comrie, p.c.). Over its three hundred years in existence, Saramaccan has evolved a new negator, such that today one (*á*) is used to negate verbal predicates and the other (*ná*) is used elsewhere (*Dí wómi á tá wáka* 'the man isn't walking', *Dí wómi ná mí tatá* 'the man is not my father') (cf. McWhorter 1996 on this development); there are a handful of lexical distinctions encoded only by tone: *dá* 'to give', *da* 'to be', etc (cf. McWhorter 1997: 87–121 on the diachrony). Features of this kind multiply in creoles as time passes, since they are natural languages subject to the same forces as all others (cf. McWhorter 2001).

But the crucial point is that creoles 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. Saramaccan features such as the negator allomorph and the epenthetic [m] are scattered ones of highly local import, and the fact remains that with the single exception of determination marking, according to what is currently known of both Tsez and Saramaccan, the grammar of Tsez is more complex (according to our metric) than that of Saramaccan in all areas. For example, morphophonological processes are richer and deeper in Tsez than

in Saramaccan. Comrie et al. (2000) describe four main rules, which are applicable across the grammar or to general word classes, and allude to other "more sporadic" ones; in contrast, all morphophonemic rules documented to date in Saramaccan are relatively "sporadic", applying to the single preposition *fu*, the pronoun *ɛ̃*, or to a subset of the pronouns, etc.

Furthermore, we must note that gender marking is not the only function served by inflection in Tsez which is not marked at all in Saramaccan. For example, with experiencer verbs, Tsez uses a lative marker:

(10) *aɬo-r* *meši* *b-esu-s.*
shepherd-LAT calf.ABS AGR-find-EVID
'The shepherd found the calf.' (Comrie et al. 2000: 19)

In contrast, Saramaccan does not mark experiencer verbs with any overt marker or strategy on a grammaticalized basis. Importantly, experiencer subject marking is not unique to inflected languages. It is common in Mande languages (Reh & Simon 1998) as well as Polynesian, as demonstrated in Maori where a group of experiencer verbs take the preposition *ki*:

(11) *ka* *piirangi* *ia* *ki* *ngaa* *mea* *katoa.*
TS want he PREP PL thing all
'He wants all the things.' (Bauer 1993: 270)

Thus the absence of overt delineation of experiencer verbs in Saramaccan is not a mere epiphenomenon of its analytic structure.

The facts are similar with evidential marking, which, like experiencer verb marking, is indeed compatible with analytic structure, as we see in the Tibeto-Burman language Lahu:

(12) *yɔ̂* *šɔ́-pɔ̄* *gà* *la* *tù* *cê.*
he tomorrow come to FUT HEARSAY
'They say that he'll arrive tomorrow.' (Matisoff 1973: 469)

Finally, while Tsez overtly expresses many concepts that Saramaccan does not, the converse does not apply: Saramaccan does not require its speakers to attend to any but the very occasional concept that Tsez does not (e.g., it marks definiteness more widely), lacking, for instance, obviative marking, inverse marking, tracking of referentiality as in languages of the Philippines, switch-reference marking, consonant mutations, clitic movement, subjunctive marking, dummy verbs, etc.

In other words, the fact that Saramaccan is hardly devoid of marked or complex features does not entail that a complexity gradient between grammars is an empirically invalid concept. With this acknowledged, it appears evident that Tsez is a more complex grammar than Saramaccan.

It must finally be observed that my observations about Saramaccan all apply as well, with minor modifications irrelevant to the basic thrust of the thesis, to the other creoles which I have designated as closest to the Creole Prototype in previous papers: Sranan, Ndjuka, Tok Pisin, Bislama, Solomon Islands Pijin, Torres Strait "Broken", Aboriginal Pidgin English, São Tomense Creole Portuguese, Principense Creole Portuguese, Annobonese Creole Portuguese, Angolar Creole Portuguese, Negerhollands Creole Dutch, Baba Malay, and Papia Kristang Creole Portuguese. The French plantation creoles (e.g., Haitian, Louisiana, Mauritian, Seychellois, Martiniquan, French Guyanais, et al.), due to contact over the centuries with French, have borrowed many French lexicalized derivation-root combinations and thus do not exemplify the Creole Prototype in the purest possible form; however, overall their grammars contrast with Tsez in the same ways as Saramaccan and the above creoles.

The Northwest Caucasian language Kabardian is even more fearsomely elaborated than Tsez (Colarusso 1992: xix):

> There are forty-eight consonants in most dialects [. . .] unusual contrasts and complex clusters abound. The phonological rules are complex and highly ordered. The result is that many Kabardian surface forms appear far removed from their underlying sources [. . .] the morphology of the language is highly complex [. . .] the verb can inflect for every noun in the sentence as well as for a range of subtle geometrical, aspectual, temporal and pragmatical features. The complex verbal use of several distinct temporal and adverbial positions, as well as the intermixing of personal indices with geometrical and pragmatical ones, strongly suggests that multiple layers of morphology and distinct morphological processes are at work. Such complexity is not restricted to the verb, but appears in word formation in general.

This sort of description is simply inconceivable of any creole language, and Colarusso refers to Kabardian and Caucasian languages in general as "extraordinarily complex by any linguistic standard" (1992: 2). Any thesis that all natural languages are equally complex automatically renders Colarusso's statement false and underinformed, when in fact it is unclear that a viable case that Mauritian Creole French or creolized Tok Pisin equal Kabardian in overall complexity would be possible.

4. Analytical grammars old and new

As seen above, much of the complexity difference between Tsez and Saramaccan is an outgrowth of the effects of inflection. Although we have also seen that there are also complexity differences independent of this factor, one might reasonably suppose at this point that the complexity difference between creoles and analytic languages would be less than that between them and inflected lan-

guages. Nevertheless, when we actually compare an older analytic grammar with a creole, we find a similar difference in complexity.

A comparison of this type is particularly urgent given a hypothesis in some current creolist work that the traits commonly found in creoles are predictable manifestations of a general analytic grammar, which implies that creoles are not synchronically delineable in any qualitative way from Sinitic, Mon-Khmer, Kwa, Polynesian, or other older analytic languages. DeGraff (1997, 1999), for example, argues that the differences between a creole grammar and that of its source language 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). DeGraff considers the association of creoles with pidgins unmotivated, and the exploration of a synchronic delineation of creoles an empirically unmotivated essentialization.

This provocative suggestion does not square with an actual comparison of older analytic grammars with creole grammars, in which case it becomes clear that creole genesis entailed a transformation of source language structures which far bypassed the relatively non-disruptive processes which German dialects underwent in becoming Yiddish. Namely, creoles are natural languages reborn from a radical reduction of their source languages into makeshift jargons. Here I will present a direct comparison of a well-described analytic grammar, the Tibeto-Burman language Lahu (data from Matisoff 1973; indicated page numbers from this source), with Saramaccan.

4.1. *Lahu and Saramaccan*

4.1.1. *Phoneme system.* The Lahu inventory of 33 segments includes palatal and post-velar stops as well as the typical bilabial, alveolar and velar, with both aspirated and unaspirated phonemic alternants in all five places of articulation; a voiced velar spirant; a central high unrounded vowel and schwa are also phonemic (1, passim).

As noted above, Saramaccan's phonemic inventory is smaller and less marked overall.

4.1.2. *Tones.* Lahu has seven lexically contrastive tones: *ca* 'look for', *cà* 'fierce', *câ* 'eat', *cā* 'feed', *cá* 'boil', *câʔ* 'string', *càʔ* 'machine' (22). Tone also often encodes causativity, as in *câ* 'eat', *cā* 'feed' (33). There are various other unsystematic tone-based distinctions between closely related words, such as *mû-* 'heaven, sky', *mu* 'be high, tall'; *phu* 'silver, money', *phû* 'price, cost' (29).

Saramaccan has a high-low tone contrast which is lexically contrastive in less than a dozen cases (*bigí* 'to begin', *bígi* 'big'; *á* 'not', *a* 'he, she, it'). (Among creoles, Saramaccan is unusual in displaying lexically or morphosyntactically contrastive tone even to this extent.) There are no semantic relationships between words encoded with tone, either in productive or fossilized fashion.

4.1.3. *Derivational morphology.* Lahu has about ten derivational markers occupying a continuum of productivity and semantic transparency. This includes agentive markers distinguished by sex (454–455); two others distinguished by sex to denote ownership or mastership, which can also be used to nominalize clauses (457); two causative markers (244–246); and a marker transforming verbs into corresponding nouns, e.g.: *u* 'to hatch', *ɔ̀-u* 'egg'; *cā* 'to sprout', *ɔ̀-cā* 'a sprout, a shoot' (68).[8]

As noted above, Saramaccan has only a few derivational markers; it has a single, gender-neutral agentive marker.

4.1.4. *Numeral classifiers.* Like many Southeast Asian languages, Lahu has several numeral classifiers, for people, animals, shapes, and more general purposes (89–92); in the absence of these the noun can be replicated itself as a classifier: *yɛ̀ tê yɛ̀* 'one house' (89).

Saramaccan has no classifiers of any kind.

4.1.5. *Grammatical relations.* Lahu has an accusative marker, *thàʔ* (155–158). In itself, marking patienthood is not necessarily more "complex" than leaving it unmarked. However, Lahu's marker is only very approximately characterized as an accusative marker, in fact occupying a wider and more idiosyncratic semantic space. For one, it is used with patients only to encode certain shades of emphasis. Furthermore, in many uses it marks not patients but other thematic role:

(13) a. *yɔ̂ qɔ̀ʔ la ve ha-pa qhà ve thàʔ leʔ*
he repeat come NOMZ month which NOMZ ACC INT
'In which month will he come back?' (157)

8. It is also worth noting that Lahu, unlike Saramaccan or any other prototype creole, has a goodly component of derivation whose meaning has drifted beyond compositionality. In some cases, an original meaning is vaguely perceptible: Matisoff (1973: 62) suggests that *-qɔ* "possibly involves the idea of enclosing, as a wrapper or receptacle": *yàʔ-qɔ* 'road', *mû-lòʔ=qɔ* 'noon, daytime', *làʔ-tɔ=qɔ* 'palm of the hand', *khi-tɔ=qɔ* 'sole of the foot', *mɔ̀ʔ-1ɔ* 'mouth'. Other cases have drifted beyond any recoverability, such as *-ni*: *cɔ̀ʔ-pɛ̄=ni* 'waist', *kī-ni* 'sweat', *pa-pa-qú-ti=ni* 'dragonfly', *fī-kôʔ=ni* 'orphan' (Matisoff 1973: 63).

b. *ɔ́-qā chi hê?-pā thà? vɨ̀ ā ve lâ?*
buffalo this Chinese ACC buy.from PERF NOMZ INT
'Did you buy this buffalo from a Chinese?' (158)

Saramaccan does not have overt markers of grammatical relations. It is not clear that overt expression of accusativity is more complex than not expressing it, but under the analysis of complexity adopted in this paper, the overt expression of a form which covers an arbitrary space partly occupied by transitivity and partly by intransitivity is more complex than both a purely accusative marker, and thus by extension, more complex than no overt expression of grammatical relations of any form.

4.1.6. *Particles.* Lahu has a number of modal and pragmatic particles central to basic expression which are conventionalized into highly particular and idiosyncratic subdivisions of semantic and pragmatic space, reminiscent in both proliferation and in particularity of meaning to those in German and Dutch. (14) exemplifies the use of just some of them (*cɛ, tí, tɛ̀, qo*), with the translations provided by myself according to the grammar and checked with the author:

(14) *kɔ́lɔ́ cɛ tí tɛ̀ qo ɔ̄, te mâ pə̀*
Thai to.the.extent.that just EMPH as-for TOP do NEG finish
tù hɛ́.
FUT probably
'If it's really only the Thai [who are doing it], they'll probably never get it finished!' (181)

It may be useful to note that this is a thoroughly ordinary sentence of Lahu, not an unusually "expressive" one, and more importantly, there is no attempt in the glossing to "exotify" this sentence for the sake of my argument. To give the reader even a preliminary grasp of the actual semantic space filled by each of these particles would require more space than this paper allows (the interested reader ought consult Matisoff's magnificent grammar). However, for example, *cɛ* does not in any sense translate simply as 'if', but in its other uses clearly makes a semantic contribution leading Matisoff to term it an "extentive" particle (170–171); *tɛ̀* is not merely a focus particle, but one with the specific pragmatic purpose of averring the truth of a proposition in forceful observation or argument in a 'See, there you go!' sense (171); etc.

Saramaccan can certainly convey all of the meanings encoded by the Lahu particles, but importantly, has not grammaticalized nearly as many items for such pragmatic uses, such that (i) pragmatic space is not overtly divided up so finely; (ii) the markers that do exist are mostly less conventionalized and less, in Lehmann's (1985: 307–309) term, obligatorified, than those in Lahu,

and thus (iii) since concatenations of several overt pragmatic indicators are not typical of Saramaccan speech, the grammar does not require control of their specific possibilities of combinability.[9]

4.1.7. *Serial verbs.* Lahu, like Saramaccan and many other creoles, makes great use of verb serialization and concatenation to express concepts usually expressed with prepositions and adverbs in English. However, within this grammatical tendency, Lahu has tightly grammaticalized many more verbs than Saramaccan (Matisoff lists 57 verbs which have evolved such conventionalized usages [212–213, 222–230, 237]), with many of their functions being more specific and more deeply abstract than those in Saramaccan.

For example, one use of the verb *là* 'to come' indicates that the verbal action is for the benefit of either 1st or 2nd person but not 3rd – *chɔ lâ*, 'chop for me/us/you' (324–330); *e* from 'to go' expresses not only concrete movement away (*lòʔ e* 'enter into'), but also extremely figurative senses of movement away such as with the verb 'to finish' to mean 'all used up': *pə̀ e* (318–319). *Gu* 'to fix' has grammaticalized into indicating that an action is done better than before, and from here to simple repetition: *gu chɨ̀ʔ* 'retie' (213). There are many grammaticalized particles used with verbs where a lexical progenitor is no longer in use; whether these are derived from verbs or not, they add often highly specific meanings to verb complexes. For example, the particle *šɛ̄* is an "anticipatory inchoative", implying that something implied beyond the utterance is just about to occur: *ɨ́-kâʔ hɛ́ šɛ̄* (water bath PART) 'Take a bath first [and then we'll eat]' (337).

Furthermore, grammaticalized serial verbs in Lahu vary as to where they occur in relation to the head verb, or within a concatenation of more than two verbs.

In Saramaccan, the serial verbs with especially grammaticalized usages constitute about a dozen usually relatively shallow instances of metaphorical extension such as *téi* 'to take' (instrumental); *gó* 'to go' (direction away), *kó* 'to come' (direction towards / becoming), *dá* 'to give' (benefactive), *kabá* 'to finish' (completive), *tooná* 'to return' (repetition), *táki* 'to say' (complementizer), and *púu* 'to pull' (movement off of) (see Veenstra 1996: 73–103 for the most comprehensive and insightful treatment to date). Some verbs have moved somewhat further along the grammaticalization cline; e.g., *dá* 'to give'

9. There does not yet exist a thorough grammar of Saramaccan, such that some machinery along these lines may have yet to be discovered. However, given how many studies of various areas of its grammar have indeed been carried out, we can be sure that Saramaccan does not have anything approaching the array of highly conventionalized modal and pragmatic particles that Lahu does; indeed, such machinery is not evidenced in recordings of running speech in the language.

also expresses causation and source: *Dí móni lási dá mi* 'The money got lost because of me' (Veenstra 1996: 165). Yet on the whole, the body of grammaticalized serials in Lahu exhibits a greater degree of metaphorical extension than Saramaccan's, and a tendency toward finer subdivision of semantic space: for example, Lahu's 'come' verb *là* is restricted in its grammaticalized use as a benefactive marker to 1st and 2nd person, while Saramaccan *kó* 'to come' applies to all persons in all of its grammaticalized usages, nor does Saramaccan limit any grammaticalized serial to particular persons. Finally, all of Saramaccan's serial verbs except causative *mbéi* 'to make' and (only partially grammaticalized) instrumental *téi* come after the head verb, and thus the use of a serial verb in this grammar does not regularly entail knowledge as to which of several possible placement patterns applies to the intended meaning.[10]

4.1.8. *Verb concatenation.* In Lahu, certain verbs occur only in concatenation with a specific other verb to convey completion. Thus *tòʔ* 'catch fire' only occurs after *tú* 'to kindle': *tú tòʔ* 'to catch fire'; *kɨ̀* 'be melted' only appears after *lɔ̄* 'cause to melt'. The negator *mâ*, which usually precedes verbs and verb compounds, intervenes between the two verbs: *tú mâ tòʔ* 'does not catch fire' (207–208).

There are no verbs with such lexically and syntactically restricted occurrence to my knowledge in Saramaccan, and no verbs which condition exceptional negator placement.

4.2. *Implications of the comparison*

Even factoring out the fact that Lahu is SOV, and despite these two grammars sharing analytic structure and various traits often found in analytic languages like verb serialization, short word length, and tone, there is a clear qualitative difference between these two languages. That difference is that Lahu is "busier", an impression which, translated into linguistic terms, comprises three

10. DeGraff (1999) misinterprets my previous observations that grammaticalization in Saramaccan and other creoles generally has yet to metaphorically extend into lexicalization as claiming that creoles lack metaphor in general. The mistake here is the misconception that opaque lexicalizations in older grammars embody metaphor in the mind of the synchronic speaker, when in fact the very emergence of lexicalization entails that the form is no longer generatable via metaphorical inference and now requires storage as an independent form. Along these lines, in the strict sense a language with little or few lexicalized forms is richer in metaphor than one where a large subset of compounds and derivation–root combinations have drifted beyond metaphoric processibility. Metaphor is central to meaning in any human language; my point is simply that older grammars display evidence of now-defunct metaphoricizations, since replaced by the ones now living in that grammar. Moreover, the metaphoricization of compounds and derivation–root combinations is an issue largely separate from that of idiom, proverb, and word play, which the Creole Prototype hypothesis neither addresses nor entails any denial of in creole languages (*contra* DeGraff 2000).

main observations. One is that Lahu has developed a more elaborate tonal system than Saramaccan's, much of this obviously traceable to segmental erosions during its long history (such as tonally encoded causativized verbs [Matisoff 1973: 32–34]). The second is that it shows evidence of the development, erosion, and subsequent regeneration of derivational morphology, another process which requires more time than creoles have yet existed in. The third, and most germane to this paper, is that Lahu has developed more fine-grained overt expressions of underlying semantic and pragmatic distinctions in many areas, as a result of having been used continuously over countless millennia.

Some readers might observe that English is also less complex than Lahu in all but one of the features cited (derivation), and thus question whether creoles are unique among natural languages in the degree to which they are less complex than Lahu and similar languages. Yet the important point is that English is more complex, according to our metric, than Lahu in a great many aspects other than those covered in Section 4.1. First, of course, English not only has eight inflections, but also irregular marking of past in strong verbs and various suppletive verb forms as well (*was*, *went*, etc.). English inverts subject and verb to form questions, and combines this with *do*-support, the latter also required not simply with negation in general, but in particular kinds of negated constructions (i.e., there is no *do*-support in *I'm not going*). English has overt markers of definiteness and indefiniteness, whose occurrence is determined by referentiality as well: *I saw a movie last night*, where movie is marked with *a* despite being presupposed to the speaker, because it is not yet known to the hearer. Neither definiteness nor referentiality are marked overtly in any grammaticalized fashion in Lahu. The subtle distinction maintained in English between the *will* future and the *going to* future is another feature which gives fine-grained and grammaticalized manifestation to a distinction lacking in Lahu as well as a great many other languages.

What is important is that I believe that it would be impossible to present an eight-point comparison of Saramaccan and Lahu in which Lahu appeared the less complex according to my definition of complexity, and I speculate that this would be true under any alternate definition of complexity as well.

Again, there is no claim here that Saramaccan, or any creole, is the "ground zero" of human expression, to the extent that such a language could be conceived. There are certainly elaborated features in Saramaccan that Lahu lacks. Tone sandhi is relatively weak in Lahu (27, passim) but rather elaborated and conventionalized in Saramaccan, for example breaking at the juncture of a verb and a following object, and in shared object serials skipping the object and continuing on the second verb (Rountree 1972). (15) gives underlying (top) and surface forms (second line); underline indicates spread tones.

(15) *mi wási koósu butá a dí sónu*
mi wási̲ koósu bu̲tá a dí sónu.
I wash clothes put in the sun
'I washed my clothes and put them in the sun.'

Along the lines of the array of pragmatic particles in Lahu, Saramaccan has conventionalized the word *seéi*, from *self*, into a highly versatile particle with a quite idiosyncratic domain of application cutting across emphasis, intensification, continuation, and delimitation (Rountree & Glock 1977: 131–135):

(16) a. *Mi seéi mbéi ɛ̃* 'I did it myself.'
b. *De gó a Fóto seéi.* 'They went all the way to Foto.'
c. *Mi gó ku mi seéi nɔ́ɔ.* 'I'm going by myself.'
d. *A tá-kái mi seéi.* 'He keeps calling me.'
e. *Mi seéi ó gó.* 'I'm going, too.'
f. *Má kísi seéi.* 'I didn't catch a single one.'
g. *A hánse seéi.* 'She's beautiful.'
h. *Gó seéi de gó.* 'They're gone for good.'
i. *Dóu seéi de dóu.* 'They just got here.'

Saramaccan also differs from Lahu, as it does from Tsez, in having definite and indefinite markers (although their expression of referentiality is less grammaticalized than in English).

The situation is similar in all creoles. For example, grammaticalization has proceeded faster in general in Angolar Creole Portuguese (and its sister creoles in the Gulf of Guinea) than in any other creoles I am aware of (this may possibly be due to the fact that these creoles are about 150 years older than New World creoles like Saramaccan). The subdivision of semantic space between the tense and aspect markers is highly complex and subtle (Maurer 1995: 67–89), with only the barest outline of the tidy form-meaning correspondences in the "prototypical creole TMA system" posited by Bickerton (1981 and elsewhere). In serial constructions, some verbs have drifted into meanings as metaphorically removed from their root meanings as is common in Lahu (often specifically reminiscent of the depth of drift evidenced by "light" verbs in Indo-Aryan; cf. Hook 1991 for characterization and diachronic analysis). A combination of 'cut' and 'put' is one way of conveying inchoativity, while 'throw' conveys completion:

(17) a. *ane kota ona pê.*
3PL cut murmur put
'They started to murmur.' (Maurer 1995: 105)

b. *a* *baga* *kampu* *ce* *ta.*
3S ravage field DEM throw
'The field was completely ravaged.' (Maurer 1995: 109)[11]

It is also relevant here, contrary to what might quite reasonably be supposed, that not all natural languages even have overt marking of categories as seemingly inextricable to human expression as tense and aspect (e.g., the Papuan language Maybrat; Dol 1999); many do not have relativization strategies overtly distinct from general subordination (e.g., many Native American languages, cf. Kalmár 1985); various languages of Southeast Asia have been argued to evidence no underlying distinction between basic constituent categories (Gil 1994). All known creoles have all of these traits and a few others not as universal in language as one might suppose, and as such, regularly depart from "the heart of language" in terms of functionally central features unmarked in some older grammars. It also bears mention that a many older languages have smaller phonemic inventories than any known creole (Polynesian being the most obvious example).

Thus our stipulation is neither that creoles were born utterly devoid of complexity nor that they have not evolved some over the centuries of their existence. Because this is not the stipulation, the following is germane: the identification of scattered exceptions in various creoles to the general tendency I have identified does not constitute a refutation of my argument. The argument refers not to binary opposition, but to degree: the crucial point is whether or not creoles tend to exhibit AS MUCH complexity as older grammars do overall. I suggest that they do not, and more to the point that creoles are MUCH less complex than all but a very few older languages. For example, while Saramaccan is more complex than Lahu in terms of tone sandhi, it is not accidental that

11. One reviewer asks whether there is proof that these constructions resulted from evolution within the creole, as opposed to having been inherited from a substrate language. Conclusive resolution of that question must await further work on Angolar's parent creole São Tomense (forthcoming as of this writing), as well as further work on the main serializing substrate language of this creole dialect complex, Edo (Bini). However, for Angolar to retain serials in which verbs' uses are so far abstracted from their core meanings would be unexpected. Only in cases where one substrate language had unusually heavy influence in the contact situation or persisted in use alongside the creole over a long period of time do we usually encounter transfers so specific and extending even to the abstract level (i.e., dialects of Melanesian Pidgin English spoken alongside Melanesian languages, or Portuguese creoles in India spoken alongside Marathi or Gujarati, or Sinhala). Where a substrate language competed with others, its contribution was generally robust but of limited idiosyncrasy. A useful comparison is Saramaccan, whose substrate was about as neatly divided between a serializing language and Bantu as São Tomense's was (Fongbe vs. Kikongo in the former case, Edo vs. Kikongo and its close relative Kimbundu in the latter). The Fongbe contribution to Saramaccan is rich (McWhorter 1996), but largely stops at highly abstract syntactic and semantic features such as more deeply grammaticalized serials (cf. Migge 1998).

in comparison to tone sandhi in its (analytic) substrate languages, Ham (1999: 87–88) notes that:

> All of the major differences between sandhi patterns in Anlo [a Gbe variety] and Saramaccan suggest that, where the choice has presented itself, Saramaccan has taken the easier road – four surface tones reduced to two, elimination of an environmental trigger, and a general capacity for left edges to block sandhi, not requiring any information about the syntactic function of the projection.[12]

Finally, it must be clear that Lahu is not a special case as analytic grammars compared to creoles go. Examining a grammar of an analytic language, the linguist familiar with a wide range of creole grammars typically encounters a number of semantic distinctions virtually or completely unknown in creoles, and a tendency for case assignment, derivational processes, or syntactic configurations to vary according to factors such as transitivity, definiteness, experientiality, interrogativity, focus, et al., in contrast to the strong tendency in creoles to display much less particularity and variation in the expression of these underlying oppositions.

4.3. *A replicable result*

It would belabor the point and tax the reader for me to present additional such cases with the detail allotted Lahu above. However, even a brief look at one more example will be useful. Maori not only lacks inflection, but unlike Lahu, also lacks even tone. Nevertheless, Maori grammar presents one feature after another whose expression is more semantically or syntactically complex than the equivalent in any creole grammar according to the metric proposed in this paper. A Polynesian language is particularly useful for our purposes in showing that these languages' grammars do not parallel their phonologies in simplicity, and place them perceptibly higher on the complexity scale than typical creoles.

Just a few examples, by no means an exhaustive list of such features in Maori, will suffice to illustrate (examples from Bauer 1993).

Maori has several different INTERROGATIVE CONSTRUCTIONS according to the grammatical status of the constituent questioned; the following is but a subset of the available strategies:

---

12. Nor is it accidental that all three times I have taught a seminar on the structure of Surinam creoles, the tone sandhi in Saramaccan has immediately elicited spontaneous and lasting interest from at least one student, having at this writing inspired three papers and served as a major plank in a dissertation. The sandhi is one of the few features in Saramaccan that is relatively elaborated and idiosyncratic from a crosslinguistic perspective; the paucity of such features in the language is a simple consequence of its recent origins in a pidgin and the brief period of time which has elapsed since then.

(18) a. Subject: topic-particle fronting
ko wai kua hoki ki te kaainga?
TOP who T/A return to the home
'Who has gone home?' (Bauer 1993: 7)
(Saramaccan: *ambɛ́ bi gó a wósu?*)

b. Object: marked for genitive and relativized
(NEARH = near hearer)
he aha taa teeraa wahine e horoi naa?
a what GEN that woman T/A clean NEARH
'What is that woman cleaning?' (Bauer 1993: 8)
(Saramaccan: *andí dí mujɛ́ɛ-dɛ́ tá limbá?*)

c. Indirect object: no movement
i paatai te maahita ki a wai?
T/A ask the teacher to ANIM who
'Who did the teacher ask?' (Bauer 1993: 10)
(Saramaccan: *ambɛ́ dí mɛ́sítɛ bi hákisi?*)

There is no creole known to me where interrogation strategies vary to anything approaching this extent according to grammatical relation.

Maori has the notorious Polynesian feature of a subtle POSSESSIVE DISTINCTION between marking with *oo* or *aa*, reminiscent of an alienable/inalienable distinction, but contingent basically upon dominance of possessor over possessee (Bauer 1993: 209–216). In addition, in marking specific possession, a prefix marks a distinction between future (*m-*) and non-future (*n-*). No creole known to me makes a grammaticalized alienable/inalienable possessive distinction (McWhorter 2001):[13]

13. According to Bettina Migge (p.c.), Ndjuka exhibits a tendency to mark alienably possessed items via juxtaposition of possessor and possessum and inalienably possessed items with the preposition *fu*. However, this gives all signs of being an internal development rather than an original feature. For one, the Ndjuka pattern does not parallel the inalienable possessive marking of its main substrate language Fongbe, making it implausible as having been transferred at the genesis of the creole. Second, the feature is undocumented in Ndjuka's ancestor creole Sranan, and unelicitable from my informants of sister creole Saramaccan, where Fongbe influence is otherwise even stronger. Finally, even in Ndjuka the feature is only partially conventionalized rather than obligatory, suggesting a change in progress.
This thesis hardly rules out that creoles not begin to accrete needless complexities as time passes, but in order to avoid unfalsifiability, it is required that it be firmly establishable that the feature is a recent innovation rather than an original feature. What would be unexpected is that creoles would display no such features, as all natural languages do. However, my claim is that creoles will display less such features than older ones. See McWhorter (2001) for details on this point.

(19) a. *n-aa Hone te kii nei.*
PRES-POSS John the key NEARS
'This key belongs to John.'
b. *m-oo Hone te hooiho raa.*
FUT-POSS John the horse that
'That horse is for John.' (Bauer 1993: 208)

In SUBORDINATE AND NON-FINITE CLAUSES, subjects of intransitive verbs are marked as possessives, while the verb itself is nominalized:

(20) a. Maori
a te tae-nga mai o ngaa moni, ka
at the arrive-NOMZ to.here GEN the.PL money T/A
hoko mai ahau i te koha maa-u.
buy hither I ACC the gift for-you
'When the money arrives, I'll buy you a present.' (Bauer 1993: 278)
b. Saramaccan
té dí móni kó, hɛn mi ó bái wan soní
when the money come then I FUT buy one thing
paká dá i.
pay give you

(Moreover, the initial preposition *a* in the Maori sentence varies according to tense; it would be *i* in the past.)

Once again, we see that analyticity alone does not explain the body of features regularly found in creoles. More precisely, analyticity does not provide an explanation for what is regularly NOT found in creoles – specifically, the high degree of elaboration inevitable in grammars tens of thousands of years old, but impossible in a grammar just a few centuries old.

4.4. *Simplification in language change versus the Creole Prototype*

Doubtlessly, certain older languages have evolved into states approaching the relatively low complexity level of creoles. However, evidence suggests that it would be formally impossible for an older language to actually attain the state of, for example, Saramaccan or Tok Pisin. Simplification is an ongoing process in older languages, as phonetic erosion and analogy exert their effects over time. However, this simplification is of course complemented by emerging complexifications as well, the very process which prevents a language from eroding into a mouthful of dust. A classic example is the erosion of Latin's future suffixes and their replacement by new ones grammaticalized from forms of *habēre* 'to have' in most of the Romance languages. A less familiar example is

in Lahu, where an erstwhile causative prefix *s-* eroded long ago, leaving devoicing of initial consonants and tonal disturbances on about a dozen verbs, e.g., *dɔ̀* 'drink', *tɔ* 'give to drink'. Meanwhile, lexical verbs such as *te* 'do, make' have been recruited into a new causative construction, which may well develop into affixes over time (Matisoff 1973: 243–244). Moreover, at all times, the complexified results of other long-ago accretions at all times remain alongside the new complexifications emerging, since features often persist in a grammar for long periods before being "brought down" by erosions and analogies.

Because this complexification is always "working against" the simplification, and at any given time only a subset of a grammar's complexifications are eroding, an older language retains at all times a degree of complexity alongside the simplifications it is undergoing. 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 zero, rather than alongside the results of tens of thousands of years of other accretions. As such, creoles tend strongly to encompass a lesser degree of complexity than any older grammar. As language change is a non-discrete and in many facets a contingent process, quite expectedly some older languages fall closer to the "creole" end of the complexity cline than others, such as some Southeast Asian languages, some Mande languages, and the Polynesian languages.

The older language known to me which comes the closest to exhibiting the degree of complexity of a typical creole is the Riau dialect of Indonesian (Gil 1994), a language with what appears to be essentially a pidgin-level syntax, even less specified than that of any creole known to me (a structural type which Gil argues to be shared by a number of Southeast Asian languages including Vietnamese). However, even this language reveals its age in having three inflections, at least some opaque derivation-root combinations, and optional numeral classifiers. It is of further interest that this language is a koine, traditionally used as an interethnic lingua franca, meaning that its particularly 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.

5. Complexity and previous investigations of creole markedness

This paper is by no means the first treatment to suggest that creoles represent a fundamental layer of natural language, unobscured by the results of millennia of phonological, syntactic, and semantic drift which make Universal Grammar such a challenge to glean in older languages (Kay & Sankoff 1974, Bickerton 1981, Seuren & Wekker 1986). However, the complexity metric proposed here

designates creoles as "fundamental" for reasons generally independent from these previous treatments, and it will be useful to make this as clear as possible fo the purposes of future discussions that this paper may engender.

5.1. *Semantic transparency?*

In the semantic area, the "one form – one meaning" constraint often attributed to creoles (e.g., Kay & Sankoff 1974, Seuren & Wekker 1986) is upon examination a treacherous one. Kihm (2000: 167) notes that the concept of 'fetch' is potentially reducible to the three semantic atoms GO TAKE COME, which is how it is expressed in the Papuan language Kalam as the serial *am d ap*. Creoles do, as a legacy of their pidgin heritage, have some tendency to use compounds in place of unitary equivalents in their lexifiers (Sranan *fes'ede* front head 'forehead'). However, a crucial point here, which Kihm does not happen to pursue but with which I presume he would agree, is that it is unclear that creoles map semantics onto lexical items to any appreciably more atomistic degree than many older languages. Vietnamese, for instance, makes heavy use of concatenations of basic verbs where both "Standard Average European" and creoles would use a single one, and forms plural pronouns via combining singular ones with a plural morpheme. As such, we find a contrast such as:

(21) a. Vietnamese
chúng tôi bắt đầu làm bài.
PL I take start do lesson
'We began to do lessons.' (Comrie 1989: 43)

b. Saramaccan
u bi bigí u lési.
we PAST begin to read
'We began to read.'

Moreover, this is not even a tendency restricted to analytic languages; Colarusso (1992: 141) notes that the Caucasian language Kabardian, for example, has relatively few lexical roots, such that "one may see the lexical semantics component at work in a way matched by few other languages and exceeded by none".[14]

Certainly no scholar has made the brute claim that creoles are semantically transparent to the ultimate possible degree; e.g., Seuren & Wekker (1986: 64) describe this as a "tendency". Yet the extent to which creoles depart from this ultimate degree is nevertheless significant, because at the end of the day it is

14. The fundamental nature of Kabardian grammar is more to the point here than whether or not the author of the quote has actually been able to check whether or not any language happens to exceed Kabardian in this capacity.

unclear that creoles are "semantically transparent" overall to any greater extent than certain older languages. Certainly, the roots of creoles in pidginization leave creoles somewhat higher on the semantic transparency scale than their source languages often are; for example, as Kihm (2000: 186) notes:

> The more I look into linguistic diversity, the more I am convinced that creole languages rank distinctively high on the isomorphism scale, despite occasional departures [. . .] The fact that a language group we know to be distinguished by the special way it came into being is characterized by these features must be significant [. . .]

However, older grammars can apparently develop similar degrees of semantic transparency in the course of diachronic development without the intermediation of any detectable break in transmission. As such, it is unlikely that any claim could stand that creoles are the world's most semantically transparent grammars, and as such, the one-form – one-meaning conception will be of little use in a principled synchronic distinction between creoles and older languages.

5.2. *Markedness versus complexity*

In the realm of syntax, the hypothesis that creoles are closer to an ontogenetic foundation than many other languages appears promising. Bickerton (most recently and summarily, 1999) has long maintained that creoles represent unmarked syntactic and semantic settings. In recent statements he has stressed the semantic aspects, but Roberts (1999) reanimates the syntactic aspect of the hypothesis with his argument that creoles exhibit various unmarked parameter settings, specifically weak feature specifications with epiphenomena such as lack of verb movement to INFL, SVO order, preverbal tense-mood-aspect particles, absence of pro-drop, and absence of complement clitic movement.

Lightfoot (1999: 167) is skeptical that parameter settings vary in terms of markedness, asking why a human grammar would tolerate a setting that was not optimal. One might indeed argue that no parameter setting is more cognitively ungainly than another, or that the added movement processes (such as V to I movement) required by some parameters considered "marked" do not constitute significant enough of a mental exertion to be any riper for elimination than unmarked settings.[15]

Yet there are deducibly reconstructable aspects of the first human language, and if we agree that Universal Grammar was set on the basis of this first grammar, then there some parameter settings which are indeed "derived" in relation

15. I am aware that the Chomskyan framework of this work is held in skepticism by many readers with different theoretical persuasions. However, the work of Bickerton and Roberts in particular has engendered substantial debate in creole studies, such that properly speaking, this paper would be incomplete without addressing the basic issues to some extent.

to what our innate linguistic endowment would presumably specify, this operating quite independently of whether or not the "derived" setting is cognitively disfavored once it emerges.

For example, observed and documented processes of language change make it clear that the main source of affixes is erstwhile free morphemes. Certainly there are cases where affixation arises via other processes, such as the reinterpretation of fortuitous homophonies or patterns exhibited by certain segments in members of a constituent class (e.g., Nichols 1992: 139–142 on the origin of some noun class markers). Even here, however, the universal principle would appear to be that affixation is not simply created out of the blue: it emerges from the grammaticalization, reanalysis, or reinterpretation of material which was not originally inflectional. It follows logically that THE FIRST LANGUAGE HAD NO AFFIXES (cf. Comrie 1992). From this, it subsequently follows that, if verb movement to I is conditioned by a manifestation of rich inflectional affixation, then the first human language had no such verb movement.

We could only escape this conclusion via identifying an inherent aspect of Universal Grammar which would have determined that, even before a grammar existed which had evolved rich morphology, that once that morphology had emerged, the specific result would predictably and inevitably be the specific one of verbs moving to acquire tense. This would also require a motivated ruling out of the verb staying in place and acquiring tense from afar. (This last point would also require demonstrating that there exist no highly inflected grammars where finite verbs do not move to I – something it is unclear that the researchers on this topic have yet investigated.[16]) Short of this, it would appear that parameters contingent upon affixation are indeed "derived" in relation to Universal Grammar – compatible with it, but less direct manifestations of its makeup than the alternate parameters. This gains further support in the fact that children do not venture spontaneous affixes during language acquisition, but instead their first attempts at language tend to lack affixation, as well as in the fact that when language "begins anew" amidst pidginization, the linguistic vehicle consistently lacks affixation entirely or exhibits it only minimally, with

16. The absence of a theory-internal motivation for V-to-I movement occasions an intriguing speculation. Short of a theory-internal reason why human cognition REQUIRES verbs to move to I when a language has rich inflection, we might suppose that the particular consequences of the "strong" agreement parameter were one of many possible configurations in "alternate universe" Universal Grammars distributed among early human populations. In other words we might propose that the V-to-I parameter setting is now universal because of chance: perhaps only one subpopulation happened to be ancestral to all of today's humans, and they had a version of Universal Grammar where the V-to-I parameter happened to have developed. The problem here, however, is that since it is difficult to see what reproductive advantage the movement of verbs to I would have conferred upon a group of humans, this would seem to force us into a kind of Lamarckian analysis of language evolution.

affixes developing only slowly even when the pidgin is creolized.[17]

Thus the controversy over whether parameters can be ranked according to markedness in terms of relative cognitive infelicity does not in itself belie that some parameters may be ontogenetically primary in relation to others. Indeed, in cases of language genesis, if certain Universal Grammar parameters tend strongly to be expressed over others, then this would appear to be strong evidence that parameters exist in this kind of relationship.

Nevertheless, returning to my specific analysis, the fact is that if creoles do reflect the "default" settings of Universal Grammar, this does not in the logical sense entail that they are simpler in the overall sense than older grammars. In fact, upon examination it becomes clear that the question of whether or not creole grammars are less complex than older ones as I have posed it (i.e., according to the metric of complexity outlined in Section 2.4) is essentially independent of markedness of parameter settings.

This becomes clear when we note that (i) a grammar could have only parameter settings considered unmarked and yet be immensely complex according to our metric, while (ii) a grammar could have all marked parameter settings and yet be quite simple according to our metric.

For example, a grammar could have no inflection, categorically overt subject pronouns, and SVO word order (to take three parameters or manifestations thereof whose markedness ranking appears to be a matter of general agreement), and nevertheless have a large number of noun class morphemes, a plethora of pragmatic particles central to basic expression, five lexically contrastive tones, deep morphophonemic processes, various dummy verbs, and movement rule asymmetries between matrix and subordinate clauses, all factors quite independent of the aforementioned parameter settings (in other words, basically a hypothetical but plausible Southeast Asian language). By our metric this would be a highly complex language, despite a syntax which a theoretical syntactician might analyze as "unmarked" – and this would remain true even if we added any number of other "unmarked" parameter settings. Conversely, a grammar could be inflectional, have pro-drop, and be OVS, but with agglutinative morphology (and thus none of the complexification that can follow from inflection), an unmarked phonology, relatively broad tense-aspect distinctions, relatively little exceptional case marking, no evidential or switch-reference markers, and no overt markers of determination (in other words, a kind of object-first Swahili without noun class markers). This would be a

17. Creoles which remain in contact with their lexifier, however, have a tendency to borrow liberally from its derivational (but not inflectional) morphology, as evidenced by French plantation creoles.

much simpler language than the first one by our metric, despite its "marked" morphosyntax.[18]

Thus while the hypothesis that creoles display unmarked UG parameter settings and my own hypothesis preliminarily give the appearance of pursuing the same or highly similar questions, in fact, the validity of either hypothesis is independent of that of the other one, because what distinguishes grammars in terms of complexity according to my definition is largely independent of the syntactician's conceptions of markedness or optimality. Thus my claim that creoles are less complex than older ones does not address whether or not creoles display unmarked parameter settings – because a maximally unmarked grammar could also be a massively complex one by my metric – and cannot be evaluated according to whether or not the weak-parameter hypothesis proves true (which, I suspect, it will in any case).

5.3. *Greenbergian markedness and complexity*

A traditional conception of markedness which complexity (as defined in this paper) does correspond to is one of more direct application to typological studies: implicational markedness as defined by Greenberg (1966a, 1966b). The further a feature falls on the right in a typical hierarchy as designated by the Greenbergian framework (e.g., marking of singular / plural / dual / trial, or masculine / feminine / neuter), the more complex it can be considered to render a grammar. This is because its presence implies that of the features to the left, thus lending the grammar more overt distinctions. The Greenbergian markedness framework also accounts for my claim that larger phonemic inventories are more complex, in that the rarer sounds imply the presence in the inventory of the more common ones (e.g., non-nasal / nasal). Creoles, as less complex overall than older grammars, tend strongly to hone to the left sides of the various implicational hierarchies outlined in Greenberg's work; indeed, if one devised a system to quantify a "score" for a grammar according to how many steps on each of these hierarchies were overtly marked in a given grammar, creoles would generally have lower numbers than almost all older languages both analytic and synthetic.

6. Evaluation

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 Universal Grammar. Certainly the random complexities of any older grammar – Kikongo's four past tenses, Russian's verbs of motion, Maori's

18. It is not insignificant in this light that Swahili displays clear evidence of heavy second-language acquisition in its history (McWhorter 1992).

nominalized intransitives – are compatible with Universal Grammar and in many ways illuminate its structure. Kikongo's past marking gives usefully overt manifestation to interactions between completivity, perfectivity, and pastness; Slavic motion verbs grammaticalize overt marking of telicity; Maori's treatment of intransitives is one of many ways (such as ergativity) that a grammar may overtly mark valence differences.

Our point is that for every grammar that does give overt, grammaticalized marking to the above distinctions, there are several others that neither have the same marking nor any equivalent mechanism: English, for instance, does not have past endings to distinguish an action performed yesterday from one performed the day before that, three ways of saying 'go' depending on whether there is a destination and whether it is a habitual activity, or single out intransitives to nominalize. Yet languages lacking such things nevertheless fulfill the needs of human speech. It follows from this that all three of the strategies in question are, in the strict sense, incidental to what is necessary to human communication. In other words, not just the occasional marginal or fossilized feature, but a GREAT DEAL of any older language "came into being for a reason, but with no purpose", as Trudgill puts it (1999: 149). Having arisen for locally driven reasons unconnected to outcome in the sense Keller (1994) describes as the operations of "the invisible hand", gender marking, multiple past tenses, nominalized intransitives in subordinate clauses, and for that matter evidential markers, subjunctive marking, and subjects marked as experiencers, are all what Lass deftly terms "linguistic male nipples" (1997: 13).

A claim of this sort may appear to smack of "linguocentrism", but in fact the analysis applies to all older grammars including the one I am writing in. An example of overspecification in English is the overt and categorical marking of definiteness and indefiniteness on singular nouns, which goes beyond the needs of a human grammar as far as encoding definiteness is concerned. This is clear from the thousands of grammars without such overt marking. The highest estimate known to me of the percentage of the world's languages that have grammaticalized articles (free or bound) at all is roughly half (Plank & Moravcsik 1996: 204), and Moravcsik (1969: 87, 93–98) estimates that only 39% of that subset have both definite and indefinite markers, and thus roughly one in five of the world's languages.[19] Typically word order and intonation play a role in encoding the distinction, but not as categorically as in English, the distinction often left to context (e.g., in Chinese and Russian).

19. Dryer (1989: 85) suggests that a third of the world's languages have articles and that 8% have both definite and indefinite articles, but his survey does not include cliticized and bound forms.

Once more, our claim is not that creoles are devoid of this kind of random accretion. In Seychellois Creole, the past marker *ti* and the completive marker *fin* interact in a highly delicate interplay; Angolar Creole Portuguese has one verb, *ba* 'to go', which surfaces as an allomorph *be* in contexts of unbounded trajectory; Saramaccan restricts its reduplication of verbs to create attributive adjectives to transitive verbs (**Dí gógó wómi* "the goed man" as in 'the man who went away'), etc. However, it is also true that where creoles display overt marking of such grammatical fundamentals, it is almost always in fashions less complex overall, as defined earlier in the paper, than in older languages: Seychellois, for instance, has two past markers, not four; Angolar has two allomorphs of one verb, not three forms each of a set of sixteen motion verbs; while intransitiveness can condition verb nominalization in Maori, Saramaccan to my knowledge only obligatorily nominalizes verbs at the syntactic level in one affective construction: *Dí woóko mi woóko!* (the working I work) 'Boy, I worked hard!'; and so on. To wit, in terms of accreted complexity (according to our metric) incidental to the core requirements of effective and even nuanced human communication, the least complex grammars in the world are all creoles.

To be precise, the claim is not that all creoles fall further towards the "simplicity" pole than any older language, since there are creoles whose social histories have lent them moderate inflection and various other elaborations. However, our claim is indeed that if all of the world's languages could be ranked on a scale of complexity, there would be a delineable subset beginning at the "simplicity" end and continuing towards the "complexity" one all of which were creoles. Perhaps the first older language to occur would be the likes of Riau Indonesian, which would be followed by semi-creoles like Réunionnais Creole French and creoles highly impacted by substrate transfer such as Sri Lanka Creole Portuguese. However, in the final analysis, there would be a healthy band of languages beginning at the "simplicity" pole which would all be creoles. This is not a function of syntactic parameter settings, and certainly not a function of culture or psychology: it is a predictable and, in the end, rather unremarkable result of the recent origins of creole languages.

Certainly, further grammatical work on creoles will add to our stock of knowledge about their structures. However, I do not believe that the reason creoles appear at this writing to be less complex overall than older grammars is merely because of the admitted paucity of thorough reference grammars of creole languages. For one, there are many more such grammars today than there were twenty years ago, several more on the way, a wealth of extended studies of several others, and a representative number of articles on most. The situation is hardly ideal, but the fact remains that for the modern creolist, the creoles for which there is little useful data available are very much the exception, rather than the rule as it was thirty years ago. In addition, while certainly

briefer grammars miss various features, it is also important that even briefer grammars of older languages readily display the kinds of accreted complexities which are strikingly rarer in creoles.

Hopefully this hypothesis will stimulate further thought on this issue. New data and revisions of my definition of complexity are to be expected. Perhaps there will prove to be a conception of complexity other than mine under which creoles could be considered more complex than older languages. Examination of this issue, however, will ideally proceed within a constant awareness of the following observation, which has motivated this paper. Consider the following group of natural languages: Sranan, Saramaccan, Ndjuka, Tok Pisin, Bislama, Solomon Islands Pijin, Torres Strait "Broken", Aboriginal Pidgin English, São Tomense Creole Portuguese, Principense Creole Portuguese, Annobonese Creole Portuguese, Angolar Creole Portuguese, Negerhollands Creole Dutch, Baba Malay, Haitian Creole, Mauritian Creole, Seychellois Creole, Martiniquan Creole, French Guianese Creole. Among these nineteen languages, according to the grammatical descriptions known to me and consultation with experts on many of them, there is neither 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), or lexically contrastive or morphosyntactic tone beyond a few isolated cases (subtract Saramaccan and there is none at all).[20] None of these factors require inflectional morphology for their occurrence in a grammar, and thus their absence is not an epiphenomenon of isolating typology.

Crucially, (i) one would find a great many of the above features in the lexifier and substrate languages that were spoken by the creators of these creoles; (ii) one would find it impossible to present nineteen of the world's 6000-odd older (i.e., non-creole) languages which lacked all of the above features; (iii) the hypothesis I have presented suggests an explanation for this observation, which I consider to be one meriting examination and explanation.

Received: 10 April 2000
Revised: 24 February 2001

University of California–Berkeley

20. Subtract the last five – French creoles which have heavily borrowed from the French lexicon – and there are no lexicalized derivation–root combinations (although as in all natural languages, such combinations have developed institutionalized [idiomatized] meanings based on metaphorical and culturally-based inference) (see McWhorter 2000 for details).

Correspondence address: Department of Linguistics, University of California–Berkeley, 1203 Dwinelle Hall, Berkeley, CA 94720, U.S.A., e-mail: johnmcw@socrates.berkeley.edu

Acknowledgments: The author would like to thank Bernard Comrie, three anonymous reviewers, Stéphane Goyette, and Mikael Parkvall for invaluable comments on the first draft of this paper, as well as the scholarly community at the Max-Planck-Institut für Evolutionäre Anthropologie in Leipzig for commentary and conversations that inestimably seasoned my thoughts on the issues the paper examines.

Abbreviations: 1PLSTAT first person plural prefix used for stative verbs, ABS absolutive, ACC accusative, ANIM animacy marker, CL noun class, CONN connective, DAT dative, DEM demonstrative, EMPH emphatic marker, ERG ergative, FOC focus, FUT future, GEN genitive, INST instrumental, INT interrogative, LAT lative, NEARH near hearer, NEARS near speaker, NEG negator, NOMZ nominalizer, PERF perfective, PL plural, PP past participle, PRES present tense, SR switch reference to different subject, SS switch reference to same subject, SUBJ subject marker, T/A tense/aspect marker, TOP topic marker, TS tensedness.

References

Andersen, Henning (1988). Center and periphery: Adoption, diffusion, and spread. In Jacek Fisiak (ed.), *Historical Dialectology: Regional and Social*, 39–83. Berlin: Mouton de Gruyter.

Bakker, Peter (1987). Reduplications in Saramaccan. In Mervyn C. Alleyne (ed.), *Studies in Saramaccan Language Structure*, 17–40. Amsterdam: Instituut voor Algemene Taalwetenschap.

Bauer, Winifred (1993). *Maori*. London: Routledge.

Bickerton, Derek (1981). *Roots of Language.* Ann Arbor: Karoma.

— (1995). *Language and Human Behavior.* Seattle: University of Washington Press.

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

Colarusso, John (1992). *A Grammar of the Kabardian Language.* Calgary: University of Calgary Press.

Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. 2nd edition. Oxford: Blackwell.

— (1992). Before complexity. In John A. Hawkins & Murray Gell-Mann (eds.), *The Evolution of Human Languages*, 193–211. New York: Perseus.

Comrie, Bernard, Maria Polinsky, & Ramazan Rajabov (2000). Tsezian languages. Manuscript.

Crystal, David (1987). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

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

— (1999). Morphology in creole genesis: a 20-minute (?*!\"%$°?!!?) prolegomenon. Paper presented at the Neuvième Colloque International des Études Créoles, Aix-en-Provence.

— (2000). Creole speakers as a biologically definable class? Morphology in language creation. Paper presented to the Language in Society Workshop, University of Chicago, January 2000.

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

Dol, Philomena (1999). A grammar of Maybrat. Doctoral dissertation, University of Leiden.

Dryer, Matthew (1989). Article–noun order. *Chicago Linguistic Society* 25: 83–97.

Edwards, John (1994). *Multilingualism*. London: Penguin.

Gil, David (1994). The structure of Riau Indonesian. *Nordic Journal of Linguistics* 17: 179–200.

Greenberg, Joseph (1966a). Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Grammar*, 73–113. Cambridge: Cambridge University Press.

— (1966b). *Language Universals, with Special Reference to Feature Hierarcies*. The Hague: Mouton.

— (1978). How does a language acquire gender markers? In Joseph H. Greenberg (ed.), *Universals of Human Language*, Volume 1, 47–81. Stanford: Stanford University Press.

Ham, William H. (1999). Tone sandhi in Saramaccan: A case of substrate transfer? *Journal of Pidgin and Creole Languages* 14: 45–91.

Henry, Alison (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting.* New York: Oxford University Press.

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.

Hook, Peter Edwin (1991). The emergence of perfective aspect in Indo-Aryan languages. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization*, 59–89. Amsterdam: Benjamins.

Kalmár, Ivan (1985). Are there really no primitive languages? In David R. Olson, Nancy Torrance, & Angela Hildyard (eds.), *Literacy, Language and Learning*, 148–166. Cambridge: Cambridge University Press.

Kay, Paul & Gillian Sankoff (1974). A language-universals approach to pidgins and creoles. In David DeCamp & Ian Hancock (eds.), *Pidgins and Creoles: Current Trends and Prospects*, 61–72. Washington: Georgetown University Press.

Keller, Rudi (1994). *On Language Change: The Invisible Hand in Language.* London: Routledge.

Kihm, Alain (1984). Is there anything like decreolization? Some ongoing changes in Bissau Creole. *York Papers in Linguistics* 11: 203–214.

— (2000). Are creole languages "perfect" languages? In John H. McWhorter (ed.), *Language Change and Language Contact*, 163–199. Amsterdam: Benjamins.

Kimball, Geoffrey D. (1991). *Koasati Grammar.* Lincoln, Nebr.: University of Nebraska Press.

Lass, Roger (1997). *Historical Linguistics and Language Change.* Cambridge: Cambridge University Press.

Lehmann, Christian (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20: 303–318.

Lightfoot, David (1999). *The Development of Language: Acquisition, Change, and Education.* Oxford: Blackwell.

Matisoff, James A. (1973). *The Grammar of Lahu.* Berkeley: University of California Press.

Maurer, Philippe (1995). *L'angolar: Un créole afro-portugais parlé à São Tomé.* Hamburg: Buske.

McWhorter, John H. (1992). *Ni* and the copula scenario in Swahili: A diachronic approach. *Diachronica* 9: 15–46.

— (1996). The diachrony of predicate negation in Saramaccan Creole: Synchronic and typological implications. *Studies in Language* 20: 285–311.

— (1997). *Towards a New Model of Creole Genesis.* New York: Lang.

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

— (2000). Defining "creole" as a synchronic term. In Ingrid Neumann-Holzschuh & Edgar Schneider (eds.), *Degrees of Restructuring in Creole Languages*, 85–123. Amsterdam: Benjamins.

— (2001). The rest of the story: Restoring pidginization to creole genesis theory. *Journal of Pidgin and Creole Languages.*

McWhorter, John H. & Mikael Parkvall (forthcoming). Pas tout à fait du français: Une étude créole. Unpublished manuscript.

Migge, Bettina M. (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.

Moravcsik, Edith (1969). Determination. *Working Papers on Language Universals* 1: 64–130. Palo Alto: Stanford University Press.

Nichols, Johanna (1992). *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.

O'Grady, William, Michael Dobrovolsky, & Mark Aronoff (1997). *Contemporary Linguistics: An Introduction*. 3rd edition. New York: St. Martin's Press.

Plank, Frans & Edith Moravcsik (1996). The Maltese article: Language-particulars and universals. *Rivista di Linguistica* 8: 183–212.

Polinsky, Maria & Eric Potsdam (2001). Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19: 583–646.

Reh, Mechthild & Christiane Simon (1998). Experiens-Konstruktionen in Mande-Sprachen. In *Experiens-Kodierung in afrikanischen Sprachen typologisch gesehen: Formen und ihre Motivierungen* (Universität Hamburg, Institut für Afrikanistik und Äthiopistik).

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

Rountree, S. Catharine (1972). Saramaccan tone in relation to intonation and grammar. *Lingua* 29: 308–325.

Rountree, S. Catharine & Naomi Glock (1977). *Saramaccan for Beginners.* Paramaribo: Summer Institute of Linguistics.

Schwegler, Armin (1998). El Palenquero. In Matthias Perl & Armin Schwegler (eds.), *América negra: Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*, 218–292. Frankfurt: Vervuert.

Seuren, Pieter & Herman Wekker (1986). Semantic transparency as a factor in creole genesis. In Pieter Muysken & Norval J. Smith (eds.), *Substrata versus Universals in Creole Genesis*, 57–70. Amsterdam: Benjamins.

Trubetzkoy, N. S. (1931). Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague* 4: 96–116.

Trudgill, Peter (1989). Contact and evolution 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.

— (1996). Dialect typology: Isolation, social network and phonological structure. In Gregory Guy, Crawford Feagin, Deborah Schiffrin, & John Baugh (eds.), *Towards a Social Science of Language*, Volume I: *Variation and Change in Language and Society*, 3–21. Amsterdam: Benjamins.

— (1999). Language contact and the function of linguistic gender. *Poznaø Studies in Contemporary Linguistics* 35: 133–152.

Veenstra, Tonjes (1996). *Serial Verbs in Saramaccan.* The Hague: Holland Academic Graphics.

Welmers, William E. (1973). *African Language Structures.* Berkeley: University of California Press.

Commentary

Creoles and the notion of simplicity in human languages
by *Claude Hagège*

1. Introduction

In the following comments, I will examine the notion of complexity (Section 2), the metric proposed by McWhorter to measure it (Section 3), the limitations of the comparison between creoles and older languages (Section 4), and the extent to which creoles may be said to be simple (Section 5). In conclusion, I will suggest a criterion for characterizing creoles as distinct from older languages (Section 6).

2. The notion of complexity

2.1. *Complexity and Universal Grammar*

According to McWhorter, a complex language is one which, if compared to a simpler one, contains more "overt signalling of [. . .] distinctions beyond communicative necessity" (Abstract). McWhorter's purpose is not to examine how languages other than creoles differ among themselves with regard to these distinctions – although this would also be an interesting study. He simply says that creoles are on the lowest level of the complexity scale. Consequently, since, in McWhorter's view, they are "unobscured by the results of millennia of [. . .] drift which make Universal Grammar such a challenge to glean in older languages" (Section 5), creoles are the most direct illustrations of Universal Grammar (UG).

If this notion involves universals of the Greenbergian kind, then it would mean that creoles contain more universal features than older languages. But McWhorter's reference to works such as Seuren & Wekker (1986) indicates that what is referred to here is Universal Grammar defined in Chomsky (1981a, 1981b) as a system of principles some of which are rigidly fixed and hence

Linguistic Typology 5 (2001), 167–387 1430–0532/2001/005-0167