

Competition and Selection in Language Evolution*

S. S. MUFWENE**

University of Chicago, Chicago, IL, USA

(Received and accepted: 10 November 2001)

The primary thesis of this paper is that selection plays a role in language evolution. Underlying this position is the assumption that a language is a Lamarckian species, a construct extrapolated from idiolects spoken by individuals who acknowledge using the same verbal code to communicate with each other. There is no perfect replication in any case of language “acquisition”, which is actually a recreation process in which the learner makes a system out of features selected from utterances of different individuals with whom he/she has interacted. In a way similar to gene recombination in biology, each learner gradually and selectively reintegrates into new system features which are often modified in the process. At the population level, the congruence of some divergent idiolectal selections is often strong enough for a language to evolve into a new communal system. A fundamental question for my hypothesis is: What principles regulate selection? I also assume hybridism in language “transmission”, which is polyploidic, as features of every idiolect originate not only in various competing idiolects, but possibly also in different dialects or languages in contact. The question about feature selection remains the same.

Keywords: Acquisition, competition, creole, ecology, evolution, feature/gene pool, gene recombination, hybridism, idiolect, Lamarckian, mixed language/system, restructuring, selection, species, transmission, Universal Grammar

1. Preliminaries

A central question we were invited to address in this forum is whether selection plays a role in language evolution.¹ In the Introduction that Professor David

Hull prepared for our panel (on “Language Change as a Selection Process”), he also addressed the question of whether hybridism plays a role in language evolution and indirectly enticed me to express my position on the subject matter. My answers to both

* This paper was originally presented at the meeting of the International Society for History, Philosophy, and Social Studies of Biology, held at Quinnipiac University, Hamden, Connecticut, July 18–22, 2001. I participated in the panel on “Language Change as a Selection Process”, organized by Professor David Hull, of Northwestern University. I am very grateful to Michel DeGraff for candid comments on the post-conference draft of this essay. I am solely responsible for the remaining shortcomings, especially where I did not follow his advice (literally).

**Corresponding author: Salikoko S. Mufwene, University of Chicago, Department of Linguistics, 1010 East 59th Street, Chicago, IL 60637 USA.

¹ The invitation I received did not specify the interpretation of “selection” that applies to language, let alone whether the term *selection* itself is applicable. The answer to the latter question is affirmative. Although speakers often make conscious choices of words, pronunciations and grammatical rules when they speak a language (which they have already mastered), the

selections that form their *idiolectal* habits (i.e. individual ways of speaking) are generally unconscious ones. They are influenced by where speakers “acquired” the language (see below regarding the quotation marks) and who they have interacted with. These factors – social environment and network of communication – are part of what I identify as the “ecology of language” (Mufwene, 2001). Ruling out the typically unsuccessful attempts of academies to influence vernacular speech, speakers are normally unaware of how communal norms emerge from such idiolectal habits to form dialectal or language-specific peculiarities. Several “ecological” factors – having to do with the ethnography of communication and the internal structure of the language at the time of its “acquisition” – apply at this communal level. All these considerations justify speaking of “natural selection.” This underscores the adequacy of the comparisons of the linguistic and biological species which I am about to submit below. Consistent with Sober (1984, Chapter 1), selection works directly on organisms (idiolects in the case of language). It is its effects that show up at the communal level of languages.

questions are affirmative. Selection plays an important part in *language evolution* interpreted literally as ‘change in the structure and/or pragmatic constraints associated with a language variety.’ Hybridism is also a normal condition not just of those languages that have been disfranchised as “mixed” ones, among them creoles (Thomason and Kaufman, 1988; Arends et al., 1995), but also of other, putatively unmixed ones. I consider hybridism to be not just a matter of the vocabulary originating from one language and the grammar from another (a common misconception about creoles), nor a matter of a 50:50 ratio of contribution to the makeup of a language. It is a more interesting condition which I associate with a “feature pool” from which idiolects and languages select materials for their systems. This paper, which extrapolates from creoles to non-creole languages, will revolve around this “feature pool” notion, which is obviously an analogue of “gene pool” in biology (Mufwene, 2001).² An important caveat about this analogy is that the transmission of linguistic features is typically with modi-

fication, unlike in biology, where the default condition of the transmission of individual genes is perfect replication and innovations arise from how they are recombined into genotypes.

According to the feature-pool approach, partners in the reproduction of a species, which I take a language to be,³ make contributions to the structural makeup of new members (idiolects in case of language), which share features on the family resemblance model. These new members select their features (typically with some modification) from the same pool, although the recombinations are never the same, and those which wind up as dominant are not always the same. In those cases where the contributors have more or less the same (kinds of) features, it is pointless to try to identify a single (dominant) donor. We then speak of congruence (Corne, 1999), i.e. a situation where similar features reinforce each other. They are favored by the ecology of the formation of the new organisms or idiolects to be dominant.⁴ However, when features are different, there is room for competition, which, as noted by

² It is difficult to explain without a lengthy digression what kinds of languages creoles (such as Gullah on the coast of Georgia and South Carolina, Jamaican Creole/Patois or Louisiana Creole) really are, contrary to what the linguistic literature has mistakenly propounded about them. More interested readers should check Chaudenson (2001) or Mufwene (2001). These vernaculars have definitely not developed from erstwhile pidgins. They developed on plantation settlement colonies from the appropriation of non-standard varieties of European languages by populations that are not (fully) of European descent. The label of *creole* assigned them is primarily social stigmatization, a kind of disfranchising act relative to other colonial varieties that developed from the same European languages during the late 17th and 18th centuries (such as North American dialects of English, and Louisiana and Québécois French dialects). Nothing in the evolution of creoles makes them unique, not even the contact and mixedness associated with the development of their structures, as these factors are evident in the evolution of any language and in any case of language diversification, going back to Proto-Indo-European, Proto-Bantu, and the like. Linguists who continue to discuss them as special cases, other than relative to the particular socioeconomic ecologies that produced the communities in which they emerged, have been misguided by the 18th–19th-century social biases that disfranchised them as less natural evolutions from the relevant European languages. See DeGraff (2001) for an apt review of the biases.

³ I argue in Mufwene (2001, Section 2.1: “Communal languages as ensembles of I-languages”) that languages as communal means of communication are similar to biological species. Like the latter, they exist only as extrapolations from their members, which are idiolects in the case of language. They also

share the following properties: 1. interindividual variation in the distribution of genes (identified with linguistic features – see below) and in the manifestation of phenotypes, 2. inheritance of genes/features (which is only partial in case of languages), and 3. differential reproduction. As they satisfy the conditions that Lewontin (1970) identifies as prerequisites for evolution, we could even speak of language evolution in the same way we can speak of species evolution. I then found it normal to go beyond analogies and speak of linguistic species, although their evolutionary properties are largely determined by their ontological peculiarities, more or less in the same way that there are differences in the conceptualizations of animal species and parasitic species, with their evolutionary peculiarities being correlated with their ontological properties and patterns of transmission. I also submit that linguistic species are more like parasitic than like animal species, and they are more of the Lamarckian than of the Darwinian kind.

⁴ According to Chaudenson (2001), congruence is an important reason why the new colonial varieties of European languages spoken by descendants of Europeans have not diverged as much from their lexifiers as creoles have. Western European languages are typologically more similar among them than any of them is with many, or most, of the African languages they came in contact with. Although such colonial varieties have in fact diverged from their metropolitan varieties, due indeed to contact, they are less divergent than creoles are. Another important reason which is obvious in Chaudenson’s work is that race segregation, which reduced the extent of verbal interactions between Europeans and non-Europeans on the plantation settlement colonies, favored a more extensive divergence of creoles, under the conditions discussed in the main text (Mufwene, 2001).

Sober (1984), is typically won on the golf model, rather than on the tennis alternative. It is more often a matter of some variants becoming (more) dominant and masking the latent ones in a language (the counterpart of recessive genes in a biological species).

I have thus far highlighted differences in contact conditions between the development of creoles and non-creoles in the colonies, not absence or presence of contact. Congruence, which favors common features among the languages in contact, is itself a form of selection, since some non-congruent variants are often left out, especially at the level of grammar and phonology, notably if the populations that do not speak the target language natively are socially integrated. In any case, whether or not there is (extensive) congruence, every new colonial variety of a European language (creole or non-creole) is contact-based, by the same competition-and-selection language evolution mechanisms from a feature pool.

I argue that selection is made possible thanks to the nature of language itself both as a species and as a complex adaptive system. As explained in note 3, it is a species because it exists only as an extrapolation from similar but varying idiolects, just like a biological species is an extrapolation from the existence of individual organisms. It is a complex adaptive system because its structure is multi-modular. There are in almost every module units (viz. sounds, morphemes, words and idioms) and usage principles (production and combinatoric ones) which compete or alternate with each other, at least in some contexts of use. All of these are units on which selection (as explained in note 1) applies while an individual develops competence in a language and learns, for example, whether the word *bucket* or *pail* will become part of their (dominant) vocabulary, whether they

will pronounce the word *neither* with the diphthong /ay/ or the monophthong /i:/ after the sound /n/, and whether they will say *I asked not to go* or *I asked to not go*.⁵ The selections an individual makes are similar to that between whether an offspring will inherit brown or blue eyes from their parents, in the sense that very often selecting some particular variants does not entail complete exclusion of the seemingly latent/recessive alternatives, which may be selected by other learners to which such speakers serve as models. A question that has received some attention in the study of the development of creoles is what principles or constraints bear on such unconscious selections into speakers' idiolects. More challenging is even the following question: Under what particular conditions do such individual selections become group selections and part of the communal speech norm? I return to these questions below, though I must alert you not to expect much about the second.

Selection is also made possible by the fact that a language is not acquired in the same way that a genotype of a biological organism is. Actually, while the latter may also be said to be transmitted wholesale, a language is not. To begin with, the phrases *language transmission* and *language acquisition* are only terms of convenience, which are basically inaccurate. We can indeed speak of inheritance of both genes and linguistic features, from parents in biology or from (other) speakers in the case of language – with of course the caveat that it is inheritance with modification in the case of language. One does not inherit the meaning of a word such as *mind* or *meaning* in the same precise way that other speakers use them, or pronounce the word *daddy* in exactly the same way as everybody else. Fortunately, when there is already some variation in the input, chances remain high that what an individual makes part of their competence remains within the range of variants considered normal or native.

However, the similarity (not identity) stops right there. The transmission of genes and their recombination into a genotype involve no reproduction and reconstruction activity on the part of the offspring. At least in the animal species, gene recombination takes place at conception and includes no active involvement on the part of the offspring. The offspring is the outcome of this recombination process! On the other hand, a language is “acquired” piecemeal, selectively, and recreated in a way that makes every

⁵ Croft (2000) uses the term *lingueme*, defined as utterance, to identify units of selection in a language. The idea of *lingueme* as a cover term is probably not that bad, although using it seems to obscure the multi-modular nature of a linguistic system, with every module using different kinds of units and usage principles. In any case, I object to equating *lingueme* with *utterance*. Speakers do not learn to reproduce other speakers' utterances. Instead, they learn (some of) the units and principles that enable those speakers to produce utterances. It is the phonemes (sounds), morphemes (minimal meaningful units), words, principles of grammar and pragmatics (usage in context) that the term *lingueme* could be used to designate, as an umbrella term.

idiolect different and imperfect replication of the units and principles of a language (hence of its system) the default and normal condition in language reproduction.⁶ The acquisition of a(nother) language is thus a creative process (Meillet, 1929; Hagège, 1993), which involves taking the target apart and putting it back together in a manner that is not identical with the starting point. (See also DeGraff, 1999a, 1999b, and 2001, section 7, for more elaborate discussions from a principles-and-parameters perspective.) No two speakers recreate the target language in exactly the same way, because they have not been exposed to the same E-languages in the first place,⁷ and also because, as much as we are all assumed to be endowed with a Universal Grammar (UG) qua biological endowment for language, there is as much variation in our abilities to process and internalize language as there is in our capacities to use our legs to walk, our vocal apparatus to laugh, or our minds to process and internalize culture, just as our anatomies are variably capacitated to digest food items.

2. The normal and natural development of creoles

My concern with the questions of hybridism and selection started with my research on the development of creoles, which is grounded in language contact.

⁶ My attention was first brought to the problem with speaking of language *transmission* by Professor Richard Lewontin (p.c., 2000). He observes that transmission in biology applies typically to cases where the genes inherited at the conception of an individual, do not change in the life span of that individual. He suggested that the term *propagation* might be more appropriate. I then reacted that the term *transmission* was as adequate for language as it is in epidemiology. However, in feedback to the same chapter read by Lewontin, Professor Michel DeGraff (p.c., 2000) pointed out that the term *transmission* was a problem even in linguistics, since it suggests that the speakers from whom learners “acquire” the target language, play an active role in its transmission, as in a classroom (the untypical cases for most languages and for native speakers). He also argued that the term would be inconsistent with the recreation part of the position that a language is partly inherited and partly recreated which I myself have also adopted in my own work since about 1996. I then responded only by stating explicitly the specific way in which I use the term in Mufwene (2001). I hope I have now caught up with the ramifications of both Lewontin’s and DeGraff’s comments.

⁷ Chomsky (1986) proposes a distinction between I-language (internalized language) and E-languages (externalized language). An I-language is the abstract system that enables

Better yet, it is grounded in population contact, since languages do not move around without population movements, and they come in contact by virtue of speakers interacting with each other, just like most parasites come in contact with each other only through the interactions of their hosts.⁸ As in Mufwene (2001), I argue below that the kinds of contacts that produced creoles can also be observed in our daily interactions and in a way all languages are also mixed systems, based on the contacts of idiolects in a community.

Although Thomason & Kaufman (1988) stipulate that, being mixed languages, creoles cannot be assigned genetic classifications, creolists do actually speak of lexifiers of creoles. These are languages from which creoles have inherited the largest proportions of their vocabularies, often in the upwards of 90%. This is why they are usually identified, for instance, as English, French or Portuguese creoles. I reject the traditional stipulation that creoles have developed their grammars from sources that either rule out inheritance from the lexifiers or treat it as only marginal, suggesting that they are made *a novo* by children, thanks to the language bioprogram (Bickerton, 1984), also identified as UG or language blueprint. I am more sympathetic with approaches that underscore both inheritance from the lexifier and influence from substrate languages. However, I submit that like in other cases of language recreation

speakers to produce utterances in a particular language (e.g. English or Kiswahili); and an E-language is the body of utterances that have been produced by the population speaking what is identified as English, Kiswahili, or any such entity. Strictly speaking, an E-language is not the language itself. It is what speakers of a particular language, which is an abstract system, produce. Every speaker is obviously exposed only to a subset of it, being limited in space and time, and having had to interact only with a subset of speakers. My position that a language is a species applies to the population of idiolects or I-languages associated with speakers who can claim to speak the same language. I also use the term *communal language*, by opposition to idiolect, for the same notion of linguistic species. See Mufwene (2001) for a more elaborate discussion.

⁸ As parasitic species, languages differ from biological parasites in not being able to have an existence that is separate from that of their speakers. They are more or less like symbionts in animal digestive systems, except that they are recreated, not inherited, properties of animal lives. Unlike other biological parasites, which colonize their hosts from external sources (e.g. pathogens), languages are made by, and co-evolve with, their hosts, their speakers. Nonetheless, they have “transmission” and evolutionary properties more similar to those of biological parasites than of animal species (Mufwene, 2001, Chapters 1 and 6).

during the “acquisition” process, inheritance from the lexifier (the target for those who developed creoles) is modified under the influence of substrate languages or of the emergent system itself. For instance, the preposition *for* can be used as an OBLIGATION modal auxiliary in English creoles, as in the following Gullah example: *Jean bin fuh come* ‘Jean had to come’ or ‘Jean was expected to come’.⁹ This extension of its usage in English was made possible in part by the fact that adjectives and prepositions can be used predicatively without a copula (the verb ‘be’ in non-existential uses), as in *Lorry (very) pretty* ‘Lorry [is] (very) pretty’ and *Fonzo in the house* ‘Fonzo [is] in the house’. One must also clearly distinguish between “origin” of a feature and “influence” on the use of that feature. Often these notions are not coextensive, as usage of a feature from the lexifier has been influenced by the way a similar feature is used in substrate languages. That is, selection applies in more than one way, which is possible where “transmission” is polyploidic (with a speaker, in case of language, “acquiring” the same feature, or similar ones, from more than one source).¹⁰

My position is motivated by the following considerations: A grammar is not independent of the units that are used in a language. It is a function of how these units, most obviously words, are used in a language (Bolinger, 1973). It is a set of generalizations over their respective behaviors in combination with like units (phonemes, morphemes or words). For instance, phonemes and morphemes combine only in specific ways in a language. In English, a word can start with the phonemic sequence /st/ as in *stand* but not with /ts/. One cannot coin a new word such as *tsand*. Yet this would be possible in German, probably under the spelling *zand*, because this language allows the sequence /ts/ word-initially. Likewise, in English and perhaps in all languages, all words must have a base morpheme, such as *code* in

encode and *decode*. While a word can consist of a base morpheme only, it cannot consist of derivative morphemes only. Thus, *ende* (intended to be pronounced as [endi]) or *deen* (intended to be pronounced as [dien], with two syllables) are not acceptable words in English if they represent combinations of the prefixes *de-* (pronounced [di] and *en-* (pronounced [en]).

Many of these constraints vary from language to language and each module has constraints specific to itself. For instance, not counting Ebonics, English syntax does not generally allow main clauses in which adjectives are used predicatively without a copula. One must say *John is taller than Mary* but not *John taller than Mary*. However, Chinese and Kwa languages (of West Africa) allow such copula-less main clauses. In language contact settings, such variation fosters competition and selection during the recreation aspect of language “acquisition”. Because such variation often also exists within the same language, between alternatives provided by different dialects or idiolects (for instance a child exposed to both educated English and Ebonics in the case of predicative adjectives), everyone “acquiring” a language in a naturalistic setting experiences some kind of competition and selection (Mufwene, 2001).

Coming back to grammar, when a system arises consisting of units and principles about how to use the units, including how to combine them into larger ones, we identify it as a language. Naturalistic “acquisition” of such a system proceeds through learning its vocabulary, apparently the most obvious of its units, where speakers can also identify smaller ones such as sounds and morphemes and with which they associate one level of meanings. Although replication is imperfect in language development, each unit is learned with an interpretation of the constraints associated with its usage, except that its “acquisition” is not necessarily faithful.

For instance, when one “acquires” the meaning of the word *father*, they also learn that it is a count noun and therefore must occur with a determiner when used in the singular, as in *my father*, and the noun phrase it forms with the determiner can function as the subject of a verb, as in *my father lives in the Congo*, or as the object of a verb, as in *somebody robbed my father*. Some speakers also learn that a verb of the same form can be derived from this noun,

⁹ The morpheme *bin* is an ANTERIOR tense marker, equivalent to PAST tense in this example, but is sometimes has the meaning of PAST PERFECT. The preposition *fuh* is pronounced with a schwa, the first vowel of *above* in American English.

¹⁰ These different inputs often vary among themselves. A number of factors then apply to determine which variant prevails in a speaker’s idiolect. Greater frequency is one of them but not the only one, as explained below in section 5.

as *John fathered more than one child with Mary*, but they also learn that *father* in this case must have a male subject, so that one would not say *Mary fathered more than one child with John*, unless they are suggesting that Mary and John actually started their lives as male and female, respectively, and agreed to change sexes and assume new names after having their children (a complicated story of course), or simply if by some other kind of surprise *Mary* names a man and *John* a woman. And such speakers also learn that the verb *father* is typically used to refer to an event but not to a process. Thus one cannot ordinarily say that *John was fathering a child with Mary*.¹¹

Combined together, all such constraints associated with uses of words and other units make up the grammar of language. Speakers who made creoles proceeded the same way that speakers of other languages do, except that the conditions under which they “acquired” the lexifier led them to modify the target units and principles more than one can observe in other cases of “language transmission,” when the feature pool contains units and principles from only one language. In the case of creoles, the heterolinguistic nature of the feature pool to which speakers were exposed sometimes enabled influences from languages other than the lexifier to prevail and therefore produced more extensive modifications of the target language. Thus, although English provides the model construction *Katie is (very) pretty*, in which the copula is required before the predicative adjective *pretty*, most English creoles have dispensed with the copula and produce

something like *Katie (very) pretty*, a construction that is patterned on some substrate languages.¹²

The modifications themselves can be overgeneralizations from something that is present in some variety of the lexifier. For instance, non-standard English provides a possessive use of *me* in *me child*. Speakers of languages in which possession is expressed by preposing the possessor noun or pronoun (with an invariant form in this case) to the head noun easily interpreted this pattern in a more general way and therefore also produced constructions such as *you chile* ‘your child’, *he chile* ‘his/her child’, *we chile* ‘our child’, and *deh/dem chile* ‘their child’, just like *Sara chile* ‘Sara’s chile’, which are attested in Gullah and other creoles. These vernaculars’ pronouns also tend to have the same forms for the subject, object and possessive function, though Gullah shows some variation.

In the same way that such more obvious congruences made possible the divergence of creoles’ systems from those of other offspring of their lexifiers, other forms of congruence played a role, too. For instance, creoles owe the selection of constructions such as *she pretty* ‘she is pretty’ and *Kim the talles* ‘Kim is the tallest’, in which an adjective is used predicatively without a copula, to the less obvious congruence of several factors: 1. the absence of a meaning associated with the copula,¹³ 2. the absence of a category Adjective in several non-European languages (Dixon, 1982), 3. the fact that in many languages that have adjectives these can be used predicatively without a copula, especially when tense is not expressed by a verbal inflection

¹¹ One can say *John was fathering children while everybody else was busy planning their retirement*, with the object in the plural, but here the progressive aspect (*was fathering*) denotes a repetition of the event over a period of time. This progressive inflection of the verb does not convey the same meaning as in *John/Mary was mothering the baby while the older kids were left to themselves*.

¹² Note that in this predicative function, the adjective *pretty* can still be modified by the intensifier *very*, which proves that it continues to function as an adjective and does not become a verb, contrary to what much of the literature on creoles has claimed. Even in English creoles one cannot modify a verb with the intensifier *very*. One cannot say *De man very/bery ron* any more than one can say *The man very runs* in any English dialect. These vernaculars really differ from their European lexifiers in not using a copula and forming a verb phrase for predication in this context. More on this below.

¹³ In English, the copula carries no particular lexical meaning, although it carries tense. One can say that the form of the copula, rather than the verb itself, carries the tense meaning. In infinitival clauses such as *Holly wants to be a doctor*, the copula carries no tense at all. So the real reason why it is used in English sentences is that this language requires a predicate phrase to be headed by a verb and this verb must be a dummy one where the principal element of predication is an adjective, preposition or noun (except where secondary predication is involved, as in *Hugh ran naked across the park*, or *I was caught off guard*, *The Chief wants him alive, not dead*). The copula carries tense simply because English also has another requirement that in finite clauses tense be expressed through a verbal inflection. Some languages, such as Chinese and those of the Kwa family, do not have these requirements. See Mufwene (1992) for such non-monolithism in English grammar and DeGraff (1992, 1998) for more on such secondary predications in English and Haitian Creole.

(*I worked* in English, but *me bin work* in creoles), and 4. the fact that in some languages what the lexifier expresses with an adjective is expressed by a verb (in the resultative aspect).

Many concurrent factors obviously influenced feature selection during the development of creoles. This is in no way different from the congruent influences that influenced the evolution of non-creole languages, regardless of whether or not contact is acknowledged as a factor. For instance, the development of relative clauses such as in *the man with whom Joyce went out* in standard English is a congruent legacy of both Old English, on the one hand, and Latin and Norman French, on the other. Since Old English, this pattern, involving Pied-Piping the preposition with the relative pronoun, has alternated with the more common alternative of colloquial and non-standard English *the man (that) Joyce went out with*, in which the preposition is stranded, the relative clause starts with the complementizer *that*, and the complementizer itself can be omitted. This option does not exist in either Latin or French. The nature of the contact between, on the one hand, English and, on the other, Latin and French, in which the latter languages were prestigious and used by the upper class and intellectual elite, explains the kind of impact that they had on its evolution. The pattern of the language of the lower class has continued in colloquial and non-standard English and that of the upper and elite classes, influenced by foreign languages, was selected for standard English. This bifurcated evolution is also consistent with other specializations in English, where a word such as *perceive* is part of standard and educated speech, whereas *see* and *feel*, whose meanings it conveys, are more general. *Perceive* would be out of place in vernacular speech.

There is actually evidence of such socially-stratified selections in English creoles. Relative clauses start with *weh* (which behaves more like a complementizer than a relative pronoun) or a null complementizer, as in *de man (weh) Joyce go out wid*. (One cannot say *de man wid weh Joyce go out*.) *Weh* appears to have been modified from the non-standard English relativizer *what* (one of whose pronunciations was [wæt]), as in *everything what you said*. Its selection into creoles reflects the fact that their lexifiers were not the standard varieties of European

languages, which probably were introduced only much later to the colonies.

We must also remember that since the Indo-European language family was first postulated, genetic classifications have been based primarily on lexical kinship. The comparative method has traditionally been applied to the lexicon, and to the sounds from which lexical items are built. Grammatical considerations have been secondary. Precluding creoles from genetic classifications because they are putatively mixed languages, when in fact all languages are mixed in some way (Hjelmslev, 1938; and see below), appears to be an uncritical endorsement of a misguided position based more on a sociohistorical view of what creoles are than on any proved differences between how they and non-creole languages have developed. Mostly history prevents linguists from retrospectively extending the label *creole* to the Romance languages, for instance.

3. Development and change by selection

Against the received established position in creole linguistics, I have argued in my work (notably Mufwene, 1996, 2001), along with Chaudenson (1992, 2001), that the development of creoles starts with the choice of a target language among the languages in contact.¹⁴ This is similar to selections made routinely by most speakers around the world in multilingual and/or multidialectal communities, viz. they select (under particular ethnographic conditions) which particular language variety they will target as their native variety or their vernacular. In those socioeconomic ecologies where creoles developed, the lexifier was that particular target, and its

¹⁴ I argue in Mufwene (2000), contra Baker (1990) and Thomason (1997), that the linguistic heterogeneity of the contact settings in which creoles developed did not preclude the identification of a particular language variety as a target. Actually, these settings are reminiscent of the heterogeneity of the linguistic stimulus in those settings involving only one language, as each idiolect and each dialect has its own peculiarities and the learner must often make choices, or at least determine which variants are dominant (against other alternatives) or which ones he/she can use under what particular ethnographic conditions. There is no particular naturalistic setting where a language is made available to the learner in a homogeneous form.

selection entailed attempting to “acquire” its grammar based on the sample of it that these populations had been exposed to.

As noted above, the “acquisition” of a language proceeds piecemeal both in its partial-inheritance and in its partial-recreation aspects. In the classic cases of language contact the language acquisition process is subject to influences from the other languages that the target is in contact with. As can be observed both in child language “acquisition” and in naturalistic adult second language “acquisition,” the process is gradual with the lexicon building up incrementally and with different components of the grammar emerging at different points in time. When one acquires a language as a non-native variety, the “acquisition” of every unit and principle is actually in competition with what one already knows. Interference from the language one already speaks, what can cumulate into substrate influence at the population and communal language level, reflects those cases where options previously known to the speaker prevail, for instance, when speakers fail to aspirate voiceless stops in stressed prevocalic environment (as in the word *pail*) or omits a copula before a predicative adjective in the present tense in English. Although such influence happens despite every effort the learner makes to speak and to command the target language in the best way they can, it is an instantiation of selection from among competing alternatives in a feature pool including both native and xenolectal options. It is one of the forms of selection in language development, one of those that account for communal language change when such modifications spread across the population, beyond the level of just a few idiolects.¹⁵

Spread from idiolect to communal language represents another aspect of selection in the life of a language. Although a language is used by a community, it is acquired individually. It is also spoken differently by every individual speaker, although all idiolectal varieties of the same language are related on the Wittgensteinian family-resemblance model. A language does not exist any more than a species does as a physical entity. It exists only by virtue of

the idiolects identified with individual speakers. Like a biological species, it is an extrapolation from its members, the idiolects. The existence of these individual varieties means variation within a language, just like variation within a species. This means that every time a new member joins a community by birth or by immigration, they face choices to make from among the models they are exposed to. The choices need not be conscious (hence the term *selection* in my work), but they occur nonetheless, dictated in part by who a learner or speaker wants to be associated with. These are selections that also take place through accommodations that speakers make to each other in their respective communicative networks. Since networks vary among themselves, very often alternatives that one abandons are still preferred by others and variation is preserved within the population of speakers. However, due to particular factors in the ecology in which a language is spoken, some variants may disappear, new ones may be introduced, and the overall pattern of variation changes. Hence, there is evolution, whose ultimate origin lies in the selections that speakers make during their communicative acts, while accommodating other speakers or exapting some forms or constructions to meet new communicative needs.

In case of creoles and other so-called “mixed” languages, these selections are more obvious because one can more easily tell when there is influence from a language other than the lexifier. However, even in such cases, language “acquisition” is still individually-based and the mind of every individual learner/speaker is an arena where different options from the same target language and/or from the languages known to him/her compete with each other. While creoles make competition a more obvious factor in the life of a language, they also highlight the reality of what takes place even in those cases where presumably only one language is targeted. In the end, it all boils down to a speaker deciding which particular options they want to make part of their idiolect, which particular units and principles prevail in a particular language, and, for the student of language evolution, whether the end result amounts to evolution in the overall system or a simple redistribution of the same variants within the population of speakers. Thus arises the challenging question of what the relationship is between (the dy-

¹⁵ Idiolects change all the time, as speakers innovate (in ways identifiable as exaptations) or accommodate each other. However, most of these changes have no effect on a communal language, as explained below.

namics of) the individual selections made by particular speakers and the group selections that determine the evolutionary trajectory of a language.

4. Hybridism in language

The foregoing helps me answer the question of hybridism in language. Creoles make it more obvious because they are quintessential cases that illustrate how influences from different languages can produce a new language (variety). However, they also show clearly that it is never a simple matter of the lexifier putatively providing the vocabulary and the rest of the system originating elsewhere. Much of the grammar also originates in the lexifier, although it is hardly preserved intact. Like in gene recombination, units and principles selected from different varieties of the lexifier are restructured into a new system. This explains why Québécois is quite a new French variety different from other non-standard French varieties and even from Louisiana French, which developed under similar sociohistorical conditions. In the case of creoles, the substrate languages often contribute by congruence, although they have also contributed with peculiarities not attested in the lexifier. However, one can observe similar restructuring processes in the development of the Romance languages, from the contact of Latin and Celtic languages, and in the development of English, originally from the contact of the Germanic languages that were brought to England among themselves and with the Celtic languages, too, before Old English came in contact with Old Norse and Norman French. Without selection-guided restructuring under contact conditions it is difficult to explain how these languages have evolved in the specific ways they have. So, the Romance languages and English are as mixed as creoles are, in that they have developed by the same language evolution formula, with of course different values assigned to variables in every case. This is true of the evolution of other languages, too, where much less is known, or admitted, regarding language contact. The history of population movements and contacts in Eurasia and Africa alone over the past 5,000 years or so suggests that we cannot adequately account for language speciation in their respective language fami-

lies (e.g. Indo-European or Bantu) without factoring in language contact.

In all these cases, selections were made from among competing alternatives and those selections account for differences between the Romance languages and Vulgar Latin, between English and continental European Germanic languages, between Bantu and Proto-Bantu, among the different Romance languages, among English dialects, and among the Bantu languages. All of these outcomes depended on the specifics of contact in particular settings and the particular ways that selection resolved or failed to resolve the relevant competition.

The whole process of speciation in genetic linguistics is a by-product of competition and selection in different contact settings, as a proto-language was moved from one location to another and thus brought in contact with other languages. Population movement and contact have usually played a role in language speciation, and the particular dynamics of idiolectal contact account for the paths of language change and the development of new dialects and/or languages. The actual arena of competition and selection in all such cases, including the development of creoles, lies discontinuously in individual speakers. Unless they interact with each other (and thus have the options of accommodating each other), there is no language contact. Part of the accommodation is made possible in the speakers' disposition to learn the other's language or even shift to it, thereby influencing its system. An accumulation of such influences can lead it to speciate into a separate language, what linguists have claimed has happened to creoles (sometimes to the surprise of their own speakers).¹⁶ Thus also, every variety is a hybrid which is best conceived of on the plant allopolyploidism model submitted by Professor Hull in his paper.¹⁷

¹⁶ Gullah speakers typically think they speak English. It is outsiders who disfranchise them, insisting that Gullah is a separate language. Almost the same is true of Ebonics, in which case it is some of the African-American elite, but not its average speakers, who claim that it is a separate language.

¹⁷ This term is applicable here in the sense of crosslinguistic congruence discussed above. In the case of creoles, the influence on one's idiolect originates not only in more than one individual speaking idiolects of the same language but also, and more critically in this case, in individuals with xenolectal features. Selected xenolectal features contributed more divergence from the lexifier.

5. Constraints on selections

The question of what constrains the selections made by speakers remains an important one. Traditional accounts in linguistics have invoked markedness principles, which determine what particular options were unmarked, i.e. (likely to be) favored/preferred, under specific conditions of competition. The competition of course takes place where there is variation along a particular phonological, morphological, syntactic or semantic parameter. For instance, given a choice between a PAST tense construction with an inflection on the verb (e.g. *-ed* in English) and one with a free periphrastic marker (e.g. *been*) that combines syntactically with any predicate (be it verbal or non-verbal), which one is likely to prevail and why? Did the fact that an English creole allows non-verbal predicates to head predicate phrases (as in *Don (ben) taller 'an me* 'Don is/was taller than me') bear on the selection of *ben/bin* as a PAST/ANTERIOR tense marker? Or was it only a question of what marker was more salient?¹⁸

However, if it were just a matter of salience in the perception of the markers, why have forms such as *dead*, *gone* and *broke* survived in creoles? Or is suppletion in the *die/dead*, *go/gone* and *break/broke* alternations part of salience, too? Then why was *went* not retained or the forms *ran* and *brought* instead of *run* and *bring*, which are the norm in English creoles? These are the kinds of questions that arise from study of English creoles in particular. What criteria for determining markedness values can one propose without being trivial and uninformative, e.g. creoles have selected the particular structural options they have because those options are unmarked – and of course the proof lies in the selections themselves. Such an explanation would of course be circular, and one must resort to factors which can explain these specific choices.

Typologists, such as Croft (1990), have suggested that there are universally unmarked parametric options. It is important to clarify that the primary

¹⁸ *Ben* is considered more salient than the inflection *-ed* on the verb. In this particular case, it is an autonomous full syllable, is not used in a reduced form, can bear an emphatic stress. The inflection has none of these peculiarities and is often affected by cluster simplification in sloppy speech. It is not uncommon even among native speakers to produce a form such as *worked* /wɔːkt/ as [wɔːk].

criterion for this position is the statistical distribution of particular structural options among the world's languages. For instance, there are fewer languages that use numeral classifiers (e.g. *head* in *two heads of cows/children* in some English dialects) than those that do not. However, the approach has an inherent problem, viz. that the proportion of languages does not necessarily correspond to the proportion of people using those options. A case in point: Chinese language varieties fall in the category of those languages that use numeral classifiers. If one adds some Indic languages to this category, it is no longer so certain that the numeral classifying system is more marked than the singulative system common in (Western) European languages. By headcount of speakers, there are definitely more users of numeral classifying systems than of singulative systems. Hence, we could also argue that the numeral classifying option is unmarked and the singulative system marked.

There is, however, another kind of question that bears more significantly on this paper: How relevant is such typological ranking to a person "acquiring" a language who does not have one or some of the typological options available to them? Is the fact that creoles of the Atlantic do not have numeral classifiers a function of the assumption that singulative systems are typologically unmarked? Or is it rather a function of the fact that the numeral classifying option was not part of the competition in the ecologies at the time of their formation? Note that, like their lexifiers, they still use constructions such as *hand/bunch of bananas*.

What do we make of the fact that Melanesian pidgins, which are structurally as complex as creoles but developed in settings where numeral classifying is the norm, have a reduced numeral classifying system. They use *pela* (from English *fellow*) with numerals and demonstratives, regardless of whether the noun would be considered count (e.g. 'cow') or mass (e.g. 'water') in English. Did not their creators simply favor an option more familiar to them, just like the makers of Atlantic creoles did? Did not they also select an option unmarked to them, as it was more normal to delimit nouns this way rather than otherwise?

Here it helps to compare the linguistic situation with a biological one. If an individual has been conceived in a society where pigmentation varies al-

most exclusively within the Negroid range, does it matter that there are other options in other societies? Would not black and kinky (rather than brown and straight) hair be the normal, common, hence unmarked, expectation? Is not it really the options available within a particular community that matter? In cases of hybridization, what principles determine whether an offspring will have blue or brown eyes, especially when the parents have different phenotypical traits on this particular parameter? Such considerations have led me to submit that markedness factors must be assessed relative to the ecology in which a language is spoken, articulating those factors that matter to speakers, such as what option is the most common, or the most transparent, or the most regular, the most salient, not semantically empty, etc. (Mufwene, 1991, 2001), bearing in mind that what may be unmarked in a particular ecology may be marked in another, and only subsets of these factors apply in different cases. These observations do not rule out the fact that some of these markedness values obtain in many typologically different languages and may be claimed to be universals. However, even universals must be verified by language-specific phenomena and perhaps the apparent universality of such cases may be grounded in the cognitive infrastructure of language itself. The question remains open. Still, what matters the most in this case is that there is often crosslinguistic variation in the ways these values are assigned to competing variants, so much so that even creoles that devel-

oped in different contact ecologies from more or less the same lexifier (e.g. non-standard colonial English in the Caribbean and the Pacific) have not always selected the same variants. [See also DeGraff (1999b) for a slightly different take on this subject.]

This is not really the right forum to carry on the discussion with which this paper is ending. I have mentioned markedness here only to show that in a way linguists have long dealt with alternatives competing for the same functions, hence with competition and selection. It is not obvious that markedness considerations account for all the selections. There are also many cases where there is no competition, like the fact that a particular language may present no variation regarding the position of the determiner in a noun phrase. However, one cannot deny the fact that competition and selection play a role in language “acquisition” and evolution.

6. Conclusions

Linguistic systems as internalized by individual speakers are complex adaptive systems. So are communal languages as extrapolations from idiolects and as they selectively reflect in their evolutions adaptations undergone by idiolects. Because they exist by virtue of being spoken by individuals, they are inherently variable, aside from not being monolithic structures.¹⁹ Variation in their “transmission” and “acquisition” is amplified in contact settings, because the target language and the other languages that it is brought in contact with make concurrent contributions to the feature pool from which the learner recreates their version of the language (viz. what becomes their idiolect). Although the units and principles are predominantly selected from the target, there is usually (congruent) influence from the other languages, and there are often borrowings. In monolingual settings, there is competition and selection simply because individual speakers contribute variably to the pool from which the learner draws the materials for their idiolect. The new idiolect is a replica of none of the extant ones. It is a hybrid produced by allopolyploidism, with influences coming from several speakers but with each speaker impacting variably on the new idiolect as it develops.

Competition and selection are thus inherent in the dynamics of language evolution. Assuming a fea-

¹⁹ This is an aspect of the architecture of grammars that I did not discuss in this paper, because it would have been digressive to do so. Interested readers can consult Mufwene (1992) and Labov (1998). Rules of a grammar often overlap, as would the pieces of a messy mosaic. For instance, the principle that introduces a relative clause with the complementizer *that* (as in *the book that I like*) is not coextensive with the rule that introduces a subordinate clause with a null complementizer (as in *the book I like*). Although these complementizers still alternate with each other in *I said (that) I like the book*, they do not in *the allegation that Bill Clinton has had several affairs*. One cannot say *the allegation Bill Clinton has had several affairs*. In a different vein, other than because of possible influence from continental European languages, it is not clear why in American English one goes to *church* or *graduate school* (without an article on the noun) but they go to *the hospital* or *the market* (with a definite article on the noun), when in both cases the destination is non-specific and indefinite, and the purpose is the activity associated with the destination. Variation arising from non-monolithism in the architecture of a grammar is itself ground for competition and selection in language “acquisition” and evolution.

ture pool from which every idiolect, and by extrapolation every new state of the communal language, draws its units and principles and recreates a new system, every language is naturally a hybrid of some sort. Those that have recently been influenced by other languages, or emerged from the contact of several languages, are more obvious cases than others. However, the history of mankind shows that every language spoken today has either been influenced by others or emerged from the contact of other languages spoken before it.

References

- ARENDS, J., MUYSKEN, P. and SMITH, N. (eds) (1995): *Pidgins and Creoles: An Introduction*. John Benjamins, Amsterdam.
- BAKER, P. (1990): Off target? Column. *Journal of Pidgin and Creole Languages* 5:107–119.
- BICKERTON, D. (1984): The language bioprogram hypothesis. *Behav. Brain Sci.* 7:173–221.
- BOLINGER, D. (1973): Getting the words in. In MCDAVID, R. Jr. and DUCKERT, A. (eds): *Lexicography in English*. New York Academy of Science, New York, pp. 8–13.
- CHAUDENSON, R. (1992): *Des îles, des hommes, des langues: essais sur la créolisation linguistique et culturelle*. L'Harmattan, Paris.
- CHAUDENSON, R. (2001): *Creolization of Language and Culture*. Routledge, London.
- CHOMSKY, N. (1986): *Knowledge of Language: Its Nature, Origin, and Use*. Praeger, New York.
- CORNE, C. (1999): *From French to Creole: The Development of New Vernaculars in the French Colonial World*. University of Westminster Press, London.
- CROFT, W. (1990): *Typology and Universals*. Cambridge University Press, Cambridge.
- CROFT, W. (2000): *Explaining Language Change: An Evolutionary Approach*. Longman, London.
- DEGRAFF, M. (1992): The syntax of predication in Haitian. In DEGRAFF, M. (ed.): *Proceedings of the 22nd Meeting of the North-Eastern Linguistics Society*. GLSA, University of Massachusetts, Amherst, pp. 103–117.
- DEGRAFF, M. (1998): Non-verbal predication in Haitian and in Irish. In DEGRAFF, M. (ed.): *Proceedings of the Sixteenth West Coast Conference on Formal Linguistics*. Center for the Study of Language and Information, Stanford, CA, pp. 113–128.
- DEGRAFF, M. (ed.) (1999): *Language Creation and Language Change: Creolization, Diachrony and Development*. MIT Press, Cambridge, MA.
- DEGRAFF, M. (1999a): Creolization, language change and language acquisition: A prolegomenon. In DEGRAFF, M. (ed.) (1999): *Language Creation and Language Change: Creolization, Diachrony and Development*. MIT Press, Cambridge, MA, pp. 1–45.
- DEGRAFF, M. (1999b): Creolization, language change and language acquisition: An epilogue. In DEGRAFF, M. (ed.) (1999): *Language Creation and Language Change: Creolization, Diachrony and Development*. MIT Press, Cambridge, MA, pp. 473–543.
- DEGRAFF, M. (2001): On the origin of creoles: A Cartesian critique of 'neo'-Darwinian linguistics. *Linguistic Typology* 5:213–310.
- DIXON, R. M. W. (1982): *Where Have All Adjectives Gone? and Other Essays in Semantics and Syntax*. Mouton, Berlin.
- HAGÈGE, C. (1993): *The Language Builder: An Essay on the Human Signature in Linguistic Morphogenesis*. John Benjamins, Amsterdam.
- HJELMSLEV, L. (1938): Etudes sur la notion de parenté linguistique. *Revue des Etudes Indo-Européennes* 1:271–286.
- HULL, D. (this issue): Species, languages, and the comparative method, pp. 17–28.
- LABOV, W. (1998) Co-existent systems in African-American vernacular English. In MUFWENE, S. S., RICKFORD, J. R., BAILEY, G. and BAUGH, J. (eds): *African-American English: Structure, History, and Use*. Routledge, London, pp. 110–153.
- LEWONTIN, R. C. (1970): The units of selection. *Annu. Rev. Ecol. Syst.* 1:1–18.
- MEILLET, A. (1929): Le développement des langues. In *Continu et discontinu*, Bloud and Gay, Paris, 119ff. Reprinted in Meillet 1951:71–83.
- MUFWENE, S. S. (1991): Pidgins, creoles, typology, and markedness. In BYRNE, F. and HUEBNER, T. (eds): *Development and Structures of Creole Languages: Essays in Honor of Derek Bickerton*. John Benjamins, Amsterdam, pp. 123–143.
- MUFWENE, S. S. (1992): Why grammars are not monolithic. In LARSON, G. et al. (eds): *The Joy of Grammar: A Festschrift in Honor of James D. McCawley*. John Benjamins, Amsterdam, pp. 225–250.
- MUFWENE, S. S. (1996): The founder principle in creole genesis. *Diachronica* 13:83–134.
- MUFWENE, S. S. (2000): Creolization is a social, not a structural, process. In NEUMANN-HOLZSCHUH, I. and SCHNEIDER, E. (eds): *Degrees of Restructuring in Creole Languages*. John Benjamins, Amsterdam, pp. 165–184.
- MUFWENE, S. S. (2001): *The Ecology of Language Evolution*. Cambridge University Press, Cambridge.
- SOBER, E. (1984): *The Nature of Selection: Evolutionary Theory in Theoretical Focus*. University of Chicago Press, Chicago.
- THOMASON, S. G. (1997): A typology of contact languages. In SPEARS, A. K. and WINFORD, D. (eds): *The Structure and Status of Pidgins and Creoles*. John Benjamins, Amsterdam, pp. 71–88.
- THOMASON, S. G. and KAUFMAN, T. (1988): *Language Contact, Creolization, and Genetic Linguistics*. University of California Press, Berkeley.