

## Oxford Handbooks Online

### **A Null Theory of Creole Formation Based on Universal Grammar**

Enoch Aboh and Michel deGraff

The Oxford Handbook of Universal Grammar

*Edited by Ian Roberts*

Print Publication Date: Dec 2016 Subject: Linguistics, Language Contact, Language Acquisition

Online Publication Date: Feb 2017 DOI: 10.1093/oxfordhb/9780199573776.013.18

### **Abstract and Keywords**

Creole languages are typically the linguistic side effects of the creation of global economies based on the forced migration and labor of enslaved Africans toiling in European colonies in the Americas. Section 1 addresses terminological and methodological preliminaries in Creole studies, including definitions of ‘Creole’ languages that contradict some of the fundamental assumptions in studies of Universal Grammar (UG). Section 2 evaluates Creole-formation hypotheses, including claims about the lesser grammatical complexity of Creoles and about an exceptional ‘Creole typology’ outside the scope of the Comparative Method in historical linguistics. Section 3 offers the sketch of a framework for a Null Theory of Creole Formation (NTC) that excludes *sui generis* stipulations about Creole formation and Creole languages and that is rooted in UG, as it applies to all languages. Section 4 concludes the paper with open-ended questions on the place of Creole formation within larger patterns of contact-induced language change.

Keywords: Creole languages, Comparative Method, complexity, typology, language acquisition, language change

---

CREOLE languages in the Caribbean are among the outcomes of peculiar historical processes linking Europe, Africa, and the Americas: these languages are the linguistic side effects of global economies based on the forced migration and labor of enslaved Africans toiling in European colonies in the Americas.<sup>1</sup>

Because the postulated processes of ‘Creole formation’ are most controversial (perhaps even more so than Universal Grammar), section 17.1 addresses terminological and methodological preliminaries. After a brief historical survey of early Creole studies, we revisit some of the initial definitions of ‘Creoles’ in order to highlight the various biases that these definitions may have introduced into linguistics from the start. Many of these

biases go against the spirit of Universal Grammar (UG). The claims to be overviewed in this section will show the persistence of certain mistaken tropes in Creole studies. These tropes force a certain degree of polemics in any state-of-the-art survey of the field, especially a survey like ours where some of the basic foundations of UG are confronted with theoretical claims with antecedents and correlates in anti-universalist views on language variation.

**(p. 402)** These polemics will take us to section 17.2, where we evaluate the hypotheses introduced in section 17.1, with a focus on recurrent claims about the relative lack of grammatical complexity in Creoles and on various attempts at establishing an exceptional ‘Creole typology’ that lies outside the scope of the comparative method in historical linguistics. In other words, Creoles are claimed as a type of language consisting of ‘orphans’ that dwell outside the family-tree model of language change (Taylor 1956; Thomason and Kaufman 1988; Bakker et al. 2011; etc.).

Section 17.3 offers the sketch of a framework, for what we call a Null Theory of Creole Formation (NTC).<sup>2</sup> This null theory does away with any *sui generis* stipulation that applies only to Creole languages. Instead it is rooted in basic assumptions and findings about UG that apply to all languages.

Section 17.4 concludes the chapter with some open-ended questions for future research on the place of Creole formation within larger patterns of contact-induced language change with both children and adults engaged in language acquisition viewed as a UG-constrained (re)construction process with, as input, socio-historically contingent Primary Linguistic Data (PLD).

## 17.1 Terminological and Methodological Preliminaries from a Historical Perspective

### 17.1.1 Basic Caveats: What’s in the Name?

Let us first clarify our objects of study and their label. In this chapter, we use the phrase *Creole languages* as an ostensive label to refer to a set of languages extensionally defined, keeping in mind Mufwene’s (2008:40–58) caveats to the effect that creolization should be taken as a socio-historical, and not a linguistic, concept. Our main objects of study in this chapter come from the set of classic Creoles: the Creole languages of the

Caribbean (see DeGraff 2009). This well-circumscribed and uncontested set of Creole languages will suffice to make the points we need to make, especially in light of our cautious epistemological stance whereby ‘we should not expect any specific sociohistorical or structural claim about any subset of languages known as “Creoles” (e.g., Caribbean Creoles or French-based Creoles) to be straightforwardly extrapolated to all other languages known as ‘Creole’ across time and across space’ (DeGraff 2009:894).<sup>3</sup> With these caveats (p. 403) in mind, we use data from Haitian Creole (HC) to make our case against various claims about Creole languages *as a class* with stipulated pan-Creole structural characteristics.

From its genesis onward, the notion Creole in linguistics and related fields (e.g., ethnography, anthropology, and cultural studies) has been shrouded in a mist of terminological and theoretical confusion (Chaudenson and Mufwene 2001; Stewart 2007; Roberts 2008). Our hunch is that this confusion is partly rooted in the fact that the concept Creole, arguably from the Portuguese *crioulo* and Spanish *criollo* (from *criar* ‘to raise, to breed’ in Spanish and Portuguese), first emerged in the 16th century, not as a linguistic term, but as a geopolitically-rooted classificatory label that acquired ethnographic significance in the midst of European imperialism in the Americas, especially Latin America (for extensive discussion, see Mufwene 1997; Chaudenson and Mufwene 2001; Palmié 2006; Stewart 2007; Roberts 2008).

The term Creole first applied to biological entities, namely flora, fauna, and humans, that were ‘raised’ in the then-recently discovered ‘New World’ though their ancestors were from the ‘Old World.’ This ‘New World,’ though new to the Europeans, was, of course, not new to the indigenous Amerindians who inhabited it prior to Columbus’s arrival. But this Caribbean world did become ‘new’ after the European colonists who laid claim to it eliminated, through disease and warfare, much of the Amerindian population there, and then brought in indentured workers from Europe and enslaved Africans as laborers to turn their New World colonies into settlements that produced immense wealth for Europe.<sup>4</sup> These enslaved laborers brought with them a wide range of typologically diverse African languages, mostly from the Niger-Congo area. The European settlers also spoke a variety of languages, even when they pledged allegiance to a single flag. It is in this milieu of conquest, global economy, and language contact that new languages emerged that were subsequently labelled as Creoles. These new varieties were then enlisted as instruments of that conquest and global economy, both through their uses as *linguae francae* and through their descriptions by European scholars whose prestige and funding relied, by and large, on the forced labor—and ultimately the dehumanization—of Creole speakers.

Consider Saint-Domingue (the colonial name of Haiti). There the French settlers spoke a range of French dialects including *patois* varieties from Normandy, Picardy, Saintonge, Poitou, Anjou, and so forth (Alleyne 1969; Brasseur 1986; Fattier 1998; (p. 404) Chaudenson and Mufwene 2001). It is in this context that new speech varieties were created that were perceived as related, but distinct from and inferior to, the French spoken by French settlers (Girod-Chantrons 1785; Moreau de Saint-Méry 1797; Ducœurjoly 1802; Descourtilz 1809). These new varieties were referred to as ‘Creole,’ on a par with other (non-indigenous) colonial phenomena (e.g., ‘Creole’ cows and ‘Creole’ rice as in note 3) that were perceived as distinct from their counterparts in Europe or Africa. These new ‘Creole’ varieties became associated with, often as an emblem, the Creole people (i.e., people born in Saint-Domingue, with non-indigenous parents—that is, with parents from Europe or Africa; but see note 4). Moreau de Saint Méry, for example, made it clear that the most fluent Creole is spoken by the Creole people. But we’re getting ahead of ourselves. So let’s first dwell on the societal uses of the term Creole since these uses preceded the strictly linguistic ones.

### 17.1.2 A Brief History of the Label ‘Creole’

Let’s first draw attention to the resemblance between, on the one hand, the conquest and language-contact milieu of the colonial Caribbean, which gave rise to Creole languages, and, on the other hand, the analogous milieu in the Roman Empire, which gave rise to the Romance languages as non-Roman tribes in various parts of Europe shifted to varieties of Latin. Such similarity will be important to keep in mind throughout this chapter. For now, there’s one basic ethnographic fact to highlight as we discuss the foundations of Creole studies: Creole people in the Caribbean were distinguished both from the indigenous inhabitants (i.e., Ameridians) and from the then relatively new arrivals from Europe and Africa. In the Caribbean, the term Creole subsequently evolved to encode various social biases related to now outdated notions of racial hierarchy contrasting Europeans to non-Europeans.

This racial hierarchy is most clearly articulated in Moreau de Saint-Méry’s (1797) description of Saint-Domingue, where the author states that ‘for all tasks, it is the Creole slaves that are preferred; their worth is always a quarter more than that of the Africans’ (1797:40). Saint-Méry (1797) further argues that Creole blacks ‘are born with physical and moral qualities that truly give them the right to be superior over Blacks that have been brought from Africa’; ‘domesticity has embellished the [Black] species’ (Moreau de Saint-Méry 1797:39). For Saint-Méry, like for many observers since then, the gold standards for humanity, cultures, languages, and so forth, are dictated by

race- and class-based hierarchies—the same hierarchies that motivated Europe’s *mission civilisatrice* in Africa and the Americas.

Thus, from its very first ethnographic usage, the term Creole already had an exceptionalist flavor attached to it. This exceptionalist flavor was carried along to the linguistic realm when the term was applied to the new speech varieties emblematic of the recently created communities in Caribbean colonies. These speech varieties were eventually attributed structural or developmental characteristics that were perceived as *sui generis* (this is the core thesis of ‘Creole Exceptionalism’). In the colonial era, (p. 405) the often explicit goal was to fit Creole languages into linguistic categories consistent with the race-related assumptions that prevailed during the Creole-formation period and were also used to justify the enslavement of Africans. The writings of Saint-Méry and of many other scholars of his and later periods mistakenly suggest that Creole languages lie somewhere between the language of civilisation spoken by the colonists and the primitive tongues spoken by enslaved Africans in the colony (see DeGraff 2005a for an overview).

### 17.1.3 Racial Hierarchies and Linguistic Structure in Creole Studies

One central factor in the early debate on the formation of Caribbean Creoles is related to the Europeans’ assumption about the Africans’ cognitive ability to acquire European languages. The numerically most important group of adults engaged in the acquisition of European languages in the colonial milieu was the Africans. In a worldview where languages were used to measure the intellectual and moral advancement of nations, the speech varieties of the enslaved Africans had to be ranked as inferior to those of the European colonists. Often this inferiority was explicitly theorized on a racial basis with African minds considered primitive and European minds advanced, as in this definition by Julien Vinson in the 1889 *Dictionnaire des Sciences Anthropologiques*:

Creole languages result from the adaptation of a language, especially some Indo-European language, to the (so to speak) phonetic and grammatical genius of a race that is linguistically inferior. The resulting language is composite, truly mixed in its vocabulary, but its grammar remains essentially Indo-European, albeit extremely simplified.

(Vinson 1889:345–346)

Note here the claim about ‘extreme simplification’ which, in various guises, has dogged Creole studies from its inception onward through the writings of linguists such as Antoine Meillet, Otto Jespersen, Leonard Bloomfield, Louis Hjelmslev, Albert Valdman, Derek

Bickerton, and Pieter Seuren (see Degraff 2001a,b, 2005a, 2009 for overviews). To this day, many linguists and other scholars from various fields still assume that Creoles are at the bottom of variously defined hierarchies of structural complexity (see e.g., Bakker et al. 2011; McWhorter 2011; Hurford 2011). This view is now reified in linguistics textbooks as well, where it is sometimes taken to an extreme as in Dixon's (2010:21) claim that '... of the well-documented creoles, none equals the complexity ... of a non-creole language.'

Contrary to the view just quoted that Creole languages are 'essentially Indo-European,' we find, among European scholars of the same period, the view that Creole languages are peculiar 'hybrids' of European and African languages. From this perspective as well, Creoles are not only structurally simpler than European languages, but also, as noted by Mufwene (2008), bad or unfitting, according to the ideology of race and (p. 406) language purity that prevailed in the 19th century. Hybrids were then considered maladaptive compared to pure species.

One oft-quoted exponent of this view is Frenchman Lucien Adam, for whom the French-derived Creoles of Guyane and Trinidad were 'Negro-Aryan dialects' created by Blacks from West Africa who 'took French words [even as they] conserved, as much as possible, the phonetics and grammar of their mother tongues' (1883:5). In Adam's scenario, the Africans cannot reproduce the grammatical properties of the European target languages: the latter are too complex for the primitive minds of the African learners who can only replicate the words of the European language. Adam was in the avant-garde of 'biolinguistics' in a loose metaphorical sense: he framed his race-based research project in an explicitly biological perspective, namely *Hybridologie Linguistique*. In this perspective, languages, like plants, hybridize. In the case of languages, the structural results of hybridization are bounded by the least complex languages due to the lower cognitive capacities of their speakers, namely speakers of African languages.

Adam is an early proponent of the still current 'substratist' view, as instantiated, for example, in the Relexification Hypothesis, according to which the Atlantic Creole languages (i.e., those that emerged around the Atlantic Ocean) embody Niger-Congo grammars spelled out with morphemes whose forms are derived from Romance or Germanic languages. Though contemporary linguists do not adhere to Adam's racial biases, many substratist theories promote Creole-formation scenarios similar to his. A case in point is Suzanne Sylvain who, though her book documented influences from both French and African languages in the formation of Haitian Creole, concluded her description with the famous description of the language as 'French cast in the mold of African syntax or ... an Ewe tongue with a French lexicon' (1936:178). Similarly Lefebvre's (1998) reformulation of Muysken's (1981) relexification hypothesis suggests

that Haitian Creole is constituted of Gbe grammar relexified with French-derived phonetic strings (see DeGraff 2002 for a critique).

In another set of popular proposals with intellectual antecedents in the 19th century, Creole languages are considered *ab ovo* linguistic creations that offer exceptional windows on the prehistoric foundations of language in the human species. Certain aspects of this line of argument go at least as far back as the 1872 book by Alfred and Auguste de Saint-Quentin on the Creole of Guyane—also in a ‘biolinguistic’ perspective that postulates a minimum of cognitive capacities and cultural characteristics among the creators of Creole languages:

[Creole] is, therefore, a spontaneous, hasty and unconscious product of the human mind, freed from any kind of intellectual culture. For this reason only, it would be remarkable to find in this language anything but a confused collection of deformed French phrases. But when one studies its structure, one is so very surprised, so very charmed by its rigor and simplicity that one wonders if the creative genius of the most knowledgeable linguists would have been able to give birth to anything that so completely reaches its goal, that imposes so little strain on memory and that calls for so little effort from those with limited intelligence. An in-depth analysis has convinced me of something that seems paradoxical: namely, if one wanted to create from (p. 407) complete scratch an all-purpose language that would allow, after only a few days of study, a clear and consistent exchange of simple ideas, one would not be able to adopt more logical and more productive structures than those found in Creole syntax.

(Saint-Quentin 1872:lviii–lix)

Unlike Vinson’s and Adam’s views sketched in this section, this *ab ovo* perspective on Creole formation, especially in some of its contemporary instantiations, draws a sharp line between the genealogy of Creole languages and that of Indo-European and African languages. The most extreme implementation of this hypothesis can be found in Derek Bickerton’s Language Bioprogram Hypothesis, which takes Creole formation to resemble the initial evolutionary steps of language in the human species, especially in the putative catastrophic evolution from the Pidgin (*qua* ‘protolanguage’) to the Creole stage under the agency of children exposed to extraordinarily impoverished PLD.

The three 19th-century views sketched here, with illustrative quotes from Saint-Quentin (1872), Adam (1883), and Vinson (1889), all make specific ‘biolinguistic’ claims about Creole languages as extraordinarily simple languages—much simpler than the European languages from which the Creoles selected their lexica. This idea of linguistic structural simplicity associated with the alleged cognitive limitations of Creole speakers runs

through the gamut of pre-20th century Creole studies. As described in DeGraff (2005a), this peculiar exceptionalist mode of thinking about languages and their speakers was part and parcel of pre-20th century 'normal' scholarship ('normal' in the sense of Kuhn 1970: 10–34).

Given the title of this *Handbook of Universal Grammar* (UG), these claims are incompatible with a theory of UG that leaves no room for grammatical distinctions to be rooted in alleged racial characteristics. UG is truly 'universal' in the sense that it entertains basic ingredients and operations (e.g., abstract grammatical features and Merge) and constraints (e.g., structure dependence) that apply to *all* human languages notwithstanding their history of formation, the race of their speakers, and so on. The data and observations in this chapter will further invalidate Creole Exceptionalism claims.

## 17.2 A Primer against Creole Exceptionalism

The belief that Creole languages manifest the most extreme structural simplicity is often related to, among other things, an alleged 'break in transmission' due to the emergence of a structurally reduced Pidgin spoken as *lingua franca*, immediately prior to their formation.<sup>5</sup> This Pidgin would constitute a bottleneck for the transmission of complex (p. 408) structures from the languages in contact. The first Creole speakers are assumed to have been the first children exposed to the Pidgin in the course of language acquisition. These children would have created the grammars of their native languages based on the Pidgin input, which allegedly explains the drastic simplicity of the emergent Creole structures (see Bloomfield 1933:472–474; Hall 1962; Bickerton 1981, 1984, 1988, 1990, 1999, 2008; and others). This is the 'Pidgin-to-Creole life-cycle' that is found in most contemporary introductory linguistics textbooks (see, e.g., O'Grady et al. 2010:503–504).

Related to this 'break in transmission' claim is the assumption that Creole languages do not bear any genealogical affiliation with any prior languages. In other words, the results of Creole formation are language varieties that are outside the branches of well-established language families. Creoles are not even genealogically affiliated to any of the languages whose contact triggered their emergence. The contemporary *locus classicus* for this claim is Thomason and Kaufman (1988), where Creole languages are considered fundamentally distinct from non-Creole languages to the extent that Creoles are strictly outside the purview of the comparative method. For Thomason and Kaufman and for many other linguists, Creoles are taken as languages without genealogical affiliation due to their 'abrupt formation.' Given the history and evidence to be overviewed in this



chapter, the popularity of this exclusionary approach to Creole languages makes ‘Creole Exceptionalism ... a set of sociohistorically rooted dogmas with foundations in (neo-)colonial power relations’ in modern linguistics (DeGraff 2005a:576).

### 17.2.1 Some Historical Background and Preliminary Data

As early as 1665 in the French Caribbean colonies of Martinique, Guadeloupe, and Marie-Galante, the Jesuit missionary Pierre Pelleprat was already comparing the verbal systems of Caribbean French-lexicon Creoles with that of French and making quotable comments about Creole structures, comments that are still rehashed by 21st-century creolists. Pelleprat’s attention was drawn to the apparent simplicity of the Creole verbal system which he attributed to the enslaved Africans’ failure to learn French:

We wait until they learn French before we start evangelizing them. It is French that they try to learn as soon as they can, in order to communicate with their masters, on whom they depend for all their needs. We adapt ourselves to their mode of speaking. They generally use the infinitive form of the verb [instead of the inflected forms—EA, MdG] ... adding a word to indicate the future or the past.... With this way of speaking, we make them understand all that we teach them. This is the method we use at the beginning of our teaching ... Death won’t care to wait until they learn French.

(Pelleprat 1655 [1965, 30–31], our translation)

(p. 409) Pelleprat’s observations offer some insights about the ways in which African adult learners in the colonial Caribbean may have reanalyzed certain verbal patterns from 17th-century French according to general and now well-documented strategies of second-language acquisition (e.g., non-retention of inflectional morphology, preference for analytical verbal periphrases over synthetic constructions for the expression of tense, mood, aspect, etc., as discussed later in this section). But Pelleprat’s and his colleagues’ 17th-century thinking was not about universal strategies of language acquisition. It was rooted in the belief that Blacks ‘lacked intelligence and were slow learners, thus required lots of patience and work from their teachers’ (Pelleprat 1665:56). Such thinking was imbued with a *mission civilisatrice* (e.g., being enslaved by the French was then described as preferable to ‘enslavement by Satan,’ that is, slavery was lauded as a means for the Africans to enjoy ‘the freedom given to God’s children,’ Pelleprat 1665:56). Therefore, it is no surprise that Creole verbal patterns were then considered as reflexes of the Africans’ inferior humanity, even though similar patterns (e.g., preference for periphrastic constructions with invariant verbal forms) are also found in popular varieties of French as described in Frei (1929) and Gougenheim (1929).

No effort was made, back then, to analyze Creole grammars as autonomous systems with their own internal logic, some of which was influenced both by French varieties and by the Africans' native languages. In contemporary terms, Pelleprat's view would be translated as the claim that Creole speech was a manifestation of early stages in second language acquisition—this view is found most recently in Plag (2008a,b) where Creoles are described as 'conventionalized interlanguages of an early stage.'

In the particular case of the Creole verbal system that attracted Pelleprat's attention, he didn't notice that some of the words that 'indicate the future or the past' are actually derived, in both distribution and interpretation, from vernacular French periphrastic verbal constructions of the 17th and 18th centuries. For example, Haitian Creole (henceforth HC) *te* for anterior marking as in *Mwen te rive anvan ou* 'I had arrived before you (SG)' is derived from forms for the imperfect of the French copula *be* such as *était* and *étais* as in *J'étais arrivé avant vous* 'I had arrived before you (SG)'. Similarly HC *ap* for progressive marking as in *Mwen t(e) ap danse* 'I was dancing' is derived from the French preposition *après* as in *J'étais après danser* 'I was dancing.'

As it turns out, verbal periphrastic constructions in French, which are very common in spoken French (Gougenheim 1929; Frei 1929), often employ forms that are either non-inflected (e.g., the infinitive as in *J'étais après danser*) or less inflected (e.g., the participle as in *J'étais arrivé*). For most French verbs (i.e., the French verbs with infinitives ending in *-er* such as *chanter* 'to sing'), both the infinitive and the past participle end with a suffix pronounced /e/, written as *-er* for the infinitive and *-é(e)(s)* for the participle. In written French, the participle of verbs in *-er* shows gender agreement (*-é* for masculine and *-ée* for feminine) and number agreement (with a word-final *-s* for the plural) but these gender and plural orthographic markings usually have no reflex in spoken French. It is thus that the suffix *-e* has become a general verbal suffix in HC via reanalysis in second language acquisition of the sort already adumbrated in Pelleprat (1665) (see DeGraff 2005b for further details and references). These reanalysis patterns, based on the French (p. 410) infinitival and participial forms, will play a major role in our account of some fascinating properties of HC clausal syntax in section 17.3.

Other HC preverbal markers include: (i) the irrealis (*a*)*va* from forms of the French verb *aller* 'to go' (e.g., *vas* in the 2SG present indicative or *va* in the 3SG present indicative); (ii) the completive marker *fin(i)* from forms of the French verb *finir* 'to finish'; (iii) the marker of recent past *sòt* from forms of the French verb *sortir* 'to leave'; (iv) the modal marker *dwe* from forms of the French verb *devoir* 'to owe'; and others. (see DeGraff 2005b, 2007 for further details). All of these preverbal markers for Tense, Mood, and Aspect (TMA), which belong to the grammatical layers of the VP domain in HC syntax, have

straightforward etyma in French morphemes in periphrastic verbal constructions whose meanings often overlap with those of the corresponding TMA+V combinations in HC—thus providing additional evidence for the genealogy of HC as a descendent of French, according to the comparative method.

Furthermore, these TMA markers in preverbal positions in HC are subject to complex combinatorics. If we consider only the three TMA markers *te*, *ap*, and *(a)va*, we already get eight possible combinations with distinct semantics for each, and HC has at least a dozen such auxiliary-like elements with complex distributions, co-occurrence restrictions, and semantic specifications (see Magloire-Holly 1981, Koopman and Lefebvre 1982, Fattier 1998, 2003, Sterlin 1999, Howe 2000, DeGraff 2005b 2007, Fon Sing 2010 for data samples and diachronic and synchronic analyses). Consider, say, the following word-order/interpretive correlations involving the tense marker *te* and the modal *dwe*: *Jinyò te dwe vini* ‘Jinyò was obliged to come’ (literally: Jinyò<sub>ANT</sub> must come) vs. *Jinyò dwe te vini* ‘It is likely that Jinyò had come’ (literally: Jinyò must<sub>ANT</sub> come). Like HC *dwe*, the French verb *devoir*, which is the etymon of HC *dwe*, is also ambiguous between an epistemic and a deontic interpretation when used in verbal periphrases, on a par with English *must*, to wit: *Jean doit être là* ‘John must be there,’ which can be either deontic or epistemic. But it is not a straightforward matter to decide, without any theoretical analysis, whether the syntax–semantics interaction in HC constructions with *dwe*, where the verb is not inflected, is any less complex than its analogs with French *devoir* where the verb is inflected.

In 1665, Pelleprat’s remarks on a French-based Creole verbal system focused on the ‘general use of the infinitive form.’ Pelleprat found this pattern lacking in comparison to French. His approach exemplifies a more general trend in Creole studies where isolated aspects of Creole languages are claimed as ‘simple’ independently of their internal workings as part of a larger complex system and independently of their apparent analogs with historically related languages (see DeGraff 2001a,b, 2005, 2009 for surveys of other examples of this approach).

In addition, the use of main verbs alongside preverbal TMA markers is, to some degree, analogous to the verbal system in the Gbe languages spoken by many of the Africans in Saint-Domingue during the formation of HC. Given this contact situation, the patterns described by Pelleprat can be better understood as general learning strategies in L2 acquisition (e.g., the role of nonnative acquisition in the diachronic emergence of new English varieties and the rise in the 16th century of non-inflected English modals (p. 411) from previously inflected main verbs). In section 17.3, we will revisit such structural and socio-historical analogs between Creole and non-Creole formation as we clarify basic methodological issues.

### 17.2.2 Empirical Issues with ‘Simplicity’ and ‘Creole Typology’ Claims

Pelleprat’s early focus on a Creole verbal system is all the more striking given that this empirical domain of inquiry has led in the 20th and 21st centuries to controversial claims about a Creole typology. In what may be the most famous such claim, the TMA markers in Creole languages, among other features, are taken as a pan-Creole manifestation of a ‘Language Bioprogram’ that surfaces relatively intact when the learner’s linguistic environment is most extremely impoverished (Bickerton 1981, 1984). Bickerton postulated that Creole TMA’s distribution and semantics, which he took to be similar across Creoles, is the manifestation of a genetically wired ‘Language Bioprogram.’ More generally, the argument implies that Creole languages constitute a particular typology, one that generally does not show any structure whose acquisition requires exposure to positive evidence from the PLD. This scenario suggests that Creoles are in some sense pristine in comparison with other languages (Bickerton 1988:274).

The most recent versions of this claim state that Creole languages as a class do not manifest complexity levels that surpass that of older, non-Creole languages (see, e.g., McWhorter 2001, 2011; Bakker et al. 2011; and Hurford 2012). But this claim is straightforwardly defeated by evidence from HC where noun phrases exhibit a complex set of syntactic and morpho-phonological properties that are absent in French and in the Gbe languages that participated in its formation: (i) a prenominal indefinite determiner and a postnominal definite determiner: *yon chat* ‘a cat’ vs. *chat la* ‘the cat’; (ii) a set of (at least) five allomorphs for the definite determiner: *la, lan, a, an, nan*. In addition, bare (i.e., determiner-less) noun phrases and noun phrases with the definite determiner manifest semantic options that are not attested in contemporary dialects of French and Gbe. (See Aboh and DeGraff 2014 for further details.) These characteristics are not outliers: DeGraff (2001b:284–285) produces a list of structural complexities that are found in Creoles, but that are not found in various non-Creoles. Likewise, Aboh and Smith (2009) and Aboh (2015) provide a variety of empirical and theoretical arguments highlighting ‘complex processes in new languages.’ One such process, which is analyzed in Aboh (2015), is agreement between the determiner and the complementizer of the relative clause modifying the noun as in *Di fisi di mi tata kisi bigi* ‘The (singular) fish that (singular) my father caught is big’ vs. *Dee fisi dee mi tata kisi bigi* ‘The (plural) fish that (plural) my father caught are big.’ Such agreement processes, which are found in non-Creole languages such as Dutch (e.g., Booij 2005:108–109), directly contradict Plag’s (2008a,b) claim that, as ‘conventionalized early interlanguages,’ Creoles lack interphrasal information exchange. (p. 412)

Proponents of the Pidgin-to-Creole Life Cycle do not usually describe any Pidgin source for Caribbean Creoles. Yet, one occasionally finds descriptions such as the following in Bickerton (2008:216) for Hawaiian Pidgin as spoken in 1887.

Mi ko kaonu polo Kukuihale, kaukau bla mi nuinui saho  
(1) Me go town large Kukuihale drink beer me plenty-plenty drunk  
'I went to the big town Kukuihale, drank beer, and got very drunk'

In introducing this example, Bickerton (2008) remarks that 'by 1887 many people in Hawaii were speaking a pidgin that mixed Hawaiian and English words indiscriminately.' The Pidgin is further described as 'word salad,' 'macaronic,' without 'any consistent grammatical structure,' a 'linguistic meltdown,' 'almost totally devoid of complex sentences,' etc. (Bickerton 2008:217–218, 223). Yet, a cursory look at this sentence shows that the words there were not jumbled together as 'word salad.' First, the speaker has access to a coordination strategy that is similar to that in English, as is evident from the translation. Here the first clause, unlike the other two, starts with an overt subject. Second, though bare, the noun phrase *kaonu polo Kukuihale* 'town large Kukuihale' does not seem to come out of free concatenation: the noun is adjacent to its modifier *large* and the noun-modifier sequence precedes the proper name. This seems a systematic grouping of the type [[N-Modifier]-Proper Name] or [N-Modifier]-[Proper Name]. A 'macaronic' sequence could have been one whereby the noun and its modifier seem arbitrarily separated such as in *kaonu Kukuihale polo* (town Kukuihale large), but this is not what this Pidgin speaker produces. In addition, it is important to realize that, in many languages, constituent structures in sequences like (1) are intimately related to prosody (absent from Bickerton's description). Assuming the right prosody, one can get a similar sequence in English:

(2) I went to a small town, Fort Valley, drank a whole lot of beer, and got very very drunk.

So once we factor in prosody, the sequence in (1) appears to be on a par with the English example in (2) at least with respect to combinatorial possibilities. Unlike in English though, (1) displays noun-modifier order. But this is nothing exceptional given that such an ordering is commonly found cross-linguistically. This is for instance the case in Gungbe and most Kwa languages (cf. Aboh and Essegbey 2010). Consider the following example from Gungbe.

Ûn yi Djègàn dâxó. Ûn nù bià dón, bò mú kpédékpedé  
(3) 1SG go Djègàn dâxó 1SG drink beer there and drunk very  
'I went to *Djègàn dâxó*. I drank beer there and got very drunk.'

As the reader can see from the gloss, what is presented as 'word salad' in (1) is not only relatively close to English structure but shows a number of noteworthy similarities with

Gungbe. Furthermore, both languages transform English *beer* into *bia*. This comparison (p. 413) of (1) with patterns in English and Gungbe suggest that the word order and phonological changes attested in (1) are made available by UG. Furthermore, in the case of Hawaii, it can be argued that those shifting to English in the corresponding language contact setting also paid attention to its word order patterns.

### 17.2.3 Methodological Issues with ‘Simplicity’ and ‘Creole Typology’ Claims

Creole simplicity has been argued largely on the basis of hypothetical pre-Creole Pidgins as lingua francas with drastically reduced and unstable structures. Yet there isn’t, to the best of our knowledge, any documentation of any such pre-Creole Pidgin in the history of the colonial Caribbean. The lack of evidence for pre-Creole Pidgins with ‘massive structural reduction’ is admitted by McWhorter (2011:30–31, 70); cf. Alleyne (1971), Chaudenson and Mufwene (2001), Bakker (2003:26) and Mufwene (2008:ch. 3) who then argues that it is ‘the *linguistic* facts [that] strongly suggest that Atlantic creoles arose as structurally reduced pidgin varieties.’ (2001:31). What are these linguistic facts?

The discussion of Bickerton’s (2008) Pidgin example in (1) already suggests that diagnosing Creole simplicity on the basis of alleged Pidgins is not a straightforward task. Yet, McWhorter (2011:31–39) proposes four tell-tale signs for Creoles’ hypothetical Pidgin ancestry: (i) generalization of the infinitive; (ii) absence of copula; (iii) no case distinctions among pronouns; (iv) preverbal placement of the clausal negation marker. In the absence of theoretically-grounded definitions, these ‘Pidgin’ characteristics appear overly vague: how do we determine synchronically whether an ‘infinitive’ has been ‘generalized’? Is such generalization of the infinitive qualitatively different from the patterns observed for instance by Frei (1929) and Gougenheim (1929) where spoken varieties of French show a strong tendency toward invariant verbal forms? What are the morphosyntactic criteria for a ‘copula’? If we define ‘infinitive’ as a least inflected form of the verb, and ‘copula’ as an overt linking morpheme between subject and certain non-verbal predicates, Vietnamese as described in Dryer and Haspelmath (2011) would qualify as a quasi-Pidgin, notwithstanding its long history. And Vietnamese seems even more ‘Pidgin’-like than the earliest documented varieties of Creole in 18th-century Haiti (see later in this section). Even more problematic is the fact that the postulated criteria in (i)–(iv) are disconfirmed by data from Pidgins that have been documented outside the Caribbean (see e.g., Bakker 2003; Thomason 2007). For example, Kenya Pidgin Swahili, Pidgin Ojibwe, Taymir Pidgin Russian, and Fanakalo show tense inflectional morphology on verbs while Pidgin Ojibwe, Central Hiri Motu Pidgin, Arafundi-Enga Pidgin, and Lingala have morphological subject agreement (Bakker 2003:20–21). Bakker (2003:20)

further remarks that ‘some cross-linguistically uncommon inflectional affixes can also be found [in Pidgins—EA, MdG], such as reciprocal, and negative past.’ Such language-contact patterns seem like reflexes of their specific ecology and are evidence against a cookie-cutter approach to Creole formation whereby all (p. 414) Pidgins must look alike. The classic Creoles of the Caribbean, having emerged from languages that are often at the low end of the cline of inflectional richness, are unsurprisingly at the low end of that cline as well—and with even fewer inflectional affixes than their source languages, given the well-documented effect of second language acquisition on inflectional paradigms.

Be that as it may, the general cross-Pidgin typology contemplated by McWhorter has no empirical basis in the history of Caribbean Creoles. Based on the linguistic evidence from a sample of Pidgins worldwide and the observation that the morphology of Pidgin seems quite distinct from that of Creoles, Bakker (2003:24) considers the possibility that Creoles need not necessarily derive from Pidgins, noting in addition that ‘[t]here are no cases where we have adequate documentation of a (non-extended) pidgin and a creole in the same area’ (2003:26) (cf. Mufwene 2008:ch. 3 for the complementary geographical distribution of Pidgins and Creoles).

And we certainly have no evidence for any Pidgin in the history of HC—a ‘radical’ Creole in the sense of Bickerton (1984) where radical Creoles are postulated to have emerged on the basis of radically reduced Pidgin input. On the contrary, whatever evidence we have about the earliest documented varieties of HC contradicts claims about a drastically reduced Pidgin as an essential ingredient in Creole formation. Consider, say, McWhorter’s claims for lack of case distinctions in Pidgins. We have documentation of 18th-century Creole varieties in Haiti that show robust case distinctions in pronouns such as nominative 1sg *mo* vs. accusative 1sg *moé*, and nominative 2sg *to* vs. accusative 2sg *toué*. Here are two examples from Ducœurjoly (1802:353) with the French translations given there: HC: **To** va bay **moué** nouvelles / French: *Tu m’en diras des nouvelles* ‘You(+sg) will give me news’ vs. HC: **Mo** te byen di **toué**/ Fr: *Je te l’avais bien dit* ‘I had told you well.’ Such case distinctions in early HC are also reported in Anonymous (1811), Sylvain (1936:62f), and in Goodman (1964:34–36). The latter shows that similar case distinctions also apply to other French-based Creoles such as Louisiana Creole and Mauritian Creole. These morphological distinctions have now disappeared in contemporary HC, thus suggesting that Creole varieties closer to French (so-called ‘acrolectal’ varieties) must have been more prevalent in the earlier stages of Creole formation, but were later replaced by varieties structurally more divergent from French (so-called basilectal varieties). This is consistent with observations about the history of other Caribbean Creoles as in Jamaica (Lalla and da Costa 1989) and Guyana (Bickerton 1996). (See note 8.)

Saramaccan is another ‘radical’ Creole (actually, the most radical Creole according to Bickerton 1984:179 and a ‘prototypical’ one according to McWhorter 1998). Yet it too manifests case distinctions. In this case, we have a nominative vs. accusative opposition in the 3rd person singular: *a* ‘3sg nominative’ vs. *en* ‘3sg accusative’ (Bickerton 1984:180; Aboh 2006a:5). This contrast too is contrary to expectations based on McWhorter’s Pidgin criteria. In a related vein, Creoles with verbal paradigms that go beyond ‘a generalization of the infinitive’ are documented in Holm (2008) and Luís (2008) with data from Portuguese Creoles that manifest inflectional verbal suffixes.

(p. 415) Once Creoles are analyzed holistically, taking into account much more than the four isolated patterns arbitrarily chosen by McWhorter, it becomes doubtful that there ever was a structureless Pidgin in their history—especially one so reduced that it would have massively blocked the transmission of features from the languages in contact into the emergent Creole. Because the so-called Pidgins are human creations, we expect them to display structural properties that are made available by UG even if these properties may seem rare cross-linguistically. In this regard, the available literature on Pidgins provides a list of seemingly ‘exotic’ features, ‘exotic’ to the extent that they are lacking in many an ‘old’ language. These features include: evidential markers in Chinese Pidgin Russian, noun-class markers in Fanagalo, Kitúba, and Lingala, tense suffixes in Kitúba and Lingala, gender marking and agreement in the Mediterranean Lingua Franca, OSV and SOV word orders in Ndjuka Trio Pidgin, lexically and morphosyntactically contrasting tones in Nubi Arabic and Lingala, etc. (see DeGraff 2001b:250f for references).

As Pidgins are second languages for the majority of their speakers, they are susceptible to structural transfers from their speakers’ native languages. In effect, such an observation entails that there is, *a priori*, no such thing as an essential ‘Pidgin’ or ‘Creole’ type of language: the structural profile of each Pidgin will, to some degree, reflect the contingent ecology of its formation, including the structures of the respective languages in contact. The evidence in Thomason (1997a) and Bakker (2003) from Pidgins with non-European ancestor languages illustrates the ways in which the specific native languages of Pidgin speakers, including certain cross-linguistically rare structural properties of said native languages, do influence the structural make-up of Pidgins. Bakker as well relates the structural profiles of Pidgins and Creoles to the respective ecology of each language contact situation.

One corollary of these observations is that, be it called ‘Pidgin’ or ‘Creole’ or ‘language change,’ the eventual outcome of language acquisition in the context of language contact, carries along various properties from the languages in contact (Müller 1998; Hulk and Müller 2000; Müller and Hulk 2001; Notley, van der Linden, and Hulk 2007; Mufwene 2008:149–153; and references cited there). What specific properties are transferred to



the new variety depends on a variety of factors: socio-historical such as population structure and dynamics, linguistic-structural such as typological variation and markedness among the languages in contact, and psycholinguistic such as saliency and transparency of available features. Given that the languages in contact are usually assumed to be 'old' languages, the outcome of language contact will inherit various features from these, and this is exactly what we see in comprehensive surveys of language contact phenomena such as those cited in DeGraff (2001:250–259). Such instances of feature transfer can thus induce various increments of *local* complexity in the outcome of language contact (Aboh 2006b, 2009; cf. DeGraff 2009: 963n8).<sup>6</sup>

(p. 416) The available evidence about the complex ecology of language contact should also help lay to rest the now-popular 'fossils of language' scenarios (along the lines of Bickerton 1990:69–71, 181–185) that liken Pidgins to some hypothetical structureless protolanguage spoken by *homo sapiens'* immediate hominid ancestors. These scenarios also liken Creoles to the earliest and most primitive incarnation of modern human language. Firstly, the linguistic ecology of Pidgin speakers—in the midst of modern and complex human languages—is radically distinct from the ecology of our hominid ancestors who, presumably, did not have competence in anything that was structurally like the grammar of any human language. In any case, Pidgin speakers as modern humans have brains/minds very much unlike those of our hominid ancestors. So, even if we were to grant the validity of the Pidgin-to-Creole cycle, this cycle would have little bearing on the transition from, say, *homo erectus* protolanguage to *homo sapiens* language, a transition that most likely would have been accompanied by some reorganization of the brain from one stage to the next (Mufwene 2008:ch. 5).

These ongoing observations about the relationship between linguistic ecology and structural complexity suggest how important it is to beware the effect of sampling biases on Creole-simplicity claims. This sampling problem, already noted by Thomason and Kaufman (1998:154), Bakker (2003:26), Mufwene (2008:143–153), and Kouwenberg (2010), especially affects those claims that try to isolate Creole languages into one small corner of linguistic typology with grammars that fit a narrowly defined uniform structural template that, in turn, is placed at the bottom of some arbitrarily defined hierarchy of complexity (McWhorter 2001, 2011; Parkvall 2008; Bakker et al. 2011; etc.).

More concretely, let's consider the basic data and method in Parkvall (2008) which, in turn, has been adduced to support the claims in Bakker et al. (2011), McWhorter (2011), and others. The main argument is that Creoles are typologically distinct from non-Creoles, with grammars that are among the world's simplest grammars.

Parkvall's database consists of 155 languages as documented in the *World Atlas of Language Structures* (WALS, Haspelmath et al. 2005) alongside 34 Pidgins and Creoles—2 from WALS and 32 of Parkvall's own choosing, 18 of them from Holm and Patrick's (2007) *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars* (CCS). The data samples in Parkvall (2008) make room for a variety of confounds. Most of the Creoles in this study are historically related to typologically similar European lexifiers (typically Germanic or Romance) with relatively little affixal morphology and with few cross-linguistically rare features, and to African substrates (mostly Niger-Congo) that fall in narrow bands of typological variation as well, as noted in, e.g., Alleyne (1980:146–180); Thomason and Kaufman (1998:154); Bakker (2003:26); Mufwene (2008:136–153); and Holm (2008:319–320). This particular selection of Creole languages constitutes an extremely biased sample from the start. Furthermore, putting such a restricted and biased Creole sample side-by-side with the much larger set of non-Creole languages in WALS, languages that come from much more diverse stocks, both genetically and typologically, makes for a tendentious comparison (Kouwenberg 2010). In effect, such biased comparison is akin to the following: comparing the heights of 15-year-old basketball male players with the heights of males of other ages from the general population, then (p. 417) concluding *erroneously* that 15-year-old males are in general taller than males of other ages. Such conclusions are nothing but an artifact of sampling biases.<sup>7</sup>

A complexity metric that is based on such a small and arbitrary set of morphosyntactic distinctions, forms, and constructions can only impose a biased artificial ranking. As often noted (e.g., in Alleyne 1980), the languages in contact during the formation of these Creole languages are in the set-union of Germanic, Romance, and Niger-Congo and have relatively similar profiles—within a relatively small window of typological variation. Given such major overlaps across sets of ancestor languages plus the well-known effect of second language acquisition on phonological and morphosyntactic paradigms (Bunsen 1864; Meillet 1958:76–101; Weinreich 1958; etc.), it is thus not surprising that the Creole sample in Parkvall (2008) shows the similarities and the ranking that it does, owing to the particular 'bits' in his complexity metrics.

### 17.2.4 Conceptual and Theoretical Issues with 'Simplicity' and 'Creole Typology' Claims

Another fundamental theoretical flaw in the 'simplicity' literature on Creoles is the absence of a rigorous and falsifiable theory of 'complexity.' Consider, for example, Creole-simplicity claims where complexity amounts to 'bit complexity' as defined in

DeGraff (2001b:265–274). Such overly simplistic metrics consist of counting *overt* markings for a relatively *small and arbitrary* set of morphological and syntactic features (see, e.g., McWhorter 2001, 2011; Parkvall 2008; Bakker et al. 2011). In effect, any language's complexity score amounts to the counting of overt distinctions (e.g., for gender, number, person, perfective, evidentiality) and on the cardinality of various sets of signals (e.g., number of vowels and consonants, number of genders), forms (e.g., suppletive ordinals, obligatory numeral classifiers) and 'constructions' (e.g., passive, antipassive, applicative, alienability distinction, difference between nominal and verbal conjunction).

The problem is that such indices for bit complexity resemble a laundry list without any theoretical justification: '[T]he differences in number of types of morphemes make no sense in terms of morphosyntactic complexity, unless they tell us exactly how overt morphemes and covert morphemes interact at the interfaces, and how they may burden or alleviate syntactic processing by virtue of being overt or covert' (Aboh and Smith 2009 :7). The problem is worsened when bit-complexity metrics are mostly based on the sort of overt morphological markings that seem relatively rare in the Germanic, Romance, and Niger-Congo languages that were in contact during the formation of Caribbean Creoles.

Parkvall (2008) defines 'a complex language [as] a language with more complex constructions' (269) with 'an expression [being] more complex than another if it involves (p. 418) more rules' (265n1). However, the notions 'construction' and 'rule' only make sense as part of a larger linguistic theory. Compare, say, the 'passive construction' in Transformational Grammar vs. its counterparts (or absence thereof) in the Minimalist framework, and then compare such a 'passive construction' to its analogs in Generalized Phrase Structure Grammar and its descendants such as Head-Driven Phrase-Structure Grammar. The complexity metric in Parkvall (2008) is thus devoid of any theoretical content as regards 'constructions' or 'rules,' and its actual complexity evaluation seems quite arbitrary. Here it's worth stressing that the sense of complexity depends, not on the data *per se*, but on the particular counting method, or absence thereof. Let's take a closer look.

The complexity score for each language in Parkvall's data set is based on a total of 53 features and constructions. The terms 'feature' and 'construction' in Parkvall are used in a strictly superficial sense, that is, without any analysis of the 'rules' that may be involved in deriving, or accounting for the properties of, said features or constructions. There is, therefore, no way to systematically compute whether 'construction' X in language Y 'involves more rules' than some (analogous?) 'construction' W in language Z. In other words, the complexity metric in Parkvall (2008) is based strictly on presence vs. absence of 'feature' or 'construction' and the counting of overt 'forms' without any 'looking under

the hood' (so to speak) of these features, constructions, or forms. As this author states: 'for all the traits listed, I consider their presence (or the presence in larger numbers) to add to the overall complexity of a language' (Parkvall 2008:270). It should now be clear that the bit-complexity markers in Parkvall (2008) do not fall into any theoretically-motivated hierarchy of complexity. These features are simply a subset of those available through *WALS* (among those features that could be counted), plus certain features that Parkvall 'happened to have access to' (Parkvall 2008:273). Even more problematic is the fact that the selected bits belong to narrow domains of morphosyntax, thus ignoring other grammatical modules and the interaction therein (Kouwenberg 2010b).

The problem is more general: any complexity metric that is stipulated without any theory of complexity that is grounded in linguistic theory or in the psycholinguistics of language acquisition or processing can too easily become a self-fulfilling prophecy based on one's subjective expectation as to what should count as less, or more, complex (cf. Hawkins 2009). Indeed, the result of any comparison will depend on which bits are included in the comparanda. So bit-complexity is too unconstrained an approach, especially in light of the fact that language is a tightly-knit computational system with intricate channels of interaction across modules. If so, then one would *a priori* expect certain sources of complexity in any given module of grammar to, potentially, interact with other sources of complexity in other modules, with various increases or decreases in complexity resulting from interaction among modules. At this rate, no theory has yet been proposed that would adequately weigh the contributions of each module of grammar, plus the contribution of their mutual interaction, to a specific grammar's overall complexity (for additional comments, see DeGraff 2001b:265–274; Aboh and Smith 2009).

In the particular case of Parkvall's claims, one must ask: Why this bias in favor of overt morphological markings with much less weight accorded to other possible sources of (p. 419) complexity such as phonology, syntax, and semantics? 'To prove the claim [that Creole grammars are overall simpler than non-Creole grammars—EA, MdG], one would need to show that for every single subdomain of grammar (not just for an eclectic range of subdomains) all creoles score lower or equal to all non-creoles' (Deutscher 2009:250; also see Faraclas and Klein 2009 and Kouwenberg 2010:372 for related comments).

### 17.2.5 The Broken Pieces of 'Break in Transmission' Claims

Linguistics textbooks often cite Pidgin-based 'broken transmission' scenarios in their definition of Creole formation. These scenarios exclude Creole languages from the class of languages with normal ancestors, that is, from the class of 'genetic languages' (i.e., those that have emerged via 'normal transmission'). The opposition here is between

creolization viewed as ‘abnormal’ vs. language change viewed as ‘normal’ (Bickerton 1988). These scenarios bring to mind the early 20th-century debate opposing Hugo Schuchardt to Max Müller about ‘mixed’ languages: Is every language ‘mixed’ to some extent (Schuchardt’s position) or is mixed-ness an aberration (Müller’s position)? These are the opening questions in Thomason and Kaufman’s (1988) book on language contact and genetic linguistics. Their answer is to admit the existence of mixed languages such as Creoles, but to argue that these languages cannot be assigned any genetic classification. This echoes Taylor’s (1956) position that Creoles are genetically ‘orphans.’ They conclude that Creole languages fall outside the purview of the traditional comparative method.

We first argue that basic principles of the comparative method cast Caribbean Creole languages along the phylogenetic branches of their European ancestors (e.g., French in the case of HC). Similar arguments go as far back as Meillet (1951, 1958), Weinreich (1958), and others (see DeGraff 2009 for an overview). These results contradict the ‘break in transmission’ claims in Creole studies and allow us to uncover basic fallacies in the use of computational phylogenetic methods à la Bakker et al. (2011) for classifying Creole languages as a unique type.

### 17.2.5.1 Creole Languages Are *Bona Fide* Genetic Languages in the Scope of the Comparative Method

Let’s consider again HC, once characterized as a ‘radical’ Creole (Bickerton 1984). The core properties we discuss in this section are characteristics of the earliest (proto-)HC varieties and, thus, cannot be dismissed as post-creolization features, which would have entered the language only recently, via contact with French long after the Creole-formation period (DeGraff 2001b:291–294, 2009:940–941).<sup>8</sup>

(p. 420) In establishing the genetic affiliation of a non-Creole language, what is usually taken as confirming evidence consists of a system of correspondences between this language and some other languages—correspondences that suggest inheritance of related features from a common ancestor (or set of ancestors). This system of correspondences must meet some ‘individual-identifying’ threshold (Nichols 1996). That is, it must contain enough ‘language-particular idiosyncratic properties,’ or ‘faits particuliers’ in Meillet’s (1951, 1958) terminology, in order to reliably rule out chance correspondences, borrowings, homologous developments, and so forth.

HC offers robust evidence of straightforward correspondences with French, and similar evidence is straightforwardly available from dictionaries and descriptive grammars for other Caribbean Creoles. Müller et al. (2010) look at lexical similarity among half of the world’s languages, including Caribbean Creoles, and the latter show systematic correspondences with their European ancestors. Such correspondences between

Caribbean Creoles and their European ancestors generously meet Nichols's 'individual-identifying' threshold (for HC, see Fattier 1998, 2003; DeGraff 2001a, 2002, 2007, 2009). For example, we find the following arrays of 'faits particuliers' in HC with systematic correspondences vis-à-vis French, including the majority of affixes, the majority of paradigmatic lexical sets (including items from Swadesh lists), all grammatical morphemes, pronouns, deictic elements, and so forth. Here is a small sample to illustrate HC items that are inherited from French *with various degrees of modification*:

- All HC cardinal numbers are derived from French: *en* '1,' *de* '2,' *twa* '3,' *kat* '4,' ... *san* '100,' ... *mil* '1,000' ... from French *un, deux, trois, quatre ... cent ... mille ...*
- All HC ordinal numbers, including the suffix /-jɛm/ and its morphophonology (sandhi, suppletion, etc.), are derived from French: *premye* '1st,' *dezyèm* '2nd,' *twazyèm* '3rd,' *katryèm* '4th,' ... *santyèm* '100th' ... *milyèm* '1,000th' ... from French *premier, deuxième, troisième, ... centième, ... millième ...*
- All HC kinship terms are derived from French: for example, *frè* 'brother,' *sè* 'sister,' *kouzen* 'cousin,' *kouzin* 'cousin (feminine)' ... from French *frère, soeur, cousin, cousine ...*
- All color terms are derived from French: *blan* 'white,' *nwa* 'black,' *rouj* 'red' ... from French *blanc, noir, rouge ...*
- All body-part terms are derived from French: *cheve* 'hair,' *zòrèy* 'ear,' *je* 'eye,' *nen* 'nose,' *bouch* 'mouth,' *dan* 'tooth,' *lang* 'tongue' ... from French *cheveux, oreille, yeux, nez, bouche, dent, langue ...* (p. 421)
- All TMA markers are derived from French: *te* ANT, *ap* PROG, *FUT*, *ava* IRREALIS, *fini* COMPLETIVE ... from French *étais/était/été* (imperfect and participle of 'to be'), *après* 'after,' *va(s)* 'go+3sg/2sg+PRES', *finir/fini(s)* 'to finish' and its various participial and finite forms ...
- All prepositions are derived from French: *nan* 'in,' *pou* 'for,' *apre* 'after,' *anvan* 'before,' *devan* 'in front of,' *dèyè* 'behind' ... from French *dans, pour, après, avant, devant, derrière ...*
- All determiners, demonstratives, etc., are derived from French: *yon* 'a,' *la* 'the,' *sa* 'this/that' from French *un, la/là, ça*.
- All pronouns are derived from French: *m(wen)* 1sg, *ou* 2sg, *li* 3sg, *nou* 1pl, 2pl, *yo* 3pl ... from French *moi, vous, lui, nous, eux ...*
- All complementizers are derived from French: *ke* 'that,' *si* 'if,' *pou* 'for' ... from French *que, si, pour ...*

- Almost all HC derivational morphemes have inherited their distribution and semantics from French—with modification, of course: for example, HC *de-* as in *deboutonnen* ‘to unbutton’ and *dezose* ‘to debone’ from French *de-* which, like HC *de-*, has inversive and privative uses.
- HC morphophonological phenomena with French ancestry such as liaison phenomena in *an Bèljik* ‘in Belgium’ vs. *ann Ayiti* ‘in Haiti’; *de zan* /de zã/ ‘two years’, *twa zan* /twa zã/ ‘three years’, *san tan* /sa tã/ ‘one hundred years’ ... (cf. the pronunciation of the HC and French ordinal and cardinal numbers above; see Cadely 2002 for further examples of HC-French correspondences in phonology)

In Nichols’ terminology, these sets would count as individual-identifying ‘lexical categories with some of their (phonologically specific) member lexemes.’ As it turns out, HC even instantiates Nichols’ example of ‘the miniparadigm of *good* and *better* ... as diagnostic of relatedness.’ To wit, HC *bon* and *miyò* straightforwardly derive from French *bon* and *meilleur*. This HC example is all the more telling in that the last vowel in *miyò* (written *millor* in Ducœurjoly’s language manual (1802:330) reflects an Old and Middle French pronunciation of the word *meilleur* as *meillor* (Nyrop 1903, II:312), a pronunciation that is partly retained as *mèlyor* in Franco-Provençal dialects (Stich 2001), thus indicating that this paradigm was inherited from French and is not a late (‘decreolization’) feature of HC. This hunch is confirmed by Ducœurjoly’s (1802) analysis in his Creole language manual: he translates French *meilleur* as Creole *miyor* and French *filleul* ‘godson’ as Creole *fillol*. Compare the latter with *fillol* in 17th-century French (Nyrop 1899, I:158), *fiyòl/fiyèl* in contemporary HC and *filyol* in contemporary Franco-Provençal (Stich 2001:585). These sound-meaning correspondences between 17th-century French and HC are further confirmed by the fact that the French agentive suffix *-eur* /œr/ often maps in HC to an alternation between *-è* /ɛ/ and *-ò* /ɔ/ as in *vòlè/vòlò* ‘thief,’ *mantè/mantò* ‘liar,’ and *flatè/flatò* ‘flatterer’ from earlier pronunciations of French *voleur*, *menteur*, and *flatteur*, pronunciations still attested in Franco-Provençal (Stich 2001:221). These HC doublets thus reflect a phonological property of early varieties of French as spoken in colonial Haiti.

(p. 422) As carefully documented in Fattier (1988, 2002, 2003), there is a great variety of related morphophonological phenomena and lexical patterns that robustly show that Haitian affixes, alongside much else in Haitian grammar, were inherited early on from colonial varieties of French (see note 168). In other words, HC emerged with *bona fide* structural ‘faits particuliers’ suggesting its genetic affiliation with French. Such systematic correspondences are incompatible with the existence of an extraordinarily reduced (and affixless) Pidgin as the immediate ancestor of HC (Alleyne 1971; see Fattier

1998, 2003 and DeGraff 2001b:291–294 for further details on the origins of HC morphophonology).

As for the whole of the HC lexicon, the vast majority of HC morphemes (bound or free) are etymologically French. Fattier's (1988) six-volume dialect atlas and subsequent publications (notably Fattier 2002, 2003) establish the French stock of the HC lexicon and morphosyntax beyond any doubt. These Creole-vs.-French correspondences are attested from the earliest documentation of proto-HC, including the passage from Pelleprat (1665) quoted earlier and the very first language manual for 18th-century learners of Creole (Ducœurjoly 1802). In the latter, we find examples with French-derived TMA markers, pronouns, articles, verbs, nouns, and so forth, that are similar to contemporary HC. The Creole examples in Ducœurjoly are all the more striking in that they are given side-by-side with their French translations. Consider these two examples from Ducœurjoly (1802: 292): *Mo va tendre ly* 'I will wait for him' (translated as French *Je vais l'attendre*) and *Yo trape nion volor* 'They have caught a thief' (translated as French *On a attrapé un voleur*; note the suffix *-or* in *volor* 'thief' from the Old and Middle French, which is identical to Franco-Provençal *volor* in Stich 2001:1316).

These examples constitute additional evidence against proposals that French-derived HC morphemes would have entered the language as 'late borrowings.' If we do take the aforementioned individual-identifying evidence at face value, then we must conclude that HC, though it shows substrate influence (e.g., from Gbe as discussed in section 17.3), is indeed genealogically related to French. This conclusion, in turn, entails that HC's French-derived lexemes are 'native and cognate until shown otherwise' (again, borrowing Nichols' terms).

Once such genealogical relatedness is taken as established ('relatedness by descent' as in the better studied *Stammbaumtheorie* branches such as in the evolution of Latin to Romance), it becomes clear that 'break in transmission' scenarios à la Bickerton, Thomason, and Kaufman, McWhorter, Bakker et al., and so forth, cannot hold. Then again, one may argue that Caribbean Creoles such as HC show more 'significant discrepancy' between their lexical- vs. grammatical-correspondences vis-à-vis their respective lexifiers than French does vis-à-vis Latin (Thomason and Kaufman 1988); argument to that effect is already undermined by Meillet's observation long ago that French is of a grammatical type distinct from Latin, even though French can be considered as duly descended from Latin according to the comparative method. When we compare HC and French using the structural parameters identified by Meillet (1958:148) to show that French 'fall[s] into a typological class that is quite remote from the structural type represented by Latin,' our comparison shows that HC and French, especially colloquial (p. 423) 17th- and 18th-century varieties, are typologically closer to



each other than French and Latin are—with respect to word order, case morphology, definite determiners, and so forth. (These arguments are taken up in more detail in DeGraff 2001b, 2009:919–922 and some of the references cited there.)<sup>9</sup>

By the same token, the notion of global Creole simplicity falls apart: though French may look ‘simpler’ than Latin on the surface (e.g., absence of nominal declensions) it developed structural devices (e.g., articles) that are absent in Latin. Similarly HC developed structural devices (e.g., a TMA system, focus marking, predicate copying, etc.) that are by and large absent in French, even though some of the basic ingredients in these innovations (e.g., the distribution and semantics of individual TMA markers) have straightforward ancestry in French (more on this in section 17.3).

### 17.2.5.2 Nonnative Acquisition in *Stammbaumtheorie* Genetic Branches

One other argument that has often been leveled against classifying Caribbean Creoles as Germanic or Romance languages is that genealogical relatedness among diachronically well-behaved Indo-European (IE) languages entails unbroken sequences of native language acquisition (‘NLA’). Recent exponents of this position include Ringe et al. (2002: 63) and Labov (2007:346).

It seems that this NLA-based position assumes that IE languages would have all evolved via unbroken NLA. But this is contradicted by the crucial role of *second* language acquisition by adults in, say, the emergence of Romance languages from Latin and in other IE cases (i.e., English) where language contact (thus, by definition, non-NLA) played a key role. In these cases (e.g., in the evolution of French from Latin), we do find, like in the HC and other Caribbean Creole cases, robust systems of lexical correspondences between the outcomes of non-NLA acquisition, on the one hand, and, on the other hand, the target/ancestor language. Like in the Caribbean Creole cases, these correspondences include paradigms of bound morphemes, paradigmatic lexical sets, and other systems of *faits particuliers* that would have been inherited from the ancestor language through instances of transmission that include non-native language acquisition by adults.

The emergence of French from Latin, initially in the context of language contact among adults, seems a good example of non-NLA in genealogical branches of Indo-European. In the history of French, like in the history of Haitian Creole, we do find documentation of language contact and second language acquisition by adults, yet we do not assume that such nonnative acquisition would exclude French from the Romance family and we do not speak of Latin lexemes being ‘borrowed’ into early French. In a related vein, we know of English dialects that descended from the English learned imperfectly by Scandinavian settlers (Kroch et al. 2000; Ringe et al. 2002). As far as (p. 424) we can tell by looking at

their phonology, lexica, and morphosyntax, such dialects of English still count as West Germanic—notwithstanding the broken sequence of NLA in their history, due to language contact.

Once we consider the aforementioned French and English cases at face value, alongside the HC case as analyzed here, then genealogical relatedness cannot be taken to be strictly coextensive with unbroken NLA sequences. French counts as a Romance language, just as English and its related dialects count as Germanic. Therefore, NLA cannot count as a deciding factor for genealogical relatedness. After all, the comparative method is about the correspondences of linguistic forms. The comparative method is not about the history of the speakers of the corresponding languages (Weinreich 1958:375), which is precisely why Meillet (1951, 1958) and Nichols (1996) warn us about correspondences that do not suggest genetic kinship (see also Mufwene 2008). Similar issues arise in the many cases of ‘indigenized varieties’ of Germanic and Romance languages in postcolonial contexts in Africa and Asia. Those varieties as well can be reasonably considered Germanic and Romance (e.g., French in West Africa or English in West Africa and India) even though they are often learnt as nonnative second languages. (See DeGraff 2009:923–929 and, especially, Weinreich 1958, Mufwene 2001:ch. 4, 2004, and Campbell and Poser 2008 for comprehensive arguments against the use of non-linguistic factors in evaluating genetic relatedness.)

The facts mentioned here in favor of the genetic affiliation of Haitian Creole (HC) with French also count as counter-evidence to other exceptionalist views on Creole formation such as Lefebvre’s Relexification Hypothesis (see DeGraff 2002, 2009 for full-fledged details of this argument.)

### **17.2.5.3 Theoretical and Empirical Issues in Computational Phylogenetics in Creole Studies**

The most prominent exemplar of such methods is the 2011 paper by Bakker et al. claiming that ‘Creoles are typologically distinct from non-Creoles.’ They use the sort of computational phylogenetic algorithms described in Dunn et al. (2008). The application of these algorithms to Creole languages is riddled with empirical and conceptual problems.

One foremost challenge is the circularity and data problems in the definition of Creoles. Bakker et al., ‘in order to avoid circularity in [their] definition,’ consider a socio-historical definition, that is, Creoles as ‘nativized or vernacularized developments of pidgins, which are makeshift languages used in some contact situations’ (p. 10). But, in absence of documentation for such Pidgins in the history of the classic Caribbean Creoles, this definition triggers the ‘data problem’ described in Bakker (2003:26): ‘There are no cases where we have adequate documentation of a (non-extended) pidgin and a creole in the

same area.' In this regard, Mufwene (2008:34–35) provides a map showing the 'geographical complementary distribution between the territories where creoles developed and those where pidgins emerged.' Given the complementary distribution of Pidgins and Creoles across the world, Bakker et al.'s treatment is circular since they attribute the putatively simple properties of Creoles to their emerging from (p. 425) hypothetical, but undocumented, Pidgins qua 'simplified forms of interethnic makeshift languages [that] were insufficient for communication' (2011:36).

Another set of methodological and theoretical problems concern Bakker et al.'s assumptions about the use of computational methods in establishing historical relatedness among languages. Though Bakker et al. rely on Dunn et al.'s framework, they misapply it to their sample of Creole languages.

Firstly, Dunn et al.'s study of the isolate Papuan languages in Island Melanesia gives clear methodological priority to the classic comparative method with its vocabulary-based sound-meaning correspondences. The comparative method remains the 'gold standard for historical linguistics' to be applied whenever cognate sets can be reasonably established within the limited time depth of the comparative method, which is estimated at some 10,000 years (Dunn et al. 2008:710–712; cf. Wichmann and Saunders 2007:378). In fact, Dunn et al. first calibrate their computational structure-based comparison of isolate Papuan languages against the prior vocabulary-based results obtained by the comparative method's 'gold standard' as applied to the Oceanic languages of the same area. Then, and only then, do they apply their structure-based computational methods to their sample of Papuan languages (Dunn et al. 2008:734). The reason why these language isolates are outside the scope of the comparative method is that they 'separated so long ago that any surface traces of cognacy have been eroded' (Dunn et al. 2008:712).

One major issue with Bakker et al.'s application of these structure-based methods to Creoles is that the latter are claimed to be among the world's youngest languages, certainly younger than the allowable 10,000-year time depth for the comparative method. Furthermore Creole languages are certainly not language isolates by any means. There's plenty of lexical evidence available to trace these languages' genealogical classification to their European sources, notwithstanding borrowings through language contact as in the documented history of other Indo-European languages. As already mentioned, some of the evidence for genealogical classification is even available in texts dating back to the early emergence of these languages (e.g., Pelleprat 1665; Ducoëujoly 1802).

It's been claimed that the networks produced by computational phylogenetic methods are 'completely objective and thus not influenced by any preconceptions and prejudices' (Bakker et al. 2011:12). As we already pointed out with regard to the bit-complexity method in Parkvall (2008), the outcome of the phylogenetic computations is,

among other things, a direct result of what features (or ‘characters,’ see Nichols and Warnow 2008) are chosen for the comparison (see also chapter 16). Of course, some finite choice has to be made when comparing languages, and the relevant sets for potential features for any language are not exhausted by available reference grammars. But the key issue here is how to ensure that the initial choices do not undermine the results of our comparisons: on what theoretical basis are small sets of features selected from specific domains of grammar? How do we ensure that certain domains (e.g., isolated areas of phonology and morphosyntax) are not assigned higher priority than other domains (e.g., the lexicon or various areas of syntax, semantics, and discourse)? The point is that, (p. 426) given the availability of a vast range of structural features to compare between any two languages, the choice of any relatively small set of features is certainly open to ‘preconceptions and prejudices.’ More generally:

[t]he choice of characters for use in a phylogenetic analysis is of great importance, and has often been one of the main issues involved in critiquing a phylogenetic analysis: which characters did the authors use, and what are the consequences of that choice?

(Nichols and Warnow 2008:769)

Nichols and Warnow’s review should be a must-read for any creolist interested in computational linguistic phylogeny. Their review concludes that ‘data selection (both of characters and languages) and the encoding of the character data have the potential to significantly impact the resultant phylogenetic estimation.’<sup>10</sup>

Dunn et al. (2008) are aware of these and related issues, and they enlist the following strategies, all of which are lacking in Bakker et al.’s comparison of Creoles with non-Creoles:

- (i)** Dunn et al.’s phylogenetic computations are based on ‘the combination of structural features from different domains of a grammar (phonology, morphology, syntax, semantics)’ (715).
- (ii)** ‘As many abstract structural features from as many parts of the grammar as possible should be investigated’ (716).
- (iii)** They used 115 features for a sample of 22 Papuan languages (pp. 728, 730), with features ranging over phonology, morphology, and syntax, with the goal of providing ‘a large body of basic features for each language, which together give a broad typological profile, regardless of whether any given feature seems typologically significant. The resultant phylogenies are thus not likely to reflect a sampling bias.’ (Also see Wichmann and Saunders 2007:383 on how areal effects introduce noise in the data when the comparison is based on small set of features.)

(iv) They ‘avoid the charge of “hand-picking” features by including in [their] sample the widest feasible range of noninterdependent typological phenomena’ (p. 733; cf. Wichmann and Saunders 2007: 376, 382, 385n7).

These methodological caveats are all flouted in Bakker et al.’s comparison of Creole versus non-Creole languages, starting with the size and the nature of the features used in the comparison. Bakker et al. use features from two previous publications: Holm and Patrick’s (2007) *Comparative Creole Syntax (CCS)* and Parkvall’s (2008) aforementioned study on Creole simplicity. One major problem is that the 97 features (p. 427) in *CCS* are massively interdependent; Holm (2007:xi) warns the reader that the *CCS* feature set is indeed ‘redundant’ (more details in the next paragraph). As for the Bakker et al. study that is based on Parkvall (2008), it uses only 43 features from limited areas of grammar to compare some 200 languages. Compare with Dunn et al.’s use of 115 features for their 22 Papuan languages. Like the *CCS* features, the features in Parkvall (2008) violate the ban against interdependency and are taken from narrow and superficial areas of morphosyntax—mostly having to do with overt morphology, as noted in Aboh (2009) and Kouwenberg (2010, 2012) (see section 17.3.3).

Let’s take a closer look at the Bakker et al. study that is based on the 97 *CCS* features. It is claimed that these features make the Creole languages in the Bakker et al. sample ‘stand out.’ But the *CCS* features were explicitly chosen because they were viewed as a good candidate set for pan-Creole features if any set could exist. So it is no surprise that such features would make these Creoles ‘stand out’: the *CCS* features had been cherry-picked with the express goal of trying to group Atlantic Creoles and Niger-Congo languages together in contrast with the Creoles’ European superstrate languages (Holm 2007:vi). Bakker et al.’s choice of the *CCS* features is therefore in contradiction with Dunn et al.’s caution against sampling biases.

There’s yet another problem in Bakker et al.’s use of *CCS* features. Holm (2007:vi) and Patrick (2007:xii) make it very clear that the feature values in each and every Creole language in the *CCS* sample only makes sense as part of ‘a set of interconnected systems.’ Things become even trickier when *CCS* readers are warned at the outset that rating particular features across *CCS* languages suffers from ‘varying definitions and operationalizations’ (Patrick 2007:xii).

A good example of this is in DeGraff’s discussion of HC in Holm and Patrick (2007), where he notes that features such as ‘adjectival verbs’ and ‘copula’ cannot be taken as stable properties that can be compared with a ‘+’ (for presence) or a ‘-’ (for absence) across languages, exactly because their underlying syntax is part of larger intricate systems (see, e.g., DeGraff 2007:103–104, 112–115). To illustrate: in HC it is the ‘absence’ of a

copula in certain constructions (e.g., those that involve adjectives as in *Jezila bèl* 'Jezila is beautiful') that may give the impression that *bèl* is a verbal predicate (e.g., on a par with the verb *danse* as in *Jezila danse* 'Mary has danced'). However, there are distributional tests that do distinguish verbs and adjectives in HC and that show that *bèl* is adjectival, not verbal. Contrast, for example, these HC comparative constructions as discussed in DeGraff (2007): *Jezila pi bèl pase Mari* 'Jezila is more beautiful than Mary' vs. *Jezila mache plis pase Mari* 'Jezila has walked more than Mary' (cf. \**Jezila plis bèl pase Mari* and \**Jezila pi mache pase Mari*). These examples show two key properties of the features 'copula' and 'adjectival verb': (i) they are interdependent; (ii) they only make sense within the larger syntax of the given language. Indeed, the distinction between verbs and adjectives in HC can only be made on diagnostics that are, to some degree, internal to HC grammar. (See Seuren 1986 for related arguments about verbs vs. adjectives in Sranan.)

Another example relates to the TMA system. Among the 97 features in Holm and Patrick (2007), 40 of them are related to the TMA and verbal system, with 24 of them (p. 428) related to temporal interpretation; among the other features, some 20 are related to the nominal system (Véronique 2009:153–154). As Holm and Patrick's (2007:vi) correctly put it:

... the logic of this set of systems is clear only in terms of itself as a totality. For this reason it is unenlightening to compare, for example, a particular tense marker in twelve different Creoles without also explaining how this tense marker fits into the overall verbal system of each language.

And comparing particular markers without explaining their functions within each language is exactly what Bakker et al. (2011) do, not for 12, but for some 200 languages. Not only do their claims on Creole type go against the spirit of Holm and Patrick's work, but they also flout Dunn et al.'s (2008) caveat that non-interdependent features from diverse domains of grammar are a *sine qua non* for reliable computational phylogenetics.<sup>11</sup>

Our point, thus far, is that no matter how powerful the computational algorithms that underlie these phylogenetic methods, the initial choice of features for comparison as well as the size and nature of the samples of features and languages will exert a key influence on the outcome of said comparisons and the reliability thereof. In this regard, it is not surprising that the Creole sample in Bakker et al. (2011) shows certain typological similarities, especially given the sampling problem that Bakker (2003) warns against. As in Parkvall (2008), the comparisons in Bakker et al. (2011) are based on samples where most of the Creoles have Germanic or Romance lexifiers or Niger-Congo substrates. For

example, of the 18 Creole languages in Holm and Patrick (2007), 15 have Germanic or Romance lexifiers and 14 include Niger-Congo languages among their substrates. The focus on narrow morphosyntactic domains for broad typological claims makes the sampling bias even worse. For example, the TMA systems of Caribbean Creoles are all influenced by typological tendencies in Niger-Congo with certain analogs in Romance and Germanic (see section 17.3). So what we may be dealing with here is the sort of *Sprachbund* phenomena that are also found across distinct phylogenetic branches (e.g., in the Balkans). In the Caribbean Creole cases, this can be termed a ‘Trans-Atlantic *Sprachbund*’ (Aboh and DeGraff 2014).<sup>12</sup>

(p. 429) Let’s now ask a broader question: why would Bakker et al. choose structural-typological features when establishing phylogenetic trees for Creole languages even though it is lexical and morphophonological features that have been used to establish phylogenetic relatedness at much greater time depth than in Creole formation, as in the history of Indo-European? Such choice introduces a methodological double standard from the start. Indeed the most successful cases of computationally-derived phylogenetic trees have been drawn on the basis of *vocabulary*, not structural-typological features (this is noted in, for example, Bakker et al. 2011:13; see also Mufwene 2003 for related issues). The reliability of phylogenetic trees is much more fragile when the comparison is based on structural-typological features alone (Nichols and Warnow 2009; Donohue et al. 2011; see chapter 16 for a defense of parameter-based phylogenies). Bakker et al. consider the structure-based phylogenetic results in Dunn et al. (2005) ‘quite remarkable’ (13) and they further argue that ‘structural features may be safely used for evolutionary studies’ (p. 21). But what is not noted is that the structure-based results reported in Dunn et al. (2005) about the Papuan languages have been argued to be incompatible with previously established subgroupings (e.g., Nichols and Warnow 2009:799). Nichols and Warnow attribute the lack of accuracy in Dunn et al. (2005) to their exclusive use of structural-typological features.

Donohue et al. (2011) have argued that phylogenetic relationships cannot be reliably established in absence of morphophonological and lexical correspondences (p. 378). They thus argue in favor of the classic comparative method and against the results of Dunn et al. (2008), and they come to the more cautious conclusion that ‘there is not a direct link between the typology-constructed tree and the linguistic phylogeny, but rather ... both of them covary according to linguistic geographic distance’ (p. 373). From this perspective, ‘the analysis of abstract typological features is a valuable detection tool in that the results serve as an accurate proxy for distance, rather than a proxy for phylogenetic results such as would result from the application of the comparative method to a group of languages’ (p. 374). Their fundamental insight (p. 378) is that it’s ‘linguistic geography,

rather than phylogenetic identity [that] determines typological clusters' in the sort of networks that are produced in Dunn et al. (2008).

These conclusions are compatible with our own observations here that the oft-noted structural similarities among Caribbean Creoles are of the quasi-*Sprachbund*<sup>13</sup> type due to their origins in contact among overlapping sets of Niger-Congo and Indo-European languages of relatively similar types, not the result of any catastrophic diachronic event such as radical pidginization.<sup>14</sup> The controversy surrounding phylogenies based on (p. 430) structural-typological features is unsurprising in light of Meillet's caveats that genetic affiliation does not track typological similarity (as in the famous Latin-to-French case already discussed in section 17.2.5.2).

These caveats highlight the recurrent double standard that has long been applied to Creole languages with claims such as:

Creoles typically show lexical continuity with their lexifiers, but only limited continuity in their structural make-up, making it strictly seen impossible to consider a creole language a genetic descendant of its lexifier.

(Bakker et al. 2011:14)

It seems worth repeating that French as well shows vis-à-vis Latin 'only limited continuity in [its] structural make-up.' Yet French is a genetic descendant of Latin even if it belongs to 'a typological class that is quite remote from the structural type represented by Latin' (Meillet 1958:148). Therefore, in establishing phylogenetic relatedness, typological-structural features cannot be taken to override the sort of lexical and morphophonological correspondences that can be reliably established in the history of Creole languages (*contra* Bakker et al. 2011:21). In the case of Creole languages (languages whose emergence is much more recent than that of most other Indo-European languages), the origins of the vocabulary items and the related cognates and morphophonological correspondences are relatively straightforward, so there seems to be no need to shy away from them.

### 17.3 A Null Theory of Creole Formation (NTC)

Let's summarize our essay so far: Creole languages have traditionally been excluded by fiat from the scope of the comparative method in historical linguistics—they are thus considered to lie outside *Stammbaumtheorie* (this dogma is noted in for example Noonan 2010:60). This exclusion is based on the belief that Creoles emerged through a break in



linguistic transmission and represent an exceptional case of language (p. 431) evolution: they emerged in the absence of normal linguistic input. Given this view, various exceptionalist theories postulate a catastrophic emergence for Creole languages as an explanation of their supposedly simple structural make-up. Our ongoing critique has suggested that these claims are all mistaken, and we argue that Caribbean Creoles duly fall in the scope of the comparative method as languages genealogically affiliated with their European ancestors, notwithstanding the documentation of quasi-*Sprachbund* phenomena due to the pervasive influence of overlapping substrate and superstrate languages such as Germanic, Romance, and subsets of Niger-Congo languages (see note 13). Such quasi-*Sprachbund* phenomena are similar to borrowing patterns that are now well documented in established *Stammbaumtheorie* branches such as Indo-European. On the latter, see Ringe et al. (2002); Nakhleh et al. (2005).

In this section we sketch a framework for a null theory of Creole formation, a theory that we believe is empirically more adequate than the popular exceptionalist claims surveyed in sections 17.1 and 17.2. Our null theory does away with any *sui generis* stipulation that applies to Creole languages only. Instead it is rooted in basic assumptions and findings about UG, that is, assumptions and findings that apply to *all* languages and to how learners acquire these languages. In this approach, the emergence of any new language or language variety in the context of language contact sheds lights on the interplay of first and second language acquisition as new grammars are built from complex and variable input. The effects of this interplay are similar across familiar cases of Creole formation such as the creation of HC and familiar cases of language change such as in the history of French and English. Our null theory undermines various traditional claims about 'Creole simplicity' and 'Creole typology' whereby Creoles are considered exceptional languages of the lowest complexity. Our approach also makes for a better integration of Creole phenomena to our general understanding of the cognitive bases for language change (see DeGraff 1999b, 2002, 2009; Mufwene 2001, 2008; Aboh 2006b, 2009, 2015) for further discussion, see also chapter 18.

### **17.3.1 The Rationale from a Language Acquisition Perspective Informed by History**

To begin with, we depart from the common view that Creoles developed as a consequence of a catastrophic break in transmission or radical transmission failure. The history of Haiti, for instance, suggests that a key factor in Creole formation was the role of language acquisition among a multilingual community involving speakers with different profiles—a situation typical of other cases of contact-induced language change

as in the history of Germanic or Romance. So we need to first take stock of these learners' profiles and their role in Creole formation.

Let's recall Pelleprat's aforementioned quote about the 17th-century French colonial Caribbean, repeated here for convenience. This quote shows that there was no break in the transmission of French as long as we keep in mind that second language acquisition (p. 432) by adults with varying exposure to the target L2 was also operative in the early history of, say, Romance languages:

We wait until they learn French before we start evangelizing them. It is French that they try to learn as soon as they can, in order to communicate with their masters, on whom they depend for all their needs. We adapt ourselves to their mode of speaking. They generally use the infinitive form of the verb [instead of the inflected forms—EA, MdG] ... adding a word to indicate the future or the past.... With this way of speaking, we make them understand all that we teach them. This is the method we use at the beginning of our teaching.... Death won't care to wait until they learn French.

(Pelleprat 1655 [1965, 30–31], our translation)

A couple of noteworthy paradoxes emerge here that are often ignored in Creole studies. Without calling it 'Creole,' Pelleprat introduces the enslaved Africans' 'mode of speaking' French as a variety that arises as these Africans try to learn French 'as soon as they can.' Pelleprat's remarks also suggest that biblical teaching and presumably other sorts of instruction were carried out in this emerging variety: 'with this way of speaking, we make them understand all that we teach them. This is the method we use at the beginning of our teaching ... Death won't care to wait until they learn French.' It appears from this citation that the mode of speaking, which was to be subsequently referred to as 'Creole,' was used as the language of instruction and was therefore accepted in what could be considered, back then, relatively formal contexts. This hypothesis is further corroborated by official declarations by no less than Napoleon Bonaparte in 1799 in Saint-Domingue Creole as reported in Denis (1935:353–354) (see Aboh 2015 for discussion):

Paris, 17 Brimer, an 10, République francé, yon et indivisible ...

Consuls la République Francé a tout zabitans Saint-Domingue ...

Qui ça vous tout yé, qui couleur vous yé, qui côté papa zote vini, nous pas gardé ça: nous savé tan seleman que zote tout libre, que zote tout égal, douvant bon Dieu et dans zyé la République. Dans tan révolution, la France voir tout plein misère, dans la même que tout monde te fere la guerre contre Français. Français

levé les ens contre les otes. Mes jordi là tout fini, tout fere paix, tout embrassé Français; tout Français zami; tout hémé gouverneman, tout obéi li ...

Signé: Bonaparte

[Our translation: Paris, 17 Brumaire, year 10 French Republic, one and indivisible....

From the consuls of the French Republic to the entire population of Saint-Domingue.

Whoever you are, whatever your skin color, wherever your ancestors are from, that does not matter to us: we only know that you are all free and all equal before God and before the Republic. During the Revolution, France experienced a lot of suffering because every other country fought against the French. The French were fighting each other. But today, all of that is over. All people have made peace. All people have (p. 433) embraced the French. All the French people are friends, they all love and obey the government.

Signed: Bonaparte ]

In a related vein, Ducœurjoly's (1803) *Manuel des habitants de Saint-Domingue* contains a language primer intended for newly arrived colonists. There it is explicitly stated, in the table of contents, that this primer is the first 'Creole dictionary, with conversations translated in French and in Creole to give an idea of this language and to make oneself understood by the Blacks.' This objective suggests that the Creole was used by almost everyone in the colony including the white colonists. Given this function of the Creole as a *lingua franca*, the enslaved Africans arriving in the colony must have considered the Creole as language acquisition target, possibly in addition to other local varieties.

Taken together, these observations, from Pelleprat to Ducœurjoly, are inconsistent with the 'break in transmission' assumption often evoked in creolistics, whereby Creole languages would emerge out of a structureless 'macaronic pidgin' (Bickerton 1981). As the alert reader would have noticed, a structureless Pidgin can hardly serve as means for proper biblical instruction and political propaganda. Instead, what these documents suggest is that the colonial precursor of HC was used in both everyday and (quasi-)official contexts by the Church, by government officials and by colonists as the one language accessible to most inhabitants of the island. Given the fact that the Creole was used as language of instruction and for political propaganda as well as commercial and plantation activities, it was in all likelihood spoken by members of the emerging Creole society in increasing numbers and with diverse degrees of fluency—both by speakers of the substrate languages and by speakers of the local varieties of French. And this popularity

of Creole in colonial Haiti was explicitly noted by the anonymous author of *Idylles et Chansons ou essais de Poésie Créole par un Habitant d'Hayiti*, who defined the 'Langue Créole' as a sort of 'corrupted French' that was 'generally spoken by the Blacks, the Creoles [i.e., the Caribbean-born—EA, MdG] and by most of the colonists in our islands in the Americas' (Anonymous 1811:2). Similar remarks about the widespread use of early Creole varieties are noted in Hazaël-Massieux (2008:42–43).

In a related vein, and as already documented in DeGraff (2001:251), quoting Schuchardt, '[t]he slaves spoke the creole not only with the Whites but also among themselves while their mother tongue was still in existence, the latter being moreover constantly revived to some extent by the continual immigration from Africa.' Historical records indicate that a significant group of such African speakers were the Gbe people from the Slave Coast (see Singler 1996; Aboh 2015). These Gbe speakers were also L2 speakers of the Creole. Like adult language learners everywhere, their native languages would certainly influence their approximation of the target language (see chapter 13). Thus arises the well-documented substrate influence in HC—and in other Caribbean Creole languages (see Mufwene 2010 for a critical overview).

Beside these two types of L2 speakers, including Europeans and Africans, there were the locally-born (i.e., the so-called Creole) children. White children probably grew up bilingual in the Creole and the local variety of French, as is evident with the Beke (p. 434) population today in Martinique, while some black children must have grown up bi- or multilingual in the Creole and some of the Niger-Congo languages. Some black children, alongside mixed race children (so-called 'mulattoes') and white children also spoke the local variety of the colonial language. These bi- and multilingual speakers would contribute their diverse innovations to the mix, as influenced by their parents' heritage languages. Finally, we should not forget those enslaved African children who were brought to the colony relatively young and became early L2 learners of the then-incipient Creole varieties.

As suggested in DeGraff (2002, 2009), the interaction between these different types of learners in the early history of HC created an 'L2–L1 acquisition cascade' whereby newly arrived L2 learners and newly-born L1 learners were exposed to PLD from a mix of Creole varieties—PLD that are partly fed by previously arrived L2 learners (often by the 'seasoned slaves'), alongside older native Creole speakers. In terms of this hypothesis, L1 learners contribute to the emerging Creole by, among other things, setting the relevant parameters of their own, native and stable, idiolects on the basis of patterns that, in turn, are influenced by L2 learners' reanalyses and innovations (see chapters 12, 13, and 18). It is thus that the early norms of Creole varieties would emerge and become more and more stable as more and more native Creole speakers would converge on overlapping sets of

parameters for their Creole idiolects.<sup>15</sup> We now look at specific examples of such patterns arising through this L2-L1 cascade.

This hypothesis is compatible with views on contact-induced diachronic changes. Meisel (2011:125f), for instance, argues that:

As far as the setting of parameters is concerned, ‘transmission failure’ is unlikely to happen in simultaneous first language acquisition. Only in successive acquisition of bilingualism might L2 learners fail to reconstruct the target grammar based on the information provided by the primary linguistic data. Alternatively, monolingual or bilingual children may develop a grammar distinct from that of the previous generation if they are exposed to an L2 variety of the target language.

In addition to substrate-influenced patterns in Creole formation, we also need to think about the fact mentioned in section 17.2 that HC grammar, including its morphophonology, syntax, and semantics, shows systematic correspondences with French, even in the domain of affixes and inflectional heads such as TMA markers, complementizers, determiners, and the like. Recall that these correspondences include diverse morphological idiosyncrasies such as affix distribution, the semantic ambiguity of certain affixes, sandhi phenomena, and morphological suppletion. As we consider the various categories of learners who contributed to the creation of HC, we may wonder who among those learners were the main channels for the inheritance of French-derived forms and structures in the emerging Creole.

(p. 435) In colonial Haiti (i.e., Saint-Domingue) the most likely groups to create early Creole varieties with robust lexical and structural French inheritance were not the field slaves on large and segregated plantations, but a socio-economically privileged group constituted by Africans on ‘homesteads’ and by others with direct and regular contact with speakers of French varieties (e.g., African and European L2 speakers of the local varieties). In this privileged group the Creole people of Saint-Domingue would stand out, with Creole taken in its ethnographic sense (i.e., locally born of non-indigenous parents). Having been born in Saint-Domingue, these Creoles were in the best position to attain native fluency in the emergent Creole varieties, alongside any available French varieties (DeGraff 2009:940–941). Also important is the fact that among the non-white population, the Creoles, many of whom were of mixed European–African descent, were generally the ones with the highest social capital. This advantage was quantified by Moreau de Saint-Méry as worth ‘a quarter more than that of the Africans.’ So their native Creole varieties, including the French inheritances therein, would have had enough prestige to subsequently spread through the population at large into communal norms of Creole speech—or, more accurately, communal norms of French-derived varieties that were

increasingly perceived as autonomous from French and symbolic of the new 'Creole' community and its emerging identity.

Alongside substrate influence and superstrate inheritance, we also need to take into account novel aspects of Creole grammars, aspects without analogs in any of the ancestor languages. Here, some fundamental considerations are in order about the general role of language acquisition in language change (see also chapter 18). Given the fact that idiolects are like linguistic fingerprints, the I-languages of any learner's models, in Creole formation and everywhere else, will unavoidably be distinct pairwise. So it's logically impossible for any learner to replicate the I-languages of *all* available models (DeGraff 2009:906–907, 914–916). Therefore, language acquisition always makes room for innovations that, in turn, either endure beyond the innovators' own I-languages or have a transient short life restricted to only the span of said I-languages or the early stages therein. Endurance or transience of any given innovation depends on whether or not these innovations, which originate in individual learners' I-languages, subsequently spread into new communal norms via their widespread language use alongside further instances of language acquisition—acquisition that is relatively *convergent* vis-à-vis the particular innovation or a substantial component therein. As the innovation is spread through larger groups of speakers, there's also the possibility that its domain of application is extended to larger and larger linguistic environments. The spread of innovations is sensitive to complex ecological, socio-historical and linguistic-structural factors (see DeGraff 2009:899–914 for some discussion, including innovation via grammaticalization, and some important caveats against the conflation of spread and innovation in analyses of Creole formation and language change; cf. Hale 1998).

The key point here is that the rise of innovations through language acquisition is not a hallmark of Creole formation alone. Innovations are part and parcel of language acquisition throughout our species (see Crain et al.'s 2006 article aptly titled 'Language acquisition is language change'). We therefore conclude that what matters in a context (p. 436) of change is, not so much whether there has been a break in transmission, but rather a combination of the following: (i) the profiles of the learner (e.g., bilingual L1 learners vs. early/late L2 learners), and (ii) the input to which various learners are exposed and in what ratios (e.g., input from native vs. nonnative speakers of the target language). In this regard, Creoles do not represent an exception: they developed in the same set of conditions that have generally led to language change.

With respect to second language acquisition (L2A), it has been argued that certain innovations in adult learners' interlanguages are due to general strategies of L2A while others are due to patterns in the learners' native languages (their L1s); see chapter 13 for discussion. Therefore substrate influence in Creole formation is par for the course.

One well-established fact in the literature on L2A is that very early interlanguages are structurally reduced and unstable. This is a stage that all adult L2 learners go through, no matter how competent and no matter the nature of their input—this includes L2 learners in the history of any language. But in any language contact situation that involves young and adult learners (e.g., the ecology of Creole formation), not all L2 learners will be starting their respective L2A path at the same time and proceeding through the very same stages at the same pace. At any moment, the patterns from early interlanguages, with their reduced structures and their L1-influenced patterns, will constitute only one subset of the total ecology, alongside patterns from later stages of L2A and from L1A. In effect, both L1A and L2A would produce a mix of native and native-like patterns based on European(-derived) varieties as acquisition targets. The structural patterns in this mix and the stochastic properties thereof would be determined by contingent demographic and socio-historical factors. Accordingly, Creole languages cannot be mere fossilized interlanguages as argued by Plag (2008a,b) (cf. Mufwene 2010, DeGraff 2009:948–958, DeGraff et al. 2013, and Aboh 2015 for critiques).

In order to appreciate the extent of L1 influence in L2A in the course of Creole formation, we first need to understand the basic facts of substrate influence, alongside superstrate inheritance, in the diachrony of particular Creole languages. As in the rest of this chapter, we resist explanations that take scope over Creole languages as a class. Instead we prefer to focus on particular case studies as they shed light on specific theoretical analyses and complexities thereof—complexities that are often obscured by exceptionalist approaches to Creole languages.

### **17.3.2 Superstrate Inheritance, Substrate Influence, and Innovation in HC Formation**

In the particular case of HC, here are a set of key facts, as surveyed in the preceding sections, that need explanation. These facts involve certain similarities with the French lexifier, and the Niger-Congo languages, most notably the Gbe languages (e.g., Fongbe, Gungbe, Ewegbe) of the Kwa family. These similarities often come with innovations, including both simplification, as in the reduction of verbal inflectional morphology, and *local* complexification, as in, e.g., complementation structures and related properties (p. 437) (see note 5). We now turn to these facts. (See Aboh 2006a,b, 2009, DeGraff 2009: 916–929, 938–940, 948, etc., for related empirical and theoretical details and for references.)

### 17.3.2.1 Reanalysis in the Clausal Domain

Since Sylvain (1936) there has been an impressive list of studies documenting both similarities and dissimilarities between French and HC, with some of the dissimilarities apparently related to substrate influence. There is not enough space here to comprehensively review all the morphosyntactic similarities between French and HC, but the following examples from DeGraff (2007:109) should suffice to make our point.

Similarly to French, HC displays a null complementizer in certain nonfinite subordinate clauses (4):

- (4)
- |    |                                   |        |        |          |       |        |        |          |
|----|-----------------------------------|--------|--------|----------|-------|--------|--------|----------|
| a. | Tout                              | moun   | vle    | ale      | nan   | syèl.  | (HC)   |          |
|    | every                             | person | want   | go       | to    | heaven |        |          |
|    | 'Everyone wants to go to heaven.' |        |        |          |       |        |        |          |
| b. | Tout                              | le     | monde  | veut     | aller | au     | ciel.  | (French) |
|    | every                             | DET    | people | want.3SG | go    | to     | heaven |          |
|    | 'Everyone wants to go to heaven.' |        |        |          |       |        |        |          |

Aside from differences in nominal morphosyntax, which are discussed in detail in Aboh and DeGraff (2012), and in finite verbal inflectional morphology, which we return to in 17.3.2.2, it appears that in (4) French and HC display a parallel morphosyntax when it comes to nonfinite complementation.<sup>16</sup> Also note that every single morpheme in (4a) finds its etymon in the French example in (4b). Such correspondences hold in every domain of HC lexicon and even grammar, including the functional domain. This evidence supports our position that HC is genealogically related to French, consistent with the traditional application of the comparative method.

Interestingly, the complementation structures in French and HC manifest different properties from the substrate Gbe languages where the nonfinite clause is obligatorily introduced by a prepositional complementizer, here in boldface. A Gungbe example is given in (5):

- (5)
- |                                   |    |     |     |      |              |    |        |
|-----------------------------------|----|-----|-----|------|--------------|----|--------|
| Mè                                | lé | kpó | wè  | jró  | <b>*(ná)</b> | yì | lón    |
| people                            | PL | all | FOC | want | PREP         | go | heaven |
| 'Everyone wants to go to heaven.' |    |     |     |      |              |    |        |

(p. 438) In a sense, this Gungbe prepositional complementizer resembles French *pour* 'for,' whose HC cognate is *pou*, as in example (6):

- (6)
- |   |      |                 |           |     |      |      |
|---|------|-----------------|-----------|-----|------|------|
| Annou                                   | vote | [ <b>pp</b> pou | [kandida  | nou | vle  | a]]. |
| let-1PL                                 | vote | for             | candidate | 1PL | want | DET  |
| 'Let's vote for the candidate we want.' |      |                 |           |     |      |      |



In addition to selecting nominal projections and projecting run-of-the-mill PPs as in (6), HC *pou*, in one of its many uses, can also select for subjectless nonfinite purpose clauses on a par with its French etymon *pour*:

- (7)
- |    |   |         |            |     |      |      |        |       |       |    |
|----|---|---------|------------|-----|------|------|--------|-------|-------|----|
| a. | Kouto                                       | sa      | a          | pa  | fèt  | pou  | koupe  | pen.  | HC    |    |
|    | knife                                       | DEM     | DET        | NEG | make | for  | cut    | bread |       |    |
|    | 'This knife is not made for cutting bread.' |         |            |     |      |      |        |       |       |    |
| b. | Ce  | couteau | n'est      | pas | fait | pour | couper | le    | pain. | Fr |
|    | DEM   | knife   | NEG-be.3SG | NEG | make | for  | cut    | DET   | bread |    |
|    | 'This knife is not made for cutting bread.' |         |            |     |      |      |        |       |       |    |

The Haitian example in (7a) displays the same word order as the French sentence in (7b). Yet DeGraff (2007) also shows that HC *pou*, unlike French *pour*, selects for full finite clauses of the types given in (8):

- (8)
- |    |  |        |        |     |       |     |        |       |       |         |        |
|----|--|--------|--------|-----|-------|-----|--------|-------|-------|---------|--------|
| a. | Kouto  | sa     | a      | pa  | fèt   | pou | li     | koupe | pen.  | HC      |        |
|    | knife  | DEM    | DET    | NEG | make  | for | 3SG    | cut   | bread |         |        |
|    | 'This knife is not made for cutting bread.'  |        |        |     |       |     |        |       |       |         |        |
| b. | Li   | te     | ale    | nan | fèt   | la  | pou    | li    | te    | ka      | fè     |
|    | 3SG  | ANT    | go     | to  | party | DET | for    | 3SG   | ANT   | capable | do     |
|    | yon  | ti     | danse, | men | lè    | li  | rive,  | pa    | te    | gen     | mizik. |
|    | DET  | little | dance  | but | when  | 3SG | arrive | NEG   | ANT   | have    | music  |
|    | '(S)he went to the party to dance a bit, but when (s)he arrived there was no music.' |        |        |     |       |     |        |       |       |         |        |

It therefore appears from the HC examples in (7) and (8) that HC *pou* as complementizer can introduce both nonfinite and finite clauses. This is not the case for complementizer *pour* in Modern French which only selects for nonfinite clauses without any overt subject.

As it turns out, there's one obvious difference between HC and French that supports our L2-L1 cascade hypothesis about the formation of HC: the verbs in our HC examples are, unlike French, devoid of affixes for TMA and subject-verb agreement. This difference can be explained in a perspective that considers L2A a key process in Creole formation and in contact-induced language change more generally, with verbal inflectional (p. 439) affixes as a frequent casualty in L2A, especially in the early stages—and much more so than in L1A (Weinreich 1953; Archibald 2000; Prévost and White 2000; Wexler 2002; Ionin and Wexler 2002).

This morphological difference between French and HC seems related to a syntactic difference over a larger domain, namely the clausal domain and the selectional properties of French *pour* vs. HC *pou*. Recall the example in (8) where HC *pou* can introduce finite

clauses. This selectional property of HC *pou* is not available for French *pour*. The only way for the latter to embed a finite clause is through the intermediacy of the complementizer *que*, as in the following example with a subjunctive:

- |  |      |     |        |    |       |      |        |     |     |
|--|------|-----|--------|----|-------|------|--------|-----|-----|
|  | Jean | a   | acheté | un | livre | pour | *(que) | son | fil |
|  | John | has | bought | a  | book  | for  | that   | his | son |
- (9) puisse                      le lire.  
 can.SUBJUNCTIVE    it    read  
 'John has bought a book so that his son can read it.'

At this point, we need to go beyond Standard Modern French, and highlight the fact that the preposition *pour* in Middle French, unlike its modern descendant, could select for nonfinite clauses with overt subjects as in ... *et lour donna rentes pour elles vivre ...* '... and gave them a stipend for them to live ...,' *Quelle fureur peut estre tant extrême ... pour l'appetit chasser la volonté?* 'What fury can be so extreme ... for hunger to chase the will?' (Nyrop 1930, VI:219; also see Frei 1929:95). Frei (1929:93–94) and Nyrop (1930, VI:220) also mention that certain modern dialects show a similar pattern, as in the following example *Apportez-moi du lait pour les enfants boire* 'Bring me milk for the children to drink' and ... *un oreille pour moi dormir et un saucisson pour moi manger* '... a bonnet for me to sleep and a sausage for me to eat.' The tonic (non-nominative) pronoun *moi* as subject of *dormir* 'to sleep' and *manger* 'to eat' indicates that the embedded subject position is not assigned nominative Case—it either receives Case from the preposition *pour* or realizes default case.

There's one key difference between HC examples such as (8b) *Li te ale nan fèt la pou li te ka fè* and these Middle French examples with overt subjects in the embedded infinitival clauses, namely the fact that the embedded clause in the HC example is finite whereas the embedded clause in the Middle French example is nonfinite. But we've already noted the fact that HC finite verbs are often homophonous with French infinitives or past participles. This pattern was even more pronounced in the 17th and 18th century when the final /r/ of French infinitives such as *finir, courir, ouvrir, voir, boire*, etc., was often silent (Nyrop 1899, I:293–294, 1903, II:62; Brunot 1906, II:273; Gougenheim 1951:30; Hazaël-Massieux 2008:19; but see Nyrop 1903, III:153–154 and Brunot and Bruneau 1913, IV:208–210 for evidence of variation). The pattern in (8b) would then emerge as the outcome of a reanalysis process: French nonfinite small clauses (e.g., as the nonfinite complement of *pour*) were reanalyzed by HC speakers as full-fledged finite clauses where the main invariant verb could co-occur with preverbal TMA markers on (p. 440) the model of the French verbal periphrases that are popular in the spoken varieties of French described in Gougenheim (1929). In effect, such a reanalysis, from nonfinite small clauses

in French to full-fledged finite clauses as complement of prepositional complementizers, led to embedding possibilities in HC that are now more complex than their analogs in both Middle and Modern French. This is yet another counterexample to the Creole-simplicity claims that are so prevalent in Creole studies.

In a related vein, let's take another look at (9) where the finite embedded clause is obligatorily introduced by the combination *pour + que*—with *que* 'that' directly selecting for the finite clause. Here there's another point of comparison with HC that may support our reanalysis scenario for the emergence of the clausal structure selected by HC *pou*. One important fact about the translation of (9) into HC, is that HC *ke*, unlike its Modern French etymon French *que*, is optional:

- (10) Jan achte yon liv pou (ke) pitit li ka li li.  
 John bought a book for that son 3sg can read 3sg  
 'John has bought a book such that his son can read it.'

The optionality of HC *ke* is found in other instances (DeGraff 2007:109):

- (11) Jinyò konnen (ke) Jezila renmen l anpil.  
 Jinyò know comp Jezila love 3sg much  
 'Jinyò knows (that) Jezila loves her a lot.'

Here too we find a contrast with Modern French where the complementizer *que* in similar examples cannot be omitted:

- (12) Jeanne sait \*(que) Jeannette l'aime beaucoup.  
 Jane knows comp Jeannette her-loves much  
 'Jane knows that Janet loves her very much.'

What about earlier varieties of French? As it turns out, complementizer *que* in these earlier varieties was optional, somewhat on a par with HC. The optionality of HC *ke* is not surprising once we consider 17th-century French where *que* was also optional (Nyrop 1930, VI:159).

To summarize: while HC and English share null complementizers, this is not the case in Modern French, and while both French *pour* and English *for* select for nonfinite clauses, only HC *pou* selects for either finite or nonfinite clauses. The selectional properties of HC *pou* can be analyzed as the outcome of reanalysis based on patterns in earlier varieties of French. If we assume that (apparent?) optionality in the realization of a complementizer induces an increase in complexity, HC is, in this *particular* respect, more complex than Modern French.

We now look at the function of HC *pou* as modality marker (see Koopman and Lefebvre 1982; Sterlin 1989). The latter use of *pou* is reminiscent of patterns in both (p. 441) French and in the Gbe languages. Consider the example in (13a) where HC indicates deontic modality, in a way similar to English ‘to’ or to the French (quasi-)modal periphrastic construction *être pour* + infinitival V as in (9b). The latter construction is documented in the history and dialectology of French (with examples both from the 17th century and from dialects spoken in Picardy, Midi, Provence and Canada in Gougenheim 1971:120–121).

- (13) a. Se Jinyò ki pou te vini. (HC)  
 FOC Jinyò COMP MOD ANT come  
 ‘It’s Jinyò who had to come.’
- b. Je suis pour me marier la semaine prochaine.  
 I am for me marry the week next  
 ‘I am to get married next week.’ (Canadian French)

We find related constructions in the Gbe languages as well (Aboh 2006a). The following Gengbe examples indicate that the dative preposition *né* in (14a) can also be used as conditional mood marker as in (14b) and injunctive mood marker as in (14c) (see Aboh 2015 for discussion):

- (14) a. Kòfi pò tòmè né Kòjò.  
 Kofi hit ear-in né Kojo  
 ‘Kofi slapped Kojo’s face.’
- b. Né Kòfi vá á mì yrò-è.  
 né Kofi come TOP 2PL call-3SG  
 ‘If/when Kofi comes, call him.’
- c. Kòjò glòn bé Kwésí né vá.  
 Kojo said that Kwesi né come  
 ‘Kojo said that Kwesi should come.’

These examples indicate that both French and the Gbe languages contributed to the emergence of modal *pou* in HC. (Aboh 2006a discusses related facts in the grammar of Saramaccan; also see Corne 1999 for related facts in a diverse range of French-based Creoles, under the umbrella label of ‘congruence.’)

Here we’ve introduced a concrete sample of HC/French (dis)similarities related to verbal inflectional morphology and the use of prepositional and finite complementizers in HC and French. The word order and semantic similarities are pervasive while the dissimilarities seem to touch on well-delimited aspects of morphosyntax such as the selectional properties and optional pronunciation of certain complementizers and the

realization of inflectional verbal morphology. These HC patterns can be analyzed as the outcome of reanalysis in combination with substrate influence via second language acquisition, thus incorporating influences from both the French superstrate and the Gbe substrates. These patterns (p. 442) contradict any break-in-transmission scenario that postulates a reduced-pidgin stage in the history of HC. Furthermore, this combination of superstrate *cum* substrate influence seems to introduce a certain amount of local complexity in the relevant domains once certain crucial assumptions are made about the relevant complexity metric (see note 6). This is another illustration that shows that ‘complexity’ cannot be determined in absence of any theory of grammar. For example, the lexical entry for *pou* needs to be made more ‘complex’ than that of its French etymon *pour* to the extent that ‘complexity’ can be measured by the inventory of combinatory possibilities or the amount of structure entailed by said combinations: the complement selected by HC *pou* can be either a nonfinite small clause or a full-fledged finite clause while that of French *pour* is a nonfinite clauses, unless *pour* first takes a CP that is headed by *que*. In effect, HC *pou* has one more option for its complement (namely, a finite TP) than French *pour* (which does not select for finite TPs). Such a local increase of ‘complexity’ given a particular set of theoretical assumptions is a running theme in our ongoing survey.

### 17.3.2.2 From V-to-T to V-in-Situ: Restructuring vs. Simplification

Now we consider the profile of verbal inflectional morphology in HC and its consequences for its morphosyntax.

In exceptionalist theories of creolization, the absence of inflectional morphology in Creole languages is taken as strong evidence for postulating some drastically reduced Pidgin or some fossilized early interlanguage as the ancestor of the Creole (e.g., Bickerton 1981, 1999; McWhorter 2001; Plag 2008a,b). Such views fail to provide any insight into the structural properties of Creoles and into the deep similarities between Creole formation and general patterns of language change.

A case in point is the loss of inflectional morphology and V-to-T in the history of English and Mainland Scandinavian languages. Consider the Early Modern English (ENE) sentences in (15a-d). In these examples, the finite verb in boldface precedes the negative marker as depicted in (15e):

- a. It serveth not. [Middle English]
- b. Wepyng and teres counforteth not dissolute laghers. (Roberts 1993)
- c. Quene Ester looked never with switch an eye. (Kroch 1989)
- (15) d. ... if man grounde not his doinges altogether upon nature. (Kroch 1989)
- e. Verb placement in ENE  
Subject ...  $V_{finite}$  ... *not/never* ...

The patterns in (15) contrast with Modern English (NE) where the inflected verb cannot precede negation, hence the ungrammaticality of (16a) as opposed to (16b).

- (16) a. \*He speaks not English.
- b. He does not speak English.

(p. 443) The contrast in (15) vs. (16) indicates that while the lexical verb can precede negation in Middle English (ME), this is impossible in NE, where lexical verbs must follow negation. In such contexts, any tense or agreement specification is expressed by pleonastic *do* as in (16b) or by a relevant auxiliary/modal. The lexical verb on the other hand occurs in its nonfinite form when it follows negation as in (17):

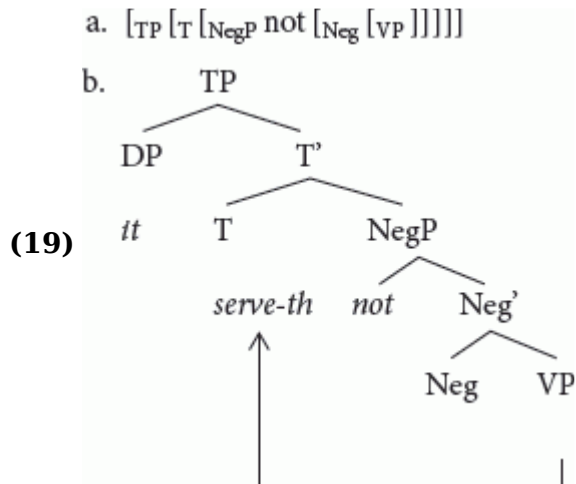
- (17) a. John will not sell his new car.
- b. John must never sell his new car.

Unlike lexical verbs, auxiliary verbs, such as *be* and *have*, precede negation in NE. The schema in (18c) describes verb placement in NE.

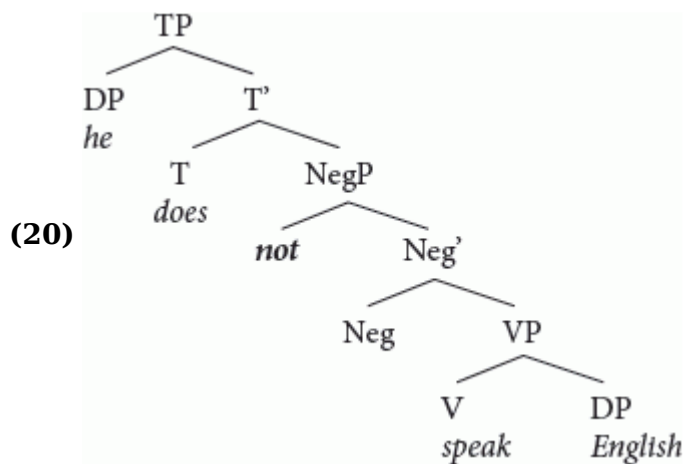
- a. John has not bought a new car.
- (18) b. John has never bought any car.
- c. Verb placement in NE  
Subject ... Aux/Mod/*do* ... *not/never* ...  $V_{nonfinite}$  ...

In the literature on diachronic changes in English, the path from the examples in (15) to those in (18) is commonly analyzed in terms of V-to-T movement vs. lack thereof. There are many competing proposals for the implementation of this basic idea (Kroch 1989; Roberts 1993, 1999; Vikner 1997; Rohrbacher 1999; Han and Kroch 2000; Bobaljik 2002; among others; see chapter 18 for discussion of models of syntactic change). Here we adopt the minimal set of assumptions without committing ourselves to any specific implementation. Starting with a clause structure similar to (19a), we take (15a) to derive from movement of the verb to T (i.e., above the negative phrase) giving the order Subject-Verb-Negation in (19b). Furthermore we take V-to-T movement to be related to the nature of verbal inflectional morphology and the licensing thereof. In addition, the

negative element *not* is analyzed as a negative adverbial (on a par with *never*) and realizes the specifier position of NegP, the functional projection responsible for the expression of sentential negation.



(p. 444) The impossibility of (16a) suggests that V-to-T movement is lost in NE. Here, the lexical verb cannot raise to T. As a consequence, sentential negation requires the presence of a pleonastic element (e.g., *do*), a modal, or a non-modal auxiliary which expresses tense or agreement specification. The sentence under (16b) is represented as follows.



ENE and NE thus differ in that the former allowed V-to-T movement unlike the latter where the lexical verb must stay within V. What also got lost in NE is a subset of verbal agreement affixes (e.g., the 2nd singular suffix *-est* as in *Thou singest*) and the possibility of stacking of tense and agreement affixes as in *Thou showedest* (Kroch 1989:238; cf. Bobaljik 2002). The loss of various verbal affixal combinations is itself understood as a reflex of language contact phenomena (e.g., second language acquisition by Scandinavian invaders; cf. Kroch et al. 2000).

With this in mind, let us now return to the diachronic path from French to HC in the context of language contact with both French and Niger-Congo languages such as Gbe. Much of the discussion here recapitulates the findings in DeGraff (1997, 2005) to which we refer the reader for a detailed discussion.

The facts about verb placement in ME are comparable to the following French data (adapted from Pollock's 1989 seminal comparative study of Romance and Germanic languages).

- (21)
- |    |                                |      |       |        |         |
|----|--------------------------------|------|-------|--------|---------|
| a. | Joujou                         | (ne) | parle | pas    | Créole. |
|    | Joujou                         | NEG  | speak | NEGADV | Creole  |
|    | 'Joujou doesn't speak Creole.' |      |       |        |         |
| b. | Joujou                         | (ne) | parle | jamais | Créole. |
|    | Joujou                         | NEG  | speak | NEGADV | Creole  |
|    | 'Joujou never speaks Creole.'  |      |       |        |         |

In French, finite lexical verbs precede negative adverbials (i.e., *pas/jamais*), as in the ME data, but unlike in NE where lexical verbs generally follow negative adverbials (i.e., *not/never*). Another difference is that Standard French, unlike English, allows the combination of the negative clitic *ne* and the negative adverb *pas*. However, this clitic is optional in most spoken French. Example (22) therefore indicates the position of the French finite verb vis-à-vis the negative adverb *pas* (comparable to English *not*).

- (22) Verb placement in French  
 Subject ... V<sub>finite</sub> ... *pas/jamais* ...

Verb placement in French is like in ME: it exhibits V-to-T movement. In addition, French is similar to ME in that it displays a larger set of inflectional suffix combinations on the verbs than NE does, including the stacking of tense and agreement suffixes as in *aim-er-ai* (love+FUT+1sg).

With these observations in mind, let us now turn to HC. As already discussed in DeGraff (1993, 1997, 2005), HC lacks V-to-T movement: the lexical verb must follow the negative marker as indicated in (23a), the equivalent of the French example (21a). The sentence in (23b) shows that a sequence comparable to that of French is ungrammatical in HC:

- (23)
- |    |                               |       |       |         |
|----|-------------------------------|-------|-------|---------|
| a. | Jinyò                         | pa    | pale  | Kreyòl. |
|    | Jinyò                         | NEG   | speak | Creole  |
|    | 'Jinyò doesn't speak Creole.' |       |       |         |
| b. | *Jinyò                        | pale  | pa    | Kreyòl. |
|    | Jinyò                         | speak | NEG   | Creole  |



In addition, the HC negative particle *pa* precedes the negative adverb *janm* ‘never’ whose etymon is French *jamais* ‘never’ (24). Both negative elements precede the verb, thus indicating that HC lexical verbs are pronounced in a low structural position.

- Jinyò pa janm pale kreyòl.  
(24) Jinyò NEG never speak Creole  
‘Jinyò never speaks Creole.’

Thus, while finite lexical verbs precede the negative elements *pas/jamais* in French, the HC equivalents *pa/janm*, precede the lexical verb, as depicted in (25):

- (25) Verb placement in HC  
Subject ... *pa/janm* ...  $V_{finite}$

The contrast between French (22) and HC (25) reminds us of the diachronic path from ME with V-to-T movement to NE without V-to-T movement. In addition, both French and ME display a set of inflectional affixal combinations that is larger than in HC and NE, respectively. Yet it must also be noted, as we already have, that popular colloquial varieties of French (i.e., the varieties that were the *terminus a quo* of (p. 446) HC) show a preference for analytic periphrastic constructions with invariant verbal forms (either infinitival or participial forms) of the sort described in Frei (1929) and Gougenheim (1929). Be that as it may, HC shows even less inflection than these nonstandard varieties of French. Building on previous discussion, we propose that the absence of V-to-T in HC, like in NE, correlates with a reduction in verbal inflectional morphology, notwithstanding a possible time lag between the two sets of phenomena (see DeGraff 1997, 2005 and references therein for additional details on the emergence of the HC patterns with special attention to their cognates in French periphrastic constructions).

For exceptionalist views of creolization, the reduction in verbal inflectional morphology in HC is evidence for structural simplification as a result of prior pidginization. Yet, the comparison with English indicates that things might not be so simple. Here we discuss two facts that suggest that in both English and HC the reduction in inflectional morphology and the absence of V-to-T movement may have triggered a local increase of complexity vis-à-vis one particular aspect of the morphosyntax of these languages.

Let’s start with English. Recall that Roberts (1985, 1993), Kroch (1989a,b) and Han and Kroch (2000), among others, have argued that the loss of V-to-T and the reduction in verbal inflectional morphology are related to the rise of *do*-support. One common explanation is that *do*-support is a morphosyntactic device that licenses inflectional specifications in absence of V-to-T movement. This can be seen in structure (20) where *do* is inserted under T where it bears tense and agreement morphology. V-to-T movement aside, there is no structural difference between the ME representation in (19b) and the

NE one in (20). The parallel between these two structures indicates that the reduction in verbal inflectional morphology and the loss of V-to-T movement does not trigger a simplification in the basic clause structure. Instead, what we see here is that absence of V-to-T movement entails a new morphosyntactic strategy, that is, *do*-support. The latter entails a new set of morphosyntactic constraints to be acquired by learners of NE (e.g., *do*-support in interrogative, emphatic constructions, negative imperatives, negative sentences; see Han and Kroch 2000 for discussion). Any approach that focuses solely on the loss of inflectional morphology from ME to NE will miss the fact that such a loss did trigger a new battery of morphosyntactic strategies that certainly needs to be taken into account in any evaluation of complexity differences between ME and NE. Similar issues arise in the comparison between French and HC.

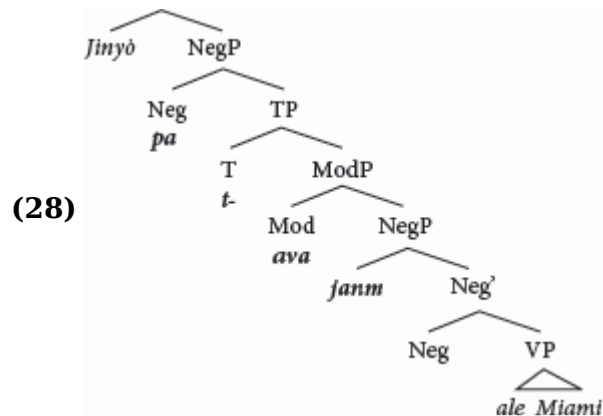
It is indeed arguable that, as in the history of English, the reduction in verbal inflectional morphology from French to HC (e.g., the disappearance of tense and agreement suffixes such as *-ais* in *Je dansais* 'I was dancing') and the concomitant loss of V-to-T movement correlates with the emergence of a series of preverbal TMA markers that fulfill similar functions as TMA suffixes in French (to wit: HC *Mwen t ap danse* 'I was dancing' with *t(e)* as the anterior marker and *ap* as the progressive marker). In this regard, the following examples show that the HC negative marker *pa* is higher in the structure than (p. 447) its English counterpart *not*: whereas the tense specification in English is realized in a position preceding *not* as in *John did not dance*, HC *pa* must precede all TMA elements, including the anterior marker *te* (cf. DeGraff 1993:63):

- (26)
- |    |   |     |     |       |     |     |        |
|----|---|-----|-----|-------|-----|-----|--------|
| a. | Jan                                       | pa  | t-  | av-   | ale | nan | mache. |
|    | John                                      | NEG | ANT | IRR   | go  | in  | market |
|    | 'John would not have gone to the market.' |     |     |       |     |     |        |
| b. | *Jan                                      | te- | pa  | (av-) | ale | nan | mache. |
|    | John                                      | ANT | NEG | IRR   | go  | in  | market |

The sentences under (27) involve negative concord where the negative marker *pa* is combined with *janm* 'never' in addition to other TMA markers. In all these examples, *pa* precedes all the TMA markers which in turn precede the verb.

- a. Jinyò pa te janm ale Miami.  
 Jinyò NEG ANT never go Miami  
 'Jinyò never went to Miami.'
- (27) b. Jinyò pa t- ava janm ale Miami.  
 Jinyò NEG ANT IRR never go Miami  
 'It was never the case that Jinyò was likely to go to Miami.'
- c. Jinyò pa t- ap janm ale Miami.  
 Jinyò NEG ANT IRR never go Miami  
 'Jinyò would have never gone to Miami.'

Examples of this sort indicate that the syntax of HC negative marker is different from that of its French etymon *pas* which follows the finite verb of its clause. According to DeGraff (1993), HC *pa* heads the higher NegP from where it dominates the series of TMA markers and verb:



(p. 448) We observe from (28) that reduction in inflectional morphology and loss of V-to-T movement do not correlate with simplification of structure. Instead, the lack of V-to-T movement in HC corresponds to a clause structure where TMA-markers license the relevant projections. The emergence of these TMA markers comes with its own increment in local complexity due to their combinatorics, the ordering constraints therein and the distinct semantics of the various combinations. Witness the subtle semantic distinctions between (27b) and (27c) and the contrast previously described in 17.2.1 with regard to *te -dwe* versus *dwe-te* sequences.

In this respect, another type of local complexity in HC concerns temporal interpretation. Indeed, the absence of V-to-T movement in HC means that the verb itself does not bear temporal specifications. Instead, these specifications are deduced from the combination of TMA markers and the lexical aspect of the verb. Put differently, temporal specification is computed based on TMA markers and *Aktionsart*. HC, like many Creoles and Niger-Congo languages (including Gbe; see (31)), displays an asymmetry between eventive/

dynamic verbs and stative verbs: when they occur without any TMA marker, eventive/dynamic verbs are interpreted as perfective, while stative verbs are interpreted as present (this asymmetry has been well known in Creole studies since Bickerton 1981). In the literature on African languages going back to the 1960s, this phenomenon has been labeled ‘factative effect’ (see Déchaine 1994 for an overview). An illustration of this effect is given in (29), taken from DeGraff (1993:77–79). As is shown by the French translation, the interpretation of a bare eventive verb (29a) is comparable to the French *passé composé* even though such sequences are commonly translated with past tense marking in English. On the other hand, bare stative verbs are interpreted as present. This is illustrated by the stative verb of the matrix clause in (29b):

- (29) a. Prèske pèsonn pa vote pou Manigat.  
 almost nobody NEG vote for Manigat  
 ‘Presque personne n’a voté pour Manigat.’ [French]  
 ‘Almost nobody voted for Manigat.’ [English]
- b. Mwen pa kwè pèsonn ap vini.  
 1SG NEG believe nobody FUT come  
 ‘I don’t believe that anybody will come.’

Contrary to French where temporal interpretation can be read off the morphology of the verbs, HC therefore requires temporal interpretation to be computed based on both the lexical aspect of verbs and the combinations of TMA markers. In turn, the combinations of TMA markers are regulated by complex constraints, and so is their semantics (Fattier 1998, 2003; Howe 2000; Fon Sing 2010; etc.). In addition, lexical aspect interacts with argument structure, and HC displays subtle interpretative nuances that are sensitive to the form and semantics of the internal argument. As is shown by the example in (30) from Déchaine (1994), a bare noun phrase with non-individuated (p. 449) generic reference allows a habitual reading while a determined noun phrase triggers a perfective reading (30b):

- (30) a. Jinyò vann chat.  
 Jinyò sell cat  
 ‘Jinyò sells cats.’
- b. Jinyò vann chat la.  
 Jinyò sell cat DET  
 ‘Jinyò sold the cat.’

So far we have surveyed an intricate set of distributional and interpretive facts in HC, all related to verbal inflectional morphology and the clausal domain. From our perspective in this chapter, popular exceptionalist claims to the effect that Creoles are the simplest languages (as, for example, in McWhorter 2001 and related work) are misleading because

they merely rely on presence versus absence of various sorts of overt morphology without taking into account the (covert) structural correlates of said morphology vs. lack thereof.

Besides that question of complexity, a question that is of interest here is whether our observations about temporal interpretations in HC have analogs in phenomena that are attested beyond Creole languages. As noted in Avolonto (1992), and Déchaine (1994), the asymmetry just described between eventive/dynamic verbs and stative verbs is also found in the Gbe languages involved in the contact situation that led to the development of HC. The following example illustrates this asymmetry in Gungbe (Aboh 2004).

- (31)
- a. Súrù sà àsé lɔ.  
Suru sell cat DET  
'Suru sold the cat.'
  - b. Súrù nyón àsé lɔ.  
Suru know cat DET  
'Suru knows the cat.'

Given these similarities between HC and the Gbe languages, we can reasonably hypothesize that the latter, as the native languages of adult learners of French or (proto-)HC, influenced the emergence of the Haitian TMA system. As a consequence, HC appears to have morphosyntactic properties in common with both French and the Gbe languages—with patterns from all the languages contributing to various local domains of complexity in the emergent Creole (e.g., in the Creole's clause structure). This leads us to the discussion of superstrate inheritance combined with substrate influence and concomitant increases of local complexity, specifically in the nominal domain. (p. 450)

### 17.3.2.3 Superstrate Inheritance, Substrate Influence, and Innovation in the Nominal Domain

We now turn to relative clauses in HC as compared to French and Gungbe. This discussion focuses on restrictive relative clauses and recapitulates some of the findings in Aboh (2006b) and Aboh and DeGraff (2014).

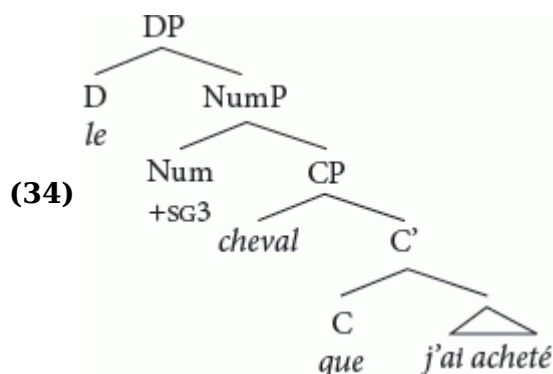
Starting with French relatives, we observe that they display the word order pattern in (32a): the relative clause follows the noun which itself is introduced by the determiner. Because Modern French does not allow singular bare noun phrases, the determiner is compulsory in relative clauses as indicated by the ungrammatical example (32b):

- (32)
- a. Le cheval que j'ai acheté  
the horse that 1SG-have bought  
'The horse that I bought'
  - b. \*Cheval que j'ai acheté  
horse that 1SG-have bought

Roodenburg (2006) shows that French allows coordinate plural nouns of the type in (33a) but even such noun phrases exclude relativization, hence the ungrammatical sentence (33b):

- (33) a. Étudiants et professeurs manifestent à Montréal.  
 students and professors demonstrate in Montreal  
 'Students and professors are demonstrating in Montreal.'
- b. \*Étudiants et professeurs qui ont manifesté à Montréal  
 students and professors who have demonstrated in Montreal

Let us assume the format of French relative clauses as in (34a). We adopt Kayne's (1994) complementation analysis of relative clauses in (34b) where the determiner introduces the relative clause and the relativized noun phrase occurs in Spec,CP. As for French DP structure, we assume that it involves at least two functional projections DP and NumP where the latter is responsible for the expression of number. We restrict ourselves to this partial representation, and we refer the reader to De Vries (2002) for a careful discussion.



With this description in mind, let us now look at relative clauses in the Gbe languages. Aboh (2002, 2004, 2005) gives ample details on relative clauses in Gungbe some of (p. 451) which we recapitulate here. In Gungbe and other Gbe languages bare noun phrases (i.e., determiner-less noun phrases) are allowed in all argument positions, unlike in Modern French and most contemporary varieties of Romance and Germanic languages. In this respect, Gbe is similar to earlier varieties of French where determiner-less nouns, up until the 17th century, could occur in argument positions as well—unlike their distribution in Modern French (see Mathieu 2009 for data and references).<sup>17</sup> The congruence of the distribution of argumental bare nouns in both Gbe and earlier varieties of French would have favored their distribution in HC as well, as described here and in Aboh and DeGraff (2014).

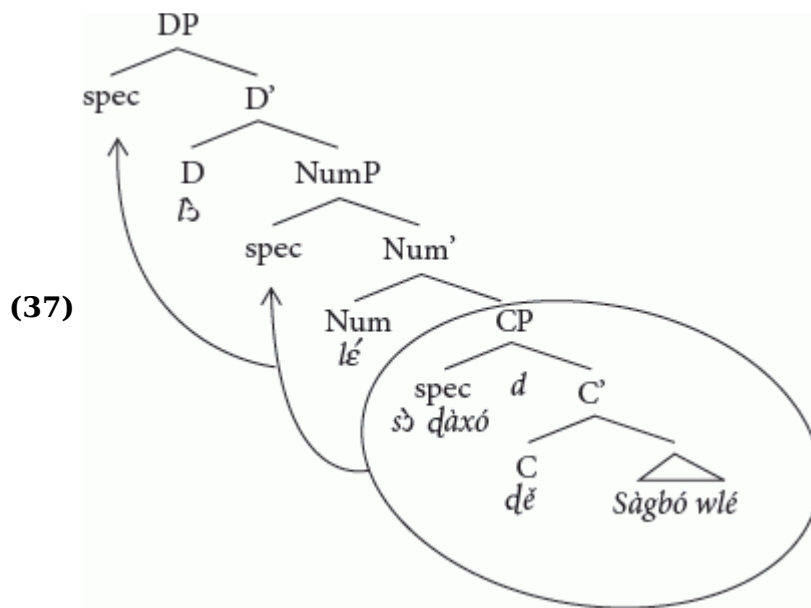
Bare noun phrases in Gbe can optionally be accompanied with nominal modifiers such as adjectives. The object noun phrase in example (35a) illustrates this. Similarly to such noun phrases, noun phrases including determiners (i.e., the specificity marker *lɔ́* and the number marker *lé*) can also occur in the same set of argument positions. This is shown by the object noun phrase in (35b). Example (35c) summarizes the sequencing of the noun phrase in Gungbe as discussed in Aboh (2002, 2004) and subsequently. In these examples, we translate the specificity-marked noun phrases as ‘N in question’ but we refer the reader to Aboh and DeGraff (2014) for detailed discussion of bare and determined noun phrases in Gungbe and Haitian Creole.

- a. Súrù xò sɔ́ (ɖàxó).  
 Suru buy horse big  
 ‘Suru bought (a)(big) horse(s).’
- (35) b. Súrù xò sɔ́ ɖàxó àwè lɔ́ lé  
 Suru buy horse big two DET PL  
 ‘Suru bought the two big horses (in question).’
- c. Noun–[Modifier]–Determiner–Number

As is clear from (35c), Gbe noun phrases display structures where the noun systematically precedes its modifiers which in turn precede the determiner and the number marker. This ordering is fixed, and the relative clause (here in square brackets) also occurs within the slot of modifiers where it is followed by the determiner. This yields the order in (36b) where the determiner associated with the head noun occurs on the right edge of the relative clause (see Aboh 2002, 2005, 2010).

- a. Súrù xò sɔ́ ɖàxó [ɖě Sàgbó wlé].  
 Suru buy horse big REL Sagbo buy  
 ‘Suru bought the big horses that Sagbo caught.’
- (36) b. Súrù xò sɔ́ ɖàxó ɖě Sàgbó wlé lɔ́ lé.  
 Suru buy horse big REL Sagbo catch DET PL  
 ‘Suru bought the big horses (in question) that Sagbo caught.’
- (p. 452)
- c. Noun–[Modifier]–[relative clause]–Determiner–Number

In accounting for the word order properties of relative clauses in Gungbe, Aboh (2002, 2005) argues that Gbe languages differ from Romance (e.g., French) and Germanic (e.g., English) because the relative clause must raise to the specifier positions of the number marker and the specificity marker as depicted in (37).



This movement serves to license specificity and number in Gbe. Without getting into the details of the analysis in Aboh (2002, 2005), what our description of relative clauses in French and Gungbe suggests is that while it is possible to postulate a similar underlying structure for both languages, Gbe languages differ from French in two crucial respects: (i) the former languages display noun-modifier sequences and (ii) they require successive movements of the relative clause CP to Spec,NumP and Spec,DP. These contrasts between Gbe and French can help us understand the formation of relative clauses in HC.

Indeed, HC is similar to Gungbe and earlier varieties of French to the extent that it allows bare nouns in all argument positions similarly to determined noun phrases. But, HC and Gbe, unlike all varieties of French, have their definite determiners follow the noun as illustrated in (38). The sequencing in (38c) shows the parallels and differences among HC, Gungbe, and French. Indeed, HC is like French in exhibiting both prenominal and postnominal adjectives though the noun in HC must precede the definite determiner, unlike the noun in French:

- (38) a. Jinyò achte (gwo) chwal (blan).  
 Jinyò bought (big) horse(s) (blan)  
 b. Jinyò achte (gwo) chwal (blan) yo.  
 (p. 453) 'Jinyò bought the big white horses (in question)'  
 c. [Modifier]–Noun–[Modifier]–Determiner

With regard to relative clauses, HC also displays bare relative clauses of the type illustrated in Gungbe.

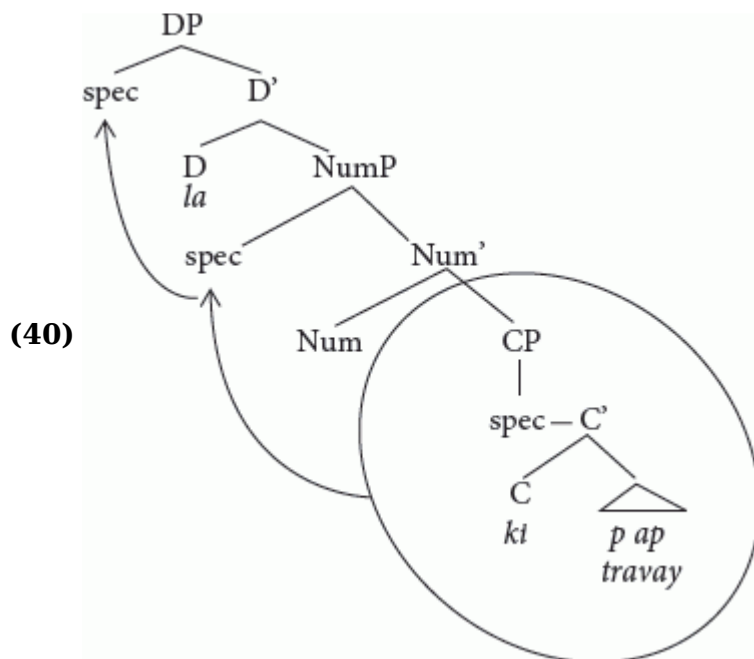


a. Moun ki p ap travay p ap manje.  
 People who do not work will not eat  
 'People who don't work will not eat'

(39) b. [Moun ki p ap travay la] p ap manje.  
 People who do not work DET will not eat  
 'The person who doesn't work will not eat.'

c. Noun – [relative clause] (determiner)

A number of properties arise that differentiate HC and French from Gungbe. Indeed, both French and HC display a prenominal indefinite article (*yon liv* 'a book') and prenominal modifiers, as in *gwo chwal* 'big horse(s),' alongside with postnominal modifiers as in *chwal blan* 'white horses' (cf. DeGraff 2007), while Gungbe (and Gbe generally) exhibit postnominal modifiers and postnominal determiners only. And it is striking that the classes of pre- and post-nominal adjectives are similar across HC and French (Sauveur 1999). On the other hand, Gungbe and HC pattern alike with regard to relative clauses. In these languages, the relative clause is to the left of the determiner that right-bounds the DP while it is to the right of the DP-initial determiner in French. Aboh (2006b) argues on this basis that HC involves a derivation similar to that in (37) where the CP relative clause raises to Spec,NumP and Spec,DP, as illustrated in (40), which represents the bracketed sequence in example (39b):



(p. 454) Aboh (2006b:239/240) argues that 'the parallels between Haitian and Gungbe determiner phrases can be regarded as an instance of pattern transmission because both languages share similar properties with regard to the function and syntax of the nominal

left peripheral elements, such as the specificity markers *lo'la* and the number markers *lé/yo.*'

The similarities between these languages are further reinforced by the fact that both languages display constructions that Aboh (2005, 2010) characterized as event relativization. In such structures, the verb is relativized but doubles such that there are two tokens of the same verb in the clause (see Glaude and Zribi-Hertz 2012 for discussion). This is unlike French (or English for that matter) where such constructions can only be formed by the noun phrase *the fact.../le fait...* underlined in the translations.

- a. [Yi dɛ Pɔl yi] wà léblánú ná mì.  
 go REL Paul go do sadness for 1SG  
 'The fact that Paul left made me sad.'
- (41) b. Yo di [pati Pol pati a] fè Elsi tris.  
 3pl say leave Paul leave DET make Elsi sad  
 'They said that the fact that Paul left makes Elsi sad.'  
 (Glaude and Zribi-Hertz 2012:84)

These constructions are structurally similar to verb focus with doubling (known as 'predicate cleft') which both HC and Gungbe display contrary to French:

- a. Yi wè Pɔl yi.  
 go FOC Paul go  
 'Paul LEFT.'
- (42) b. Se pati Pol pati.  
 FOC leave Paul leave  
 'Paul LEFT.'

The constructions in (42) differ from the event relativization in (41) in two ways. Firstly, they lack the factive reading that is highlighted by the 'the fact that ...' translation in (41), and they involve a focus marker that is absent in (41): *se* in HC and *wè* in Gungbe. We will not discuss the morphosyntax of these constructions here, and the reader is referred to DeGraff (1995), Aboh and Dyakonova (2009), Glaude and Zribi-Hertz (2012), and references therein. Assuming that these constructions are properties of the left periphery, we therefore reach the conclusion that HC and Gungbe share significant properties with regard to the morphosyntax of their left periphery: the left periphery of the clause and the left periphery of the noun phrase. Yet HC resembles French when it comes to the ordering of nominal modifiers vis-à-vis the head noun. (p. 455)

As is clear from this discussion, the emergence of such a 'hybrid' system in HC grammar necessarily comes with its share of local complexification such as: (i) the relation between verb focus with doubling (or predicate cleft) and relativization; (ii) the relation between

left peripheral markers of the nominal domain and those of the clausal domain, which in turn allows the nominal specificity marker to be used as a clausal determiner both in Gungbe and in HC. The HC examples are taken from Glaude and Zribi-Hertz (2012:91n10):

- [context]: Pou ki sa ou leve? 'Why are you getting up?'
- (43) a. M ale.  
 1SG go  
 'I'm going.' 18
- b. M ale a.  
 1SG go DET  
 'I'm going, as you knew I would.'

In Aboh's Gungbe, the clausal determiner occurs in contexts where it is introduced by the relative marker *dě* as shown in (44) (see Aboh 2004, 2005 for discussion):

- [context] There is some noise outside:
- A: Été wè jò?  
 what FOC happen  
 'What happened?'
- (44) B: *dě* ùn *dò* ná wè *dò* Súrù ná gò ló  
 REL 1SG say FUT 2SG that Suru FUT return DET  
 à má *dì*, bò yì kpón hòntò  
 2SG NEG believe COORD go look outside  
 à mòn-è flén.  
 2SG.FUT look-3SG there  
 'When I told you that Suru will return (as you certainly remember) you didn't believe me. You can look outside and you will see him there.'

As discussed in Aboh and DeGraff (2014) these examples show that the licensing properties of certain determiners in these languages cut across the morphosyntax of the nominal and clausal domains in intricate ways.

(p. 456) What this description shows is that HC displays a morphosyntax that shares properties of both French and Gbe. Yet, not only does the HC pattern represent an innovation, but it also involves apparent cases of local complexification as compared to both French and Gbe. Indeed the learners of HC must acquire a complex of DP- and CP-related morphosyntactic properties that are not found in French or in Gbe—for example, the presence of both pre- and postnominal modifiers and both pre- and postnominal articles (such patterns seem marked from the perspective of Greenbergian universals; see chapter 15). The overall point here is that the specificities and intricacies of HC

grammar, once considered as part of a system, cannot be captured—and certainly not analyzed—in any approach that takes Creole languages to all belong to the same type, with Creole formation boiling down to simplification *qua* reduction of morphological markers.

## 17.4 Creole Formation as Normal Language Change: A Recursive L2A-L1A Cascade

At this stage of the chapter, the question facing us is: How did the respective contributions of the diverse learners in the language contact ecology of the colonial Caribbean contribute to the formation of the relatively stable and uniform sets of I-languages that now go by the label Creole languages?

If we focus on the particular case of Haiti, it is now well established that HC is a relatively homogeneous language, notwithstanding the dialectal differences across geographical areas (Fattier 1998). Yet, given the history of HC, it is expected that the first proto-Creole varieties in Caribbean colonies would have manifested the structural influences of a variety of substrate languages. What would have prevailed in the earliest stages of HC is a set of structurally distinct proto-HC varieties, each showing primary influence from a specific set of substrate languages, depending on the ethnic composition of the corresponding area. One fact that is revealed in Fattier's (1998) extensive dialect atlas for Haitian Creole is that, despite class- and region-based variations, HC is relatively uniform, especially in its morphosyntax. What's striking is that the documented dialectal differences seem largely orthogonal to the inter-substrate differences that would have prevailed at the earliest stages of HC formation. In other words, in Haiti today one would be hard pressed to identify, say, a Gbe-influenced HC dialect vs. a Bantu-influenced HC dialect. What we do find are Gbe-influenced patterns (e.g., postnominal determiners) and Bantu-influenced patterns (e.g., morphemes with Congo cognates) in all dialects of HC.

In light of these observations, it thus appears that L2A did play a key role in Creole formation, with both the native languages of the L2 learners and general strategy of L2A influencing the shapes of their respective interlanguages and the ultimate outcome of Creole formation. Our hunch is that L2A plays a similar role in other instances of (p. 457) language change, as in the history of English (see Kroch et al. 2000 and chapter 18). L1A would have also played a key role in Creole formation, as it does in other instances of language change: the Caribbean-born (Creole) children would have created stable and relatively homogeneous I-languages such that any prior substrate-influenced cross-dialectal differences would have been leveled off through successive L1A by larger and

larger groups of Creole children. The latter, no matter the languages spoken by their parents, would have created their own Creole I-languages (I-Creoles in the terminology introduced in DeGraff 1999a:8-9). The emergence of these I-Creoles in the minds of these early Creole (i.e., Caribbean-born) speakers was conditioned by PLD containing proto-Creole patterns influenced by a diverse set of substrate and superstrate languages and by mutual accommodation across boundaries of these diverse heritage languages. These languages were the L1s of the older non-Creole generations—be they speakers of Niger-Congo languages or speakers of French(-derived) varieties, including proto-HC varieties. It is through successive L1A instances by Creole children that patterns influenced by specific substrate languages would have spread throughout the population at large. And it is also through such L1A that the proto-Creole varieties would acquire stable norms as natively spoken varieties by larger and larger groups of native speakers—Creole speakers with increasing socio-political influence. Thus arises the L2A-L1A cascade in Creole formation.

Similar homogenization processes (or ‘normalization’ in the terminology of Chaudenson and Mufwene 2001) have been documented in real time by Newport (1999) and Kegl et al. (1999). These two studies convincingly show the capacity of children to regularize certain patterns in their PLD. A caveat is in order: we do not consider these two studies to be replicas of Creole-formation scenarios and we do not commit ourselves to the structural details and analyses in these studies: the socio-historical circumstances in Newport and Kegl et al.’s sign language studies differ greatly from what obtained in the case of Caribbean Creole formation, and the nature of the input and output in the sign-language and Creole cases is also different in some crucial aspects—partly due to differences in modalities (spoken vs. signed). But what these studies help us evaluate is the role of children vs. adults when exposed to language input that seems unstable and non-native to varying degrees (see DeGraff 1999b:483-487 for related caveats; cf. Mufwene 2008:ch. 5 for implications vis-à-vis the emergence of communal norms at the population level). Indeed, Newport and Kegl et al. focus on learners of sign languages who are creating their L1s from PLD that is nonfluent and unstable. Such PLD does not provide evidence of certain combinations—in, for example, the morphosyntax for TMA marking. Furthermore, the PLD patterns show inconsistent variability. What the children in these studies did is to process this unusually sparse and inconsistent PLD in order to create a stable system with certain combinations that were missing in the PLD. Similar patterns of regularization by children are documented in S. J. Roberts’ (1999) study of the Hawaiian Creole TMA system (see DeGraff 2009:912-914, 934-936, 945). Such studies give further evidence in favor of a particular role of L1A in the L2A-L1A cascade that we’re positing here as crucial to Creole formation.

(p. 458) Though we use the metaphor of a ‘cascade,’ it may be more appropriate to speak of a ‘recursive cascade’ or a ‘series of overlapping cascades’ where the utterances produced by both L1 and L2 learners feed into the PLD for subsequent L1 and L2 learners, and then the latter’s utterances in turn feed the PLD of newly born L1 learners and newly arrived L2 learners, and so on. It is through these ‘recursive L2A–L1A cascades’ that certain patterns among the output of L1-influenced interlanguages become selected, through prior competition, as key triggers for the subsequent setting of stable properties in the I-Creoles (see Mufwene 2008:ch. 7 for a discussion of the complex ecological factors—psycholinguistic, structural, typological, social, and demographic—that may count toward the comparative weighting of patterns in competition in the course of language change, including Creole formation). The fact that the setting of (internal) properties in the Creole I-languages is based on (external) patterns in necessarily heterogeneous PLD automatically creates room for: (i) the appearance of substrate transfer; (ii) individual-level internal innovations such as reanalysis (or ‘selection with modification’ in Mufwene’s terms).

In section 17.3, we identified phenomena within the clausal and nominal left peripheries of HC (e.g., the emergence of prepositional and modal complementizers and determiners in HC) where patterns emerged in HC based on reanalysis of superstrate patterns with influence from certain substrate patterns (for related ideas in a different framework, see Mufwene 2008, which has inspired some of our own work). In terms of current cartographic views (Rizzi 1997; Aboh 2004) these layers in the clausal and nominal domains represent interfaces between, on the one hand, the predicate and its extended projections and, on the other hand, the discourse. Given this characterization, our discussion suggests that these zones of ‘interface’ (e.g., in the left periphery of the nominal and clausal domains) are more open to innovations based on apparent ‘recombination’ of superstrate and substrate properties (Aboh 2006b). When the parameters to be set involved these interface zones, it’s as if learners, as they processed mutually conflicting input from the PLD (input influenced by L1s with distinct parameter settings—for example, with respect to word order and semantics in the DP), converged on a ‘third way’ with an emergent grammar whose output appears to combine in a novel way certain patterns from the source languages (see Aboh and DeGraff 2012 for a DP-related case study). Our approach thus lends itself to identifying grammatical areas where Creoles innovate new parametric values and where local complexification arises as a result of PLD that are unusually complex due to the language contact situation. As far as we know, this is a novel approach to Creole formation to the extent that its basic UG-based assumptions and its faithfulness to historical details make it prone to identify, and to account for, such areas of local complexification, alongside potential areas that may seem ‘simple’ due to certain superficial consequences of adult learners’ strategies.

## Notes:

(<sup>1</sup>) The names of the co-authors are listed in alphabetical order. This chapter is the outcome of ongoing and long-term collaboration. We bear equal responsibility for the strengths as well as shortcomings of the chapter. Part of this research was supported by Aboh's 2011–2012 fellowship at the Netherlands Institute for Advanced Study in the Humanities and Social Sciences. We are grateful to the editor for inviting us to contribute to this handbook, and we are deeply indebted to Trevor Bass, Bob Berwick, Noam Chomsky, Morris Halle, and Salikoko Mufwene for most thorough and constructive discussions. Salikoko deserves a special medal of friendship for his support, intellectual and otherwise, throughout our long ongoing task of understanding what Creole languages alongside their source languages and other sorts of comparative data can teach us about language contact, language acquisition, and language change and how all of this relates to the human language capacity.

(<sup>2</sup>) This 'null theory' label for our framework was suggested by Beatrice Santorini (p.c., July 2009) with reference to Guglielmo Cinque's (1993) 'null theory of phrase and compound stress.' Cinque's 'null theory' dispenses with language-specific provisos for stress. Similarly, and as pointed out by Santorini, our views about Creole formation make superfluous any Creole-specific proviso for Creole formation.

(<sup>3</sup>) As in DeGraff (2009:894), we 'consider it a fallacy to a priori expect any specific structural or developmental feature of a given Creole (e.g., Haitian Creole in the Caribbean) to necessarily have an analogue in some other Creole (e.g., Reunionese Creole in the Indian Ocean) "simply" because both languages have been called "Creole".' This is in keeping with Mufwene's (2001:138, 2004:460) notion that 'Creole' should be considered, at best, a socio-historical label with blurry boundaries. Consider, for example, the fact (to be discussed in the main text) that the term originated, not with language, but with people (e.g., the 'Creole' people of the Caribbean) and other living species ('Creole' cows and 'Creole' rice; i.e., varieties of cows and rice that are indigenous to the New World). Then again, there are 'Creole people' (e.g., in Cuba) who never spoke any language called 'Creole.' (See Mufwene 1997, Palmié 1996, Mufwene 1997, Roberts 2008 for thorough discussions.)

(<sup>4</sup>) After slavery was abolished, certain plantation owners turned to Asia and focused also on some parts of Africa, notably Nigeria and the collapsing Kongo Kingdom, from where they could import contract laborers. Because we focus on the early years of Creole formation, we do not consider the effect of these latecomers on the emergent language.

(<sup>5</sup>) This communal sense of ‘Pidgin’ qua *lingua franca* is distinct from the use of ‘Pidgin’ in DeGraff (1999a,b, 2009) where the label is used with an individual and internal sense as a cover term for the (early) interlanguages of adult learners in the language-contact setting of Creole formation. In the framework that we sketch in this chapter, such interlanguages, at various stages of development (from early to advanced), do play a key role in Creole formation and in all other instances of contact-induced language change (see sections 17.3 and 17.4).

(<sup>6</sup>) We stress the ‘local’ in ‘local complexity’ here (and, elsewhere, in ‘local simplicity’) to highlight our skepticism vis-à-vis various claims that certain languages can be *in toto* more (or less) complex than others.

(<sup>7</sup>) The example with 15-year-old basketball players is due to Trevor Bass and is discussed in greater detail in DeGraff et al. 2013. (Thank you, Trevor, for this example and for much more.)

(<sup>8</sup>) The very concept of ‘decreolization’ in the history of Caribbean Creoles is problematic to the extent that the ‘acrolectal’ varieties that are often taken to be the results of decreolization (i.e., those varieties that are structurally closest to the European lexifier) would have been among the earliest to emerge, then, to subsequently co-exist with the more ‘basilectal’ varieties (i.e., those that are structurally most removed from the European lexifier). In other words, the earliest Creole varieties were closer to the lexifier than the later ones. This is carefully documented in Lalla and D’Costa (1989) (also see Alleyne 1971; Bickerton 1996; and Chaudenson and Mufwene 2001). In the case of Haiti, the contact with French was reduced to a minimum after independence in 1804, with most Creole speakers being monolingual and having little, if any, contact with French speakers (DeGraff 2001b:229–232).

(<sup>9</sup>) Wichmann and Holman (2010) provides extensive empirical and quantitative observations, alongside methodological caveats, about the relationship, and lack thereof, between structural similarity and genealogical relatedness—observations and caveats that raise further doubt on the reliability of Bakker et al.’s claims about typology and relatedness. See also chapter 16.

(<sup>10</sup>) ‘In addition, the characters must then be coded for each language—a step that is itself critically important as it is often here that mistakes are made, even by trained linguists...’ (Nichols and Warnow 2008:769). See Kouwenberg (2010, 2012), Parkvall (2012), Fon Sing and Leoue (2012), and DeGraff et al. (2013) for examples of such mistakes in the coding of characters.



(<sup>11</sup>) Bakker et al.'s (2011) claims become even more worrisome when it's discovered that their empirical generalizations are logically incompatible, contradict one another, contradict well-known facts from both Creole and non-Creole languages, contradict data documented in their own publications, and contradict some of their own references, including *WALS* (see DeGraff et al. 2013).

(<sup>12</sup>) Our '*Sprachbund*' concept is different from the 'scattered *Sprachbund*' in Kihm (2011), where Pidgins and Creoles are postulated to share *Sprachbund*-like similarities due to their common emergence via untutored second language acquisition (e.g., the Basic Variety as discussed in Klein and Perdue 1997). Instead, trans-Atlantic *Sprachbund*, as used here, is based on attested features in the specific languages that were in contact in the colonial Caribbean and whose speakers participated in Creole formation. As we've mentioned earlier, this would mean that Creoles will display distinct structural properties depending on the ecology they emerged in. As it turns out, some of the *Sprachbund* features highlighted in Kihm's study of Guinée-Bissau Kriyol and Nubi Arabic (e.g., inflectional morphology for passives and for plural number) are absent in Atlantic Creoles (see Kihm 2011:53, 17, 74, 80–81). Therefore, such features cannot be claimed as pan-Creole *Sprachbund* features. For related caveats (e.g., regarding Klein and Perdue's Basic Variety as an explanation to pan-Creole similarities), see DeGraff (2011b:249–250), also see note 10.

(<sup>13</sup>) The term 'quasi' in 'quasi-*Sprachbund*' is used advisedly: not to suggest that Caribbean Creoles have converged to similar patterns via mutual borrowings, but only to highlight the presence of similar patterns that seem due to their ancestry in a relatively narrow band of typologically similar languages from Romance, Germanic, and Niger-Congo languages.

(<sup>14</sup>) Donohue et al.'s concept 'linguistic geography' is misunderstood by Daval-Markussen and Bakker (2012:90) who take their results to 'go against the conclusions of Donohue et al. (2011), who claim that the various clusterings observable in phylogenetic networks are due to the effects of areality and geography rather than to genealogy.' Donohue et al. (2011:369) defines 'linguistic geography [as] the network of contact and diffusion that postdates a proto-language, in *most cases* corresponding to geographic distance' [emphasis added]. More generally, linguistic geography can be defined as 'the spatially measured network of social interactions,' which, in turn, entail 'the diffusion (*spatial or social*) of linguistic traits' [emphases added]. Such diffusion is compatible with the sort of quasi *Sprachbund* effects that, in the case of Caribbean Creoles, were caused by overlapping sets of Niger-Congo substrates and Germanic or Romance superstrates. It is important to note that such (long-distance) areal effects via socio-historically determined diffusion of certain typological features are compatible with genealogical

relationships as determined by the comparative method. According to the latter (and as argued in the main text) Caribbean Creoles are genealogically related to their respective European lexifiers, contrary to Daval-Markussen and Bakker.

(<sup>15</sup>) Here we use the term ‘parameters’ as a shorthand, as we remain agnostic on the ontology of these ‘parameters’ and the actual path and mechanics for the setting of ‘parameters’ in the course of language acquisition (see chapters 11, 14, and 16 for relevant discussion).

(<sup>16</sup>) See Koopman and Lefebvre (1982) and Mufwene and Dijkhoff (1989) for discussions of whether Creoles morphologically distinguish between finite and nonfinite verbs. Here we take ‘finiteness’ to have syntactic correlates. So a language like HC can still show finite vs. nonfinite *syntactic* distinctions (in terms of, say, structure and binding domains) independently of morphological distinctions (or lack thereof) on verbal forms (see DeGraff 2009:67f for observations and references).

(<sup>17</sup>) Mathieu’s (2009) analysis makes bare NPs contingent on the availability of interpretable  $\varphi$ -features on the N. These  $\varphi$ -features are, in turn, related to overt agreement features on N. The latter are present in Old French but absent in both HC and Gbe whose morphological profiles seem to constitute a challenge for Mathieu’s analysis.

(<sup>18</sup>) The HC definite determiner *la* has allomorphs *lan*, *nan*, *a*, *an*: *chat la* ‘the cat’, *chanm lan* ‘the bedroom,’ *machin nan* ‘the car,’ *dan an* ‘the tooth,’ *bra a* ‘the arm.’ The French etymon is the deictic locative adverbial and discourse particle *là* in Spoken French as in *T’as vu ce chat-là là* ‘Did you see that cat there, yeah?’ (with the first *là* as locational deictic adverbial and the second *là* as discourse particle).

### **Enoch Aboh**

Enoch O. Aboh is Professor of Linguistics and Learnability at the University of Amsterdam. He explores issues of learnability of human languages with a special focus on theoretical syntax as related to the discourse-syntax interface, language creation and language change.

### **Michel deGraff**

Michel DeGraff is Professor of Linguistics at MIT and Director of the ‘MIT-Haiti Initiative’ funded by the U.S. National Science Foundation. He is also a founding member of Haiti’s Haitian Creole Academy. He studies Creole languages, focusing on his native Haitian Creole. His research deepens our understanding of the history and structures of Creole languages. His analyses show that Creole languages, often described as ‘exceptional’ or ‘lesser,’ are fundamentally on a par with non-Creole

languages in terms of historical development, grammatical structures, and expressive capacity. His research projects bear on social justice as well. In DeGraff's vision, Creole languages and other so-called 'local' languages constitute a necessary ingredient for sustainable development, equal opportunity, and dignified citizenship for their speakers—a position that is often undermined by theoretical claims that contribute to the marginalization of these languages, especially in education and administration.

