1 *The problem of creole genesis and linguistic theory*

This book addresses the cognitive processes hypothesised to account for the properties of creole languages. It presents a theory of creole genesis based on processes otherwise observed to play a role in language genesis and in language change in general. It is intended as a contribution to our understanding of the mechanisms by which the properties of the source languages of a creole manifest themselves in the creole in the way they do. This chapter summarises the salient features of creole languages which any theory of creole genesis must be able to account for and situates this book with respect to other approaches to the problem. Section 1.3 introduces the theoretical framework within which these problems are addressed. Section 1.4 presents the hypothesis underlying the research. Section 1.5 discusses the scope and limitations of this book.

1.1 The complex problem of pidgin and creole genesis

The history and structure of pidgin and creole languages are characterised by the following features.[1] First, as was pointed out by Whinnom (1971), these languages are only developed in multilingual communities. Whinnom argues that, in bilingual communities, the speakers of one group will eventually learn the language of the other group.[2]

Second, communities where pidgin and creole languages emerge generally involve several substratum languages spoken by the majority of the population and a superstratum language spoken by a relatively small but economically powerful social group. Crucially, the substratum community does not have one common language. This situation creates the need for a lingua franca (see e.g. Hymes 1971a; Foley 1988), not only to permit communication between the speakers of the substratum languages and of the superstratum language, but also to permit the speakers of the substratum languages to communicate among themselves (see e.g. Foley 1988; Singler 1988: 47; Thomason and Kaufman 1991).

Third, in communities where creole languages emerge, speakers of the substratum languages generally have very little access to the superstratum language (see Thomason and Kaufman 1991). As Foley (1988: 163) puts it: 'the language of the dominant group is not easily made available to the members of the subordinate group(s)'. In fact, as has been pointed out on several occasions in the literature, creoles that most resemble their superstratum languages were created in communities where the speakers of the substratum languages had relatively

more access to the superstratum community. Creoles that are more radical (i.e. less like the superstratum language) come from communities where language learners had very little access to the superstratum community (see e.g. Bickerton 1977: 55; Baker and Corne 1982; Andersen 1983a; Thomason and Kaufman 1991; Baker 1993; Valdman 1978, 1993). For example, as is argued in Valdman (1993), Louisiana creole is closer to French than Haitian is because the substratum speakers had more access to French in Louisiana than the African population had in Haiti. Baker and Corne (1982) also discuss this issue on the basis of data from Mauritius and Reunion creoles. On Reunion, French native speakers outnumbered substratum speakers during the formative period of the creole, and Reunion creole grammar displays a significant number of French grammatical categories. By contrast, during the formative period of Mauritius creole, the proportion of native French speakers was much lower, and thus the West African speakers had a much stronger input into the creole.

A fourth point is that, ordinarily, languages change gradually. Within the span of several generations, speakers of innovative and conservative dialects are able to communicate, even though, over the course of centuries, a new language may evolve (see Lightfoot 1979). By contrast, creole languages are created in a short span of time (see e.g. Hall 1958; Voorhoeve 1973; Alleyne 1966; Chaudenson 1977, 1993; Bickerton 1984). This observation dates back to Van Name (1869–70: 123, cited in Goodman 1964: 135): 'Under ordinary conditions these changes proceed at so slow a pace as to be appreciable only at considerable periods of time, but here two or three generations have sufficed for a complete transformation.' Hesseling (1933: xi) further reassesses this point in the following terms:

> The genesis of human language is a psychological problem that no single language will ever solve, but from creole one can best learn how a given language emerges from old data and develops, because here something takes shape at a high speed, in a past recognisable to us, something which is the product, in other cases, of many centuries, with a very obscure past in its background.

Thus, in contrast to regular linguistic change, creole languages diverge abruptly from their source languages (see Thomason and Kaufman 1991), so that, within one or two generations, a different language is created. Hancock (1987: 265) claims that: 'most of the principal characteristics that each creole is now associated with were established during the first twenty-five years or so of the settlement of the region in which it came to be spoken'. Hymes (1971a), Mintz (1971) and Ferraz (1983) suggest that a creole can develop within fifty years or less. Singler (1996) is of the opinion that it takes sixty to eighty years for a creole to form. Whatever the outcome of this issue may be, creole languages constitute a unique case of accelerated linguistic change when compared with regular cases of linguistic change.

Fifth, creole languages tend to be isolating languages. This observation goes back to Schuchardt (1979) and Hesseling (1933: xvi). It is also found in Hagège (1985: 39). But it was Mufwene (1986, 1990, 1991) who clearly established this property of creole languages and the problem it poses for scholars who work on

creole genesis. Indeed, Mufwene has documented the fact that this tendency appears to hold even when the contributing languages are not isolating languages. For example, Mufwene (1986) shows that Kituba, a creole language that has emerged almost exclusively from contact among agglutinative Bantu languages, is an isolating language. 'Kituba has selected Kikongo's seemingly marked periphrastic alternative over the more common and apparently unmarked agglutinating system' (Mufwene 1990: 12).

Sixth, it has long been noted in the literature that creole languages are mixed languages in that they derive some of their properties from those of the substratum languages and some from those of the superstratum language (see e.g. Alleyne 1966, 1981; Holm 1988). Moreover, several scholars have noticed that the type of mix we find in radical creoles is not random. For example, Adam (1883: 47) states that:

> J'ose avancer . . . que les soi-disant patois de la Guyane et de la Trinidad constituent des dialectes négro–aryens. J'entends par là que les nègres guinéens, transportés dans ces colonies, ont pris au français ses mots, mais qu'ayant conservé dans la mesure du possible, leur phonétique et leur grammaire maternelles . . . Une telle formation est à coup sûr hybride . . . La grammaire n'est autre que la grammaire générale des langues de la Guinée.
>
> [I go so far as to claim . . . that the so-called patois of Guyana and Trinidad constitute Negro–Aryan dialects. By that I mean that the Guinean Negroes who were transported to the colonies adopted the words of French but, as much as possible, kept the phonetics and grammar of their mother tongues . . . Such a formation is clearly hybrid . . . The grammar is no different from the general grammar of the languages of Guinea.]

Speaking of Haitian creole, Sylvain (1936: 178) observes that: 'Nous sommes en présence d'un français coulé dans le moule de la syntaxe africaine, ou . . . d'une langue éwé à vocabulaire français.' [We are in the presence of a French that has been cast in the mould of African syntax or . . . of an Ewe language with a French vocabulary.] Similarly, in his extensive study of French-based creoles, Goodman (1964) observes, over and over again, that particular lexical items in the creoles have a phonological representation similar to a French expression but that these creole lexical items share properties with corresponding lexical items in the African substratum languages. On the basis of data drawn from Djuka, Huttar (1971: 684) also remarks that 'the use of morphemes borrowed by a pidgin or a creole language . . . from a European language often diverges from the use of the source morpheme in the source language' and often corresponds to the use of the corresponding word in the substratum languages. Voorhoeve (1973) makes a similar remark on the basis of Sranan and Saramaccan data. These observations suggest that creole languages are not formed by an arbitrary mixture of the properties of the languages present at the time they are being created. The general pattern that seems to emerge from the observations reported above is the following: while the forms of the lexical entries of a radical creole are derived from the superstratum language, the syntactic and semantic properties

of these lexical entries follow the pattern of the substratum languages. This raises the question of what the process which generates such a division of properties could be. The answer to this question is the main topic of this book.

Any theory of creole genesis must account for the properties of these languages. Therefore, as has been pointed out in Lefebvre and Lumsden (1989a), an optimal theory of creole genesis must account for the fact that creole languages emerge in multilingual contexts where there is a need for a lingua franca and where the speakers of the substratum languages have little access to the superstratum language. It must account for the fact that creole languages tend to be isolating languages even when they emerge from contact situations involving only agglutinative languages. It must also account for the fact that creole languages manifest properties of both their superstratum and substratum languages and explain why these properties are divided as they are.[3] In this book, the problem of creole genesis is addressed on the basis of an in-depth study of the genesis of Haitian creole, a typical example of a radical creole (see Bickerton 1984).

Pidgins and creoles have long been considered as separate entities on the basis of the following two sets of criteria. While pidgins have been defined as reduced codes, creoles have been defined as expanded versions of these reduced codes (see e.g. Hymes 1971b). Also, while pidgins have been found to always constitute the second language of the speakers who use them, a creole is often considered to be a pidgin that has become the first language of a new generation of speakers (see Kay and Sankoff 1974). In more recent literature, the distinction between pidgins and creoles has been levelled out in view of the fact that there are some pidgins (still used as a second language) that have been shown to have expanded in the same way as languages known as creoles (see e.g. Mühlhäusler 1980, 1986a, for an extensive discussion of this point). Hancock (1980a: 64) states: 'I prefer not to acknowledge a distinction between *pidgin* and *creole*, and to consider *stabilisation* more significant than *nativisation* in creole language formation.' Similarly, Mufwene (1990: 2) uses the term creole to refer 'to varieties traditionally called creoles but also to those called pidgins that serve as vernaculars or primary means of communication for at least a portion of their speakers'. Moreover, in recent literature in the field, scholars have started referring to pidgins and creoles as PCS, suggesting that they fall into a single category. Furthermore, as will be seen in chapter 2, pidgin and creole languages cannot be distinguished on the basis of the processes which play a role in their formation (see also Woolford 1983, for a general discussion of this point). Indeed, the processes hypothesised to play a role in the formation and development of human languages apply to both pidgins and creoles. This is a major drawback to Bickerton's (1977, 1981, 1984) Language Bioprogram Hypothesis of creole genesis, which crucially requires that pidgins and creoles be different entities formed by different processes. Since this book is about the processes involved in the genesis of these languages and since these languages cannot be distinguished on the basis of these processes, I will not make any distinction between them.

1.2 The perspective of this book

In this book, the problem of pidgin and creole genesis is cast within the framework of the cognitive processes otherwise known to play a role in language genesis and change in general. This general perspective is akin to Van Name's (1869–70: 123) claim that the type of changes undergone in creole genesis are no different in kind from those observed in regular cases of linguistic change: 'the changes which they [the creole dialects] have passed through are not essentially different in kind, and hardly greater in extent than those, for instance, which separate the French from the Latin, but from the greater violence of the forces at work they have been far more rapid'. The major processes hypothesised to be involved in the genesis of pidgin/creole lexicons are relexification, reanalysis and dialect levelling. These processes can be argued to play a significant role in language genesis and language change in general (see chapter 2). It is also hypothesised that the creators of the pidgin/creole use the parametric values of their native languages in establishing those of the language that they are creating and the semantic principles of their own grammar in concatenating morphemes and words. This approach compares with others as follows.

The presence of substratum features in pidgin and creole lexicons has traditionally been considered to result from *calquing* (see e.g. Keesing 1988) or *transfer* (see e.g. Naro 1978; Andersen 1980, 1983b; Mufwene 1990, 1993c; Siegel 1995). In this book, it is argued that such cases constitute examples of relexification when lexical properties are involved. Parametric values are hypothesised to be set on the basis of those in the substratum languages and to be carried over into the creole by its creators. The same hypothesis applies to semantic interpretation (see chapter 2).

The problem of pidgin and creole genesis has traditionally been addressed from the point of view of simplification, or reduction, and expansion (see e.g. Hymes 1971a, 1971b). Pidgins and creoles have traditionally been viewed as reduced or simplified codes when compared with their superstratum languages. Such a view, however, has been challenged by Alleyne (1966: 281), among other researchers, on the basis of a comparison between a creole language and its contributing languages.

> Dans l'histoire de la morphologie, est-il permis de partir du système de flexions français et de ne voir dans les créoles français qu'une réduction ou une simplification de ce système amenant des 'pertes' ou des 'disparitions' des flexions françaises? Ou bien notre point de départ devrait-il être la morpho-syntaxe ouest-africaine, qui est caractérisée par l'invariabilité du mot, donc par l'absence de flexions?
>
> [In the history of its morphology, is it permissible to start from the French inflectional system and to see in the French creoles only a reduction or simplification of this system, resulting in the 'loss' or 'disappearance' of French inflections? Or should we start with West African morphosyntax, which is characterised by invariable words, i.e. by the absence of inflections?]

The two questions posed in the quotation from Alleyne above stress the fact that the notion of simplification/reduction arises only when creoles are compared with their superstratum languages. These notions do not have the same relevance, however, when creoles are compared with their substratum languages (see also Brousseau, Filipovich and Lefebvre 1989). The approach taken in this book builds on the second alternative raised by Alleyne. The cognitive process of relexification will be shown to account for the link that exists between the morphosyntax of a creole and that of its substratum languages. The notion of expansion, as referred to in Hymes (1971a), corresponds to the result of the process of reanalysis, claimed to play a role in the development of pidgins and creoles.

In the last twenty years, discussions of creole genesis have centred around the debate over three main approaches to the problem (see Muysken and Smith 1986b): the universalist approach (see e.g. Bickerton 1981, 1984, 1986; Seuren and Wekker 1986), the superstratist approach (see e.g. Chaudenson 1993) and the substratist approach (see e.g. Alleyne 1981; Holm 1988).[4] The universalist approach does not account for the fact that creole lexicons manifest the properties of their source languages in the way they do (see section 1.1). Furthermore, creole languages are not uniform; like other natural languages, they manifest language-specific features, as is extensively documented in Muysken (1988b). The universalist approach does not account for the variation that exists between creoles. The superstratist approach raises a problem best stated by Mufwene (1996: 166): 'One of the problems with the superstrate hypothesis is the absence of any explanation for why creoles lexified by European languages do not correspond to any particular dialect of their lexifiers.' Finally, the problem with the substratist approach has been stated by Hall (1958), who points out that creoles in general have retained very few, if any, *visible* features of their substratum languages. The perspective adopted in this book isolates the discussion of pidgin and creole genesis from these approaches in addressing the problem from the point of view of the processes at work in their formation. Furthermore, it resolves the problems with all three approaches; indeed, the nature of the processes hypothesised to play a role in the formation of these languages will be shown to predict the respective contributions to pidgins or creoles of the languages involved in their formation.

Traditional accounts of creole genesis have generally addressed the problem using the notion of *language*. Given the nature of the mix found in creole languages, however, it is necessary to distinguish between a particular form and its properties. This requirement has been clearly stated by Alleyne (1966: 282) on the basis of the status of temporal and aspectual morphemes in Haitian.

> Il est aisé de voir que ce système diffère beaucoup de celui des verbes français, dans lequel les distinctions temporelles sont beaucoup plus importantes que les distinctions d'aspect. Par contre, les langues ouest-africaines font preuve de systèmes verbaux du même genre que celui du créole français, et il serait évidemment plus valable d'attribuer au système verbal du créole une origine

africaine plutôt qu'une origine française. Mais les particules qui marquent les aspects et les temps dans le système verbal créole semblent toutes dérivées de mots ou de groupes de mots français. Ainsi *ap*(*é*)<*après*; *té* (particule du passé)<*était*; *ava*, *va*, *a* (futur)<*avoir* (ou peut-être *va* ou *avant*, d'après d'autres hypothèses); *fèk*<*fait que* (*mwè fèk kumãsè* 'je viens de commencer'). Donc si notre point de départ réside dans les formes et non pas dans le système, nous serions amenés à chercher les moyens par lesquels le sens et la fonction grammaticale de certaines formes françaises ont pu évoluer ou se transformer dans le créole. Cela peut fausser nos perspectives et nous plonger dans des hypothèses irrationnelles pour exprimer le changement de fonction de *après* en créole.

[It is easy to see that this system differs greatly from the French verb system, in which tense distinctions are far more important than distinctions of aspect. On the other hand, the West African languages have a verb system of the same kind as that in the French creole, and it would obviously be much more valid to attribute an African rather than a French origin to the creole verb system. But the particles that mark aspects and tenses in the creole verb system all seem to be derived from French words or groups of words. Thus, *ap*(*é*)<*après*;[5] *té* (past particle)<*était*; *ava*, *va*, *a* (future)<*avoir* (or maybe *va* or *avant*, according to other hypotheses); *fèk*<*fait que* (*mwè fèk kumãsè* 'I just began'). Thus, if our starting-point resides in the forms, rather than the system, we shall be led to look for the means whereby the meaning and grammatical function of certain French forms evolved or were transformed in the creole. This can distort our point of view and immerse us in irrational hypotheses to account for *après*'s change of function in the creole.]

The situation described above calls for a theory of grammar which allows forms and functions to be manipulated independently. Models developed within the framework of generative grammar do allow for such analyses, for they provide a modular approach to the various components that define a grammar. In this approach, each module is independent from the others. Hence, phonological representations may be treated independently from the semantic and syntactic properties that define the functions of particular lexical entries. This general approach provides a tool to address the problem posed by Alleyne.

1.3 Linguistic theory

The last thirty years have seen a significant shift in the focus of linguistic theory from E(xternal) language to I(nternal) language. E-language stands for the neogrammarians' and structuralists' view that a language is a habit system assumed to be overdetermined by the available evidence. I-language refers to the generativists' view that a language is 'some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker–hearer' (Chomsky 1986: 22).

While traditional accounts of the genesis of creoles have addressed the questions posed by their origin from the point of view of E-language, the account proposed in this book takes them up from the point of view of I-language. As Chomsky (1986: 3) puts it:

> Generative grammar . . . is concerned with those aspects of form and meaning that are determined by the *language faculty*, which is understood to be a particular component of the human mind. The nature of this faculty is the subject matter of a general theory of linguistic structure that aims to discover the framework of principles and elements common to attainable human languages; this theory is now often called *universal grammar* (UG) . . . UG may be regarded as a characterisation of the genetically determined language faculty. One may think of this faculty as a *language acquisition device*, an innate component of the human mind that yields a particular language through interaction with presented experience, a device that converts experience into a system of knowledge attained: knowledge of one or another language.

Since creole languages are natural languages, it must be the case that the properties of creole languages follow from the more general properties of the cognitive system which are pertinent to the configuration of natural languages and to the transmission/acquisition of language in general.

The theory of principles and parameters (see Chomsky 1981, 1986 and related work) holds that natural languages are basically similar. In this model, those properties of language that are universal are formulated in terms of universal principles of grammar. The properties that are language-specific are hypothesised to be located in the lexicon, the syntactic parameters and the interpretive component of the grammar. This model constitutes a most useful tool for addressing the problem of creole genesis for it provides us with a principled division between language universals and language-specific features. Thus, on this approach, universals of language will be manifested in creoles in the same way as in any other natural languages. What is specific to a particular creole will be found in the components of the grammar that allow for variation between languages. Thus, using such a model provides us with a tool to identify areas where the creole can diverge from or resemble its source languages. A comparison of the language-specific features of a particular creole with corresponding features in its contributing languages should tell us the source of those features. Likewise, this model provides us with a tool to address the problem of variation between creoles.

The mentalist approach to grammar and lexicon allows for the manipulation of semantic and syntactic information independently of phonological representations, and it provides us with the appropriate tool to discuss transmission/acquisition in contexts of creole genesis in terms of the transmission/acquisition of grammar in spite of the fact that, in these cases, a new language has been created.

Finally, the mentalist approach to grammar and the lexicon defines the object of inquiry (and hence, the methodology) with regard to a creole language as follows: (1) What does a creole speaker know about the grammar and the lexicon of his/her language which enables him/her to produce and understand utterances in this language? (2) Abstracting away from the phonological representations of the lexical entries of the various languages involved in the genesis of a creole language, how does this knowledge compare with the knowledge speakers of the creole's source languages have of their grammars and lexicons? These two questions constitute the central core of the research reported on in this book.

1.4 The hypothesis

The general hypothesis tested by the research reported on here is that the creators of a creole language, adult native speakers of the substratum languages, use the properties of their native lexicons, the parametric values and semantic interpretation rules of their native grammars in creating the creole. Creole lexical entries are mainly created by the process of relexification.[6] Two other processes fed by the output of relexification, dialect levelling and reanalysis, also play a role in the development of the creole (see Koopman and Lefebvre 1981; Lefebvre 1984, 1993a; Lefebvre and Lumsden 1994b; Lumsden and Lefebvre 1994).

Relexification is a mental process defined as follows by Muysken (1981a: 61): 'Given the concept of lexical entry, relexification can be defined as the process of vocabulary substitution in which the only information adopted from the target language in the lexical entry is the phonological representation.' In testing the role of relexification in creole genesis, we have adopted the strong position that all the lexical entries listed in the lexicon could, in principle, undergo relexification (see e.g. Lefebvre and Kaye 1985–9 Projects; Brousseau, Filipovich and Lefebvre 1989; Lefebvre and Lumsden 1989a). Thus, based on a theory of the lexicon which, in addition to listing major category lexical items, lists functional category lexical items, productive derivational affixes and idiosyncratic expressions such as unpredictable compounds (see e.g. Lieber 1980, 1992), all these lexical entries[7] should, in principle, undergo relexification, within the limits imposed by the definition of the process (see chapter 2). This book presents evidence that all these types of lexical entries do, in fact, undergo relexification. The relexification hypothesis predicts that the lexical entries of the creole will have the semantic and syntactic properties of the corresponding lexical entries in the substratum languages and phonological forms derived from phonetic strings found in the superstratum language. To a great extent, this is exactly what we do find.

The idea that relexification plays a role in pidgin and creole genesis is not a new one. For example, Stewart (1962), Whinnom (1977), Voorhoeve (1973) and others have long claimed that this is so. At one point, Muysken (1981a: 77) also proposed that relexification plays a role in the formation of these languages: 'If it is the case that the Caribbean creoles show numerous African survivals in their syntax and semantics, then I think we can argue that it is not interference which led to these survivals, but relexification.'[8]

The research presented here has gone further than previous studies in several ways. First, our research has improved the formal characterisation of how superstratum data are processed in relexification (see chapter 2). Second, our theory provides a clear statement of how relexification applies in the case of functional category lexical entries and derivational affixes. Such a theory has never been proposed in the past, for the general assumption was that functional categories and derivational affixes do not undergo relexification. For example, Muysken (1988a: 15) claimed that functional lexical entries do not undergo relexification. '[Functional categories] do not have a meaning outside the linguistic system that

they are part of, since their meanings are paradigmatically defined within that linguistic system. So when you relexify a system of function words, automatically the semantic organisation of the target language comes in, and the result is at best a compromise between source and target language systems.' Muysken (1988c) also claimed that affixes and clitics may not undergo relexification either. This view is compatible with a theory of the lexicon where functional category items do not constitute lexical entries. On this approach to the lexicon, the functional category lexical items of a creole must be hypothesised to have evolved through reanalysis only, as is extensively discussed in Lefebvre (1984). In a theory where functional category items and derivational affixes are listed in the lexicon, however, such lexical entries are, in principle, eligible for relexification. The third difference between this project and previous research is that we were able to gather the resources to test this hypothesis from a global perspective (see chapter 3). To the best of my knowledge, this is the first time in the history of creole studies that such a large enterprise has been undertaken.

It is claimed here that, in creole genesis, the process of relexification is used by speakers of the substratum languages as the main tool for acquiring a second language, the superstratum language. The account proposed in this book is a further development of the second language acquisition theory of creole genesis (see Lefebvre and Lumsden 1989a). For example, Alleyne (1971, 1981), Schumann (1978), Valdman (1980), Andersen (1980), Mufwene (1990), Thomason and Kaufman (1991), and Chaudenson (1993) have proposed that pidgin/creole languages constitute a crystallised incomplete stage of second language acquisition. Without relexification, however, this approach to creole genesis does not explain why creole languages have crystallised in the way they have (see Lefebvre 1984, to appear a; Lefebvre and Lumsden 1989a). It is argued that the relexification hypothesis does explain why creole lexicons reflect the properties of both their superstratum and substratum source languages in the way they do.

By definition, relexification is a mental process that is available to speakers who are in possession of a mature lexicon. The relexified lexicons constitute the first instantiation of a new language: the early creole. Hence, according to the relexification hypothesis, a creole is not created by children who are deprived of a model for language, as is advocated by Bickerton (1981, 1984). Rather, it is created by speakers who already have a mature lexicon. This claim is compatible with the fact that the lexical entries of the relexified lexicons reproduce the semantic and syntactic properties of the substratum languages. On the basis of both historical and linguistic facts involving the genesis of Haitian creole, it is argued that creole languages must be created by adult speakers with a mature lexicon.

On the one hand, relexification is a mental process and hence it is an individual activity. On the other hand, situations where creoles are created typically involve several substratum languages. Consequently, and as has been pointed out by Lumsden and Lefebvre (1994), the new lexicons may present differences which, by hypothesis, should reflect the differences between the original lexicons.