

1 The study of pidgin and creole languages

Pieter Muysken and Norval Smith

1.1 Introduction

This book is concerned with pidgin and creole languages. This statement might well give the impression that we know precisely what is meant by these terms. In fact they are the subject of much debate. Creolists agree neither about the precise definition of the terms **pidgin** and **creole**, nor about the status of a number of languages that have been claimed to be pidgins or creoles. Mixed languages, introduced in chapter 4, have generally not been mentioned at all.

To turn first to **pidgin languages**, it is generally agreed that in essence these represent speech-forms which do not have native speakers, and are therefore primarily used as a means of communication among people who do not share a common language. The degree of development and sophistication attained by such a pidgin depends on the type and intensity of communicative interaction among the its users. Mühlhäusler (1986) makes three basic distinctions amongst speech-forms that creolists have referred to as pidgins – (rather unstable) **jargons**, **stable pidgins**, and **expanded pidgins** (see further chapter 3).

To turn to **creole languages** (or just **creoles**), one vital difference from pidgins is that pidgins do not have native speakers, while creoles do. This is not always an easy distinction to make, as one aspect of the worldwide increase in linguistic conformity, and the concomitant reduction in linguistic diversity, is that extended pidgins are beginning to acquire native speakers. This has happened for instance with Tok Pisin, Nigerian Pidgin English, and Sango (Central African Republic), to name but three cases. In particular this has tended to occur in urban environments, where speakers from different ethnic groups have daily contact with each other. The pidgin then becomes the town language. The children of mixed marriages frequently grow up speaking the home language – the pidgin – as their native language.

1.2 Historical linguistics and the definition of a creole

A creole language can be defined as a language that has come into existence at a point in time that can be established fairly precisely. Non-creole languages are assumed (often in the absence of detailed knowledge of their precise development) to have emerged gradually.

So Archaic Latin developed into Classical Latin, the popular variety of which in turn developed into Vulgar Latin, which among other things developed into Old French, which developed into Middle French, which in turn developed into Modern French. While some stages of this development involved more radical changes in the language than others, we can claim with some justification to be able to trace the line of development from Modern French back to Archaic Latin – the earliest recorded stage of Latin, with on the whole little difficulty. Before that we have to rely on linguistic reconstruction, but once again it is fairly obvious that Latin is a typical Indo-European language, and can thus be safely assumed to have developed from Proto-Indo-European, through the intermediate stages of possibly Proto-Italo-Celtic and certainly Proto-Italic. Proto-Indo-European itself may have been spoken somewhere in Southern Russia (an anachronistic term, of course) around 5000 B.C.

This kind of statement we can definitely not make when talking about creole languages. These exhibit an abrupt break in the course of their historical development. So we cannot say that Sranan (the major English-lexifier creole of Surinam; see chapter 18) derives in any gradual fashion from Early Modern English – its most obvious immediate historical precursor. Even a cursory comparison of Early Modern English with the earliest forms of Sranan (first recorded in 1718) will make it abundantly obvious that we are dealing with two completely different forms of speech. There is no conceivable way that Early Modern English could have developed into the very different Sranan in the available 70 or so years. Even the phonological developments required would be extreme, not to speak of the wholesale changes that would have had to have taken place in the syntax.

So creole languages are different from ordinary languages in that we can say that they came into existence at some point in time. Applying the techniques of historical linguistics to creoles is therefore not simple, and in addition presupposes answering the question of which languages the creole should be compared with: the language which provided the lexicon, or the language(s) which were responsible for most aspects of grammatical structure – inasmuch as it is possible to identify these.

It is clear in fact that creole languages develop as the result of ‘linguistic violence’ (and, as we shall see, frequently social violence too). In other words, we have to reckon with a break in the natural development of the language, the natural transmission of a language from generation to generation. The parents of the first speakers of Sranan were not English speakers at all, but speakers of various African languages, and what is more important, they did not grow up in an environment where English was the norm. How **creolization**, the development of a creole language, takes place, or at least what the various theories are concerning how it takes place, we cannot really go into at this juncture – this is a controversial matter that will be dealt with in chapters 8 through 11, and briefly below.

What is clear is that creole languages are not in the slightest qualitatively distinguishable

from other spoken languages. Many of them tend to have certain features in common, but creolists are divided as to the interpretation of this fact, and a language like Chinese resembles many creole languages in its grammar. This means that before we can claim a language to be a creole, we need to know something about its history, either linguistic or social, and preferably both. As we know comparatively little about the detailed development of most languages in the world, and virtually nothing of the history of most ethnic groups, this inevitably means that there may be many unrecognized creole languages around the world.

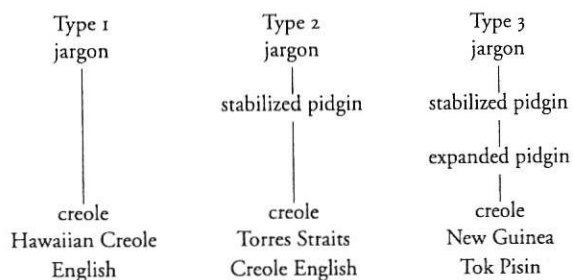
One problem in the identification of particular languages as creoles is caused by the not unusual circumstance that creoles tend to be spoken in the same geographical regions as the languages that provide the greater portion of their lexica (their donor languages, or **lexifier** languages). In some cases we find a **continuum** of speech-forms varying from the creole at one end of the spectrum (the **basilect**), through intermediate forms (**mesolectal** varieties), to the lexifier language (the **acrolect**). Sometimes speech-forms exist which apparently represent cases where either the original mesolect has survived, while the basilectal creole, and sometimes also the original lexifier language have not. Such cases may be referred to as **post-creoles**. Other cases seem rather to involve partial creolization, or influence from a creolized form of the same language. These languages may be termed **semi-creoles** or **creoloids**. Afrikaans seems likely to have been the result of some such process. While linguists would not in general wish to recognize this language as being a full creole, many aspects of Afrikaans are reminiscent of the things that happen during creolization. Other cases of putative creoloids are American Black English, and at least some forms of Brazilian Portuguese.

A quite different situation involving an ‘intermediate’ status is the case of the **mixed languages**. This type which has until now been the object of comparatively little study, involves cases where two languages clearly make a significant contribution to language – frequently one language provides the content words, and another the grammar. Here there is not necessarily any question of simplification. A well-known case of this type to be studied – *Media Lengua* (lit. ‘middle language’) (Muysken 1981b) – is spoken in Ecuador, and involves Spanish lexical items, combined with basically Quechua syntax, morphology, and phonology. Bakker (1992) has referred to this kind of situation as **language intertwining**. We refer the reader to chapter 4. This whole subject has just started to be studied in any detail. Sometimes a creole involves substantial mixture at all levels of language structure. A case in point is Berbice Dutch Creole, described in chapter 19.

Other cases where languages have become simplified to some extent are of **lingua francas** (not the *Lingua Franca* of the Mediterranean) and **koines**. These come into existence under similar circumstances – one speech-form becomes widely used by non-native speakers, undergoing a degree of simplification. Here, the process seems to be gradual – in other

words, no linguistic or social violence is involved. We speak of a lingua franca when speakers of various different languages are involved, and of a koine when the dialects of a single language are involved.

In chapter 26 there is an annotated list of languages where these distinctions and some further ones are used to classify over 500 languages and dialects. To complicate matters speech forms may change in status over time. Various scenarios or *life-cycles* (cf. Hall 1966, who used the term somewhat differently) have been proposed for the development of creoles. Mühlhäusler (1986) presents three such scenarios:



As will be argued in chapter 3, however, not all jargons or pidgins are part of such a life-cycle, and neither can we show that all creoles had a jargon or pidgin stage. It is in this respect that mixed languages display an important difference from creoles. On the one hand, mixed languages did come into existence at a particular moment in time, on the other hand they were formed from ordinary languages with native speakers – there was no jargon or pidgin phase.

1.3 Distribution of pidgins and creoles

The question of the distribution of pidgin and creole languages is one of the growth areas in linguistics. Because of their mixed character these speech varieties have frequently not been accorded the status of language. The frequent prejudice against their recognition as proper linguistic systems has meant that lists of the world's languages, produced up till fairly recently, tended to ignore these speech varieties. While many linguists, and sometimes educationalists, recognize the fact of their existence, this is by no means universally the case. The effect of this is that new creoles and pidgins are continually being added to the lists of such languages.

Recognition has come quickest for those creoles and pidgins (partially) based on Euro-

pean colonial languages, although even here we may be certain that some languages remain undiscovered. In the case of creoles and pidgins not involving a European base linguists have been faced with the above-mentioned problem that the history of very many languages is very poorly known. And as we will discover time and time again in the course of this book, a knowledge of the history of a language is often essential for determining its creole or pidgin status, or the lack of this. This means that creoles that came into existence hundreds of years ago may only be recognized as such in modern times.

The small size of many creole-speaking communities also militates against their recognition. A small linguistic community will more easily be assumed to represent a (deviant) dialect of a larger language than a large one will. Small communities also get overlooked more easily. So the Wutun 'dialect' of Qinghai province, China has been recognized as involving a problem in classification by Chinese scholars for quite some time. This mixed Amdo Tibetan-Kansu Mongol-Chinese language has certainly been in existence for several hundred years. It had been variously claimed to be Chinese, Monguor and Tibetan. Its essentially mixed status was first recognized by Chen (1982). The fact, however, that the language has only 2500 speakers in five villages has not helped it to appear in any list or classification of the world's languages. For instance, it does not appear in the 11th edition of *Ethnologue* (Grimes 1988).

We have cited the question of prejudice above. This is especially relevant in the case of pidgins. Pidgins, by their very nature, tend towards instability, both in terms of linguistic system, and in terms of their function. If they do not belong to the small group of pidgins that become standardized, or nativized, or both, they may well disappear completely when the social need that caused them to come into existence passes. An event so trivial as the disruption of a market may make a particular pidgin redundant. Population movements may have the same effect. So the *raison d'être* of the Pidgin Russian spoken in Harbin, Manchuria, between Russians and Chinese, disappeared when most of the Russians left in the fifties.

1.4 History of pidgin and creole studies

Why should there be a field of pidgin and creole language studies? Since the group of languages as a whole are not genetically related, nor spoken in the same area, the languages must be considered to have something else in common, in order to be meaningfully studied as a group. In the field there is an implicit assumption that the creole languages share some property that calls for an explanatory theory.

The earliest written sources for many creoles date from the 18th century, when missionaries started writing dictionaries, and translating religious texts into the languages of the slaves.

The first time the term 'creole' was applied to a language was 1739, in the Virgin Islands, when the very youthful Dutch-lexifier creole Negerhollands was referred to as *carriolsche* by a Moravian missionary (Stein 1987). The first grammar of a creole was written in the Virgin Islands by J.M. Magens, a scion of a local planter family (1770). In addition to missionaries, travellers or other laymen occasionally wrote brief dialogues etc. in the local creoles, at that time generally referred to as Negro-English, Negro-Dutch, etc. There are reasonable historical records for a number of creole languages, including Negerhollands, Sranan and Saramaccan (Surinam), Mauritian Creole, and Jamaican. These allow us to study the historical development of the creole languages (see chapter 10).

Creole studies originated as a systematic field of research over a century ago, with Schuchardt's (1842-1927) important series of articles. These started as an attempt to account for a more complex set of developments in the history of the Romance languages than was possible in the Neogrammarian preoccupation with the regularity of sound change. Hesselting's (1880-1941) work originally started out from an explanation of the developments in Greek, from the early dialects through *koine* Greek under the Roman Empire, to Byzantine and modern Greek. Both scholars found it necessary to allow for more complex types of linguistic change: mixture, simplification, reanalysis, and the complexity of their analyses characterizes modern creole studies as well.

Until 1965 the field remained, however, rather marginal. Creole languages were studied by a few enthusiastic historical linguists – usually Anglicists or Romanists, fieldworkers with an adventurous bent, or folklorists ahead of their time. Now the study of creole languages has moved to the center of linguistic research, a research program with universalist theoretical pretensions, half-way between theoretical linguistics and sociolinguistics. Reasons for this development are manifold, but include the political and cultural emancipation of certain parts of the Caribbean (most notably Jamaica), an interest in Afro-American culture, particularly in the U.S., and a partial reorientation of linguistic research.

1.5 Theories of origin in creole studies and theoretical linguistics

The main research effort in pidgin and creole studies has been to find a principled explanation for the genesis of the languages involved. There is an implicit assumption that the creole languages share some property that calls for an explanatory theory. What property this is depends on the theory concerned. Any of four properties are assumed to play a role:

(1) Creole languages are often assumed to be more **alike** than other languages. As we will see, creoles share many structural features, and many researchers believe that these resemblances cannot be simply due to the similarity between the languages of western Europe, or accidental.

(2) Creole languages are often assumed to be more **simple** than other languages. There is a wide-spread belief that creole languages are not just morphologically, but also syntactically and phonologically simpler than other languages.

(3) Creole languages are often assumed to have more **mixed** grammars than other languages. Many people have drawn parallels between language and biology, when thinking of creoles. It is assumed that just as many speakers of creole languages have 'mixed' African, European, Asian and in some cases Amerindian ancestry, the languages they speak are likewise simply a combination of a bit of European vocabulary with some African or Asian syntax and semantics.

(4) Pidgin and creole languages are often assumed to exhibit much more **internal variability** than other languages. They are assumed to be highly dynamic language systems and often coexist with their lexifier languages in the same speech community.

These assumptions play a role in the various theories of creole origin that have been proposed. The theories of origin have been developed in part to explain the assumed similarity, simplicity, mixing, and variability of the creole languages. We have chosen to group these theories into four categories, in chapters 8-11. Here we will briefly summarize the principal hypotheses put forward. References will be provided in the relevant chapters.

1.5.1 The European input

Some models attempt to trace the properties of the pidgins and creoles back to specific antecedents in Europe (see further chapter 8). The **Portuguese mono(-)genesis** model has undergone several modifications. Crucial to all of these is the existence of a trade language with a predominantly Portuguese lexicon, used in the 15th to 18th centuries by traders, slave raiders, and merchants from throughout the then incipient colonial societies. The monogenetic theory holds that the slaves learned the Portuguese Pidgin in the slave camps, trading forts, and slave ships of their early captivity, and then took this language, really no more than a jargon, with them to the plantations. The different creole languages as we know them are based on this jargon, but have replaced the Portuguese words with words from other European languages. The supposed similarity of the creole languages is due of course to the underlying Portuguese jargon, and their simplicity to the simplicity of this jargon.

The **restricted monogenesis** hypothesis is less ambitious. It is mostly limited to the English and French-lexifier creole languages of the Atlantic and Indian Ocean, and proceeds from the idea that there was a jargon or pidgin spoken along the coast of West Africa that later formed the primary source for a wide range of creoles. The common features of these creoles are then assumed to be due to these early pidgins.

The **European dialect origin** hypothesis holds that creoles essentially developed from non-standard dialects of the colonial languages in an ordinary way, and are the result of

migration by dialect speakers to the newly founded colonies, compounded with the existence of a strongly dialectal **nautical language**. In this theory, similarities between creoles hold only for those derived from a single colonial language; creoles may be simple because the non-standard varieties were simpler than the written national standard.

In other approaches, processes involving the transformation of the European languages play a central role, through imperfect second language learning or the reduction of speech directed at foreigners. The **baby talk** or **foreigner talk** theory is similar to the imperfect second language learning theory in postulating that creoles are frozen (i.e. fossilized) stages in the second language learning sequence. The difference lies in the fact that in the baby talk theory the responsibility for the simplification is shifted from the learners to the speakers of European languages, who provide a simplified model. The similarity between creoles would be due, in this view, to universal properties of the simplified input. The type of evidence adherents of the baby talk hypothesis are looking for thus includes simplifications made by native speakers, not by learners, in pidgins, such as the use of infinitives.

In the **imperfect second language learning** theory creoles are the crystallization of some stage in the developmental sequence of second language acquisition. The speakers of the proto-creole simply did not have sufficient access to the model, and had to make up an approximative system. In this view the fact that creoles are simple is due to the simplification inherent in the second language learning process. For some adherents of this view the possible similarities among the creole languages are due to universal properties of the learning process.

1.5.2 The Non-European input

The **Afro-genesis** model really deals mostly with the creole languages spoken in the Atlantic region: West Africa and the Caribbean, and postulates that these languages have emerged through the relexification by the slaves of the West African languages, the so-called **substrate** languages, under influence of the European colonial languages (see chapter 9). An alternative explanation is in terms of the transfer of African language structures in the process of learning the colonial lexifier languages. The similarity of the languages involved is due, in this model, to the fact that they share the same African linguistic features, mixed together with the lexicon of the European languages. The main problems with the Afro-genesis model in its strict version are the large number of structural differences between West African languages and creoles on the one hand, and the linguistic differences among the various West African languages themselves on the other. What has been claimed to save the hypothesis is that in the process of relexification certain syntactic and semantic properties of European lexical items were incorporated as well.

1.5.3 Developmental approaches

Many researchers study pidgins and creoles from a developmental perspective, as **gradually** evolving and continuously changing systems rather than as stable systems that emerged rapidly. Within this approach, expansion of pidgins through their continued use and growth in functional domain is stressed above strictly grammatical or cognitive aspects. In chapter 11 we return to various developmental approaches.

The **common social context** theory adopts a such strictly functional perspective: the slave plantations imposed similar communicative requirements on the slaves, newly arrived, and lacking a common language in many cases. The commonality of the communicative requirements led to the formation of a series of fairly similar makeshift communicative systems, which then stabilized and became creoles.

1.5.4 Universalist approaches

Universalist models stress the intervention of a specific general process during the transmission of language from generation to generation and from speaker to speaker (see chapter 11). The process invoked varies: a general tendency towards semantic transparency, first language learning driven by universal processes, or general processes of discourse organization.

The **semantic transparency** theory is not a full-blown genesis theory, but simply claims that the structure of creole languages directly reflects universal semantic structures. The fact that they are alike, in this view, is due to the fact that the semantic structures are universal. They are simple because the semantic structures involved are fairly directly mapped onto surface structures, eschewing any very complex transformational derivation. An example of this may be the fact that creole languages have separate tense/mood/aspect particles, which reflect separate logical operators, rather than incorporating tense, etc. into the inflection of the verb.

The **bioprogram** theory claims that creoles are inventions of the children growing up on the newly founded plantations. Around them they only heard pidgins spoken, without enough structure to function as natural languages, and they used their own innate linguistic capacities to transform the pidgin input from their parents into a full-fledged language. Creole languages are similar because the innate linguistic capacity utilized is universal, and they are simple because they reflect the most basic language structures. One feature shared by all creoles that would derive from the innate capacity is the system of pre-verbal tense/mood/aspect particles. Not only do they seem limited in the creole languages to a particular set of meanings, but they also seem always to occur in a particular order. The system of tense/mood/aspect particles, its interpretation and its ordering would directly reflect universal aspects of the human language capacity.

1.5.5 Theoretical implications

In all these models or theories notions such as **alike**, **simple**, **mixed**, and **variable** play a role. They are in fact taken for granted, assumed to be what requires to be explained, and therefore not called into question. The contribution that the study of creole languages can make, in our view, to grammatical theory is that it can help to elucidate these four concepts 'alike', 'simple', 'mixed', and 'variable'. All four turn out to be relevant to the central concerns of modern grammatical theory. In order to help us understand this, let us examine the concepts involved more closely.

When we say that languages *x* and *y* are more alike than *y* and *z*, we are claiming in fact that in the total (abstract) **variation space** allowed for by the human language capacity *x* and *y* are closer than *y* and *z*. Consequently, the claim that the creole languages are more alike than other languages implies a clustering in the variation space. If we think of the variation space as defined by **parameter theory** (as in recent work by Chomsky and others), trying to develop a notion of 'alike' really boils down to developing a theory of parameters, parameters along which similarities and differences between natural languages can be defined.

Consider now the concept of simplicity. The idea that creole languages are simple has been taken to mean two things. On one level it has meant that creole languages do not have a rich morphology, on another that the overall grammar of creole languages is less complex than that of other languages. Both interpretations are relevant to grammatical theory. The idea that **absence of morphology** is related to grammatical simplicity needs to be evaluated in the context of contemporary research into morphology/syntax interactions, and the grammatical status of **inflection of INFL** (Chomsky 1982; Rizzi 1982, and others) and of case marking (Stowell, 1981). Even more importantly, the idea that the creole languages are not grammatically complex in general only makes sense if one has a theory of grammatical complexity to fall back on, and this brings in **markedness theory**.

Consider next the notion of mixing. Mixing implies that elements from one language are combined with elements from another, and this in turn calls into question the cohesion of the grammatical systems involved. The tighter a particular subsystem (e.g. the vowel system, or the system of referential expressions) is organized, the less amenable it will be to restructuring under borrowing. Tightness of organization in modern grammatical theory is conceptualized in terms of **modularity theory**: the grammar is organized into a set of internally structured but externally independent modules, the interaction of which leads to the final grammatical output. For this reason, the notion of mixing is important: it forces us to think about which parts of the grammar are tightly organized, and hence about the notion of modularity.

Tightness of organization or cohesion may have either a paradigmatic dimension, in

terms of the hierarchical organization of **feature systems**, or a syntagmatic dimension, in terms perhaps of the notion of **government** (Chomsky 1981) as a central principle of syntactic organization.

An important group of creole researchers has focused on the dynamic and variable aspects of language (Sankoff 1982; Bickerton 1975; Rickford 1987). While linguists working in terms of the paradigm of generative grammar tend to abstract away from variation and change, focusing on the universal and invariable aspects of linguistic competence, many creolists have tended to put **variation and change** at the center of attention: only by studying the changes that languages undergo and the ways in which these changes are manifested in the speech community can we find out about the phenomenon of language. Pidgin and creole languages form a natural field of study for these researchers, precisely because they present so much internal variation and because they tend to change so rapidly. The extent of variation present (and this is particularly relevant for pidgins) again raises the questions mentioned above with respect to the internal cohesion of a grammatical system and how parameters determine the way languages vary.

Keeping this in mind, then, the contribution of pidgin and creole studies to linguistic theory is clear. We have come to grips with one or more of the core notions of grammatical theory:

alike:	parameter theory
simple:	morphology/syntax interactions markedness theory
mixed:	modularity
variable:	parameter theory, modularity

Studying creole languages implies a constant confrontation with these notions, and helps one to develop a vocabulary capable of dealing with them.

Further reading

The primary source for documentation on the different pidgins and creoles is still Reinecke's monumental bibliography (1975). There are a number of introductions to pidgin and creole studies on the market, including Hall (1966), Todd (1974; 1990), Mühlhäusler (1986), with much information about the Pacific, Holm (1988), strong on the history of the field, and Romaine (1988), strong on links with psycholinguistic research. In French we have Valdman (1978). In addition there is a large number of collections of articles, of which Hymes (1971), Valdman (1977), and Valdman and Highfield (1981) are the most general in scope.

Useful monographs by single authors are: Bickerton (1981), which contains a highly

readable exposition of the bioprogram hypothesis; Alleyne (1981), which documents the Afro-genesis hypothesis with a wealth of detail; and Sankoff (1980), which presents the view that the structure of creole languages is finely attuned to their functional requirements with a number of insightful articles. There are two specialized journals, *Journal of Pidgin and Creole Languages* and *Études Créoles*. In addition there is a newsletter, *The Carrier Pidgin*.

2 The socio-historical background of creoles

Jacques Arends

2.1 Three types of creole

It has been argued by some creolists that creoles cannot be defined as a distinct group of languages on typological, intra-linguistic, grounds (e.g. Muysken 1988). If this is true, the question arises whether there are any external, extra-linguistic, criteria according to which they can be grouped together in one category. The criterion that comes to mind most readily is that of the social history of these languages. Are creoles characterized by a particular social history, a social history that is common to all of these languages and that is not shared by any other group of languages? This question cannot be satisfactorily answered at the moment, simply because the external history of many creoles still has to be written, but there are strong indications that indeed in many cases there are a number of striking similarities among the historical processes through which these languages came into being. One of these concerns the fact that many creoles arose in the context of the European colonial expansion from the sixteenth century onwards. In many cases this expansion was accompanied by a specific type of economy, which had as its most characteristic feature the exploitation of relatively large agricultural units, plantations, for the production of largely new products such as sugar, coffee, and tobacco, for the European markets. Plantations, however, were not the only situations that gave rise to creolization. Therefore, before we go on to explore the commonalities in the external histories of creoles in general, we will first briefly discuss three different types of creoles that can be distinguished according to differences in their external histories.

According to their external history the following three types of creole have been distinguished: **plantation creoles**, **fort creoles**, and **maroon creoles** (Bickerton 1988). In addition, a fourth type may be distinguished: creolized versions of pidgins have emerged, e.g. in New Guinea and northern Australia. In the Atlantic area, plantations were worked by large numbers of African slaves, who were purchased along the western coast of Africa from Senegal to Angola. In the initial stage of colonization Amerindian slaves were also used, as well as indentured laborers – poor Europeans who were contracted for a specified number of years. In the case of the Pacific and the Indian Ocean, slavery was usually not the primary means of acquiring a labor force. Indentured workers from India, China, Japan, the Philippines and the South-West Pacific, were recruited to work on the plantations in Mauritius, Queensland (Australia), and Hawaii.

It is not only in the plantations, however, that creole languages arose. Apart from the plantation creoles, which emerged in the Caribbean (e.g. in Jamaica, Haiti, Guyana, Surinam), in West Africa (e.g. on the islands of Annobon and São Tomé off the West African coast), and perhaps in the southern parts of North America as well, a number of creoles developed at the so-called forts, the fortified posts along the West African coast, from which the Europeans deployed their commercial activities. In the forts some medium of communication must have been used, both among Africans from different linguistic backgrounds and between Africans and Europeans. More importantly, however, interethnic communication extended to the forts' surroundings where European men (so-called *lançados*) were living in mixed households with African women, with whom they spoke some kind of contact language. In the course of time these contact languages were expanded into creoles, in particular by the children that were born into these households. One of these is the alleged 'Guinea Coast Creole English', which, according to Hancock (1986), arose out of the interaction between English and African speakers in the settlements in Upper Guinea (Sierra Leone and surrounding areas) and which may have formed the basis of the Caribbean English-lexifier creoles.

A third type of socio-historical context that has given rise to the genesis of creoles is *marronage*, which refers to the fact that slaves escaped from the plantations and subsequently formed their own communities in the interior in relative isolation from the rest of the colony. Maroon communities developed in several parts of the New World (Jamaica, Colombia, Surinam) and in Africa as well (São Tomé). While most of these communities have been absorbed by the mainstream culture of the societies within which they existed, the Surinam maroons, who are distributed over several tribes, have preserved their own traditions and their languages up to the present day. But since these languages probably developed out of plantation creoles, we should not expect to find important structural differences between the two. What may have caused some divergence, however, is the fact that the maroon creoles developed in relative isolation from the metropolitan, European, language. This issue has as yet not been explored in any detail.

In Surinam, two maroon creole languages can be distinguished. One, consisting of the dialects spoken by the Saramaccan and Matawai tribes, is a 'mixed' creole, with two European lexifier languages, English and Portuguese (see chapter 14). In this respect it is clearly different from the coastal creole, Sranan, whose basic lexicon is English-lexifier, just like that of the other maroon creole language, spoken by the Ndjuka, Aluku, Paramaccan, and Kwinti tribes. While some of the dialects (e.g. Matawai and Paramaccan) have hardly been studied at all, Saramaccan has attracted the special attention of many creolists, who regard this language as the most pure or *radical* creole language extant today. According to these scholars Saramaccan, due to its supposedly rapid formation and its subsequent isolation from

other languages, has preserved its original creole character in a purer state than other known creoles (but cf. Alleyne 1980 for a different view). Other maroon creoles, outside Surinam, include Palenquero (Colombia) and Angolar (São Tomé), which are still spoken today. Finally, remnants of another maroon creole have been found in the 'Maroon Spirit Possession Language' of Jamaica (Bilby 1983). This is not employed in ordinary situations, but it is used by people when they are possessed during religious ceremonies, to talk to the spirits of those of their ancestors who were born in Jamaica.

The three-way division made here at least to some extent cuts across the distinction between *endogenous* and *exogenous* creoles. This distinction was made by Chaudenson (1977) in order to distinguish between creoles that arose in areas where the native languages of the creolizing population were spoken (e.g. some African creoles, such as Kituba) and those that did not, since they involved the massive relocation of the creolizing population (e.g. the creoles that arose in the New World). The distinction is especially important with respect to the potential role of the substrate in creole genesis: a creole that arose in an area where its substrate speakers had ample opportunity to continue speaking their native language(s) next to the emerging creole is bound to show more substrate influence than one that did not (cf. Singler 1988).

2.2 Colonial expansion and the slave trade

The history of European expansion and the concomitant slave trade cannot be adequately described here, but it cannot be excluded entirely either since it constitutes the socio-historical matrix in which creolization took place. Therefore, in what follows a brief outline will be given of this history as far as it concerns the Atlantic area. The main European nations involved in the colonial expansion were Spain, Portugal, France, Britain and the Netherlands. While the Spaniards and the Portuguese were the first to actually found settlements in the New World during the sixteenth century, they were followed by the others a century later.

During the entire slave trade period some ten million Africans were captured and deported to the Americas (Curtin 1969). Many of these did not survive: some died during captivity in one of the forts along the African coast, before they had even embarked on their *middle passage*, the journey to the New World. Others perished during transport as a result of disease or other causes related to the poor conditions on the slave ships. Of those who did arrive in the New World, many died after a relatively short period in the colony: in 18th-century Surinam the life expectancy upon arrival was somewhere between five and ten years.

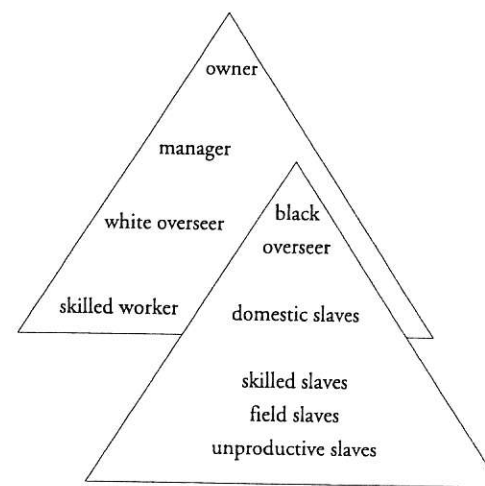
As far as the geographical origins and demographic behavior of these Africans is concerned, much remains to be discovered by historical research. In the case of Surinam

extremely detailed information has been made available by the historian Postma (1990), whose findings are based on archival documents concerning the Dutch Atlantic slave trade. These findings relate to such variables as port of embarkation, dates of departure from Africa and arrival in the New World, age and gender distribution, and the numbers of slaves that were embarked and disembarked. While we cannot go into this in any detail, let us summarize Postma's main findings, as an example of what historical research can contribute to the study of creole genesis. Between 1650 and 1815 the Dutch shipped some 200,000 Africans to Surinam. At Emancipation, in 1863 – more than 200 years after importation began – the black population still numbered no more than some 36,000. This shows that during the entire period of slavery there was a very substantial population reduction, due to an exceptionally high death rate and an exceptionally low birth rate, while at the same time the rate of immigration was very high. As a result of this, the normal situation whereby a language is acquired natively through transmission from one generation to the next, with second language learning being only marginal, was completely disturbed. Although the precise linguistic consequences of this for creole genesis are not entirely clear, it seems evident that the role of demography should be taken into account.

As far as the geographical background of the Surinam slaves is concerned, Postma provides a wealth of interesting information. From his figures it can be inferred (Arends to appear b) that during the slave trade period there have been substantial variations in the areas from which the Dutch purchased their slaves. The general picture that emerges from these figures is that over the entire period (1650–1815) the **Windward Coast** (the area stretching from Sierra Leone to Ivory Coast) served as the main supplier of Surinam slaves. However, this area started to play this role only from 1740 onwards. During the first 70 years of slave importation (1650–1720, the formative period of Sranan and Saramaccan) the **Slave Coast** (Togo, Benin) and the **Loango** area (Gabon, Congo, Zaire, Angola) supplied more than 90% of all slaves imported into Surinam. In the intervening period (1720–1740), the **Gold Coast** (Ghana) served as the main supplier of slaves. Obviously, such a finding has important consequences for the investigation of African survivals in the Surinam creoles. Thus, it seems safe to assume that in the formation of these creoles, Windward Coast languages, such as Mande and Wolof, despite their ultimate overall numerical dominance can only have played a minor role, whereas languages spoken along the Slave Coast, such as Gbe, and in the Loango area, such as Kikongo, are much more relevant in this respect. This shows that detailed historical-demographic research may drastically reduce the set of relevant substrate languages for any given creole (see further chapter 9). Other colonies for which similar work has been done include Mauritius (Baker 1982), Cayenne (Jennings to appear), and Martinique and Guadeloupe (Singler 1992b).

2.3 The plantation system

Having discussed some relevant socio-historical factors at the macro-level, it may be useful to go in some detail into a social aspect of creolization at the micro-level, i.e. the social structure of the plantation. After all, the plantation must have been the main locus of creolization. While most of what follows refers to Surinam (largely based on Van Stipriaan 1993), it has some relevance for other creole societies too. The stereotypical image of a plantation colony as a severely dichotomized society, with a small number of whites holding power over large numbers of African slaves, needs some adjustment on the basis of what is known about how plantation life was socially structured. Although there certainly was a wide social, cultural and economic gap between the small white section of a plantation's population and the numerically dominant slave force (with ratios reaching 50:1 and more), the actual situation may have been a bit more complex than it might appear at first sight. The figure below charts the social stratification according to the division of labor on a typical Surinam plantation.



Within the black population there was a division of labor between field slaves (who on average formed no more than around 50% of a plantation's black work force), house slaves, slave craftsmen, and slaves performing various other tasks, such as hunting and fishing. These differences in function correlated not only with differences in status and power within the black community, but also with the amount of linguistic interaction with whites. A

special function was that of the so-called **creole mama**, a black woman – usually elderly – who took care of the younger children. She is assumed to have provided an important model for the acquisition of language by these children, beside their parents and other relatives.

Apart from these, every slave community had one or more black overseers, who occupied an intermediary position in the power structure, between the white master and the black workforce. To him was delegated the execution of punishment and allocation of tasks, as also the decision on when sick slaves were fit to work. In addition to this, there is some evidence that he was also a religious leader in the black community. In many cases, quite surprisingly, the black overseer even seems to have been in a more powerful position than the white overseer. Probably these differences in power and status between different groups of slaves were reflected in their language use, just like in any other society, but, unfortunately, this cannot be empirically verified, due to the absence of documentary evidence.

Apart from this, there must also have been considerable differences in the quality and quantity of contacts between different groups of blacks on the one hand and whites on the other. Thus, among the blacks the overseer probably had, if not the most regular, the most elaborate verbal interactions with the whites, due to the necessity of discussing the technical details of plantation management. In descending order of frequency and intensity of contact with whites, the black overseer was probably followed first by the domestic slaves, then by the slaves who had special tasks and, finally, by the field slaves and the unproductive slaves. Although it is impossible to reconstruct the linguistic consequences this may have had, it still seems useful to be aware of the fact that the stereotypical image of plantation society as a strictly dichotomous one is an idealization, and that the actual situation was much more complex.

In the course of time, a group intermediate between the black and white populations of the plantations developed, consisting of mulattoes (coloreds), who were the results of sexual relationships between white men and black women. This was a privileged group, whose members were often sent to town in order to serve as house slaves, or, when they were recognized by their fathers, bought free. Together with manumitted black slaves, these mulattoes formed a growing intermediate group, between the small group of whites and the large mass of slaves. This adds to a further refinement of the image of slave society as multi-stratal rather than bi-stratal.

One other linguistically relevant feature of slave society is the fact that the black population consisted of two groups, the **bozals** or **salt water slaves**, those who had been born in Africa, and the **creoles**, those who had been born in the colony. Linguistically speaking, this difference is reflected in the fact that the former arrived in the colony speaking one or more African language(s), whereas the latter acquired their first language(s) (a creole, an African language, some version of the metropolitan language) in the colony. Although very

little is known about differences in L1 versus L2 use of emergent creoles, it seems safe to assume, on the basis of what is known about L1 and L2 acquisition in general, that such differences were present. As to the L2 acquisition by newly arrived slaves, it is known that, as part of a general **seasoning** process, they were assigned to an experienced slave, whose tasks included introducing new slaves to the local language, i.e. the creole. This means that the model for language acquisition was largely provided by blacks, not whites. In colonies where, due to the demographic factors referred to in section 2.2, the number of creole slaves increased quite slowly, the task of seasoning must often have been performed by African-born rather than locally-born slaves. In other words, in these cases the model for the acquisition of the creole as a second language by the African-born slaves would be a second, not a first language version of that creole. But again, it is extremely difficult to estimate the influence this may have had on the emerging creole.

Finally, two points have to be mentioned. First, the slave population was not always as powerless as is often assumed. Because of their sheer numbers, the blacks were a constant threat to the whites, who could only control them by the use of force. The most important forms of resistance used by the slaves were rebellions, strikes and **marronnage**. Second, despite their lack of freedom, slaves were much more mobile than has often been assumed. Sexual relationships, funerals, festivities, trade and other activities provided opportunities for contacts outside the plantation (Muyrers 1993). The linguistic impact of these factors, however, can as yet not be established.

2.4 Demography

Although some demographic issues have been briefly touched upon in section 2.2 above, these issues deserve a more detailed discussion in relation to creolization. This is so for at least two reasons. First, certain demographic factors constitute preconditions for the process of creolization to be able to happen in the first place. For instance, models which assume that creoles are the result of L1 acquisition (or rather creation) by the first generation(s) of locally-born children, such as Bickerton's Language Bioprogram Hypothesis (1981, 1984), presuppose the presence of a sufficient number of these children during the formative period of the creole, say the first twenty-five or fifty years after initial colonization. Demographic data on some colonies, such as Jamaica and Surinam (Singler 1986), seem to show that this condition was not fulfilled there during this period. Thus, demographic findings may provide extralinguistic counter-evidence for some hypotheses about creolization.

Second, there are some demographic factors which, although they do not constitute preconditions for creolization, are nevertheless important because they may determine the nature of the creolization process. After all, our view of creolization is still largely based on

20th-century evidence, even though creolization is essentially a historical phenomenon. That is to say, while creolization is an event of the past for all creoles except those that are in the process of being formed today, almost anything that has been hypothesized about it has been based on knowledge of the outcome of that process (the present day creole language), not on knowledge about the process itself. One of the demographic factors that may contribute to our understanding of creolization concerns the development of the quantitative proportion between the black and white parts of the population, especially during the initial period of colonization. This is a relevant factor because the presence of models speaking the dominant language is a necessary condition for the transmission of that language to L2 learners. And although L2-speaking models were certainly present in all creolization situations that we know of, the availability of these models for L2-learning Africans may have differed widely, both among different locations, and over time in a single location.

It is for this reason that Baker (1982), who was the first to draw attention to the importance of demographic factors for creolization, introduced the term **Event 1**, referring to the point in time when numerical parity between the black and white parts of the population is reached. It is hypothesized that in the period between the beginning of colonization (which might be termed **Event 0**) and **Event 1**, there are sufficient L2 speakers present for every L2 learner to have adequate access to the language, i.e. to actually learn it, as opposed to having to revert to speaking a pidgin. After **Event 1**, when increasing numbers of slaves are imported while the number of whites does not grow in proportion to these, it becomes increasingly difficult for L2 learners to gain access to native speakers to learn the language from. Obviously, the issue of access to L2 models is not just a quantitative matter; as was noted in section 2.3, also the quality of the interaction between whites and blacks may have differed widely. Although differential access must undoubtedly have had a linguistic impact, at the moment not enough is known about it to specify this in any detail.

Another important demographic event, distinguished by Baker, is **Event 2**, which refers to the point in time when the number of locally born blacks, or creoles, reaches numerical parity with the total number of whites. (The third of Baker's Events, **Event 3**, which is not relevant for the present discussion, refers to the point in time when the immigration of substrate speakers stops.) **Event 2** may be interpreted as the point in time at which the black population has creolized or nativized to such an extent that creolization of language may occur. In other words, at **Event 2** the native black population has acquired the 'critical mass' necessary for creolization to take place. This does not mean, of course, that creolization necessarily has to take place, only that a condition for it is fulfilled. Whether creolization will actually occur, depends, among other things, on other demographic factors, such as the rate of post-**Event-2** slave imports and demographic developments within the white population.

It should be borne in mind that, although **Events 1** and **2** are historical occasions that can in principle be established empirically, the reason for their introduction is a heuristic, not an empirical one. In other words, it is not an established fact that precisely these historical events correlate with specific linguistic events such as the onset of the creolization process. Rather, the identification of **Events 1** and **2** provides a means of getting at the relationship between demographic development and linguistic process more precisely. Much more research, however, is needed to gain a better insight into the significance of demographic factors for creole genesis.

The first, and up to now, only creolist who has attempted to formalize the importance of certain demographic factors for the processes of pidginization and creolization, is Bickerton (1984), who developed a **Pidginization Index (PI)**, a demographic measure for the degree of pidginization. The degree of pidginization is defined by Bickerton in terms of structural distance from the lexifier language: the higher the **PI**, the smaller the distance. The **PI** takes into account three demographic factors: **Y**, **P**, and **R**. **Y** is the number of years between **Event 0** and **Event 1**; **P** is the number of substrate speakers at **Event 1**; and **R** is the average annual import of substrate speakers after **Event 1**. The formula thus reads:

$$PI = Y \times P / R$$

Of course, this formula should not be construed as, nor is it intended as, in any way an absolute quantitative measure of pidginization. What it does attempt, is to relate the degree of pidginization to the demographic development of a particular colony. Thus, a black population that grows slowly until **Event 1** (i.e. where **Y** is high) will yield a language with a relatively high **PI**. Similarly, a high rate of post-**Event 1** importation of slaves (i.e. where **R** is high) will yield a relatively low **PI**. In the former case a colonial version of the metropolitan language will probably emerge, while in the latter the genesis of a pidgin is more likely. Although the **PI** has been severely criticized (Singler 1990), it may still serve a useful function, namely as a first heuristic in trying to come to grips with a number of complex and elusive extra-linguistic factors, which most creolists agree are of crucial importance for creolization but which are seldom dealt with in any systematic manner.

Another factor to be reckoned with is the degree of linguistic homogeneity of both the black and white populations. While in some colonies the bulk of the white population all spoke the same European language, there were others (Surinam, Virgin Islands) where a variety of European languages was spoken. More importantly, perhaps, the degree of the homogeneity of the substrate also differed widely. An extreme case is represented by Berbice, where a single African language, Eastern Ijò, is assumed to have been spoken by most of the blacks. In other cases, the African substrate was far more heterogeneous. As shown by

Singler (1988), the homogeneity of the substrate is an important factor in determining the degree of substrate influence in creole genesis.

A concept, borrowed from population genetics, which has been introduced recently into the historical demographics of creole genesis (Mufwene 1993a), is that of the **founder principle**. This principle is supposed to account for the disproportionately strong influence the founder population of a settlement may have on the genetic make-up of the population at later stages. The idea is that, similarly, the language(s) of a colony's founder population, both European and non-European, may have had a disproportionately strong influence on the creole language(s) of that colony. Although the idea itself is interesting, its value cannot be established until more exact and more reliable data about the founder populations of a number of creole societies have become available. Summarizing, it is clear that much more research is needed before the exact impact of demographic factors on creolization can be established in a systematic way.

Further reading

Rens (1953) is a social history of the Surinam creoles. Mintz & Price (1992) is a concise but excellent introduction to the social history of African-American culture, while Van Stipriaan (1993) provides a detailed study of the development of Surinam's plantation system. Chaudenson (1992) is an interesting and wide-ranging survey of linguistic and cultural creolization processes, with particular reference to the French-lexifier creoles. The history of the Saramaccan maroons is told in Price (1983). The classic work on the Atlantic slave trade is Curtin (1969), while Postma (1990) is a quantitative assessment of the Dutch participation in that trade.

3 Pidgins

Peter Bakker

3.1 Introduction

Pidgins are languages lexically derived from other languages, but which are structurally simplified, especially in their morphology. They come into being where people need to communicate but do not have a language in common. Pidgins have no (or few) first language speakers, they are the subject of language learning, they have structural norms, they are used by two or more groups, and they are usually unintelligible for speakers of the language from which the lexicon derives.

In most studies of pidgin and creole languages, pidgins fare rather poorly. Too often, they are assumed to be simple versions of creoles, or it is stated that creoles are just pidgins which suddenly acquire typical creole-like structural properties upon becoming mother tongues. Furthermore, forms of Pidgin English of the Pacific, especially New Guinea, are often given as examples, but these are not unambiguous examples of pidgins, as they may be both first and second languages and have been spoken for many generations. Hence, they share both pidgin and creole language properties. Pidgins undergo structural expansion when their use is extended to many domains. In this chapter, we want to discuss as wide a variety of pidgins as possible, focusing on pidgins which never became native languages, and never extended pidgins. These extended pidgins resemble creoles. Hence, Pacific and West African Pidgin English will not be the focus of our concern. We will argue that (a) pidgins are structurally strikingly different from creoles, (b) pidgins may have considerably complex morphology and (c) pidgins are very often based on the local language rather than on the colonial one.

The etymology of the word 'pidgin' was a subject of debate, but this has been settled recently. In Hancock (1979) several etymologies were discussed, but more recent research establishes the Chinese Pidgin English pronunciation of the English word *business* as its source (see Baker & Mühlhäusler 1990: 93), particularly because of its use in a popular Chinese Pidgin English phrase-book in Chinese characters in the early 1900's (Shi 1992). The word, spelled 'pigeon', was already used in 1807 for Chinese Pidgin English, and it was only many decades later that it became used as a generic term for all pidgins. Until then, the term **jargon** was commonly used for pidgins in some areas, as is still clear from all the North American pidgins which are called 'Jargons'. Europeans also used the term *lingua*