

theory with a general theory of social structure (J. Milroy 1992) may help us understand which changes are likely and why. Change is then actuated as it spreads from the innovator to other social groups, where it will be accepted by the early adopters, then diffuse to the other members. Gradually, the variation involved becomes orderly, the variants acquire social meaning, and change becomes apparent in the system – at which rather late stage it becomes detectable by sociolinguistic observation. Earlier stages are not yet observable or altogether explicable, and previous studies claiming to observe actuation have probably, at best, been noticing early adoption. We have, then, a framework for the transmission of gradual, socially motivated language change, and can observe and measure its progress. Actuation, the initial spread of an incipient change by innovators through weak ties with other groups, remains almost as mysterious and unattainable as ever.

10 Pidgins and creoles

10.1 Introduction

The pidgin English as spoken in these days is about the most atrocious form of speech perhaps one could find in any corner of the globe. It is neither one thing nor the other. Consisting of a mixture of Samoan and Chinese here and there, with an occasional word of Malayan, it is conglomeration truly worthy of the Tower of Babel

(*Rabaul Times*, 16 October 1925; quoted Romaine 1988: 10).

This quotation, taken from a newspaper published in Papua New Guinea, specifically involves Tok Pisin, the pidgin language of New Guinea which is now becoming a creole and which has been recognised as one of the official languages of the country. However, the criticisms contained in it are characteristic of earlier attitudes to pidgin and creole languages, both inside and outside linguistics. Popularly, pidgins and creoles are often still written off as 'inferior, haphazard, broken, bastardized versions of older, longer established languages' (Todd 1990: 1); and although the first systematic studies of pidgins and creoles were carried out in the 1880s by Hugo Schuchardt, these languages were subsequently shunned by linguists, who saw them as aberrant and uninteresting, for the best part of a century. The first conference on pidgin and creole studies took place only in 1959, and most of the key ideas in the field date from this time or later. The study of language birth is therefore a new, expanding, and often controversial one.

It may be useful to begin with some preliminary definitions. Approximately 200 pidgin and creole languages are spoken today, mostly in West Africa, the Caribbean and the South Pacific. A pidgin is essentially a contact language, developed in a situation where different groups of people require some means of communication but lack any common language. Crucially, a pidgin is nobody's first language; all its speakers learn it as adults as a second or further language, and all have native languages of their own. Pidgins are typically restricted in function, since they are used only in contact situations; for instance, in communication between masters

and slaves on former European-owned plantations in the Caribbean, or between the slaves themselves, who were often separated from others of their own linguistic group to lessen the chance of revolt. In more modern times, pidgins often arise in areas where many disparate languages are spoken – Papua New Guinea has one of the densest concentrations of indigenous languages of any area, and Tok Pisin has become very successful there as a *lingua franca*. Along with this restriction in function goes a restriction in form, as we shall see in the next section. Finally, pidgins will generally have at least three parent languages. One will typically be a European language, since pidgins arose predominantly in the context of European colonisation; this language is known as the superstrate, and was typically spoken by the dominant social group, and consequently more socially prestigious in the community. Two or more substrate languages will also be involved: these are typically less prestigious and are often indigenous, although in the Caribbean, for instance, they were typically introduced to the area by the slave population.

It is probably worth emphasising here that pidgins are not lazy, corrupted or debased forms of either the superstrate or the substrate languages, nor are they haphazard mixtures. It is quite possible for *ad hoc*, mixed and highly simplified systems to arise given a particularly pressing need for communication; these can be heard, for instance, in exchanges between taxi drivers and foreign tourists. However, these one-off, unsystematic varieties are not pidgins, but jargons, which lack fixed norms; pidgins represent a later stage in the process and have acquired rules of their own and a degree of systematicity. They can be learned as languages in their own right, albeit rather restricted ones. Pidgins therefore do not only simplify the grammars of their superstrate and substrate languages, but also restructure them to produce a new linguistic system.

A pidgin becomes a creole when it is adopted as the native language of a speech community. For instance, children may be born to parents who have no common language other than the pidgin; or perhaps to slaves who had been separated from others of their own linguistic group, and therefore had to rely on the pidgin to communicate. Such children would be exposed to the pidgin as their first linguistic input, and would acquire it as a native language. Under these circumstances, the new creole will have to fulfil many more functions than its pidgin ancestor, which was used only in particular situations by speakers who had a native language to fall back on; and creoles very quickly become expanded and elaborated in all areas of their grammar. Quite where these new, expanded structures come from is one of the most interesting and controversial areas of current 'creolistics', to which we shall return in some detail in 10.5.

Before going on to consider the structures of pidgins and creoles in more

depth, readers should be warned that there is no real agreement in the current literature as to exactly what pidgins and creoles are, or even whether it is profitable to make a distinction between them (Mühlhäusler 1986). For present purposes, I shall accept the pidgin versus creole dichotomy as a useful working definition, with the added notion of a pidgin-creole 'life-cycle'; that is, I assume that systems may proceed along a continuum from jargon to stabilised pidgin to expanded pidgin, and that creolisation may affect a pidgin at any stage, or not at all.

10.2 Pidgins

10.2.1 Theories of origin

There is considerable disagreement in the literature over how pidgin languages arise, the five most popular theories probably being nautical jargon, independent parallel development, substratum theory, monogenesis with relexification, and baby talk / foreigner talk. I shall review these hypotheses very briefly.

The nautical jargon theory assumes that pidgins are derived from the *lingua franca* used by the crews of ships, presumably through trading and other contacts. This hypothesis might explain some lexical similarities among pidgins; but it is hard to extend this to cover the numerous structural similarities. Furthermore, as Romaine (1988) notes, it is hardly surprising that pidgins have nautical elements in their lexicon, since the majority are spoken near the sea.

On the other hand, as Hall argues, each pidgin might have evolved independently; they are similar in structure because they are restructurings of similar languages, with predominantly European superstrates and African substrates, which develop in similar social and physical conditions. As we shall see in 10.5, Aitchison (1987, 1989a) views creolisation similarly, arguing that particular types of expansion are characteristic of creoles because pidgins are such simple systems: 'given a very simple initial state, there are only a few logically possible ways of extending it. Faced with this limited choice, different pidgins are likely to choose similar options' (1987: 22).

An early account of pidgin origins, which is currently undergoing something of a revival, is the idea that the superstrate or lexifier language contributes the vocabulary to a pidgin, while its grammar comes directly from the substrate(s). Mühlhäusler (1986) argues that this view is partly politically motivated; for instance, it is important for present-day Black Caribbeans to recognise and affirm their linguistic and cultural heritage from their African ancestors. However, substrate influence on pidgins rarely involves all and only the grammar. For instance (Romaine 1988),

256 Pidgins and creoles

constructions in various French-based pidgins and creoles attributed to African languages could also be traced to seventeenth century dialects of French, or might be quite natural developments from earlier French. In fact, many pidgin constructions and lexical items have two or more possible sources, which may not have conflicted, but rather conspired to include the particular structure in the pidgin. Thus, for Tok Pisin *meri* 'woman', we have English *Mary* and *marry*, and also *mari* 'pretty' from the indigenous language Tolai; while Jamaican Creole *doti* 'dirty' may be derived from English *dirty*, West African *Twi doti* 'dirty', or both, with each source reinforcing the other (Mühlhäusler 1986, and see 10.5 below).

The fourth theory, monogenesis, holds that all current pidgins are descended ultimately from Sabir, a fifteenth century proto-pidgin with a Portuguese superstrate which was used in trading and in the first Portuguese colonising expeditions to India, West Africa and the Far East. Sabir itself is sometimes said to be a relic of the original *Lingua Franca*, a medieval language used by Mediterranean traders and by the Crusaders. Sabir was supposedly acquired by the indigenous populations in trading areas; if later colonisation occurred, it would be relexified, with words from the locally dominant language replacing the original Portuguese forms. There is certainly some evidence for relexification: for instance, many Romance loans in English have supplanted inherited Germanic forms, while Voorhoeve (1973) claims that two of the creoles of Surinam, Sranan and Saramaccan, show varying degrees of relexification of Portuguese and towards English. However, the possibility of relexification does not prove monogenesis. This hypothesis does have the advantage of accounting for the Portuguese-derived words which occur very regularly in almost all pidgins and creoles, two examples being the verb 'to know' often of the form *sabi* or *savi*, and the word for 'small' or 'a child', which is frequently *pikin*, *pikni* or *pikini*. On the other hand, French and Spanish influence on West Indian creoles might equally well account for the presence of *savvy* in Early Trinidadian Creole, for instance; and structural similarities are not so readily explained as lexical ones. Finally, this theory cannot account for the existence of non-European-based pidgins which nonetheless share structural similarities with pidgins which do have European superstrates. In the absence of positive evidence for the spread and relexification of Sabir, we must treat this explanation with some caution.

Finally, the baby talk and foreigner talk accounts both relate pidgin origin to second language acquisition; either non-European, indigenous people learned an imperfect version of the target, superstrate language, or the European colonisers simplified their own language to make it easier for the substrate speakers to learn. These entirely compatible explanations

may be rather enlightening, provided that we dispose of the unjustified racist notion, thankfully becoming rarer now, that non-Europeans are intellectually incapable of mastering the subtleties of a European language. It is overwhelmingly likely that substrate speakers acquire a pidgin simply because they are not exposed to the superstrate sufficiently often or for sufficiently long for it to serve as an adequate model: 'pidginization is second-language learning with restricted input' (Bickerton 1977: 49). Access to the superstrate is generally restricted for socioeconomic reasons. For instance, slaves arriving on plantations early may have had more opportunity to acquire the superstrate, since there would be a higher proportion of Europeans to non-Europeans; but gradually, the substrate population would outnumber the superstrate speakers, diluting access to the superstrate (Muysken and Bickerton 1988). Alternatively, the superstrate speakers may not wish their language to be learned by outsiders: thus, the Chinook Indians produced a simplified version when talking to non-Chinooks (Foley 1988); the resulting Chinook Jargon is a pidgin with Chinook, Nootka, French and English as its sources. Naro (1978) provides a further example of deliberate simplification by superstrate speakers, claiming that Portuguese speakers in the sixteenth century taught a simplified version of Portuguese to Africans they brought back from early voyages to train as interpreters. However, this account is controversial (see Thomason and Kaufman 1988).

Simplifications initiated by superstrate speakers may therefore be a contributing factor in pidginisation. Pidgins and temporary foreigner-talk systems used in communicating with speakers of other languages frequently show radical simplifications of very similar kinds, which might result from universal, innate rules for simplifying grammars. It is certainly true that we all have the capacity for simplifying our native language structures, and that we do so by using sub-conscious, recurrent strategies; this is obvious when, say, an English-speaker addresses someone with a very limited command of English, or in 'motherese', the simplified language which carers use to babies and young children. This phenomenon is not restricted to Indo-European languages; for instance, in Lamso, spoken in Cameroon, there is special vocabulary used only to and by children, and tonal contrasts are reduced when addressing a child (Todd 1990). Such pre-programmed simplification strategies represent one possible explanation of the structural features shared by so many pidgins. However, Bickerton (1977: 55) denies the involvement of universals in pidginisation, pointing out that 'pidgins are usually described as lacking inflections, lacking articles, lacking markers of tense and aspect, lacking sentence-embedding, nominalization, allomorphic variation etc., etc., and people conclude from such definitions that all pidgins are more or less the

same; it would, in fact, be as reasonable to suppose that a brick is the same as a cabbage because neither has legs, wings, fur, feathers, independent locomotion, etc., etc.' We now proceed to consider some of these supposed similarities.

10.2.2 *Pidgin structure*

As we have already seen, the term 'pidgin' is in fact a cover-term for languages at various points on a continuum. The first phase involves the formation of jargons, which are extremely simple, *ad hoc* systems of communication, with very restricted structural resources, great variation, and sentences generally no more than one or two words long. These may in time acquire particular norms, become rather more complex, and extend into other domains of usage, whereupon they become stable pidgins. Finally, stable pidgins may expand, acquiring more morphology and more complex syntax, and a characteristically faster tempo of speech; such expanded pidgins are already partly creolised in linguistic terms, although they are not yet properly creoles since they lack native speakers. In what follows, I shall primarily be discussing stable pidgins.

'A pidgin represents a language which has been stripped of everything but the bare essentials necessary for communication' (Romaine 1988: 24). In other words, pidgins are characteristically used only in particular domains and to fulfil certain restricted linguistic functions. To be more precise, they are frequently used in the directive function, to get people to perform particular tasks, and in the referential function, which involves describing some situation to achieve a particular end. However, they are rarely if ever used in the interactional function, which involves the promotion of social cohesion; expressively, to indicate inner states or abstract ideas; poetically, in the creation of literature; or metalinguistically, to talk about language (Foley 1988). These restrictions mean that pidgins typically lack stylistic options, puns and metaphors, and have few sociolinguistic markers, such as politeness phenomena. These restrictions in function and in context typically accompany a restriction in linguistic form, which itself falls into two categories: pidgins are reduced, in that they have a curtailed referential capacity, making certain meanings difficult or impossible to convey; and they are simplified, in that regularity in the grammar is increased (Mühlhäusler 1986).

First, the lexicon of a pidgin is characteristically reduced by comparison with the superstrate and substrate languages; Romaine (1988: 34) estimates that 'normal' languages have approximately 25–30,000 lexical items, while Tok Pisin has around 1,500. However, each word in the pidgin will have a wider range of meanings than is the case in the superstrate or

substrates, or will cover a larger semantic domain. Thus, Cameroon Pidgin has two words for the semantic field of animals, *bif*, which is edible, and *bushbif*, which is likely to eat you; each has generic rather than specific reference. In Tok Pisin, *gras* means 'grass', but also more generally 'something which grows somewhere', as shown in (1).

- (1) *gras bilong het*
grass belong head = 'hair'
gras bilong mous
grass belong mouth = 'moustache'
gras bilong pisin
grass belong bird = 'feathers'
gras bilong solwara
grass belong saltwater = 'seaweed' ...

Words are often multifunctional, acting as nouns, verbs and adjectives; and there is no compounding, so that the expression of complex ideas requires a good deal of circumlocution and periphrasis, as shown in (1) and (2).

- (2) *liklik brum bilong klinim tit* = 'toothbrush'
bikpela box yu faitem i singaut = 'piano'
(possibly apocryphal; see Mühlhäusler 1986: 26)

Pidgins rarely exhibit any inflectional morphology, so that no marking for gender, case, number, tense and so on generally occurs. For instance, Yimas is a Papua New Guinean language, which has given rise to a pidgin Yimas, used in trading with the neighbouring Arafundi tribe (Foley 1988). Yimas is inflecting, while Yimas Pidgin is virtually isolating: Yimas has seven tense distinctions, while Yimas Pidgin has two; Yimas has three numbers and more than a dozen noun classes, while Yimas Pidgin has no inflectional marking for either category and marks plurality using *kundamin* 'two' or *manba* 'many' after the noun; Yimas verbs have affixes indicating the person, number and noun class of the subject and the direct and indirect object, while Yimas Pidgin has no such inflections and marks only the indirect object using the postposition *namban* 'towards'. Yimas Pidgin is now dying out under the influence of the more successful and socially prestigious Tok Pisin (Foley 1988), which also lacks inflections for number, as shown in (3), again using numerals and quantifiers.

- (3) *pik* 'pig' / 'pigs'
tripela pik 'three pigs'
planti pik 'many pigs'

Since pidgins lack inflectional morphology, words will typically be invariant, with no allomorphy and no irregular forms; this maximises iconicity, making the connection of form and meaning as transparent as

possible, and is probably of considerable assistance to the non-native second-language learner. In short, the hearer will find it relatively easy to decode the language, since exceptions and variation need not be coped with (although she may have to rely on context for certain cues as grammatical categories will not be made linguistically overt); while the speaker will also benefit from the minimal grammar and maximal reliance on vocabulary, since words are easier to acquire than grammar (Foley 1988).

Pidgins also tend to have a fixed, invariable word order, which is characteristically SVO. For instance, although German has SVO order in main clauses and SOV in subordinate clauses, Rabaul Creole German, also spoken in Papua New Guinea, has consistent SVO, as in (4).

- (4) *Wenn der Baby weinen, der Mama muss aufpicken.*
 'when the baby cry, the mother must pick up'
 (Romaine 1988: 30)

The example in (4) also illustrates the earlier points on morphology; the verbs *weinen* and *aufpicken* appear in the infinitive, and the nouns *Baby* and *Mama* appear with the invariable article *der*, which does not alter as its German equivalent does according to gender, number and case. Syntax is also extremely simple in pidgins, with invariant word order for statements, questions and negatives, and no complex syntactic phenomena such as embedding or relative clauses.

Finally, the phonology of pidgins is rather simple, with usually five vowels, typically /i e a o u/, or fewer, and no length distinction. The consonant system is also fairly small; for example, Pidgin Zulu, or Fanagalo, has reduced the three clicks of Zulu to [k], while Tok Pisin has replaced English /s f t/ with the single fricative [s]. Pidgins also tend to have rigidly CV syllable structure, so that loans from English into Tok Pisin, such as *salt* and *parliament*, emerge as *sol* and *palamen*. In African pidgins, tonal distinctions are also characteristically simplified or lost. Finally, words are short, usually mono- or bi-syllabic, and speech tempo is very slow.

10.3 Creoles

10.3.1 Introduction

Creolisation is the linguistic inverse of pidginisation: while pidginisation involves reduction and simplification, creolisation is characterised by expansion and elaboration. A pidgin becomes a creole when it acquires native speakers; these will generally be the children of pidgin speakers who are exposed to the pidgin as the medium of communication in the home, and who therefore have the pidgin as their primary linguistic input.

Although the pidgin was adequate for the parental generation, who used it only in specific circumstances and otherwise had recourse to a non-pidgin native language, it appears to be too restricted a system for the children, who need a native language to fulfil not only the directive and referential functions of language, but also the interactional, expressive, poetic and metalinguistic roles which pidgins do not play. Functional and linguistic expansion then typically go together; as the creole becomes the native language of a community, and is gradually extended into new domains, its linguistic resources also increase (Foley 1988). The degree of expansion and elaboration varies from creole to creole, and depends on the status of the pidgin when creolisation begins to take place: most expansion will be necessary when a creole develops from a jargon, as with Hawaiian Creole English; rather less is required for a stable pidgin, which already has certain norms of usage, like Torres Straits Creole English; and least expansion is typical of creoles which develop from expanded pidgins, which are already well on their way to creoles in the linguistic sense – this is probably the most common type of creolisation, examples being Tok Pisin, and Bislama, a creole spoken in Vanuatu and previously called Beach-la-Mar (Crowley 1991). We shall return in 10.5 to the vexed question of where these additions, expansions and elaborations come from, and quite how they become available to creole-acquiring children; first, however, we need a better idea of the types of processes characteristically involved in creolisation. Just as there are numerous characteristics which seem to be common to most or all pidgins, so unrelated creoles also share certain features and processes, some of which will be outlined below. In general, these contribute to the flexibility of the language, creating stylistic options and increasing variation in the system.

10.3.2 Phonology

Relatively little attention has so far been paid to phonological creole universals, although Romaine (1988: 63) suggests that creoles typically have CV structure with few or no consonant clusters, and select their vowels from the restricted set /i e a o u/ and their consonants from the set /p t k b d g f s m n l r w j/. The speed of speech tends to increase, and fewer words in the sentence carry stress, so that pidgin *Mi gó lóng háus* ('I go to the house' / 'I go home'), with main stresses on the verb and noun and secondary stresses on the pronoun and preposition, becomes creole *Mi gò háus*, where the preposition *lóng* has been significantly reduced and only one main stress remains, on the final noun (Sankoff and Laberge 1974). This phonological reduction contributes to the formation of morphology; for instance, Tok Pisin *baimbai*, from English *by and by*, is reduced in the

creole to *bai*, which is placed before the main verb of the sentence and typically lacks stress, and which is increasingly interpreted as a prefix (Sankoff and Laberge 1974).

10.3.3 Lexicon

Hancock (1980) identifies twelve processes by which the lexical resources of creoles and other languages can be increased, arguing that 'there is a universality of lexical process shared by all languages; the difference between these in creole and non-creole languages is one of degree' (Hancock 1980: 84). For instance, all languages seem to coin words, although none do so very frequently; English has examples like *gobbled-gook*, while Trinidadian Creole English has [bobolups] 'fat, ungainly'. A strategy used more in creoles than in non-creole languages is reduplication. The phonological reduction typical of pidgins can lead to widespread homophony, for instance in Krio where *was* meant both 'wash' and 'wasp', and *san* both 'sun' and 'sand' (Todd 1990); this is resolved in the creole using reduplication, which gives *was was* 'wash' and *was was was* 'wasp', and *sansan* 'sand'. Also in Krio, *tu* means 'two' but *tutu*, 'by twos', while in Pijin, *drai* is 'dry', but *draidrai* describes unpalatable food.

Words may also undergo semantic extension (see Chapter 7), adding a new meaning, so that Trinidadian Creole English [batn] means 'button' and 'pimple'; or may shift their meaning, losing an older one and gaining a new one, as with English *computer*, which in the seventeenth century meant 'clerk, accountant', and Krio [tomok] 'chest', which derives from English *stomach* – in Krio, 'stomach' is [bele]. Finally, creoles, like other languages, may borrow words; for instance, Trinidadian Creole English has borrowed [susu] 'mutual savings system' from Yoruba [esúsú].

10.3.4 Morphology

As we saw in 10.2, pidgins typically lack morphological marking; for instance, Tok Pisin nouns may be recognised as plural according to the context, or by the addition of a numeral or quantifier, but have no overt plural marker. In creolised Tok Pisin, however, speakers productively insert *ol* before a noun to indicate plurality, increasing redundancy when a numeral or quantifier also appears (Mühlhäusler 1986). Similarly, pidgin Tok Pisin has periphrastic causative constructions formed using the auxiliary verb *mekim*, as in *Yu mekim sam wara i boil* 'You make some water boil', which are now being replaced in the creole by shorter constructions with a transitive / causative suffix *-im* on the main verb, giving *Yu boylim wara* (Foley 1988: 178). Very frequently, new affixes are

the result of grammaticalisation: for example, as we saw in Chapter 6, Tok Pisin *save* 'to know' has become a habituality marker *sa* (Aitchison 1989b). This particle now frequently coalesces with the main verb, which it often immediately precedes, being increasingly interpreted as a prefix.

Earlier circumlocutions also give way to compounding in creoles. Pidgin Tok Pisin relied on circumlocutions like *man bilong save* 'man belong know', hence 'expert' and *meri bilong hambak* 'woman belong handbag', 'promiscuous woman', which have now been replaced in the creole by *saveman* and *hambakmeri*. Similarly, *ai bilong mi i laik slip* 'my eyes like sleep' = 'I'm sleepy' becomes *mi aislip nau*, and *yau bilong em i pas* 'his ears are closed' = 'he's deaf' becomes *em i yaupas* (Foley 1988: 177). This process reduces iconicity, as single words no longer necessarily express single concepts, reversing the trend towards maximal iconicity found in pidginisation.

10.3.5 Syntax

While pidgins lack sentence-embedding, and have only main clauses, constructions with embedded subordinate clauses tend to develop in creoles. For instance, as shown in (5), creolised Tok Pisin introduces complement clauses using *olsem*, best translated as 'that'.

(5) Pidgin Tok Pisin:

Mi no save. Ol i wokim dispela haus.
Creole Tok Pisin:

Mi no save olsem ol i wokim dispela haus.
'I didn't know that they built this house'
(Foley 1988: 179)

However, embedding is by no means the only syntactic innovation characteristic of creolisation: Bickerton (1981) gives a list of twelve features of grammar common to many creoles, although this set of properties is not exhaustive, and additional or rival features have been suggested (see Romaine 1988, Chapter 2). For instance, many creoles have the same lexical item for the existential, meaning 'there is', and the possessive, meaning 'have' (it is *get* in Hawaiian and Guyanese Creole, and *ge* in Haitian Creole), although apparently none of the superstrate languages have this feature. Creoles have no syntactic difference between statements and questions, although they do have question words, which tend to be bipartite, with the first part derived from a superstrate question word: for example, Guyanese Creole has *wisaid* 'which side' = 'where', and *wa mek* from 'what makes' = 'why'; Haitian Creole has *ki koté* from French 'which side' = 'where'; and Tok Pisin has *wanem* 'what name' = 'what'. Creoles also rarely use passive constructions, and those that do

(8)

Dispela man ia, lek bilong en idai ia,
'this man, whose leg was injured,
em istap insait nau.
stayed inside'
(Sankoff and Brown 1976: 632)

Ia is derived from English *here*, and seems to have been used first as a place adverb, then as a demonstrative equivalent to English *this*, *that*, and finally in relative clauses. Sankoff and Brown (1976) also suggest that, although the lexical material involved is from English, some substrate reinforcement may have occurred, since the indigenous Austronesian language Buang has a particle *ken* which is used as a place adverb, a demonstrative, and a relativiser; and similar patterns recur in other Melanesian languages including Iai, Nguna, Tasiko, Uripiv and Tangoa. Again, as with pidginisation, we see a mixture of substrate and superstrate influences.

10.4 The general relevance of pidgins and creoles

Pidgins and creoles were for a long time considered 'marginal' or 'special' languages. To call them this was to label them and the processes by which they developed as marginal to linguistic theory.

(Traugott 1977: 70)

If we accept this earlier point of view, we might see pidgins and creoles as interesting objects of study in their own right, more because of their unique (and, in some opinions, freakish) characteristics than despite them; but we are unlikely to regard them as central to historical linguistics. However, the recent sociolinguistic perspective in historical linguistics (see Chapter 9) has produced an increasing awareness of these languages among historical linguists; indeed, 'from being the special and the marginal case, pidginization and creolization have become, for some linguists, the test case for any theory of change' (Traugott 1977: 70).

There are various reasons why historical linguists might wish to study pidgins and creoles, and more specifically pidginisation and creolisation. First, and most trivially, these processes provide conclusive proof against Bloomfield's view that it is methodologically impossible to observe language change in progress: if we can actually see new languages being born, and trace their subsequent development, then we are certainly observing change! Change also seems to proceed more quickly during pidginisation and creolisation than elsewhere, so that developments which can only be suggested for the history of other languages can be directly observed or at least recovered in pidgins and creoles, without resorting to apparent-time studies; this may allow the confirmation of various hypotheses on the nature and spread of change. However, the processes

exist, like the *get* passive of Guyanese Creole, tend to be extremely marginal, or are recent borrowings from the superstrate, or both.

All creoles seem to have a system, as shown in (6), whereby a definite article is used for presupposed-specific Noun Phrases (which refer to a particular entity which the speaker assumes the hearer knows about); an indefinite article for asserted-specific Noun Phrases (which have a particular referent which the speaker is introducing to the hearer); and zero to mark non-specific Noun Phrases.

(6) Papiamentu (Bickerton 1981: 53)

mi tin e buki 'I have the book'
mi tin e bukinan 'I have the books'
mi tin un buki 'I have a book'
mi tin buki 'I have books'
buki ta caru 'books are expensive'

The majority of creoles also have three preverbal particles to express distinctions of tense, mood and aspect, which appear in the order TMA. The tense marker typically means [+ anterior], or roughly speaking past tense, while the unmarked verb stem means non-past; the mood marker expresses [+ irrealis], marking phenomena which have not yet happened, either because they are imaginary or conjectural, or in the future or conditional, while the verb stem indicates realised predicates; and the aspect marker signals [+ non-punctual], marking progressive or habitual actions, while the verb stem alone shows completed events. The forms used as particles differ from creole to creole, with Hawaiian Creole *bin*, *go*, *siei*, Guyanese Creole *bin*, *sa/go*, *a*, Haitian Creole *te*, *ava*, *ape*, but the order of particles and their meaning, both alone and in combination, is remarkably consistent.

Creoles tend to mark unrealised complements using a particular particle, while realised complements, which describe an action which actually took place, are marked by zero or a different complementiser. In Sranan (see (7)), unrealised complements are introduced with *foe*, and realised ones with zero; in Mauritian Creole, realised complements have *al* and unrealised ones, *pu al*. This distinction is realised in few if any non-creole languages.

(7) Sranan (Bickerton 1981: 60)

a teki a nefi foe koti a brede, ma no koti en
'he took a knife to cut the bread, but he didn't cut it'
***a teki a nefi koti a brede, ma no koti en*

Finally, creoles seem to be developing relative markers, although some, including Hawaiian Creole English, currently lack them. Creole Tok Pisin, for instance, is reported (Sankoff and Brown 1976) as using *ia* to bracket relative clauses, as shown in (8).

found in pidginisation and creolisation are only likely to be relevant to general theories of change if they are of the same type as changes which affect other languages.

Trudgill (1983: Chapter 5) attempts to draw parallels between the changes characteristic of pidginisation and creolisation, and those found elsewhere. He suggests a general division of changes into two types, which he labels 'natural' and 'non-natural': 'natural' changes are those which are 'liable to occur in all linguistic systems, at all times, without external stimulus, because of the inherent nature of linguistic systems themselves', while 'non-natural' changes are generally the result of language contact. Trudgill proposes that 'natural' changes are likely to include grammaticalisation; sound changes traditionally ascribed to ease of articulation, such as assimilation; and increases in redundancy, such as the double marking of the definite article in Norwegian seen in *den store mannen*, literally 'the big man-the'. These processes involve a general move from analytic to synthetic structure. On the other hand, 'non-natural' changes include reductions in morphological marking for case, an increase in the use of prepositions, the reduction of conjugational classes for verbs and declensional classes for nouns, increased use of periphrasis and reduced use of inflected forms, and the development of fixed word order; all these involve a shift from synthetic to analytic structure.

Trudgill's central hypothesis is that languages which have undergone a good deal of linguistic contact will exhibit more change, faster change, and specifically more 'non-natural' changes than related languages which have not experienced such contact. For instance, Faroese is a more isolated language in sociolinguistic and geographical terms than its relative Norwegian, and Faroese is more conservative in its inflectional system. Faroese retains marking for three cases in nouns and pronouns, while Norwegian has only two cases and only for pronouns; Faroese has several verb conjugations which Norwegian has reduced to one; Faroese maintains three noun declensions per gender while Norwegian again has only one; and Faroese has eleven inflected verbal forms to Norwegian's five.

Trudgill goes on to propose that pidgins, as high-contact languages, exhibit predominantly 'non-natural' changes, while creoles show mainly 'natural' ones. This seems to be broadly borne out by the evidence in previous sections; pidgins do manifest a dramatic shift towards analytic structure, and a reduction to the point of complete loss of inflectional morphology, while creolisation is characterised by grammaticalisation and increased redundancy, which lead to a move back towards syntheticity. There are, however, some problems with Trudgill's proposals. One concerns the implied categorisation of creoles as non-contact languages; creoles do seem to borrow at least some of their novel structures from the

superstrate and substrate languages of their ancestor pidgins, and are generally spoken in areas with many indigenous languages – indeed, this may promote the development of a creole as a *lingua franca*, as was the case for Tok Pisin in Papua New Guinea. Furthermore, Trudgill suggests that 'natural' change will tend to be slower than 'non-natural' change, but this is patently not the case in creolisation, which may take place, with all its attendant expansion and elaboration, in as little as a single generation – less time than some pidgins take to develop! Finally, it may not be so easy as Trudgill's dichotomy indicates to make a clear distinction of 'natural' and 'non-natural' changes, since a change which is already occurring in a language may be reinforced by a model in another language with which it is in contact; and it seems reasonable to suppose that a particular change may be the result of language-internal phenomena on one occasion, but motivated by contact on another. However, it seems less outlandish to refer to changes as sometimes internally and sometimes externally motivated, than to call the same process sometimes 'natural' and sometimes 'non-natural'; this may seem like quibbling about terminology, but terms like 'unnatural' in historical linguistics have often led to judgements of 'undesirable' or the like (see Chapter 12), and 'natural' and 'non-natural' are arguably just too close to 'natural' and 'unnatural' for comfort.

If we at least accept that certain types of processes are characteristic both of pidgins and high-contact languages, while others operate frequently in creoles and low-contact languages, we can regard pidginisation and creolisation 'not as different in kind from other processes of language change, but as extreme cases of the hybridization that goes on in language all the time' (Traugott 1977: 74). This sort of connection of pidginisation and creolisation with other cases of linguistic change has been interpreted in several ways. The most literal is the idea that pidginisation and creolisation themselves may have formed part of the history of languages which would not now be classified as pidgins or creoles. For instance, there is a view that Middle English, due to the extensive admixture of French into English during the period following the Norman Conquest, became a creole. Certainly Middle English shows a large number of 'non-natural' changes, involving the loss of inflectional morphology, but these were arguably already under way in late Old English, where scribes seem to become increasingly confused about the appropriate spelling for inflectional endings; in any case, such changes are more characteristic of pidginisation than creolisation. If Middle English were really a creole, then we would expect it to have developed not only a relatively fixed word order, and periphrastic constructions, which it has, but also grammaticalised inflections to replace the ones lost earlier, which it has not.

The problem is that claims for a creole stage in English, or in French and the other Romance languages between Vulgar Latin and the earliest Romance texts, are hard to prove, largely because appropriate data from the periods concerned tend to be lacking. Diagnostics of prior pidginisation and creolisation might be developed, for instance by taking a set of features and creolisation might be developed, for instance by taking a set of features and assessing which of these appear in certain modern or attested languages; languages with more than a certain number of these features would be classified as having creole ancestry. Romaine (1988) considers thirty pan-creole features, and discovers that Old Japanese has approximately one-third of these while Modern English has very few: however, we should be cautious before using such evidence to debunk the claim that Middle English was a creole while asserting that the ancestor of Old Japanese was one, since English may simply have undergone rapid change away from the creole state in the period since Middle English.

Some linguists have reacted to these difficulties by coining terms like creoloid, which, as Holm (1988-9: 10) suggests, 'has been used for so many different kinds of vaguely creole-like languages that its usefulness has become rather limited'. Holm himself uses the term for languages with some creole features, which nonetheless seem not to have actually undergone creolisation; while Trudgill (1983) applies it to languages like Middle English, Norwegian and Afrikaans, which have no pidgin in their history, but have undergone many contact-induced changes and are now relatively easy for adult second-language learners to acquire, compared to their linguistic ancestors. Unfortunately, creoloids and the like are rather hard to distinguish from so-called mixed languages, which have borrowed structures from other, often unrelated languages through linguistic contact; and since almost all languages have borrowed some features at some time or another, we are in danger of identifying almost all languages as having greater affinities with pidgins and creoles than they actually *do* have. The connection is simply the property of being languages used in situations involving communication between linguistic groups, which facilitate the leakage of linguistic features from one system to another. It seems doubtful whether we need more terms for this phenomenon than those, like contact and borrowing, which we have already. In short, pidginisation and creolisation share certain processes and features with other cases of language change, but that does not make those cases equivalent to pidginisation and creolisation; the sociolinguistic context and the output are clearly different.

Another claim of affinity between pidgins, creoles and other languages involves genetic relationship, and the traditional family tree representation of this. In fact, Schuchardt in the late nineteenth century first became

interested in pidgins and creoles because he felt they represented a challenge to the genetic classification of languages and the doctrine of the regularity of sound change, both elements of Neogrammarian theory which he attempted to refute. Proponents of monogenesis might suggest a separate pidgin/creole family, descended from the proto-pidgin Sabir, but this tree will necessarily be too static to show the waves of contact and relexification which form a crucial part of the monogenetic hypothesis, and will also fail to show the influence of either the superstrate or the substrate languages. However, it is equally difficult to incorporate pidgins and creoles into established family trees, because of the indirect relationship between the superstrate, the substrates, and the resulting pidgins and creoles: should Haitian Creole be a daughter of French, or a sister, or an independent branch in the Romance tree? Similarly, is Tok Pisin Indo-European, like its superstrate, or Austronesian like its substrates? Faced with questions like these, the family tree model develops glaring inadequacies. One way of dealing with these is to extend the family metaphor, as Taylor (1956: 413) does when he asserts that 'languages originating in a pidgin or jargon, while genetically "orphans", may be said to have two "foster-parents": one that provides the basic morphological and/or syntactic pattern, and another from which the fundamental vocabulary is taken'. Setting aside the difficulty of including 'foster-parents' in the conventional family tree, this view incorporates the idea, now largely discredited, that the superstrate language is responsible only or primarily for pidgin and creole vocabulary, while the substrates supply grammar: it now seems that both superstrate and substrates donate lexical and grammatical structures, and often patterns from one will reinforce material from another.

An alternative strategy is adopted by Thomason and Kaufman (1988), who claim that pidgins and creoles are non-genetic languages; although their contributing ancestors generally belonged to some linguistic family, these languages themselves have arisen through a process of discontinuous transmission and can no longer be classified as members of any family group. Thomason and Kaufman extend this set of non-genetic languages to cover those which have undergone extensive structural borrowing, such as Mbugu or Ma'a, which has effectively ceased to be a Cushitic language due to the volume of borrowings from Bantu, but which cannot really be classified as a Bantu language either. However, again, it is hard to draw the line: is English, with its widespread lexical borrowing from Romance, or Rumanian, which has borrowed heavily from Slavic languages, non-genetic? As Holm (1988-9: 14-15) points out, French and Middle English do show clear signs of the effects of language contact relative to Latin and Old English, but there is no evidence of any major break in transmission;

generation-to-generation acquisition seems to have continued uninterrupted, with speakers believing that they spoke the same language as their predecessors, whereas this is not the case for pidgins and creoles.

Finally, another productive line of research, though not one which has been much pursued as yet, involves the search for similarities between pidginisation and second language acquisition, and creolisation and first language acquisition respectively. If, as Bickerton claims, pidginisation is second language acquisition with restricted input, while creolisation is first language acquisition with restricted input, then such similarities certainly ought to exist; but one caveat should be imposed. Historical linguists have so far had little success in assuming that children learning language initiate language change, since there is very limited overlap between changes and characteristics of child language (Drachmann 1978). Invoking the child seems to make much more sense within particular linguistic theories, like Andersen's work on abduction (see Chapter 4), or Lightfoot's on syntactic change (see Chapter 5). How the child learns therefore seems more relevant to the historical linguist than the performance errors she makes. With this in mind, we can proceed to an ambitious, acquisition-based account of creolisation, which recognises creoles as unique, but also as a key to language change, language learning and linguistic evolution.

10.5 The Language Bioprogram Hypothesis

Bickerton (1977, 1981, 1984, 1990; Muysken and Bickerton 1988), the originator of the Language Bioprogram Hypothesis, aims ultimately to answer three questions, which he sees as intrinsically interconnected. First, he is concerned to discover how creoles originate; then, how children acquire language; and finally, how language evolved: and in his view, the key to this whole ambitious enterprise is the study of creolisation.

Bickerton notes that creole languages tend to share particular features of grammar; most notably, perhaps, almost all creoles have a system of three preverbal particles which mark tense, mood and aspect distinctions (see 10.3 above). The typical expansions of creole structure, and especially this TMA system, have often been ascribed to substrate influence or monogenesis. However, Bickerton claims that the similarities found across creoles cannot be due entirely to these factors. For instance, he notes that Caribbean creoles and Hawaiian Creole, which certainly do not share any substrate, have much the same TMA system; to maintain a substratist explanation, we would have to assume that all substrates had the same features, and this is simply not the case. More specifically, substratists may find a potential source for some creole feature in a particular African language, but fail to demonstrate that the speakers of that language were

in the right place at the right time and in sufficient numbers to transmit the construction into the creole. For instance (Bickerton 1984), a possible source for the creole TMA system in some creoles is the language Fon, but the number of Fon speakers in, for instance, late-eighteenth-century Haiti was rather small, probably in the region of 0.2 per cent. Similarly, to uphold monogenesis, we would have to assume that the TMA system in, for instance, Hawaiian Creole was inherited from some previous contact proto-pidgin; but Bickerton claims that the relevant constructions do not occur in Hawaiian Pidgin.

The usual difficulty with claims of this sort is the sheer impossibility of substantiating them; once a language has creolised, it tends to die out as a pidgin, and it is therefore almost impossible to find both pidgin and creole speakers at the same time, unless the language is still undergoing creolisation. However, Bickerton (1984) notes that Hawaii is probably the latest site of European colonisation, with the original pidgin being formed during the period 1900–20, and creolisation occurring subsequently. He interviewed speakers aged around seventy to ninety, who had arrived in Hawaii in the period from 1907–30, and found that they predominantly spoke a rather rudimentary version of Hawaiian Pidgin, which shows little consistency in syntax and is often heavily influenced by the speaker's native language. Bickerton argues that, even if the language of these speakers has changed since the 1900–20 period, perhaps under the influence of subsequent generations of creole speakers, it is likely to have become more rather than less complex – which suggests that early Hawaiian Pidgin was even more rudimentary than the surviving pidgin, if anything. There seems little possibility that this extremely rudimentary system supplied patterns for the more complex constructions which later appear in the creole; and yet children born in Hawaii, who were exposed to the pidgin as linguistic input, seem to have developed structures like those in (9).

- (9) a. *dei gon get naif pok you*
'they will stab you with a knife'
– speaker born 1896, showing verb serialisation
- b. *dei wen go up dea in da mawning go plaen*
'they went up there in the morning to plant (things)'
– speaker born 1896
pipl no laik tekam fo go wok
'people don't want to have him go to work (for them)'
– speaker born 1901
Shows embedding, plus *go* for realised, *fo* for unrealised complements.
(Bickerton 1984: 175)

It follows that speakers exposed to pidgin as children have acquired a more complex, creole system, while those learning as adults acquired only a very simple and restricted pidgin. Moreover, patterns for the new

constructions which emerge in the creole often have no clear sources, in the superstrate, the substrates or the pidgin itself; and to support monogenesis we would have to argue that the importers of the proto-creole to Hawaii must have taught or otherwise transmitted it to all children born in Hawaii and no adults. We cannot therefore rely on monogenesis or substrate influence to account for creolisation fully, and must seek an alternative explanation in terms of language acquisition.

However, children with a pidgin as their primary linguistic input are faced with a task rather different from 'normal' acquisition: their model is impoverished, unstable, and inadequate. As Bickerton (1981: 5) puts it, 'children of pidgin-speaking parents have as input something which may be adequate for emergency use, but which is quite unfit to serve as anyone's primary tongue; which, by reason of its variability, does not present even the little it offers in a form which would permit anyone to learn it; and which the parent, with the best will in the world, cannot teach, since that parent knows no more of the language than the child (and will pretty soon know less)'. Bickerton's answer is to suppose that these pidgin-learning children can 'invent' structures for which there is no evidence in the speech of the previous generation; that is, 'there can be rules of language that are not derived from any linguistic input' (Bickerton 1977: 65). Once we understand this sort of innovation, we may also be closer to understanding 'normal' cases of language acquisition, since a single explanation should ideally cover both creole and non-creole learning; if we propose two separate strategies, the child would have to work out what sort of data she was receiving to decide which strategy to use.

It would, however, be equally unreasonable to expect very young children to invent such complex constructions without help; and in any case, if no guidance were available, we would logically expect different children, and different creoles, to come up with different strategies and structures. As we have seen, this is precisely what we do not find; instead, unrelated, geographically distant creoles exhibit strikingly similar patterns. Bickerton therefore proposes that children are provided with some assistance in 'inventing' new creole constructions, and that this help comes in the form of the bioprogram, a neurally encoded, genetically transmitted set of instructions which specify certain semantic and syntactic features. These, Bickerton argues, were the first features to develop in linguistic evolution, and are still the first to be given linguistic expression by *any* child learning language. Similar syntactic structures and semantic distinctions will therefore appear in the early language of all children whether their primary linguistic data is French, English, Japanese, Dyirbal or Hawaiian Pidgin. However, over centuries or millennia of linguistic change, older established languages have often developed patterns of their own which

fail to match, or directly contradict, those specified in the bioprogram. In such cases, children will first produce bioprogram features, but will subsequently, due to overt correction by adults or simply greater exposure to data, replace these with the structures and distinctions of their own languages. Only in creoles, where the preceding pidgin is likely to be too impoverished and restricted to contradict the bioprogram, will bioprogram features emerge and remain as an integral part of the resulting adult language; hence the importance of creoles in discovering the initial features of evolving language and child language. For instance, child English and German often lack inverted order in questions, a feature which all creoles retain. Similarly, the construction *a gon full Angela bucket* ('I'm going to fill Angela's bucket') (Bickerton 1984: 185), with an adjective used as a verb, is attested in child English but would have precisely the same form and meaning in Guyanese Creole. Finally, in the domain of semantics, Bickerton (1981) reports an experiment by Maratsos conducted to see whether English-speaking children could discern non-specific uses of the indefinite article (*I want a dog* - any dog) from specific instances (*I have a dog* - a particular dog). His findings were that three-year-olds were 90 per cent efficient in grasping this distinction, although English gives very few overt cues to the appropriate interpretation. As we have already seen, creoles tend to mark this distinction overtly, using the indefinite article for asserted specific noun phrases, but zero in non-specific cases. In short, (Bickerton 1984: 185), many apparent errors in child language 'would have been grammatical if the child had been learning a creole language'.

We shall return below to a specification of the contents of the bioprogram, but it should first be noted that not all creoles are equal when it comes to the expression of bioprogram features. In his earlier work, Bickerton defined creoles as languages arising out of a pidgin which had existed for no more than one generation, and developing in a population where no more than 20 per cent spoke the superstrate, with the remaining 80 per cent being speakers of various substrate languages; his idea here is to concentrate on cases where (Bickerton 1981: 4) 'the human linguistic capacity is stretched to the uttermost'. Aitchison (1983) suggests that this restriction represents a convenient excuse for Bickerton to concentrate on the two creoles which he knows best, namely Hawaiian Creole and Guyanese Creole, ignoring creoles like Tok Pisin which might provide counter-examples. This allegation may be slightly unfair, since it is obviously impossible to consider all creoles, and the particular restriction Bickerton adopts, emphasising the poverty of the input available to the child, is probably as valid as any; when he has stated his case, it is then the task of other linguists to confront him with conflicting data from other creoles. The fact that Bickerton's criteria include the two creoles with

which he is most familiar might cause some faint unease, but he does not by any means restrict his data to these two languages; in his survey of features of creole grammar (Bickerton 1981: Chapter 2), for instance, Bickerton considers evidence from Guyanese Creole, Hawaiian Creole, Papiamentu, Seychelles Creole, Haitian Creole, Lesser Antillean Creole, Saramaccan, Sranan, Jamaican Creole, Mauritian Creole, Crioulo, Tok Pisin and various others; and although he has not done fieldwork on all of these himself, few of us can afford to be surprised or critical on that account. In more recent work (Bickerton 1984), Bickerton has in any case relaxed his original criteria for creole status, and instead sets up a scale of creoles, from those nearest to the bioprogram to those furthest away. Creoles are ordered on this hierarchy on linguistic and demographic grounds; broadly speaking, the more disruptive the circumstances of acquisition of the creole, the more bioprogram features it will display. The degree of disruption will depend on the amount of access which adults learning the pidgin have to the superstrate, which determines the degree of restriction of the pidgin, and the impoverishment of the data available to creole-acquiring children; and the amount of access to the superstrate in turn depends on the proportion of superstrate to substrate speakers in the community, the period between the start of immigration and the point at which substrate speakers outnumber superstrate speakers, and the rate of population increase thereafter, which all contribute to the dilution of the model. On this account, the creole nearest to the bioprogram should be Saramaccan, which was spoken initially by a group of escaped slaves, who had previously had little contact with the superstrate and who, after their escape, had access to neither the superstrate nor to any common native language. Sranan, which has English as its superstrate but had this withdrawn soon afterwards when Surinam was ceded to the Dutch, will be a little further from the bioprogram; and Haitian Creole, with its French superstrate withdrawn rather later in its history, will be further away still. Furthest away from the bioprogram will be creoles like Réunion Creole, which existed as a pidgin for a long period before the substrate population outstripped the superstrate speakers, and perhaps Tok Pisin.

So far, we have been discussing the bioprogram without explicitly specifying its contents. In fact, Bickerton argues that the bioprogram contains at least a restricted Phrase-Structure Grammar and four semantic distinctions. It is therefore unlike Chomsky's Universal Grammar, which is predominantly syntactic in nature. However, Chomsky's and Bickerton's approaches are not necessarily entirely incompatible: it might be, for instance, that Chomsky's Language Acquisition Device, which helps the child build a grammar of his native language, has as one component the bioprogram, which specifies the minimal structure required.

In terms of syntax, Bickerton (1984: 178) argues that 'creole similarities

stem from a single substantive grammar consisting of a very restricted set of categories and processes, which ... constitute part, or all, of the human species-specific capacity for syntax'. Bickerton suggests the rudimentary syntax in (10). COMP is a complementiser, which is taken to be empty but may be filled optionally by question words; INFL is the home of auxiliaries; Det. is shorthand for Determiner, the class including articles; items in brackets are optional; and the commas in each rewrite rule indicate that the order is not fixed, so that either Adjective Noun or Noun Adjective might be grammatical.

- (10) $S \rightarrow \text{COMP}, S$
 $S \rightarrow N^3, \text{INFL}, V^3$
 $N^3 \rightarrow \left\{ \begin{array}{l} S \\ (\text{Det.}), N^2 \end{array} \right\}$
 $N^2 \rightarrow (\text{Numeral}), N^1$
 $N^1 \rightarrow (\text{Adj.}), N$
 $V^3 \rightarrow V^2, (S)$
 $V^2 \rightarrow V^1, (N^3)$
 $V^1 \rightarrow V, (N^3)$

In addition to this set of Phrase-Structure Rules, Bickerton suggests a single movement rule, which will allow N^3 or V to be moved to a vacant site; the vacant site really has to be COMP, and this strategy is used in fronting constituents for emphasis, a technique extremely common in creoles.

This simple syntax is unable to generate nonfinite clauses (such as English *To see Bill is impossible*), or Verb Phrases containing further Verb Phrases (like *They all persuaded Mary to leave*). However, it is very nearly able to generate the structures of Saramaccan, which lacks such complex constructions and which is reputedly the creole closest to the bioprogram. Two modifications will be required to generate Saramaccan adequately. First, the commas must be removed from the rules in (10), so that the orders SVO and Determiner - Numeral - Adjective - Noun become obligatory. Second, the bioprogram syntax does not mention prepositions, and Saramaccan has two, *a* and *ku*, so that the syntax slightly under-generates. However, as Bickerton (1984: 179) notes, the small number of prepositions characteristically found in creoles may explain the need for serial verbs in constructions like that in (11), where non-creole languages would tend to use a preposition. With these minor alterations, the suggested bioprogram syntax can generate Saramaccan; and it certainly captures the core constructions which seem to be found in other creoles.

- (11) Saramaccan
dée o-tei faka tjoko unu
 'they MODAL-take knife stab you (pl.)'
 (Bickerton 1984: 179)

Although this syntactic component is rudimentary, it is rather specific, and this may be problematic, since not all linguists would be happy with the idea that a set of rules of this sort could be innately specified. Bickerton's more recent work may be more promising in this respect; he has proposed that, within the principles and parameters model (see Chapters 5 and 6 above), the syntactic component of the bioprogram might be interpreted as providing a default setting for each parameter. This would be overridden only if the child heard conflicting data in the linguistic input.

Bickerton (1981, 1984) claims that the bioprogram also includes the four semantic distinctions listed in (12).

- (12)
1. Specific / non-specific
 - specific noun phrases marked by determiner; non-specific unmarked.
 2. Tense - [$\pm$ anterior]
 - topic time unmarked; times prior to topic time marked by a preverbal particle.
 3. Mood - [$\pm$ irrealis]
 - realised events unmarked; unrealised events marked by a preverbal particle.
 4. Aspect - [$\pm$ punctual]
 - completed or one-off actions unmarked; habitual or progressive actions marked by a preverbal particle.

As we have already seen, creoles consistently tend to have such particles, when they co-occur, in the order TMA; and combinations of particles also tend to have consistent meanings across creoles. In terms of the syntax in (10), the three TMA particles will be generated as part of INFL.

In short, then, Bickerton proposes that all human beings have as part of their genetic inheritance a linguistic bioprogram, which contains basic syntactic and semantic information. The constructions and distinctions specified in the bioprogram are assumed to be those which developed first during human evolution of language (although here we are reduced, as Bickerton himself admits, to speculation). Furthermore, they will tend to surface during child language acquisition. However, children learning an established language will encounter novel constructions, built up during centuries of linguistic change, which may contradict bioprogram instructions; and they will tend to select the forms and distinctions they hear in the data around them. On the other hand, children having a pidgin as their primary linguistic input will have no access to features contradicting the bioprogram; on the contrary, they will hear no evidence at all for many syntactic and semantic features, and will have to rely on the bioprogram to help them build their grammar. 'In other words, the human species comes equipped ... with the capacity to reconstitute language itself should the

normal generation-to-generation transmission of input data be interrupted or distorted by extralinguistic forces' (Muysken and Bickerton 1988: 282). Bioprogram features will therefore appear consistently in creoles, with most occurring when the preceding pidgin has been most impoverished and access to superstrate and substrate languages has been least available. Creoles consequently become central to linguistics, as the only reliable source of information on features which are crucial in linguistic evolution and acquisition.

It is scarcely surprising that a proposal as radical as Bickerton's should have become controversial, particularly since we cannot hope, at least at present, for conclusive evidence of the existence of the bioprogram, which is unobservable except by its alleged results. This gives the bioprogram much the same status as Chomsky's equally unobservable language faculty or Language Acquisition Device, and opens it to much the same set of criticisms: notably, Romaine (1988) accuses Bickerton of circularity in his claim that the TMA distinctions found in creoles are basic and part of the bioprogram, when the only evidence is their occurrence in creoles. Direct evidence on linguistic evolution seems unlikely to be forthcoming, and we are consequently stuck with what we can observe in current or attested systems, making such protests of circularity almost inevitable. However, one promising line of research might be to consider other areas of language acquisition where, as in creolisation, the normal method of transmission is impaired or interrupted. Romaine (1988: 297ff) reports on some work on the acquisition of sign language, where again there is weak adult-to-child transmission and the child's sign language is therefore an emerging system based on relatively little evidence. In fact, sign language does seem to share many characteristics with creoles; for instance, both lack copulas and passives, use adjectives as verbs, and have a distinctive pattern of article use. Such support for the notion of the bioprogram from a non-creole source might defuse accusations of circularity.

The main opponent of Bickerton's ideas, however, is Aitchison, who begins her review of his 1981 book with a quotation from the philosopher Whitehead, who asserts that 'It is more important that a proposition be interesting than that it be true' (Aitchison 1983: 83). We have already mentioned Aitchison's criticism of Bickerton's restricted definition of creoles and his use of data predominantly from Hawaiian and Guyanese Creole, although these points were partly answered above. Aitchison also criticises Bickerton's assertion that novel constructions found in Hawaiian Creole were 'invented' by children with help from the bioprogram. She holds that the changes involved, such as the reanalysis of main verbs as auxiliaries, are simply natural tendencies which could be found in any language, and which do not require special explanation. She also argues

that, if Bickerton's data were more extensive, it might show traces of these allegedly new constructions in Hawaiian Pidgin. The second point is essentially uncontested, since the variety really no longer exists; however, the answer to the former criticism might be that the bioprogram is implicated in linguistic change regardless of the language concerned – if it exists, it is, after all, common to all of us.

Aitchison's main criticism concerns Bickerton's evidence from child language acquisition. She argues that he makes use of rather limited experimental data, and does not take sufficient account of variation in child language or in child language theory, suggesting that 'Bickerton's pronouncements on child language show both naivety and wishful thinking. His standard mode of argument is to take one or two well-known papers on a topic, and to assume that these represent the consensus of opinion in the field. If necessary, he reinterprets them to suit his purpose' (Aitchison 1983: 91). It is true that Bickerton is not an expert on language acquisition, and nor does he claim to be one; on the other hand, if we restricted inter-disciplinary work to those with acknowledged expertise in all the fields being investigated, there would be precious little of it around. Bickerton may not be fully up-to-date with acquisition studies, or fully aware of the controversies in the area; but this does not provide proof against the bioprogram. If anything, it provides an opportunity for someone who *does* have expertise in language acquisition to impartially re-examine Bickerton's claims. Finally, Bickerton may have recast some of the observations he found on child language in terms of the bioprogram, but it is dubious to suggest that he therefore reinterpreted results to suit himself: after all, he is hardly likely to have found mention of bioprogram phenomena or links with creoles in child language studies carried out before he had suggested the existence of a bioprogram.

Aitchison herself regards appeals to innateness as a last resort, arguing that 'an obsession with innateness can lead to a neglect of other possible reasons behind language universals' (Aitchison 1987: 14). She favours a commonsense approach, based on the idea that 'one fairly obvious observation about human life is that, for most problems, there are usually a finite number of possible solutions. And quite often, the same solution will be rediscovered by generation after generation' (Aitchison 1987: 17). Consequently, similarities among creoles may simply result from the fact that pidgins, as extremely restricted linguistic systems, have very few options for expansion; in extending their resources, they are therefore likely to select similar strategies. For instance, Aitchison notes that many creoles are SVO, but with leftward movement rules and a morphological marker distinguishing transitive from intransitive verbs. Tok Pisin, in its pidgin stage, was SVO; to focus a noun phrase, its speakers have various

restricted options, such as changing the word order, repeating the constituent, coining some special particle, or using intonation. Both word order change and repetition will alter the fixed SVO order.

In creole Tok Pisin, the solution is to repeat the focussed noun phrase, but in a reduced form, as a pronoun, giving the order NP Pronoun V NP, or SSVVO. This has subsequently been generalised to the Object Noun Phrase, which can also be fronted; and the repeated object pronoun has resulted in an *-im* suffix on transitive verbs. Finally, fronting has now been extended to adverbs, pronouns, *wh*-words, and other forms.

Aitchison argues that all these leftward movements of constituents produce the same sort of surface pattern, and that they all reinforce one another, leading to a proliferation of constructions of the same type. We therefore see a type of linguistic conspiracy, whereby one process produces a certain surface structure, which becomes the target for subsequent processes. A language in this situation is involved in a 'snowball' effect, where rules back one another up and push the language in a certain direction; and in the case of pidgins, where the starting point is so restricted, the options available will be few, so that 'different pidgins will be caught in more or less the same snowball, giving the impression that there is some preordained blueprint. In fact, the phenomenon may represent the rediscovery of the same type of rules, given certain basic parsing and memory abilities' (Aitchison 1987:26).

Aitchison (1989a) also uses this notion of preferred pathways to argue against Bickerton's notion that primitive, general categories tend to be generated in creole development, then subsequently differentiated. She considers the TMA system of Tok Pisin, and attempts to show why certain options were chosen and others rejected during the development of this system. Her data are from six young women, aged seventeen–twenty, who are first-generation creole speakers. These informants used *bin* as the particle marking pastness or [+ anterior], and *pinis* to mean 'after'. However, Aitchison claims that *bin* and *pinis* were formerly both possible markers of pastness, which have diverged and specialised in their meaning. Once *bin* was accepted as the pastness marker, it was used more frequently and redundantly, and located just before the main verb. Finally, it began to be combined with other particles, initially by one young, fluent, fast speaker who claimed not to know her parents' languages well. The habituality particle *sa* and the irrealis / future marker *bai* developed in a similar, gradual way. In all cases, the apparently neat creole pattern is the result of 'the gradual simplification of an originally highly variable and fairly messy situation' (Aitchison 1989a: 168). The eventual system is favoured by various principles: these include generality (the particles which 'win' are those which can be used most frequently); overlapping

(each particle has a number of meanings); and elimination of pointless variety.

However, this account is not entirely explanatory since we do not know why these particular principles should exert such a strong influence on linguistic development. Aitchison (1987, 1989a) talks hopefully of 'general cognitive factors' and of eventually linking these principles with properties of the human mind. It is tempting to say that this brings her closer to Bickerton than she might like to be; after all, a bioprogram could well turn out to be principle-based, and if we need a name for some undefined something in the human mind which helps us in certain circumstances to select certain linguistic options, we might as well call it the bioprogram as anything else – so long as we do not think that having a name for it means no more investigation is needed.

The validity of Bickerton's Bioprogram Hypothesis must be left, for the moment, as an open question. However, it is worth stressing that the bioprogram is not intrinsically incompatible with other explanations of creole features, such as claims of substrate or superstrate influence – after all, the bioprogram may tell the child to make some semantic distinction or create some syntactic structure, but the child then needs to find linguistic material to clothe it in. Arguably, then, substratist, superstratist and universalist explanations are all compatible with one another. This sort of all-embracing attitude is symptomatic of a recent trend in creole studies, which may represent the next step forward. For instance, Thomason and Kaufman (1988: 164–65) argue that adults may well contribute something to an emerging creole as well as their children, in that linguistic material in the creole frequently comes from the pidgin or the adults' native languages. When structures in different substrates coincide, these will be especially likely to be introduced into the creole; the bioprogram is here seen as a last-resort explanation, to be invoked when the relevant substrate structures conflict or no evidence for a particular structure is available.

We should be careful, however, not to return by this route to a radically substratist approach. Of course, substrate influence can sometimes be observed in grammar and vocabulary (such as Tok Pisin *kaikai* 'food' which has no superstrate source). However, not all creole grammar is substrate-based; for instance, Crowley (1991) notes that Bislama, the creole language of Vanuatu and a relative of Tok Pisin, has certain structures which can be ascribed to various indigenous languages, but cautions against an over-enthusiastic interpretation of this information, arguing that 'substratum has very clearly not been the *only* factor involved in the evolution of Bislama. Modern Melanesian Pidgin is certainly not relexified Tolai, Kwaio, or Nakanamanga, by any stretch of the imagination' (Crowley 1991: 387).

It is also crucial that superstrate influence should be properly taken into account: but it cannot be demonstrated if we look only at the modern standard version of the superstrate language concerned – modern Standard British English, for instance, with an RP accent. Instead, we must try to ascertain when English speakers, for instance, arrived in the Caribbean or in New Guinea, which areas they came from, and therefore what dialects they are likely to have spoken. A recent study of this kind (Singh 1991) demonstrates the influence of Early Modern English dialects on Trinidadian Creole. For instance, the universal relative form in the creole is *that*, as shown in (13), just as it was in sixteenth and seventeenth century Scots varieties, and still is in modern non-standard Scots dialects.

(13) Trinidadian Creole

The man in the car that get bounce, he foot get break.
I know the girl that did get shoot.

(Singh 1991: 22)

Early Modern English also had an uninflected genitive form, lacking final [s], for words ending in [s] or before words starting in [s], as shown in (14). This was extended into certain formations, such as *ladybird* (from *Our Lady's bird*) and *lady chapel* (from *Our Lady's chapel*). This tendency remains in the creole (see (15)).

(14) *for conscience sake*
for God sake

(15) Trinidadian Creole

Christmas Day is Our Lord day.
The doctor son does come here plenty.

(Singh 1991: 22)

Finally, Trinidadian Creole uses *does be* as a habitual marker, as shown in (16); this construction is reminiscent of structures like those in (17), which are found in modern Irish dialects, and (18), which were current in South Western England until at least 1898. Constructions like those in (19) are still very common in the varieties of English spoken in the South Welsh valleys (Mari Jones, personal communication).

(16) Trinidadian Creole

I does be there everyday.
He does be playing the fool all the time.
He does work on the estate.

(Singh 1991: 18)

(17) Irish dialects

He does be late for dinner sometimes.
They does be fighting among other.

(18) *She do be strict with us gals. (Oxfordshire)*
The childer do be laffen at me. (Cornwall)
 (from Rickford 1986)

- (19) *She do be quite good at that.*
He do be less talkative since his wife died.

The *does be* habitual may have spread from Irish English into the creole, since many Irish indentured workers migrated from Barbados to Trinidad in the nineteenth century. However, the Irish at this period may well have spoken Gaelic rather than English, and may have learned English in the Caribbean, introducing features such as *does be* by analogy with similar constructions in their native Gaelic. This habitual construction may also have been reinforced by parallel forms in South Western English varieties, which were spoken by many of the Trinidadian overseers and land-owners. The Irish, as indentured servants, would have had contact with these superstrate speakers, and perhaps passed on the *does be* construction to the slaves, who were predominantly speakers of West African languages, many of which also have habitual markers. Singh (1991) is at pains to point out that the impetus for introducing a habitual marker into the creole at all may have come from the bioprogram, or some universal demand to have this distinction marked; but that this universal motivation for the category does not rule out the possibility that its expression comes from the superstrate (which would be English for the Irish, and Irish English for the Africans).

This view of mutual reinforcement is probably best represented by Crowley (1991: 387–8), who argues that 'the more potential sources a form or a construction can be construed as having, the more chance there is that this form or construction will be incorporated as part of the final product' – that is, the creole. For instance, in Bislama, a form *se* is being introduced as a complementiser and a copula, in examples like those in (20), and Crowley traces this form to both English *say* and French *c'est*, both languages which have been important in the formation of Bislama.

- (20) Bislama
 a. *Namba blong yu se 3093?*
 'Is your number 3093?'
 b. *Oli stanemap wan komiti se disasta komiti.*
 'A committee was formed as a disaster committee'
 (Crowley 1991: 398)

Similarly, *be*, from colloquial French *mais > bais*, is becoming a subordinator and a copula in Bislama. However, some indigenous languages of Vanuatu have a very similar form as a copula, including North Ambrym *be* and Paama *ve*, leading here to a good case for mutually supporting superstrate and substrate influence.

Crowley's conclusion is that the bioprogram alone may be replaced by 'a consensus that what is most likely to 'survive' in a radically altered contact language is a set of features combining substratum and super-

stratum features, as well as other features that develop independently, for a variety of reasons, including universal pressures' (Crowley 1991: 385). Such harmonious consensus seems a distant hope at the present stage of creolistics, but may represent a goal towards which historical and linguistic data on creoles and creolisation, and attempting to include pidgins and creoles in our theories of language change. However, we must face the fact that no single, monolithic theory will easily cope with the complex phenomenon of creolisation; instead (Singh 1991: 4), 'the ultimate theory of creolisation ... may have to be as much of a creole as the language it describes'.