EXERCISES

1. Papua New Guineans using English as a second language occasionally make errors such as the following in their speech:

Standard English	Papua New Guinea English
hibiscus	*hibiscuit*
pandanus	*panda nuts*
lingua franca	*lingo franco*

   (Another example of the same thing, but involving only a spelling change rather than a change in pronunciation, is the change from *surname* to *sir name*.) What factor is responsible for these unpredictable phonetic changes in the English of those people who might say these things?

2. People for whom English is their first language normally pronounce the word 'gibberish' as /ʤɪbərɪʃ/ and 'gesture' as /ʤestʃə/. What factors might be responsible for the very common pronunciation of these two words by Papua New Guineans as /gɪbərɪʃ/ and /gestʃə/ respectively?

3. The English word *ambassador*, when copied into Tok Pisin, would normally have become /embesada/. Some speakers actually say /embesirep/ instead. Can you say why?

FURTHER READING

1. John Samuel Kenyon 'Spelling Pronunciation', in Anderson and Stageberg (eds) *Introductory Readings in Language*, pp. 248–54.
2. Eugene Nida 'Analogical Change', in Anderson and Stageberg (eds) *Introductory Readings in Language*, pp. 86–92.
3. Leonard Bloomfield *Language*, Chapter 23 'Analogic Change', pp. 404–24.
4. Theodora Bynon *Historical Linguistics*, Chapter 4 'The Neogrammarian Postulates and Dialect Geography', pp. 173–97.
5. Otto Jespersen *Language: Its Nature, Development and Origin*, Chapters 1 to 4 'History of Linguistic Science', pp. 19–102.
6. Hans Henrich Hock *Principles of Historical Linguistics*, Chapter 15 'Linguistic Contact: Dialectology', pp. 426–71.

CHAPTER TWELVE

LANGUAGE CONTACT

There are many bilingual and multilingual societies in the world. Canada is officially bilingual, with both English and French functioning at the national level. Switzerland is officially quadrilingual, functioning in German, French, Italian, and Romansh. Other nations are more complex in their linguistic make-up, such as the former Soviet Union, India, or Indonesia, where there are hundreds of separate languages spoken. The most complex nations in the world in terms of their linguistic composition are the small Melanesian countries. Papua New Guinea boasts over 800 distinct languages, spoken by a population slightly larger than that of New Zealand (i.e. about three and a half million people). Nearby Vanuatu has only a hundred or so languages, but its population is much smaller, with the total number of people scarcely reaching 140,000!

However, just because a society is multilingual or bilingual does not necessarily mean that there is a great deal of language contact, as we can speak of language contact only when there are significant numbers of individual members of the society who are bilingual or multilingual. While Belgium recognises both Flemish and French as official languages, there is relatively little language contact as 85 per cent of the population is monolingual in either Flemish or French, and does not speak the language of the other group.

For genuine language contact to occur, there must be significant numbers of people who operate in two (or more) languages. But in world terms, monolingualism is relatively rare. This may come as a surprise to some people, especially to people from Western industrialised societies. There is a standard joke among migrants to Australia that goes like this:

Q. What is a person who speaks three languages?
A. Trilingual.
Q. What is a person who speaks two languages?
A. Bilingual.

Q. *What is a person who speaks one language?*
A. *Australian.*

People from Vanuatu generally speak two, three, four, and sometimes even more languages fluently, and they often find it incomprehensible that the average Anglo-Celtic Australian or Pākehā New Zealander speaks only English. In this chapter, I will explore some of the linguistic consequences of language contact in societies such as those of Melanesia and elsewhere where multilingualism is a fact of everyday life.

Up to now in this volume, I have frequently referred in passing to the results of language contact, though this has almost always involved discussion of language change that has involved lexical change as a result of new words being copied into the lexicon from other languages. In this chapter, however, I will be looking not so much at how languages can influence each other lexically, but at how the whole phonological or grammatical system of a language can be influenced by that of another language.

12.1 CONVERGENCE

When you hear somebody speaking and their first language is not English, it is generally very easy to recognise that he or she is not a native speaker of English. There are usually a number of tell-tale signs that indicate not only that the person is not a native speaker of English, but also what that person's first language actually is. By this I mean that it is often possible to recognise from the way somebody speaks English whether he or she is a speaker of French, German, Italian, Chinese, Japanese, Russian, or whatever other language. Typically, people carry over features from their first language into another language that they learn later in life, and we hear this at the phonological level as a *foreign accent*, and at the grammatical level as *learner errors*. However, it is not just among people who are learning a second language that one language can influence another. Even among people who can be considered to be fluently bilingual — that is, people who have been speaking two languages regularly and fluently from early childhood — we find that features of one language can cross over into the way that person uses the other language. The influence of one of the linguistic systems of an individual on the other linguistic system of that individual is referred to in general as *interference*.

Interference can occur in the phonological system of a language, in its semantics, or in its grammar. Phonological interference simply means the carrying over of the phonological features of one language into the other as an accent of some kind. This might involve the incorrect transfer of the distribution of the allophones of a particular phoneme into the other language in such a way that the phonological system of that language is violated. For example, the English of a Japanese-English bilingual who says *rots of ruck* instead of *lots of luck* has been influenced by interference from the fact that in Japanese there is no phonemic contrast between /l/ and /r/ as there is in English.

To illustrate grammatical interference, examine the sentence below which contains a relative clause. Sentences such as these are often produced by school children in Vanuatu who are learning English:

This is the book which I read it yesterday.

To a native speaker of English, this sentence contains an obvious error, namely the use of the pronoun *it* after the verb *read* in the relative clause. English grammar contains a general rule which deletes any reference to noun phrases in a relative clause that have already been mentioned in the sentence. Since the book has already been mentioned, there is no need — according to the rules of English grammar — to refer to it again, which is why we just say this:

This is the book which I read yesterday.

However, relative clauses in the first languages of children in Vanuatu schools typically require that the noun phrase be mentioned again in sentences such as these by means of some kind of a pronominal copy after the verb. To illustrate this kind of construction, I will give an example from one of these languages, i.e. Paamese:

Tu:s keke na-les-i naŋaneh keiek:
book which I-read-it yesterday this
'This is the book which I read yesterday.'

In the example above, you can see that in Paamese it is necessary to include an object pronoun referring to the book after the verb (in the form of the pronominal suffix /-i/). A speaker of Paamese who fails to delete the pronoun in sentences such as these in English is engaging in grammatical interference from his or her first language.

Semantic interference can also be referred to as *semantic copying*, as *loan translation*, or as *calquing*. A *calque* (or a *semantic copy* or a *loan translation*) is when we do not copy a lexical item as such from one language into another, but when just the meanings are transferred from one language to the other, while at the same time we use the corresponding forms of the original language. The term *hot dog* as a name for a kind of fast food originated in English, but in French in Québec the same thing is referred to as a *chien chaud. Chien*, of course, is the French word for 'dog' and *chaud* is the word for 'hot'. Thus, we can say that *chien chaud* is a calque based on English 'hot dog'.

If somebody from Vanuatu were to say the following, this would also be a calque on that person's first language:

He sat there and just listened to his kava.

Since kava is a drink, in English it is not something that we can 'listen' to. However, the verb that means 'listen to' in Vanuatu languages is also used to refer to the quiet contemplation of the rather delightful effect that a few servings of kava has on the system, and people from Vanuatu speaking English sometimes form a calque on the basis of their first language to produce sentences such as the one that I have just given.

As I said at the beginning of this chapter, I do not plan to enter into a great deal of discussion about lexical interference (or lexical copying, or borrowing) between languages, as this has been covered elsewhere in this volume (most notably in Sections 1.3, 7.4, and 8.2). However, I would like to mention at this point that, while the introduction of lexical items from one language into another does not necessarily affect the structure of the language that is receiving the new material, it is also possible that introduced lexical items can affect the phonology and the grammatical system of a language. In Chapter 4, I showed how words originating from English which have been introduced into the Motu language of Papua New Guinea now show signs of disrupting the previous complementary distribution between [t] and [s] and are in fact causing a phonemic split to take place in the modern language. It is also possible for completely new sounds to be introduced into a language via words copied from other languages. Bahasa Indonesia originally had no voiced velar fricative at all, either as a separate phoneme or as an allophone of some other phoneme. However, with the introduction of large numbers of words of Arabic origin into the everyday vocabulary of the language, we can now show evidence of phonemic contrast between /g/ and /ɣ/ in this language.

It is also possible for words from other languages to introduce new grammatical patterns into a language. To a very minor extent this has happened in English, as some words of foreign origin have kept their original plurals, e.g.

	Singular	Plural
Greek	phenomenon	phenomena
	criterion	criteria
Latin	datum	data
	index	indices
	cactus	cacti
Italian	lingua franca	lingue franche
Hebrew	kibbutz	kibbutzim

It is very rare for bound morphemes to be incorporated into the general grammar of another language, so it is unlikely that any of these patterns for

---

the formation of plurals will spread beyond the words that originally introduced the patterns in the first place. In fact, most nouns of foreign origin are quickly adapted to the rules of the language anyway. So, the plural of *atlas* in English is now *atlases*, and not *atlantes* as we might have expected on the basis of the morphological behaviour of the word in its original Greek.

While the example that I just gave involved the influence of one language on another in the area of morphology, it is possible for lexical copying to influence higher levels of grammar as well. In Paamese, all verbs are required to carry prefixes which indicate the pronominal category of the subject, as well as a variety of tense and mood categories. So, from the root /lob/ 'run' (which cannot occur without any prefixes), we can derive the following inflected forms (among many others):

naloh	'I ran'
niloh	'I will run'
koloh	'you ran'
kiloh	'you will run'
aloh	'they ran'

However, verbs such as /staːt/ 'start', /raːu/ 'argue' (from 'row') and /riːt/ 'read' that are borrowed from English are not permitted to carry any prefixes, and so a new grammatical construction evolved just to handle these new forms. There is a verb of the form /vi/ in Paamese which functions as a copula in sentences such as the following:

Inau	na-vi:	meahos.
I	I-am	man
'I am a man.'		

The only kinds of words that could originally follow the verb /vi/ in Paamese were nouns in equational sentences such as the above. However, in the modern language, verbs introduced from English have also been incorporated into the same grammatical construction, and the prefixes which would ordinarily have been attached directly to the verb root are now attached instead after the preceding copula, as in the following examples:

navi: staːt	'I started'
kovi: raːu	'you argued'
avi: riːt	'they read'

You should note that in these examples, while a new pattern in Paamese grammar has emerged as a result of new words coming into the language, this pattern has not come from English. It is in fact a brand new pattern that has emerged out of the existing structural resources of Paamese as a way of coping with introduced vocabulary that speakers somehow felt did not 'fit' the language properly.

It is absolutely clear that languages can influence each other lexically (and, through lexical introductions, also to some extent grammatically), and it is just as clear that a speaker's first language can influence the way he or she speaks another language at all levels of language (i.e. in the phonology, the grammar, and the semantic system). However, there has been considerable debate in recent years on the question of whether one language as a whole can really influence another language as a whole (as against individual speakers of the language).

There is a significant body of literature on the subject of linguistic *diffusion* and *convergence*, which is based on the assumption that languages can and do influence each other. The term *diffusion* is used to refer to the spread of a particular linguistic feature from one language to another (or, indeed, to several other languages). One example of diffusion that is often referred to is the spread of the uvular [ʁ] in the languages of Europe. This is the kind of sound that you are taught to produce when you are learning to pronounce French words such as *rare* 'rare', *rire* 'laugh', and so on. Originally, these words were pronounced in French with an alveolar pronunciation, just as we find in Italian today. However, it appears that in the 1600s, speakers of French in Paris began to pronounce their *r* sounds as uvulars rather than as alveolar sounds. This change then spread to other language areas in Europe, and people in Copenhagen (in Denmark) were apparently doing the same thing in Danish by about 1780. The uvular pronunciation of *r* is now common in French, German, and Danish, and it is also used in some areas where Dutch, Norwegian, and Swedish are spoken. The following map suggests that the spread of the uvular *r* has hopped from city to city, and that it has then radiated out from the cities to the surrounding rural areas.

Albanian, Bulgarian, Romanian, and Greek, all spoken in the Balkans area of Europe, are only fairly distantly related to each other within the Indo-European language family. However, these languages share certain grammatical features that do not appear to be derived from their respective protolanguages. One of these features is the use of a special complex sentence construction instead of the infinitive construction to express meanings such as 'I want to leave'. All of these languages express this meaning instead by a construction that translates literally as something like 'I want that I should leave'. The following examples show that while the words that are used to express this meaning are quite different in these four distantly related languages, the grammatical construction is basically the same:

Albanian	Due	te	shkue.
Bulgarian	Iskam	da	otida.
Romanian	Veau	sa	plec.
Greek	Thelo	na	pao.
	I-want	that	I-should-leave
	'I want to leave.'		

This similarity between these four languages is not something that we would have predicted from Proto Indo-European, and the suggestion is that these four languages have *converged*, or come to resemble each other structurally as a result of a long period of linguistic contact and mutual interference.

Languages which have come to resemble each other as a result of linguistic convergence in this way are said to belong to *linguistic areas*, and the features that have diffused among the languages that belong to such an area are called *areal features*. Thus, in the case of the languages that I have just described, we could refer to the Balkans as a linguistic area (or sometimes as a *Sprachbund*, to use a word of German origin), and the special construction that I illustrated above would be called an *areal feature*. Linguistic areas can be recognised in a number of different parts of the world. Chinese, Thai, and Vietnamese all belong to a linguistic area, as all have developed phonemic tone distinctions. The Indo-European and the Dravidian languages of the Indian subcontinent have developed widespread retroflex consonants, which set them apart as a linguistic area, and a number of Bantu languages and Kalahari languages in southern Africa also constitute a linguistic area which is characterised by the presence of rather unusual click consonants.

A linguistic area can be characterised by shared phonological features, as well as grammatical features, as illustrated by the example given above of the construction in the Balkans linguistic area. In Section 7.2, I referred to the possibility that **SOV** word order has diffused from some non-Austronesian languages in Central Province in Papua New Guinea to the Austronesian languages, resulting in a linguistic area characterised by **SOV** syntax. Some scholars who have described both the Austronesian and the non-Austronesian languages of parts of the West New Britain province of Papua New Guinea

have argued that syntactic convergence among these languages has been even more thorough than this, involving quite a number of different syntactic constructions. For many sentences, it seems that speakers of a number of different Austronesian and non-Austronesian languages in this area map their own words onto grammatical constructions that are almost identical. In fact, the same constructions are also found in Tok Pisin, even though this language is lexically derived mostly from English:

Non-Austronesian

Anêm	Ezim	o-mên	da-kîn

Austronesian

Mouk	Eliep	max	na-nas
Aria	Bile	me	ne-nenes
Tourai	Bile	me	na-nes
Lamogai	Bile	me	ne-nes
Lusi	Vua	i-nama	na-sono
Kove	Vua	i-nama	na-sono
Kabana	Bua	i-nam	na-sono
Kilenge	Vua	i-mai	na-sono
Amara	Eilep	i-me	a-nas

Tok Pisin	Buai	i kam	mi kaikai
	betel nut	it-come	I-chew

'Hand me some betel nut to chew.'

The diffusion of grammatical features in this way has caused some linguists to question further the validity and basic assumptions of the whole comparative method. Some languages appear to have undergone so much diffusion in the lexicon and the grammar that it can be difficult to decide which protolanguage they are derived from. According to the comparative method as I have described it in this volume, it is possible for a language to be derived from only a single protolanguage, yet some linguists have found it necessary to speak of *mixed languages*, which seem to derive from two different protolanguages at once.

Linguists tend to be thankful that such cases appear to be fairly rare. However, where such languages exist, they often produce much heated discussion as different scholars come down in support of undeniable membership in one language family or another, and yet others argue that such either/or conclusions do not accurately reflect the genuinely indeterminate nature of the language. One example of such a situation involves the languages of the Reef-Santa Cruz islands in the Solomon Islands of Melanesia, where there has been debate as to whether these are basically Austronesian languages that have been heavily influenced by non-Austronesian languages, or whether

they are non-Austronesian languages that have been heavily influenced by Austronesian languages.

Despite the fact that areal studies of languages frequently refer to linguistic convergence, and scholars often speak of the 'borrowing' of features at all levels of language, there are some linguists who are reluctant to accept the possibility of syntactic copying between languages. While accepting the obvious fact that lexical copying occurs, as well as the possibility that individual words can bring certain morphological characteristics with them into another language, some linguists argue that grammatical patterns as such cannot be copied, or if they are, that this happens only in the rarest of circumstances. Facts which are often quoted as evidence of syntactic copying, these scholars argue, often turn out to have quite different explanations.

For instance, it is fairly frequently stated that Québec French is changing, not only lexically, but also syntactically, in the direction of the dominant English language, and this tendency is widely condemned by purist Québécois. In English it is possible to end a sentence with a preposition (despite the claims of the prescriptive grammarians among us), as in the following:

That's the girl I go out with.

French differs from English in that it is not possible to end a sentence with the corresponding preposition *avec* 'with', and in order to express the same meaning, the sentence must be organised differently, as indicated below:

C'est	la fille	avec qui	je sors.
that-is	the girl	with who	I go-out

'That's the girl with whom I go out.'

In the French that is spoken in Québec, however, sentences of the following type, which closely parallel the English construction, are frequently heard:

C'est	la fille	que je sors	avec.
that-is	the girl	that I go-out	with

'That's the girl I go out with.'

Despite the close structural similarity between the English and the French patterns in those examples, we cannot assume that, merely because there are structural similarities between the two languages, one is necessarily derived from the other. Historical research reveals that there is in fact written evidence of the stranding of *avec* without a following pronoun in French going back about 600 years (which was well before French and English came into contact in Québec!). The same pattern is apparently still preserved in some French dialects in France that have not been in contact with English, and even in some other Romance languages, which suggests that the pattern

goes back even further in time. Another point to consider is that, while we can strand any preposition in English without a following pronoun, this is possible in French only with the longer prepositions. With very short prepositions such as à 'to', this construction never occurs. So, note that the following is not possible in French:

*C'est la fille que j'ai parlé à.
that-is the girl that I-past speak to
'That's the girl I spoke to.'

12.2 LANGUAGE GENESIS — PIDGINS AND CREOLES

According to the model of language change that I have presented in this volume, every language is derived as a result of gradual change from a single language that was spoken in the past. However, there is one category of languages that appears to have evolved under rather special circumstances — the languages that are known as *pidgin* languages and *creole* languages. When speakers of several different languages come into contact in a situation where there is an urgent need to communicate and there is little social opportunity to learn whatever happens to be the dominant language, and where no other language predominates in terms of numbers of speakers, what often happens is that a pidgin language develops. The pidgin that forms has a vocabulary that derives largely from the dominant language, but the vocabulary is very much reduced in size. The grammar of a pidgin language is radically different from that of the dominant language, and typically involves much greater regularity than the grammar of the dominant language, as well as less redundancy. A pidgin language also tends to have only free morphemes with very few bound morphemes. In addition to these purely linguistic features, a pidgin language is generally used only as a second language by all of its speakers.

Pidgin languages have evolved frequently and in many different parts of the world when the contact circumstances have been ripe for their formation. When Melanesian labourers were taken by English-speaking Europeans from what are now Vanuatu, Solomon Islands, and Papua New Guinea in the nineteenth century to work on sugarcane plantations in Queensland and Samoa, the circumstances for the formation of a pidgin based on English vocabulary were ideal. There were speakers of large numbers of different languages working together under European overseers. Very rapidly a new pidgin language came into existence.

This language is still spoken in slightly different forms in Papua New Guinea (where it is known as Tok Pisin), Solomon Islands (where it is known as Pijin), and Vanuatu (where it is known as Bislama). Although between 80 and 90 per cent of the vocabulary is derived from English, there is also a sizeable proportion of words that come from a variety of different local languages. Some words of German origin have also found their way into Tok Pisin, while a significant number of words of French origin are found in Bislama. A fluent speaker of Melanesian Pidgin (which is how we can refer generically to these three dialects) cannot be understood by someone who speaks only English, and Melanesians who speak their variety of Pidgin cannot understand speakers of English unless they learn it in school. By all criteria, therefore, Melanesian Pidgin is a new and distinct language with its own phonology, grammar, and lexicon.

As an illustration of what a pidgin language is like, I will refer to Tok Pisin. As I have already indicated, the vocabulary of this language is largely of English origin, in this case about 80 per cent, though the words have been phonologically restructured to fit Melanesian sound systems, for example:

dok	'dog'
aus	'house'
rot	'road'
ren	'rain'
trausis	'trousers'

Of the remaining 20 per cent of the lexicon, most comes from the languages of the New Britain and New Ireland people who were the original labourers on the Samoan plantations. So, we find words such as the following:

kakaruk	'chicken'
kiau	'egg'
buai	'betel nut'
kunai	'long grass'
kulau	'drinking coconut'

The small number of remaining words in Tok Pisin do not come from English or from local languages, but from a variety of other sources. Such words include the following:

rausim	'take out'	From German *heraus* 'get out'.
beten	'pray'	From German *beten* 'pray'.
pater	'priest'	From Latin *pater* 'father'.
binatan	'insect'	From Malay *binatang* 'animal'.
pikinini	'child'	From Portuguese *pequenho* 'small'.
kanaka	'bumpkin'	From Hawaiian *kanaka* 'man'.
kaikai	'eat'	From Māori (or other Polynesian) *kai* 'eat'.

The vocabulary of Tok Pisin is also clearly 'reduced' with respect to that of English as well as that of Melanesian languages. This language lacks the vocabulary that we have in English to discuss many concepts in law, science, and technology, and it also lacks much of the vocabulary that is present in

Melanesian languages to name different parts of the natural environment, especially some of the rarer flora and fauna, as well as cultural practices.

Grammatically, if you compare Tok Pisin with English, you will find that Tok Pisin is much simpler in its structure, in that it is much more regular. For example, while English has many unpredictable past tense forms for verbs, Tok Pisin verbs are the same in all their forms. So, while in English we have to learn the past tense forms of the following verbs separately, verbs in Tok Pisin exist in only a single invariant form:

Present	Past
bring	brought
ring	rang
string	strung
ping	pinged

Differences in tense and aspect, which are sometimes marked in English by suffixes to the verb, are marked in Tok Pisin by independent grammatical words, for example:

Em   i   toktok.
(s)he predicate talk
'(S)he talks.'

Em   i   bin   toktok.
(s)he predicate past talk
'(S)he talked.'

Tok Pisin grammar also differs from that of English in that it has far less redundancy built into its grammatical system. For example, in English, plural marking is expressed in a variety of different ways in a sentence, often in more than one way at once. For instance, it can be marked in the following ways:

(i) by a separate form of the noun, i.e. *dog* vs. *dogs, child* vs. *children, man* vs. *men, woman* vs. *women.*

(ii) by a difference in the form of a preceding demonstrative, i.e. *this* vs. *these, that* vs. *those.*

(iii) by a separate form of the verb, i.e. *am* vs. *are, is* vs. *are, does* vs. *do.*

So, in the sentence below, the idea of plural is expressed in three separate places, as shown by the contrasting singular form:

Those women are singing.
This woman is singing.

In Tok Pisin, however, the idea of plural is expressed only once in the sentence, and even then it is optional. We can say the following to refer to one woman or to many women:

Dispela   meri   i   singsing i   stap.
this/these woman/women predicate sing predicate continuous
'This/these woman/women is/are singing.'

If you specifically want to mark the fact that there is more than one woman involved, you can use the plural marker *ol* at the front of the noun phrase, but you will note that none of the other words in the sentence are marked in any way:

Ol   dispela meri   i   singsing i   stap.
plural these   women   predicate sing   predicate continuous
'These women are singing.'

Pidgin languages can be formed in any situation where the contact circumstances are right. There are pidgin languages in which the lexicon is derived predominantly from Spanish, French, Portuguese, and Dutch in various parts of the world. It is not necessary that the lexicon of a pidgin should be derived only from European languages, as there are also cases where pidgins have been formed out of non-European languages. In the Pacific, for instance, we find Hiri Motu which is widely spoken in Papua today, and this language is based on the vocabulary of the vernacular Motu language of the Port Moresby area. When outside labourers were introduced into Fiji, the resultant pidgin was not based on the vocabulary of English, but that of Fijian.

I mentioned at the beginning of this section that pidgin and creole languages tend to avoid bound morphemes, but the Tok Pisin examples do not illustrate this very well because English is a language that has relatively few prefixes and suffixes, at least when compared with many other languages of the world. In order to illustrate this point, and also to illustrate what a pidgin that is derived from a non-Indo-European language looks like, I will now give some examples from Hiri Motu and compare these with the vernacular Motu from which it is lexically derived. Some of the differences between these two languages involve the following points:

(a) Objects to verbs in vernacular Motu are expressed as suffixes to the verb, and these have the following shapes:

	Singular	Plural
First	-gu	inclusive -da
		exclusive -mai
Second	-mu	-mui
Third	-(i)a	-dia

In pidgin Motu (or Hiri Motu), objects are expressed by full form pronouns that have the same form as the subject pronouns. The grammatical difference between subject and object is shown by the position of

the form in the sentence. The full form pronouns are the same in both vernacular and pidgin Motu, i.e.

	Singular	Plural
First	lau	inclusive ita
		exclusive ai
Second	oi	umui
Third	ia	idia

(b) Subjects to verbs are marked in vernacular Motu as prefixes to the verb, and the forms of these prefixes are as follows:

	Singular	Plural
First	na-	inclusive ta-
		exclusive a-
Second	o-	o-
Third	e-	e-

In pidgin Motu, subjects are expressed by placing the full pronoun in the subject position of the sentence and there is no further subject marking on the verb.

(c) To make a verb negative in vernacular Motu, there is a different set of subject markers from those that are used in the affirmative, as given above. The negative prefixes are as follows:

	Singular	Plural
First	asina-	inclusive asita-
		exclusive asia-
Second	to-	asio-
Third	se-	asie-

In pidgin Motu, negation is marked by placing the free form /lasi/ after the verb phrase. The word /lasi/ also occurs in vernacular Motu, where it is a word meaning 'no'.

The following examples are presented to show the difference between vernacular Motu and pidgin Motu. The two languages are not mutually intelligible, even though most of the words that occur in pidgin Motu are derived directly from roots that are used in vernacular Motu:

Vernacular Motu	Pidgin Motu
Ia e-ita-mu.	Oi ia itaia.
'(s)he (s)he-see-you	you (s)he see
'(S)he saw you.'	'(S)he saw you.'
Asi-na-rakatani-mu.	Oi lau rakatania lasi.
not-I-leave-you	you I leave not
'I didn't leave you.'	I didn't leave you.'

It is possible for a pidgin language eventually to replace the original vernaculars that in a sense 'spawned' it. This did not happen in the case of Melanesians who worked in Queensland and Samoa, as they generally just went to work overseas for three years and then returned home to their families who had stayed behind. However, if significant numbers of women are also brought on to the plantations and people are prevented by slavery from ever returning home, it is inevitable that the men and women from different language groups will marry and have children with whom there will be little option but to speak in the pidgin. This is what happened in the Caribbean among African slaves. While the multitude of African languages initially resulted in the slaves developing a pidgin based on English vocabulary, the following generations of slaves grew up speaking only the pidgin, and the original African languages ceased to be passed on at all. When a pidgin language replaces the original vernaculars that caused it to evolve in this way, we say that the pidgin has become a *creole*. Thus, to this day there are many creole languages in the Caribbean which evolved out of pidgins that were originally spoken on the slave plantations. In the South Pacific, pidgins have become creoles in some parts of northern Australia as the pidgins used by Aborigines have come to replace the original vernaculars.

Linguists have drawn a distinction in the past between pidgins and creoles because they have argued that there are structural differences between the two. Being only a contact language, a pidgin has generally been seen as a very basic sort of language indeed, with the smallest possible lexicon, as well as a very rudimentary grammar. However, once a pidgin becomes the mother tongue of a community, it is generally assumed that it undergoes rapid lexical and structural expansion in order to meet the normal needs of a community of native speakers.

The study of pidgin and creole languages has only relatively recently moved into mainstream linguistics. Pidgin languages have traditionally been regarded as poor imitations of 'real' languages, with no structure of their own. For instance, the governor of the colony of Papua in the early twentieth century called the Melanesian Pidgin of the time a 'most atrocious form of speech'. The typical view in the past was that a pidgin was nothing more than broken English, and many people still mistakenly refer to the language in this way, even though it has its own grammar, and is not mutually intelligible with English. The result of prejudices such as these has been that most linguists did not take pidgins and creoles as being worthy of serious study until the last few decades.

However, the tables have now turned, and pidgins and creoles are now seen by many as being central to an understanding of how languages change. So keen have some linguists been to see exactly how a pidgin originates that there was recently a proposal to produce a pidgin artificially and observe its formation. The idea was that a number of people who spoke quite different languages would be brought together in one place and 'fed' a number of basic words that they could use to communicate with other participants in the

experiment. The researchers would then remain in the background for a period of months to see what kinds of grammatical patterns emerged when these words were put together as an indication of how serious some linguists were in searching for answers on the question of the origin of pidgins and creoles.[1]

Pidgin and creole languages have aroused this kind of interest because linguists are keen to find out how these languages acquire their structures. You may have noticed that up to this point I have spoken about pidgins and creoles having a predominantly English (or French, or Spanish, or Motu) vocabulary, yet they are still mutually unintelligible with the languages from which their vocabularies are derived. This suggests that pidgins and creoles are structurally very different from their *lexifier* languages (i.e. the languages from which their vocabularies are derived), and this is a point that I think you will appreciate from the examples of Tok Pisin structure that I presented earlier in this section (as well as in earlier chapters of this volume).

Many linguists have been struck by the fact that pidgin and creole languages often show strong parallels in their structure with their *substrate* languages rather than their *superstrate* languages. The term *superstrate* (or *superordinate language*) is used to refer to the dominant language in the contact situations in which a pidgin or creole language develops. In the case of Tok Pisin, for example, English is clearly the superstrate language. The *substrate*, on the other hand, refers to the vernaculars of the people who actually develop a pidgin or creole. In the case of Tok Pisin, the substrate languages would be the various vernaculars of the New Britain and New Ireland labourers who were originally taken to work in Samoa and Queensland in the nineteenth century. While the grammar of Tok Pisin is clearly different from that of English, it seems that when we examine many of the points of difference between English and Tok Pisin, we can find structural parallels with the substrate languages. For instance, the form *i* that occurs in the examples above as a 'predicate marker' corresponds roughly, in shape and in function to a morpheme *i* that is found in Tolai (and many other of the substrate languages), for example:

Tolai
To   Pipira   i   vana.
article Pipira predicate go
'Pipira is going.'

Tok Pisin
Pipira i   go.
Pipira predicate go
'Pipira is going.'

[1] The proposers of this project intended to use Papua New Guineans who spoke only their vernaculars and to throw them together in a situation that they could not have been expected to see the relevance of or fully understand. Although the possible advances to scientific knowledge from this project were great, it was justifiably refused permission by authorities in Papua New Guinea for its exploitative and inhumane aspects.

The existence of two separate forms of the first person non-singular pronoun in Melanesian vernaculars is also paralleled in the structure of the Melanesian Pidgin pronoun system but not in that of English. In Tok Pisin, there are two separate pronouns corresponding to the single form 'we' in English. Firstly, there is *yumi* which means 'we' when you are including the person you are speaking to (i.e. the so-called inclusive pronoun). Secondly, there is the form *mipela* which means 'we' when you are excluding the person you are speaking to (i.e. the so-called exclusive pronoun). This distinction is widespread in the substrate languages for Melanesian Pidgin, but English grammar does not make the distinction (and sometimes English-speakers even find it hard to use the pronouns *yumi* and *mipela* correctly).

The existence of such structural parallels between pidgins and creoles and their substrate languages has led many scholars to argue that pidgins and creoles are mixed languages in the sense that they derive their lexicons from the superstrate, while their grammars come predominantly from the substrate. If this interpretation is correct, then pidgin and creole languages differ dramatically in their genesis from other languages as they have multiple ancestors rather than a single ancestor. According to such a view, it would be impossible to classify Tok Pisin either as an Austronesian language or as an Indo-European language as it contains significant elements from both language families. (You will remember that I referred to the possible existence of mixed languages also in Section 12.1.)

You will also remember from the preceding section that some scholars today do not accept that languages can easily influence each other structurally. Linguists who hold this point of view sometimes extend this even to pidgin and creole languages, arguing that the existence of parallels in structure between pidgins and creoles and their substrate languages is not necessarily evidence that a pidgin has been structurally influenced by the substrate and they argue that other factors may also be involved. For instance, it could be equally argued that the 'predicate marker' *i* that I described earlier in Tok Pisin does not derive from the substrate at all, but that it derives from the English pronoun 'he' which may have been repeated after the subject noun phrase. Thus, *Pipira i go* 'Pipira is going' is not necessarily derived from the Tolai construction, but from a pre-pidgin 'broken English' sentence of the form *Pipira he goes* (and sentences of this type do sometimes occur when people are learning English as a second language).

Scholars who deny any significant impact of substrate structural patterns in the development of a pidgin or a creole language tend to point instead to what they see as the remarkable structural similarities between pidgin and creole languages that have radically different histories and even different lexical aspects.

source languages. For instance, if you compare the grammatical structure of a simple intransitive sentence in Tok Pisin with the corresponding sentence in Haitian Creole spoken in the Caribbean (which has French as its lexifier language), you find that there are remarkable similarities between the two. Compare the following two sentences in these two languages:

Tok Pisin   Em   no   bin   save.
Haitian Creole   Li   pa   te   konè.
                 (s)he not   past   know
                 '(S)he did not know.'

You can see that, although the words in these two languages are quite different in their shape, reflecting their different origins, the order in which the words occur is exactly the same.

This becomes even more significant if we compare the corresponding sentences in their respective lexifier languages. In English, the structure of the sentence (S)he did not know involves the following facts:

(a) The first element is the subject pronoun.
(b) The second element is the verb do which is put there to carry the tense marking. In this case, the tense is past, so the verb appears in the form did.
(c) The third element is the negative marker not (which optionally appears reduced in form to the suffix -n't).
(d) The fourth element is the verb know which occurs in the infinitive form, i.e. it does not take any suffixes for tense as this is already in the form did.

The corresponding French phrase Il/elle ne connaissait pas, however, has the following quite different structure:

(a) The first element is again a subject pronoun, of the form il 'he' or elle 'she'.
(b) The second element is the form ne, which marks the verb as being negative.
(c) The third element is the verb root connaiss- 'know'.
(d) Attached to this verb is the suffix -ait which marks the verb as being in the past tense, as well as agreeing with the subject il/elle.
(e) The final element is the form pas which, in conjunction with ne before the verb, also marks the negative.

Thus, the structures of the English and French sentences can be summarised as follows:

| English | SUBJECT | DO+TENSE | NEGATIVE | VERB |
| French | SUBJECT | NEGATIVE | VERB+TENSE | NEGATIVE |

The question that we need to ask ourselves now is this: if the structures of the two English and French are so different, how is it that the structures of the two

pidgin and creole languages that are derived from them are so similar? Both Tok Pisin and Haitian Creole share the following basic structure in these sentences:

| SUBJECT | NEGATIVE | TENSE | VERB |

The two pidgin languages are closer in structure to each other than either is to French or to English. Clearly, this cannot be because of the influence of the superstrate languages, as English and French are quite different from each other. We cannot put this down to similarities in the substrate languages either, as these are the languages of New Britain and New Ireland in the case of Tok Pisin, and West African languages in the case of Haitian Creole, and these languages are quite different from each other.

One explanation that has been proposed in the past to explain facts such as these was that speakers of all languages are born with some kind of basic idea about how to simplify their language in situations where it is necessary, typically in language contact situations. This means that we all have some kind of ready-made instructions in our heads that tell us how to simplify our languages and to speak a kind of basic, understandable language where all we have to learn is the vocabulary. The reason why Tok Pisin and Haitian Creole exhibit such similarities is that people in both places share this basic set of instructions about how to simplify language.

Despite the existence of similarities such as this between Tok Pisin and Haitian Creole, it has become apparent that pidgin languages exhibit many differences as well as similarities. The apparently remarkable similarity between these two languages that you have just seen may in fact not be as significant as it appears. If we accept that both English and French structures are going to have their bound forms eliminated as well as grammatical morphemes in whatever pidgin emerges in order to express this meaning, redundancy reduced, it is almost certain that we will end up with four morphemes in whatever pidgin emerges in order to express this meaning. Given that the basic word order in both English, French, and the two sets of substrate languages is SVO, it is again predictable that the subject pronoun would end up coming before the verb. The verb in both English and French is the final element in the verb phrase in these clauses, so again it should not be a great surprise to find the other morphemes marking negation and tense occurring before it. The only real surprise is the relative ordering of the negative and tense marker in Tok Pisin and Haitian Creole, but with just this single similarity, we could suggest that this is due to mere chance.

Attempts to find shared structural characteristics among all pidgins and creoles have failed to reveal anything that is absolutely consistent for every case, and attention has since turned specifically to creoles. Pidgins, it is now felt, are less likely to show up any kinds of features common to all languages, because pidgins are by definition nobody's mother tongue. This means that there is always the possibility that substrate patterns could interfere with patterns derived from features that might be common to all languages. If parallel features develop among creoles, however, presumably this cannot be

due to substrate interference as the speakers of such languages do not know any other languages. The prediction is that as a pidgin becomes a creole, it will expand structurally (as well as lexically). We should therefore be able to examine the structures of creoles in order to find out how it is that languages world-wide undergo creolisation. However, initial studies of the process of creolisation produced disappointing results. It has turned out, in comparisons between people who speak Tok Pisin as a second language and the increasing number who are growing up speaking it as their first language, that there are very few real differences in how the two groups speak the language.

The terms *pidgin* and *creole* are actually quite difficult to apply to particular situations when they are defined as they are in this section (even though these are the definitions that are given in almost all standard textbooks on the subject). As I have just indicated, in Papua New Guinea today, Tok Pisin is spoken as a second (or third, or fourth) language by the majority of the population, but a sizeable minority of urban Papua New Guineans, typically those whose parents come from different parts of the country and who speak different vernaculars, are now growing up speaking Tok Pisin as their first language. Do we say that Tok Pisin is a pidgin, or a creole, or a pidgin that is becoming a creole? The fact that there are no major differences in the speech of those who speak it as a 'pidgin' and those who speak it as a 'creole' makes the distinction seem almost pointless.

Some linguists have avoided this problem by redefining what a creole is. For them, the term *creole* should be used only to refer to a contact language which has developed over a very short period of time and which has developed out of a pidgin language that had not yet had time to acquire a stable structure. This would exclude Tok Pisin, for example, because it has had a stable structure for several generations. Linguists who adopt this revised definition find the study of these sorts of situations more interesting, because if the pidgin has not yet acquired a stable structure, *and* if the members of the following generation grow up with only this unstable pidgin as their input, any stable rules that develop in different creolising situations presumably reflect features of the human linguistic capabilities that we are all born with. What we are talking about here is the *Language Bioprogram Hypothesis* which has been formulated during the last decade and a half by scholars such as Derek Bickerton. According to this hypothesis, there are certain grammatical features that are predestined to emerge in the kinds of creolising situations that I am referring to.

Of course this is just a hypothesis at this stage, and the issue is currently being hotly debated among scholars, with some taking the view that creole languages ultimately derive much of their structure from their substrate languages, and other scholars supporting Bickerton's view that some kind of genetic structural predisposition is involved. The matter is nowhere near resolved, but one thing that has happened is that pidgin and creole languages have shifted from the periphery of linguistics to a central issue in historical linguistics.

12.3 ESOTEROGENY AND EXOTEROGENY

Quite apart from any individual scholar's position on the issue of substrate versus bioprogram, there is still the burning issue of how pidgins and creoles should be handled in a family tree model of language change. Pidginisation is generally regarded as a somewhat exceptional case in the evolution of languages. 'Normal' languages can be said to be descended from another language which is clearly recognisable as its ancestor. French, for example, is descended from Latin, and Samoan is descended from Proto Polynesian. But where do pidgin languages fit into the comparative method? Melanesian Pidgin does not have an ancestor in the same sense in which French has Latin as its ancestor. In 1840, Melanesian Pidgin did not exist, but by the 1860s, it was widely spoken in some parts of Melanesia, and had already spread to other areas where Melanesians had been taken as labourers, such as parts of Queensland in Australia.

What language is Melanesian Pidgin descended from? The family tree model breaks down when it comes to pidgin languages, because in a sense they spring out of nowhere! In an effort to force pidgin languages into the family tree model, some linguists might be tempted to classify Melanesian Pidgin as a Germanic language, and to place it in a subgroup along with English (as a kind of daughter language of English). Certainly the lexicon of Melanesian Pidgin is largely derived from English, but it is much harder to say that its grammar is derived from English grammar. Although there are many features of the grammar of Melanesian Pidgin which seem to derive from Austronesian languages, few linguists would go so far as to draw a family tree of the Austronesian languages with Melanesian Pidgin as one of the branches. After all, there are no systematic sound correspondences between Melanesian Pidgin and other Austronesian languages, as its lexicon is largely derived from English.

So pidgin languages tend to be either ignored, or placed in the 'too hard' basket by traditional comparative linguists. However, what happens when a language undergoes pidginisation is in some ways little different from what happens when 'normal' languages undergo ordinary linguistic change. There is nothing unusual about the kinds of sound changes that take place between a pidgin language and its lexical source language.

Many of the grammatical changes that take place in the development of a pidgin or creole language are also similar to the kinds of grammatical changes that take place in more 'normal' kinds of languages, and there are some linguists who do not see pidginisation as representing a special case in the study of language change at all. Rather than being problematic and peripheral to historical linguists, some would prefer to see pidginisation as being central to the study of language change. They would also argue that it is necessary to dismantle the family tree model that finds pidgin languages so difficult to incorporate. Many of the changes that took place when modern English evolved out of Old English after the Norman invasion in 1066 directly parallel

the kinds of grammatical changes that take place when a pidgin language is formed. If you compare the grammar of modern Dutch and modern Afrikaans, it is tempting to describe Afrikaans as pidginised Dutch, as its grammar is certainly simplified and more regular than that of Dutch. In fact, perhaps all languages involve some kind of structural 'pidginisation' as they change!

Of course, the process of pidginisation cannot account for *all* language change. If this were true, then the vast majority of the world's languages would be creoles, which means that all languages should be structurally simple. However, while languages such as English and Afrikaans have undergone considerable simplification, there are other Indo-European languages, such as Lithuanian, which have retained their grammatical complexity. Speakers of modern Icelandic find it relatively easy to read thousand-year-old Norse sagas, which their close relatives the Norwegians can no longer read unless they learn Old Norse as a foreign language. This means that in addition to processes akin to pidginisation, there are other languages that change in the opposite way, building additional complexity into their languages.

The linguist William Thurston introduces a distinction between esoteric and exoteric languages as a way of explaining this difference. An *esoteric* language is one that is used primarily for intra-group communication, and which sets a group off from surrounding groups. Such languages tend to become increasingly complex as they are transmitted from generation to generation as they are subject to a number of functional pressures. Phonological efficiency is developed at the expense of morphological transparency, which means that there is likely to be a greater number of portmanteau morphemes, and a greater amount of allomorphic variation. Such languages typically develop suppletive morphological marking, and the lexicon makes an increasingly fine set of semantic distinctions. Originally optionally marked categories become grammaticalised. Outsiders typically find an esoteric language difficult to learn, which means that it functions even more efficiently as a marker of identity.

An *exoteric* language, on the other hand, is one that is also used for inter-group communication. Given the kinds of circumstances in which such languages are used, there will be many people for whom intelligibility rather than grammaticality is the primary concern. Such languages tend to develop in ways that make them easier to learn. Changes in exoteric languages are therefore likely to be in the opposite direction to those that are characteristic of developments in esoteric languages.

In introductory linguistics courses, linguists — sometimes even myself — often make the point to students that 'all languages are of equal complexity'. We do that for a purpose, as there is a temptation for people from technologically advanced cultures to think that languages such as English or Chinese might be more sophisticated than languages such as those of the Australian Aborigines. It is for this reason that Edward Sapir said in 1921 that:

We know of no people that is not possessed of a fully developed language. The lowliest South African Bushman speaks in the forms of a rich symbolic system that is in essence perfectly comparable to the speech of the cultivated Frenchman.

While all languages are amazingly complex, languages are in fact not all *equally* complex. Perhaps we linguists would be better in touch with reality if we were to say instead: the level of complexity of languages is in no way related to the level of technological complexity of the culture of their speakers.

Of course, if we are going to allow that languages can differ in their degrees of complexity, we need to offer some kind of absolute definition of what constitutes linguistic simplicity, as vague feelings are not going to be enough to go by. We also have to avoid the possibility that a particular pattern in Polish may be relatively complex for me as, say, an English-speaker, though a speaker of Russian may find it quite unchallenging, simply because the two languages are structurally similar to begin with. According to William Thurston, a language that approximates to the following characteristics can be described as simple:

(a) There is an approximation to a one-to-one correspondence between form and meaning.

(b) There is little stylistic or sociolinguistic variation.

(c) There are relatively few grammatically marked distinctions.

(d) There are relatively few bound morphemes, and these morphemes exhibit little allomorphic variation or suppletion.

(e) Grammatical constructions would approximate towards one function per pattern, and the patterns would apply regularly, i.e. without special exceptions.

(f) At the level of the lexicon, there would be few opaque idioms.

Whether a language is simple or complex is obviously relative rather than absolute. This means that it is possible to place languages on a continuum between two extremes. Comparing closely related German, Dutch and Afrikaans, for example, it is clear that in most respects these languages are structurally very similar, except that the Dutch inflectional system is simpler than that of German, while that of Afrikaans is simpler than that of Dutch. In this case, then, we clearly have a cline of structural simplicity: German → Dutch → Afrikaans.

The evolution of new languages can therefore be said to involve two different kinds of processes: exoterogeny and esoterogeny. *Exoterogeny* results in the development of a new exoteric language. In the most extreme example of this kind of process, words are simply taken from another language and mapped in the simplest possible way onto the phonological, syntactic and semantic patterns of the language that the community already speaks. *Esoterogeny*, on the other hand, involves the development of a new

exoteric language, in which diversification proceeds in the direction of greater complexity.

Given that there is a continuum between a simpler and more complex language, the distinction between exoterogeny and esoterogeny is obviously also relative rather than absolute. A relatively simple exoteric language may develop, and if this language then comes to be emblematic of a single community, it may develop a new set of exoteric features. Some languages may only be called on relatively infrequently for use as intergroup languages, so the need for an easy language may be less urgent than in another situation where there is frequent use made of a language as a lingua franca. In such circumstances, we could expect that esoterogeny will proceed to differing extents.

The view that languages may be subject to pressures to move towards exoteric or esoteric types offers many historical linguists much to think about, and it certainly offers some serious challenges to those who insist on a rigid application of the family tree model of linguistic diversification. The family tree model does not cope well with notions such as 'mixed languages'. This kind of model, however, accepts that language mixing is a real phenomenon.

It is likely that the distinction between esoterogeny and exoterogeny will be ignored by some. It will be actively disputed by others, who will continue to push languages such as Tok Pisin into the 'too hard' basket. Others may be more open to accepting these kinds of views, though there are possibly still some questions that remain to be resolved.

While the correspondence between local emblematicity and esoterogeny on the one hand and use as a lingua franca and exoterogeny on the other is appealing, and many examples can be presented to make the correspondence appear convincing, we probably have some way to go before we can say that we have explained all instances of linguistic diversification according to this model. For instance, while it may be possible to invoke such explanations in the exogenetic development of Afrikaans out of Dutch — given what we know of the multilingual situation of the early Dutch settlers in South Africa — it would be much more difficult to account for the apparently exogenetic development of Dutch out of an earlier pattern that was more like that which we find in German.

And as soon as somebody puts up a boldly stated and very general theory, there will always be apparent counterexamples appearing in the literature. On the island of Erromango in Vanuatu, for example, there were originally several distinct — though closely related — languages, possibly as many as six. Around the time of first colonial contact — or possibly even before this — in the mid-1800s, one of these languages, known as Sye, came to be adopted as a general lingua franca, while the other languages became 'vernacularised', i.e. they came to function solely as intra-group languages. Given this scenario, we would expect that Sye should be the grammatically simplest of these languages, while the other languages should show signs of structural complexity. In fact, the reverse seems to be the case, with Sye being

one of the more complex Oceanic languages, while the more restricted Ura language has undergone considerable morphological simplification in its patterns for expressing possession, as well as the ways in which it marks verbal objects.

12.4 LANGUAGE DEATH

In Chapter 1 I referred to the fact that a language can die. Language death is something that is almost always associated with language contact. The only situation in which a language may die without language contact taking place is in the comparatively rare situation in which an entire speech community is wiped out by a massive calamity such as a volcanic eruption, a military slaughter, or an epidemic. Such things have unfortunately happened in the past. Oral tradition in central Vanuatu tells of the once large island of Kuwae which was shattered by a volcanic cataclysm into the much smaller present-day islands of Tongoa and the Shepherd Islands. This massive eruption must have killed large numbers of people. Oral tradition records that, although a small number of people from Kuwae survived this holocaust, when the new, smaller islands were resettled by people from the nearby larger island of Efate, they brought with them their own language, which explains why the people from these islands speak a dialect of the Efate language to this day. Presumably the original language of Kuwae disappeared with the death of the last survivors of the eruption. The history of Aboriginal Australia is full of accounts of the extermination of whole communities of Aboriginal people by European settlers, often by the most inhuman methods such as the deliberate introduction of smallpox, or by vicious shooting sprees. Again, unknown numbers of languages disappeared from the record with the disappearance of their speakers.

Tragic as such circumstances are, they are of primary interest to scholars of history rather than linguistics. Language death typically occurs in much less catastrophic circumstances, and arises as a result of language contact over an extended period of time. When speakers of two languages come into contact and speakers of one of the two languages have power over speakers of the other language, either by force of social prestige or by demographic dominance, it is possible for speakers of the socially weaker language to abandon their language in favour of the dominant language. This has taken place in many parts of the world in the past, and is probably accelerating today as languages like English and French become increasingly dominant world-wide through the power of education, government, and the mass media.

Many Australian languages have disappeared, not because their speakers were exterminated, but because the generations of the past either chose to or were forced to speak to their children in English. Only about 1 per cent of Hawaiians today speak Hawaiian, the remainder having shifted to English, and Māori in New Zealand has shown signs of going the same way, with only

about 10 per cent of Māori people today speaking the ancestral language. The languages in some parts of Papua New Guinea (especially in the Sepik) are under pressure, not from English, but from Tok Pisin. In Europe also, minority languages are under pressure from larger languages. Irish, Scots Gaelic, and Welsh are all under pressure from English; Friesian is under pressure from Dutch; and Breton is under pressure from French.

A description of the social circumstances surrounding the death of a language belongs in a volume on sociolinguistics, so in this book I will concentrate not on what causes a language to die, but on what happens to the language itself as it dies. Before I can do this, however, *some* discussion of what causes a language to die is necessary. To do this, I will outline what has happened in the history of Māori in New Zealand. From the time of the original settlement of Aotearoa (known to the outside world as New Zealand) about 1000 years ago, the Māori had uncontested control over their territory, and their language functioned as part of their flourishing culture. In the beginning of the nineteenth century, European settlers began arriving, initially in small numbers, and from the second half of the nineteenth century, in an increasing flood. The Māori lost much of their land to the settlers and quickly came under the military control of the settlers, and later also under their political and economic control. However, the Māori remained a largely rural rather than an urban people, living together in communities, and their language continued to flourish, even though their children learned English at school.

A major social change occurred after the Second World War as many Māori began moving from the rural areas to the cities and towns in order to get jobs. Without a fluent command of English, it was difficult to get jobs, and parents saw it as benefiting their children if they refused to speak to them in Māori and insisted on only English in the home. The next generation that grew up in towns therefore tended to learn only a little Māori (possibly from their grandparents who spoke little English), or none at all. Of course, the children of this generation who are today's teenagers and young adults have also grown up speaking nothing but English. It is probably not completely accurate to say that the Māori language began to die; rather, it began to commit suicide. The result is that today, about 90 per cent of Māori speak only the language of their original conquerors.[2]

In communities where it is recognised that a language is in a precarious situation, the remaining fluent speakers frequently comment on the fact that younger generations no longer speak the language 'properly'. Fluent speakers of Māori, for example, point to overwhelming lexical interference from the

[2] Many Māori see the possible loss of their language as a threat to their cultural identity and are taking steps to ensure that the language does not disappear. Older speakers of Māori are now being involved in special childcare centres and preschools known as *kōhanga reo* (literally 'language nest') in which only Māori is used. Thousands of children are now growing up as fluent speakers of Māori.

dominant language, confusion of grammatical distinctions, and poor command of the stylistic repertoire. The same sorts of changes are found all over the world in situations where languages are showing signs of being replaced by other languages. Sometimes the older generations will attempt to correct the mistakes of younger *partial speakers* (i.e. those whose command of the language has suffered as a result of language shift taking place). This can, of course, cause partial speakers to become embarrassed and to avoid using the language with older people for fear of further correction. Rather than improving the chances of the language surviving, this may even make it less likely that it will survive, especially if it is a very small one as in the case of Australian Aboriginal languages.

One language that is recognised as being near to extinction is the Dyirbal language of the coast of northern Queensland in Australia. While the older people are recognised as being able to speak the language 'correctly', the younger generations have grown up either speaking no Dyirbal at all (using only English), or speaking a kind of Dyirbal that everybody recognises to be 'corrupted' in some way. At the simplest level, this involves the frequent use of words (and phrases) of English (or Pidgin) origin for which there are established Dyirbal words. Sometimes the younger people may have forgotten the original Dyirbal word, though in other cases they may use an English word even though they do know the corresponding word in Dyirbal.

The use of words of foreign origin is not necessarily a sign of imminent language death. If it were, then English with its huge number of borrowed words should be a prime example of a dying language. Instead, the enthusiasm with which English has accepted new vocabulary is generally taken as a sign of its extreme vitality. However, the speech of younger people in the Dyirbal community is also grammatically quite different from that of the older people. Younger speakers are reducing the morphological complexity of the language by eliminating some suffixes. The grammatical functions that were originally expressed by these suffixes are now often expressed by free forms that are derived from English, as in the following examples:

Old People's Dyirbal
Ban ḍugumbil ɲina-ɲu jugu-ŋga.
feminine woman sit-nonfuture log-on

Young People's Dyirbal
Ban ḍugumbil ɲina-ɲu on jugu.
feminine woman sit-nonfuture on log

'The woman sat on a log.'

Grammatical constructions that are very different from those of English are also particularly subject to change. The Dyirbal that is spoken by the older people has a very free word order, similar to what I described for Latin in

Section 7.1. This is possible because all noun phrases in Dyirbal are obligatorily marked by suffixes which indicate clearly which is the subject and which is the object. However, younger speakers tend to leave the ergative suffix off nouns that function as the subjects of transitive verbs, and distinguish the subject and object noun phrases by using a fixed SVO word order as in English. (See Section 7.1 again for more detailed discussion of ergativity.) Instead of using the ergative form /buliman-du/ 'policeman (ergative)', a younger person might produce a sentence such as the following, with no suffix at all on the subject — which an older person would judge as ungrammatical:

Buliman   ŋanba-n   ban   buladi.
policeman ask-nonfuture feminine two
'The policeman asked those two (women).'

As the language comes under increasing pressure from English, we can expect that there will be greater influence of English vocabulary and structural patterns on the language. Some speakers are already producing sentences that are basically English even though they still contain fragments of Dyirbal, such as the following:

They bin gunimariŋu but they never bin find-im.
'They looked for him but they didn't find him.'

The Dyirbal verb /gunimariŋu/ 'look for' occurs in a sentence with an English subject, and with the past tense marker /bin/ that derives from the earlier Aboriginal Pidgin. The object suffix /-im/ on the verb *find* also derives from the pidgin.

The question of just what happens to the grammar as a language dies has begun to arouse considerable interest among scholars of language change. In one sense what has happened to the vocabulary and the grammar of Dyirbal is quite unexceptional. The incorporation of vocabulary from one language into another is, as we have seen, a perfectly normal aspect of language change. The kinds of grammatical changes that are taking place are also not radically different in nature from the kinds of changes that take place in situations of ordinary language change. In Chapter 7, I indicated that inflecting and agglutinating languages often evolve into isolating languages, and that morphological irregularities in languages tend to be eliminated by the process of analogy. Just as an accusative language can, over time, acquire an ergative structure, so the shift of Dyirbal from an ergative structure to an accusative structure marked by word order rather than by case suffixes is again perfectly within the bounds of normal language change.

What is exceptional in the case of Dyirbal is that the changes are happening on such a massive scale and in such a short period of time. These structural changes have all taken place within the space of 25 years. English

began to undergo very similar sorts of changes from around the time of contact after the Viking and Norman invasions, but it took several centuries for this to happen. Another difference between the kinds of changes that are taking place with Dyirbal and those which happened in the history of English is that in Dyirbal, a change that results in the loss of some aspect of the grammar (such as the loss of the locative suffix /-ŋga/ in one of the examples above) has not been compensated for by a corresponding development somewhere else in the language. What has happened is that speakers have simply taken over the corresponding English form to express this function, i.e. the preposition *on*.

This has led some scholars to suggest that what happens when a language dies is somehow similar to what happens when a pidgin language comes into existence, except that events take place in the opposite order. Just as a pidgin in the early stages of its formation involves grammatical reduction and both structural and lexical variability, so too does a dying language. A pidgin also has a reduced stylistic repertoire compared to a 'normal' language, and we find the same thing with a language that is dying. Others have pointed instead to parallels in reverse between language death and creolisation. Just as a pidgin that undergoes creolisation is said to expand structurally, lexically, and stylistically, a dying language experiences structural, lexical, and stylistic reduction.

Although it need only take a break in transmission between a single generation for a language to be doomed, it is possible for features of an old language to be maintained over a relatively long period. In Section 9.4 I talked about how languages can sometimes change to allow a local group to mark its separate identity in some way. Although the examples that I gave there came from a Papua New Guinea context, this is not something that is restricted just to these languages. Very often when an ethnic group switches from one language to another, people develop ways of marking their ethnicity through their new language. There is a variety of English in America that is typically associated with Blacks as against Whites. Māori in New Zealand can often be distinguished from Pākehā by the way they speak English. Books about the history of Tasmanian Aborigines point out that the last fully-descended Tasmanian died in 1876, and her language died with her, yet some of the 4000 or so people in Tasmania today who are of Aboriginal descent (and who proudly identify themselves as Tasmanian Aborigines) still use the occasional word of Aboriginal origin in their speech. We can expect that while Dyirbal is doomed as a distinct language, the succeeding generations of people who belong to this community will continue to sprinkle their English with individual words of Dyirbal origin, even though there will be little or no evidence of Dyirbal grammatical structures.

READING GUIDE QUESTIONS

1. What is interference as distinct from diffusion?
2. What is calquing?
3. Can phonemes be copied from one language to another?
4. How can morphemes from one language enter another?
5. What is the difference between convergence and diffusion?
6. What is a linguistic area?
7. To what extent is syntactic copying possible?
8. What are the characteristics of a pidgin?
9. What is a creole?
10. What is the difference between a superordinate and a substrate language?
11. What is the Language Bioprogam Hypothesis?
12. What is meant by language death?
13. What parallels are there between pidginisation and creolisation on the one hand and language death on the other?

EXERCISES

1. Examine the data below from two different languages, one of which is vernacular Fijian and the other the pidginised form of Fijian that emerged on plantations in Fiji during the last century:

Language A	Language B	
na noqu vale	na vale koyau	'my house'
na nomu veiniu	na veiniu koiko	'your plantation'
na nona koro	na koro kokoya	'his/her village'
na nodra vale	na vale koratou	'their house'
na nodratou veiniu	na veiniu koratou	'their (three) plantation'
na nodrau koro	na koro koratou	'their (two) village'
na nomudrau bilo	na bilo kemudou	'your (two) cup'
na nomuni vale	na vale kemudou	'your (many) house'
na nomudou veiniu	na veiniu kemudou	'your (three) plantation'
na noda vosa	na vosa keitou	'our (many inclusive) language'
na neimami vosa	na vosa keitou	'our (many exclusive) language'
na meirau wai	na wai keitou	'our (two exclusive) water'
na meitou bia	na bia keitou	'our (three exclusive) beer'
na meimami bia	na bia keitou	'our (many exclusive) beer'
na medaru bia	na bia keitou	'our (two inclusive) beer'
na medatou wisiki	na wisiki keitou	'our (three inclusive) whisky'
na meda wisiki	na wisiki keitou	'our (many inclusive) whisky'

na tamaqu	na tamana koyau	'my father'
na tamamu	na tamana koiko	'your father'
na tamana	na tamana kokoya	'his/her father'
na ligada	na ligana keitou	'our (many inclusive) hands'
na ligadra	na ligana koratou	'their hands'

(Note that in the Fijian orthography that is used in these examples, the symbol *q* is used to represent a prenasalised voiced velar stop, which is phonetically [ŋg].)

(a) Which of these two languages is vernacular Fijian and which is the pidgin form of Fijian? What are the structural features which enable you to say this?

(b) Give the equivalents of the following phrases in both vernacular Fijian and pidgin Fijian:

your hand
your (two) father
your (many) father
his/her water
their (two) whisky
their (three) beer
our (many exclusive) house
their language
our (many inclusive) plantation
my cup

(c) Consider the following additional forms in Language A:

na tinamu 'your mother'
na tinana 'his/her mother'

On the basis of this information, give the following in both vernacular Fijian and pidgin Fijian:

our (many inclusive) mother
their mother
your (two) mother
your (many) mother
our (three exclusive) mother

2. Examine the following forms in Haitian Creole and in French:

French	Haitian Creole	
je suis malade	m malad	'I am sick'
ils sont malades	yo malad	'they (masculine) are sick'

elles sont malades	yo malad	'they (feminine) are sick'
j'étais malade	m te malad	'I was sick'
ils étaient malades	yo te malad	'they were sick'
nous acheterons	n ap achte	'we will buy'
je vais	m ale	'I am going'
vous irez	u ap ale	'you (plural) will go'
tu iras	u ap ale	'you (singular) will go'
il a couru	li te kuri	'he has run'
il est bon	li bien	'he is good'
elle va	li ale	'she is going'
je suis allé	me te ale	'I have gone'
il a acheté	li te achte	'he has bought'

(a) What are the features of Haitian Creole by which we can recognise that it has undergone the process of pidginisation?

(b) How would you express the following in French?

yo ap malad		
m achte		
n te kuri		
yo te bien		
u ap ale		

3. Compare the following forms from Bandjalang language of northern New South Wales (in Australia) as it was spoken by people who learned the language in the early twentieth century and people who learned to speak it in the late nineteenth century. In what way are the changes that have taken place similar to the changes involved in the formation of pidgins and creoles?

Older People's Bandjalang	Later Generation Bandjalang	
gala gibirga:	gala gibirga:	'mahogany tree'
ga:ɲu gibi:ɳbilga:	ga:ɲu gibirga:	'mahogany trees'
gala bunawga:	gala bunawga:	'bloodwood tree'
ga:ɲu buna:ɳbilga:	ga:ɲu bunawga:	'bloodwood trees'
gala barbamga:	gala barbamga:	'spotted gum tree'
ga:ɲu barba:ɳbilga:	ga:ɲu barbamga:	'spotted gum trees'
gala buḏe:ga	gala buḏe:ga	'Moreton Bay fig tree'
ga:ɲu buḏe:ɳbilga:	ga:ɲu buḏe:ga	'Moreton Bay fig trees'
gala bilaŋga:	gala bilaŋga:	'oak tree'
ga:ɲu bila:ɳbiŋga:	gala bilaŋga:	'oak trees'
gala ŋarulga:	gala ŋarulga:	'box tree'
ga:ɲu ŋaru:ɳbiŋga:	ga:ɲu ŋarulga:	'box trees'
gala jigam	gala jigam	'piece of meat'
ga:ɲu jigambil	ga:ɲu jigam	'pieces of meat'
gala wura:ɳ	gala wura:ɳ	'leaf'
ga:ɲu wura:ɳbil	ga:ɲu wura:ɳ	'leaves'
gala ḏinaŋ	gala ḏinaŋ	'foot'
ga:ɲu ḏinaɳbil	ga:ɲu ḏinaŋ	'feet'
gala deberdebe:r	gala deberdebe:r	'plover'
ga:ɲu deberdebe:rgan	ga:ɲu deberdebe:r	'plovers'
gala bagawaŋ	gala bagawaŋ	'leatherhead bird'
ga:ɲu bagawaŋgan	ga:ɲu bagawaŋ	'leatherhead birds'
gala muḏumḏar	gala muḏumḏar	'son'
ga:ɲu muḏumgir	ga:ɲu muḏumḏar	'sons'
gala baniḏar	gala baniḏar	'father'
ga:ɲu baniḏar	gala baniḏar	'fathers'
gala balun	gala balun	'river'
ga:ɲu balungali	ga:ɲu balun	'rivers'
gala bagul	gala bagul	'canoe'
ga:ɲu bagulgali	ga:ɲu bagul	'canoes'
gala daba:j	gala daba:j	'dog'
ga:ɲu daba:jgali	ga:ɲu daba:j	'dogs'
gala muru	gala muru	'nose'
ga:ɲu murugali	ga:ɲu muru	'noses'
gala dubaj	gala dubaj	'woman'
ga:ɲu dubaymir	gala dubay	'women'
gala wagaŋ	gala wagaŋ	'catfish'
ga:ɲu wagaŋmir	ga:ɲu wagaŋ	'catfishes'
gala bajgal	gala bajgal	'man'
ga:ɲu bajgalbajga:l	ga:ɲu bajgal	'men'
gala dugun	gala dugun	'mountain'
ga:ɲu dugundugu:n	ga:ɲu dugun	'mountains'
gala bargan	gala bargan	'boomerang'
ga:ɲu bargan	ga:ɲu bargan	'boomerangs'
gala mundu	gala mundu	'stomach'
ga:ɲu mundu	ga:ɲu mundu	'stomachs'
gala jamba:	gala jamba:	'carpet snake'
ga:ɲu jamba:	ga:ɲu jamba:	'carpet snakes'
gala ŋa:wun	gala ŋa:wun	'wood duck'
ga:ɲu ŋa:wun	ga:ɲu ŋa:wun	'wood ducks'