Chapter 7: Contact Languages I: Pidgins and Creoles

‘The Negroes who established themselves on the Djuka Creek two centuries ago found Trio Indians living on the Tapanahoni. They maintained continuing relations with them....The trade dialect shows clear traces of these circumstances. It consists almost entirely of words borrowed from Trio or from Negro English’ (Verslag der Toemoekhoemak-expeditie, by C.H. De Goeje, 1908).

‘The Nez Perces used two distinct languages, the proper and the Jargon, which differ so much that, knowing one, a stranger could not understand the other. The Jargon is the slave language, originating with the prisoners of war, who are captured in battle from the various neighboring tribes and who were made slaves; their different languages, mixing with that of their masters, formed a jargon....The Jargon in this tribe was used in conversing with the servants and the court language on all other occasions’ (Ka-Mi-Akin: Last Hero of the Yakimas, 2nd edn., by A.J. Splawn, 1944, p. 490).

The Delaware Indians ‘rather design to conceal their language from us than to properly communicate it, except in things which happen in daily trade; saying that it is sufficient for us to understand them in that; and then they speak only half sentences, shortened words...; and all things which have only a rude resemblance to each other, they frequently call by the same name’ (Narratives of New Netherland 1609-1664, by J. Franklin Jameson, 1909, p. 128, quoting a comment made by the Dutch missionary Jonas Michaëlius in August 1628).
The list of language contact typologies at the beginning of Chapter 4 had three entries under the heading ‘extreme language mixture’: pidgins, creoles, and bilingual mixed languages. All of these are types of contact languages; they are three separate phenomena, but pidgins and creoles go together naturally, in contradistinction to bilingual mixed languages. The distinction is basically the same as the one between contact-induced change with and without imperfect learning: pidgin/creole genesis is akin (though not identical) to shift-induced interference, and bilingual mixed language genesis is akin to, and in effect actually is, borrowing. That is, pidgins and creoles develop in social contexts where few or no members of the groups in contact are bilingual or multilingual in each other’s language(s), while imperfect learning plays no significant role in the development of bilingual mixed languages. This characterization is oversimplified, as we will see below (and in Chapter 8), because contact-induced language change is not the only component of contact-language genesis; but the dichotomy between two basic categories of contact language is valid in spite of the oversimplification. Because of this dichotomy, pidgins and creoles will be discussed together in this chapter, and bilingual mixed languages will be covered in Chapter 8.

Before beginning any detailed discussions, we need a definition of ‘contact language’. There is unfortunately no uniform definition of the term in the scholarly literature, but there are two main kinds of usage. Some people apply it always and only to languages of wider communication, i.e. to all LINGUA FRANCAS. Under that definition, any language that is used for intergroup communication is a contact language—including not only pidgins and creoles, but also non-pidgin/non-creole languages like English, which is certainly the most widely used lingua franca in the modern world. (See Chapter 2 for examples of other lingua francas around the world.)

In this book a different definition will be used. Here, a contact language is any new language that arises in a contact situation. Linguistically, a contact language is identifiable
by the fact that its lexicon and grammatical structures cannot all be traced back primarily to the same source language; they are therefore mixed languages in the technical historical linguistic sense: they did not arise primarily through descent with modification from a single earlier language. By definition, therefore, contact languages are not members of any language family and thus belong in no family tree—except perhaps as the ancestor of a language family: a contact language has no single parent language in the historical linguist’s usual sense, but it may have descendants. In other words, I’m defining ‘contact language’ on the basis of the type(s) of historical connections to other languages.

It follows from this definition that contact languages needn’t be lingua francas at all, and in fact some aren’t. Pidgins and creoles emerge in contexts in which people from different linguistic backgrounds need to talk to each other regularly, and they are therefore lingua francas in origin; but bilingual mixed languages are in-group languages, not languages of wider communication. We’ll return to some of the implications of this definition of contact language below, and again in Chapter 8.

The organization of this chapter is as follows. We’ll begin in the next section with definitions of pidgins and creoles, and then move on to their history and current or attested former distribution around the world. In the next section we’ll consider two common beliefs, partly true and partly false, about pidgins and creoles: that they’re maximally simple (especially pidgins) and that they’re all alike (especially creoles). The last major section covers the major origin theories for pidgins and creoles—historical scenarios that have been proposed and defended for their emergence.

What, when, where?

The first step in investigating pidgins and creoles as a contact phenomenon is to place
them in historical and geographical context. Again, we need to begin with definitions. As with just about everything else in the field of pidgin/creole studies, all the available definitions of ‘pidgin’ and ‘creole’ are controversial. I’ll present and discuss the definitions that seem most workable to me, but readers who want to explore the controversy can find references to other definitions in the ‘Sources and further reading’ section at the end of the chapter.

Let’s start with pidgins. Traditionally, a pidgin is a language that arises in a new contact situation involving more than two linguistic groups; the groups have no shared language—that is, no single language is widely known among the groups in contact—and they need to communicate regularly, but for limited purposes, such as trade. For some combination of social, economic, and political reasons, they do not learn each other’s languages, but instead develop a pidgin, with vocabulary drawn typically (though not always) from one of the languages in contact. The new pidgin’s grammar doesn’t come from any one language; instead, it is a kind of cross-language compromise of the grammars of the languages in contact, with more or less (usually more) influence from universals of second-language learning: in particular, ease of learning helps to determine the linguistic structure of a pidgin. The process by which a pidgin is created is “negotiation”, as outlined in Chapter 6.

Adopting this view of pidgins carries several implications. One is that a pidgin is nobody’s native language: pidgins are always spoken as second languages (or third, or fourth, or...) and are typically used for limited purposes of intergroup communication. A second implication is that, thanks to their limited social functions, pidgins have less linguistic material than non-pidgin languages—fewer words, limited grammatical and stylistic resources in syntax and discourse. In addition, for reasons that are probably connected mainly to ease of learning,
pidgins tend to lack elaborate morphological structures.

A creole, by contrast, is the native language of a speech community. Like pidgins, creoles develop in contact situations that typically involve more than two languages; also like pidgins, they typically draw their lexicon, but not their grammar, primarily from a single language, the lexifier language. (The terminology for labeling the lexifier language varies: in particular, the terms English-lexifier creoles and English-based creoles are used interchangeably.) The grammar of a creole, like the grammar of a pidgin, is a cross-language compromise of the languages of its creators, who may or may not include native speakers of the lexifier language. (That is: the lexifier language has to be available to pidgin or creole creators at least to the extent that they take most of the new language’s vocabulary from it; but L1 speakers of the lexifier language aren’t always involved in “negotiating” the structure of the contact language itself.) In fact, some creoles are nativized pidgins. Such a creole was originally a pidgin which later became the main language of a speech community, learned as a first language by children and used for general community activities; its linguistic resources expanded with the expansion in its spheres of usage. Other creoles seem never to have gone through a pidgin stage at all, but to have developed gradually by increasing divergence from the lexifier language, and still others apparently arose abruptly, also without going through a well-defined pidgin stage. We’ll examine all these possibilities below.

In the present context, what’s important is that the definitions just given apply to some, but not clearly to all, pidgins and creoles. It may be reasonable to say that the pidgins and creoles that best fit these definitions are prototypical—that they are the basic type, and that other kinds of pidgins and creoles diverge in one or more respects from this fundamental type. This is reasonable, of course, only to the extent that one agrees with the basic definition; some scholars would propose quite different definitions. Much of the controversy stems from competing theories of how pidgins and especially creoles arise in the first place—
theories of pidgin and creole genesis—so we’ll postpone consideration of the controversies to the section below on pidgin/creole genesis.

Meanwhile, here are two ways in which some pidgins and creoles differ from the picture supplied by the definitions. First, there are a few two-language pidgins and creoles. One is described in typical non-linguist’s language in a quotation at the beginning of the chapter: Ndyuka-Trio Pidgin, spoken in Suriname in northern South America, is an example of a two-language pidgin. It probably dates from the 18th century, when many slaves escaped from coastal plantations, fled into the interior of the country, formed a creole-speaking community, and later came into contact with the Trio Indians of southern Suriname. Its main function is the classic one for pidgins: it is used primarily for trade between the Ndyukas (whose language is an English-lexifier creole) and the Trios. Berbice Dutch Creole is a rare example of a creole that arose among speakers of just two languages. Like Ndyuka-Trio Pidgin, it is (or was—only four or five speakers remained as of 1993) spoken in northern South America, but in Guyana rather than Suriname. It was created by speakers of Dutch and the West African language Eastern Ijo after Berbice was founded in 1627 as a Dutch colony; it became a British colony in 1796. Another two-language creole is Baba Malay, which is spoken by about 15,000 ethnic Chinese and ethnically mixed people in Melaka and Singapore. Lexically based on the Austronesian language Malay, Baba Malay has as its other component Hokkien Chinese.

Second, some pidgins expand in both functions and linguistic resources without, or at least before, being learned as a first language by a population of speakers. An example is Tok Pisin (the name means ‘pidgin language’, and its components ultimately come from English talk and pidgin), now one of the official languages of Papua New Guinea: although it is currently a creole by the above definition, with a population of native speakers, Tok Pisin had expanded into all spheres of New Guinea public life long before it acquired native
speakers; and its transition to a native language therefore involved little or no significant
linguistic change at all. Expansion without nativization is hardly surprising, given the right
social circumstances; but the requisite social circumstances obtain relatively rarely, so that
the classic trade pidgin, restricted in functions and in linguistic material, remains the pro-
totype for pidgins. Some scholars prefer to use the term ‘creole’ for an expanded pidgin
like Tok Pisin before it had native speakers; others (including me) prefer to stay with the
traditional definitions, while recognizing that the correlation between linguistic expansion
and nativization will not hold in every single case.

Another point needs to be made before we turn to the ‘when’ and ‘where’ of the section
title. A notion that is implicit in most of the pidgin/creole literature and explicit in some
of it is that pidgins and creoles arise because speakers of the various other languages in the
new contact situation have too little access to the lexifier language to learn it ‘properly’.
Certainly imperfect learning plays a role in the genesis process, as noted above, but lack of
access is too simplistic a notion, as it’s usually presented. For one thing, there is no reason to
assume that the lexifier language is always a target language in the sense that speakers of the
other languages strive or even want to learn it; in at least one case, that of Chinese Pidgin
English, Chinese speakers are said to have created it in part because they were unwilling to
learn the inferior foreigners’ language, English—and even more unwilling to permit English
or other foreigners to learn Chinese.

For another thing, even if they would like to learn the lexifier language as a whole,
learners are sometimes actively prevented from doing so by speakers of the lexifier language.
Two of the quotations at the beginning of this chapter illustrate slightly different motives for
withholding the lexifier language from would-be learners: the Nez Percé Jargon was clearly
used by the Nez Percé people as one means of keeping their slaves in a separated position
in the community, so that the slaves could not understand their masters when the masters
were talking among themselves. As in some other Northwest communities, speaking Jargon may also have been one way in which slaves were expected to make their inferior position clear to everyone.

The Delaware people’s attitude reflects a situation that is found here and there around the world: they deliberately withheld their full language from outsiders (at least from Europeans) in order to maintain social distance between themselves and others, to prevent others from getting too close to their culture. It was in this context that the now-extinct Pidgin Delaware arose, probably in the 17th century. Similar comments have been made by other frustrated outsiders in other contexts. Here are two more examples, one from fieldworkers among Hamer speakers in Ethiopia and one from an early missionary’s experience among the Motus in what is now Papua New Guinea:

‘For the next seven months we lived in Hamer villages without any interpreter or intermediary between ourselves and the Hamer... at the end of the seven months, we felt we had achieved a working knowledge of Hamer.... Today we realize that the language which we had learned... was a kind of “Pidgin Hamer” which is used only for and by policemen, traders, and non-Hamer settlers. In the past year we have succeeded in having our Hamer friends and companions talk to us in proper Hamer.’

‘The first references of any sort to any “unusual” language spoken by the Motu is that contained in references to W.G. Lawes’ early attempts to learn Motu from villagers in Port Moresby harbour where he and his wife first settled in 1874. According to several recent reports, the Motu were never keen on teaching him their “true” language but instead attempted to communicate with him and later to teach him “a simplified form of their language”.... However, it was not until
some time later that his son, Frank, who played with the boys of the village and learned the “true” language from them drew his father’s attention to the deception. Even so it was only with difficulty that Lawes was able to learn the true language, because many of the villagers were still opposed to imparting the knowledge to strangers.’

When and where?

How long ago did pidgin and creole languages first arise, and where are they spoken now and in the recent past? The first question can’t be answered definitively: we have little or no hope of discovering very old pidgin or creole languages, both because mentions of languages of any kind are sparse in ancient documents and because pidgens and creoles have been marginal languages for most of their history, thanks to their universal status (until very recently) as unwritten languages.

In most of the literature on the history of pidgins and creoles, the famous Lingua Franca has pride of place. This was an Italian-lexifier pidgin that arose sometime during the Crusades, which were carried out from 1095 CE until the mid-15th century. It was used for intergroup communication in the lands bordering the Mediterranean Sea; in the western Mediterranean it acquired much Spanish lexicon as well. Some scholars believe that it underlies at least some later pidgins and creoles elsewhere. This view has been especially popular, in a global form encompassing most or all pidgins and creoles, among scholars who believe that pidgens and creoles are rare phenomena, confined largely to areas of European exploration, trade, and colonialism. But others, probably by now the majority of specialists, believe that such contact languages are likely to have arisen rather frequently all over the world, and that the reason so few are reported from outside the range of European activities
results from the fact that most histories have been written by Europeans and their New World descendants.

In any case, the Lingua Franca is not in fact the earliest documented pidgin or creole language. That distinction goes to an 11th-century Arabic-lexifier pidgin that was reported by the great Arab geographer Abu ‘Ubayd al-Bakrī (ca. 1028-1094 CE), whose magnum opus, *Roads and kingdoms* (in Arabic, *al-Masālik wa ‘l-Mamālik*), was completed in 1068. Al-Bakrī himself did not undertake travels he describes in his book; he was born in Andalusia in southern Spain and spent his entire life there. He gathered stories from travelers and used secondary sources as well. He frequently prefaces a story with ‘someone told me’, as he does in his report of what appears to have been an Arabic-based pidgin:

‘Someone told me that a dignitary from the people of Aswan used to travel a lot. One day he reached a small town called Maridi. Upon his return, he said to the prince of the believers [the ruler in Islamic faith, presumably in this case the Fatimid Caliph of Egypt], “Sir, may God give you plenty of good and honour your face, here is my case! Its goal is to preserve and spread the word of God. The Blacks have mutilated our beautiful language and spoiled its eloquence with their twisted tongues. During my visit, Sir—may God protect you—only God’s guidance helped me escape the dangers and understand their miserable Arabic. Sometimes, and may God forgive me if I did wrong, I could only laugh at what they called Arabic; and may God forgive me if I call what they uttered Arabic.’

These prefatory remarks are followed by a ten-sentence story given to illustrate the “bad Arabic” spoken by the people of Maridi; in this story the lexicon is almost entirely Arabic but the grammar is not like that of any ancient or modern variety of Arabic. The structural features do, however, match those of modern Arabic-lexifier pidgins in a number of respects,
so it seems safe to conclude that the short text represents a pidginized form of Arabic, though of course there is no way to tell whether it was a stable pidgin language or not.

The traveler’s comment about ‘Blacks’ twisted tongues’ shows that racist assumptions about how and why pidgens and creoles arise go back to the earliest recorded example; similar comments are all too easy to find in quite recent literature as well. All such remarks reflect a profound misunderstanding of processes of pidgin and creole genesis: even if we assume that the inhabitants of Maridi wanted to learn Arabic itself rather than just constructing a pidgin to use in conversing with Arab traders, it’s hardly likely that anyone, of any color, could achieve fluency in a language after hearing it spoken briefly by an occasional traveler once a year or so. (It’s not certain exactly where Maridi was, but the village apparently had no significant permanent Arab population; it was probably either in what is now the Sudan or in what is now Mauritania.)

Thanks to al-Bakrī’s account of Maridi Arabic, we know that pidgens were spoken during the early Middle Ages in at least two places, Africa (probably either in the Sahara Desert or the Sudan) and the Mediterranean. There may have been many other pidgens and even creoles all over the world by then, but the next ones we have any knowledge of arose during the Age of Exploration, which was inaugurated by Prince Henry the Navigator of Portugal (1394-1460). Prince Henry sent his first expedition down around the coast of Africa in about 1418, and many other exploring expeditions along the West African coast. Europeans reached the mouth of the Congo River in 1482; in 1488 they rounded the Cape of Good Hope at the southern tip of Africa, and soon after 1515 they reached the port of Canton in southeastern China. From there Europeans spread to the entire Pacific, and they crossed both the Pacific and Atlantic to the New World. Traders and colonists followed (or, in some cases, accompanied) the explorers; soon the European presence was world-wide. Meanwhile, Arab traders and conquerors also expanded far beyond the Arabian Peninsula, to northern
and then sub-Saharan Africa and eastward into Asia, extending their reach as far east as Canton in China.

All these activities eventually gave rise to pidgins and creoles all over the world, especially—given the sea routes that most explorers followed—on and near the world’s seacoasts. Maridi Arabic was apparently not near a coast, but the Lingua Franca was. There are Arabic-lexifier pidgins on islands near the Arabian Peninsula as well as in the interior of northeastern Africa, but most known pidgins and creoles are connected with European expansion and trade, and of these the great majority have European lexifier languages. In Africa there are several Portuguese-based contact languages, among them West African Pidgin Portuguese (15th-18th centuries, but poorly attested), Cape Verde Creole (Crioulo), and São Tomense Creole (spoken on an island in the Gulf of Guinea); in Asia there are several others, e.g. Sri Lanka Creole Portuguese and (in Macao/Hong Kong) Macanese Creole Portuguese.

English-lexifier pidgins and creoles are more numerous and more wide-spread; in fact, English is by a large margin the most frequent lexifier language for known pidgins and creoles (although this large margin results in part from the fact that numerous Portuguese-lexifier pidgins and creoles have become extinct). Here are a few examples, chosen to illustrate the geographical range: in the New World, American Indian Pidgin English in northeastern US, the African-American creole Gullah in coastal southeastern US, Jamaican Creole English in the Caribbean, Miskito Coast Creole English in Nicaragua, Sranan and Ndyuka (both are English-lexifier creoles) in Suriname; in Africa, Cameroonian Pidgin English; in the Pacific, Chinese Pidgin English, Tok Pisin in Papua New Guinea, Hawaiian Pidgin English (now nearly extinct), Hawaiian Creole English, Pitcairnese (a creole), Kriyol in the Northern Territory of Australia. Some of these are trade pidgins (American Indian Pidgin English, Cameroonian Pidgin English, Chinese Pidgin English, probably Hawaiian Pidgin English); others emerged out of the Atlantic slave trade (Gullah, Jamaican Creole English, Miskito
Coast Creole English, Sranan, Ndyuka), and still others arose when free, or semi-free, workers were imported to plantations (Tok Pisin, Hawaiian Creole English, and perhaps Kriyol—all three of these are creolized pidgins). One creole, Pitcairnese, stands out from all the rest, and indeed from all other known creole languages, because we know precisely both the date on which the stage was set for its formation and the backgrounds of its creators: it developed after nine mutineers, having evicted Captain Bligh from the *Bounty* in 1790, settled on Pitcairn Island with sixteen speakers of the Polynesian language Tahitian. (It’s possible that a few of the sixteen Polynesians instead spoke languages closely related to Tahitian, and most of the male Polynesians died or were killed too soon after the group arrived on Pitcairn to have contributed much to the emerging creole. Still, we have much more information about the starting point for this creole than for any other.)

Similar lists of languages and circumstances can be given for pidgins and creoles with French, Spanish, and Dutch lexicons, though these are neither as numerous nor as wide-ranging geographically as the English-lexifier languages. There are also a few German- and Russian-based pidgins and creoles, for instance the Pidgin German that was once spoken in New Guinea and the Chinese Pidgin Russian that was used for trade on the Russo-Chinese border, e.g. in Kjaxta, south of Lake Baikal. One trading pidgin, Russenorsk, was used in northern Europe in trade between Russians, Norwegians, and (to a lesser extent) Samis (Lapps), Low German speakers, and Swedes. Two Basque-lexifier pidgins were used for trade in the Atlantic, one in Iceland and one in Newfoundland.

Most known pidgins and creoles with non-European lexifier languages also arose as a direct result of European expansion. In the northeastern United States, for instance, Pidgin Delaware was used between European traders and missionaries, on the one hand, and Delaware Indians on the other. Pidgin Inuit (Eskimo) was used for trade with Europeans in Alaska; Sango—a pidgin-turned-creole with the Ubangian language Ngbandi as its lexi-
fier language, spoken in the Central African Republic—apparently arose as a pidgin among local Africans and then African workers who accompanied Frenchmen and Belgians as they traveled up the Ubangi River; the now-extinct Nootka Jargon (actually, probably, a pidgin language) developed on Vancouver Island in what is now southwestern Canada, between European explorers and Nootkas; and Fanagaló, a Zulu-lexifier pidgin, arose (like several other Bantu-lexifier mining pidgins) in Europeans’ African mines, specifically, in this case, in South Africa.

There may have been, and there may even still be, a great many pidgins and perhaps creoles that have nothing to do with Europeans or European expansion. As noted above, our chances of finding out about such contact languages are relatively slim: in most cases most of the speakers would have been illiterate, so the languages can attract scholars’ attention only if they happen to be noticed and identified by outsiders. Nevertheless, although we can’t know how many there have been, we do have evidence and direct attestations of quite a few of these languages. Bazaar Malay, for instance, is used in Malaysia for intergroup communication in the market squares. Hiri Motu (also called Police Motu), which takes its lexicon from the Melanesian language Motu and which arose through trade relations with neighboring groups, is one of the major national languages of Papua New Guinea. Chinook Jargon, a pidgin once spoken widely in the Pacific Northwest of the US and neighboring British Columbia, probably arose before European contact through trade with Native neighbors and (as with the Nez Percé Jargon mentioned above) in communication between the Lower Chinook people and their slaves. The Nez Percé Jargon itself, if it was a full-fledged pidgin language, also falls into this category. The Pidgin Inuit that was used in trade with Athabaskan speakers is not directly attested, but apparently it differed greatly from the Pidgin Inuit that developed in European-Inuit contacts.

A final comment on the distribution of pidgins and creoles: contrary to what common
sense might suggest, it isn’t safe to assume that the presence of one contact language ensures
the absence of one or more others. Sometimes, to be sure, overlap seems to be avoided. It is a
striking fact that the pidgin Chinook Jargon stopped spreading eastward at about the same
point that the Plains Indian Sign Language—which fulfilled a pidgin’s functions—stopped
spreading westward; in northwestern Montana, for example, the Montana Salish people
knew the sign language but not Chinook Jargon, while the tribes just west of Montana knew
Chinook Jargon but didn’t use the sign language.

Nevertheless, there are several known instances (and possibly many unknown instances)
of co-existing pidgins in the same areas, and there are also cases of co-existing, co-territorial
pidgins and creoles. During the 17th century, Pidgin Delaware and American Indian Pidgin
English co-existed in parts of their territory in the northeastern US, for instance, and Pidgin
Hawaiian and Hawaiian Pidgin English apparently overlapped in Hawaii in the 19th century.
In their trade relations (though perhaps not with the same trading partners) Motu speakers
used both Hiri Motu and two or more pidgins based lexically on non-Austronesian languages
(Toaripi Hiri Trading Pidgin, Koriki Hiri; hiri in the names of these languages comes from
the term for the Motus’ annual trading voyages along the New Guinea coast). In maps
in the Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas, Pidgin Fiji and Pidgin Hindustani overlap in parts of Fiji (vol. 1, Map 30), as do Pidgin Chukchi and Pidgin Russian in northern Siberia (Map 110), and
Inuit-Russian Pidgin and Alaskan Russian Pidgin in Alaska (Map 110). Other examples of
overlapping pidgin territories can also be found. And perhaps it shouldn’t be surprising that
the same set of languages in contact can give rise to more than one lingua franca: as we have
seen, there are different motives for developing a pidgin, and it’s easy to imagine more than
one motive being present in the same groups of speakers, or even the same motive pushing
the groups to form and then use different pidgins.
As for pidgins co-existing with creoles, every pidgin that became a creole by nativization must have passed through a stage in which some people spoke the language natively while others spoke it as a second language. So, for instance, through most of the 20th century Hawaiian Pidgin English was spoken side by side with Hawaiian Creole English; the same is true of Tok Pisin, although (unlike the Hawaiian languages) its creolization process was strictly social, with no significant linguistic consequences, because it had already expanded linguistically long before it acquired native speakers.

**Pidgins and creoles are not maximally simple and not all alike**

The title of this section may be a bit puzzling: why should anyone expect pidgins and creoles to be either super-simple or identical? The two expectations are connected and belief in them is common, but it is asymmetrical for pidgins and creoles. Pidgins (but not creoles) are widely believed to be maximally simple, but both pidgins and creoles are widely believed to be very similar to each other structurally. Each of these widespread beliefs stems from two main sources.

Both beliefs surely arose in the first place because pidgins in particular seemed to early investigators to be ‘reduced’—simplified—versions of their lexifier languages, and because of perceived similarities in the structures of pidgin and creole languages. Early specialists noticed a number of similarities immediately and found them important enough to motivate the hypothesis that all these contact languages are in fact alike and simple. The features that are most often proposed as pidgin/creole universals are a lack of morphology, a lack of ‘exotic’ sounds and complex consonant clusters, SVO word order, and, in creoles only, a particular distribution of particles indicating tense, mood, and aspect (TMA). Or, to put it more generally, since all these features (except perhaps the TMA patterns) are universally
unmarked, the typical hypothesis is that pidgins and (at least if they emerge abruptly, rather than evolving from well-established pidgins) creoles specialize in unmarked linguistic features.

This hypothesis, based as it is on the empirical evidence of perceived similarities, makes good intuitive sense too when one is making predictions about the linguistic outcome of processes of pidgin/creole genesis. First, for the maximal simplicity claim, there’s the standard view of the social conditions under which a classic trade pidgin arises: new contact situation, no shared language, need for intergroup communication, very limited communicative functions (like the stereotypical “you give me gun, me give you beaver pelts”). Under these conditions, one might reasonably expect that an emergent pidgin will have only the most rudimentary structure and a minimal vocabulary. That doesn’t, of course, entail a prediction that creoles will also have rudimentary structure and lexicon: a creole, by definition, serves as the main language of a speech community, so it ought to have all the lexical and structural resources of any other language—and all known creoles do in fact meet this expectation.

The nature of a new pidgin’s expected rudimentary structure is in part predicted by the second main reason for the belief in similarity across all pidgin and creole languages. This is essentially the reason outlined in Chapter 6 under Mechanism #4, “Negotiation”: when nobody in a contact situation knows anybody else’s language—that is, in a classic pidgin or (abrupt) creole genesis situation—guesses about what the other person will understand are most likely to be accurate when they involve universally unmarked features, both because all the languages in contact are likely to share numerous unmarked features and because unmarked features are easier to learn than marked features. The idea, then, is that if all pidgins and creoles develop through communication among groups that share no common language, their structures will all have predominantly unmarked features; and since unmarked features are supposed to be universal, this in turn should mean that all pidgins and creoles will be
structurally similar. Pidgins should be similar to each other in their rudimentary structures, and creoles, though not limited in their range of structures, should resemble each other closely.

Neither prediction holds up completely to close scrutiny, however. They are certainly not totally wrong; pidgins that are used for limited functions do have less overall morphosyntactic structure than any of the languages whose speakers created them, and markedness does clearly play an important role in the linguistic outcome of pidgin and creole genesis processes. Most pidgins and creoles either lack morphology entirely or have very limited morphological resources compared with those of their lexifier and other input languages. Morphology also tends to be extremely regular when it does exist in pidgins and creoles, without the widespread irregularities that are so very common (to the distress of students of foreign languages) in other languages’ morphological systems.

Nevertheless, in their strong form, both predictions are false. The empirical basis of the hypotheses turns out to be misleadingly limited, primarily to languages that have western European lexifier languages and (for creoles) primarily to languages that have related and often typologically similar languages of western Africa as their substrate languages. The western European lexifier languages are typologically similar in a number of respects, as almost all of them are both genetically related to each other and members of the Standard Average European linguistic area. So on typological grounds alone, it’s not surprising that Caribbean creoles share a sizable number of structural features. In any case, once one looks beyond the most-studied pidgins and creoles, and especially at those with non-European lexifiers, it is immediately clear that pidgins and creoles display much more structural diversity than specialists used to envision.

Some of the problems with the common-sense predictions about pidgin/creole structures have to do with processes of pidgin and creole genesis, and these will be discussed in the
next section. Some obvious general points can be made here, though. For one thing, the communicative purposes for which pidgins are used are usually not confined to basic trade transactions; either from the beginning or soon after the pidgin crystallizes, a pidgin is very likely to serve at least a few other purposes as well, for instance for domestic conversations between slaves and masters. The structure of a pidgin at the time of crystallization may therefore not be minimal. (A vexing practical problem is that determining what structural features a pidgin or creole had when it first arose ranges from difficult to impossible, because the language that is available for study is all too likely to have changed between its time of origin and its time of attestation. Most known creoles have a presumed history of 300-500 years, ample time to allow for structural changes, and even some pidgins are 200 years old or more. But using the standard methods of historical linguistics makes it possible to support hypotheses about initial pidgin and creole structures—except in those cases for which we have relatively early attestations, such as Sranan.)

In addition, even the most rudimentary pidgin structures might not be similar, because the creators’ native languages play a role in the selection of features that go into the emerging pidgin. To justify the title of this section, therefore, let’s look at some examples of non-simple, non-similar pidgin and creole structures.

One recurring claim about these contact languages is that their phonological structures are simple, consisting of unmarked consonants and vowels arranged in simple syllables with at most one consonant at each end; it is often proposed that pidgins and creoles tend to have CV structure, with no syllable-final consonants at all. This fits with the markedness prediction, since CV is generally considered to be the least marked syllable type cross-linguistically. The consonant phonemes of Chinook Jargon, together with the pidgin’s highly marked consonant clusters, offers perhaps the most striking violation of this prediction; and since it’s unlike the consonant system of any other known pidgin or creole, it also violates the all-structures-equal
prediction about pidgins and creoles. The system does, however, make areal sense in view of the fact that its features match those of the Pacific Northwest Sprachbund which was outlined in Chapter 5. Table 1 gives the basic consonant inventory (each item in the table is a single unit phoneme; columns 3-5 are affricates).

```
p  t  ts  tš  k  kʷ  q  qʷ  ?
p’ t’ tl ts’ (tš’) k’ kʷ’ q’ qʷ’
b  d  g
l  s  š  x  xʷ  X  Xʷ
m  n  (N)
r  l
w  y
```

**Table 1: Chinook Jargon consonant phonemes.**

As with all pidgins (indeed, as with all languages, period), there was some variation among Chinook Jargon speakers: many Native American speakers of the pidgin didn’t use either /b/ or /d/, for instance, and only the most talented and perceptive Whites used glottalization, uvulars, the /ts/ affricates, and the lateral fricative and affricates. But there is enough consistency in 19th- and 20th-century documentation of the pidgin to justify this elaborate inventory, with such marked (and non-Standard Average European) consonants as the ones most White speakers of the pidgin didn’t learn. Similarly, there is ample evidence to support the claim that Chinook Jargon had unusual non-SAE-ish consonant clusters like /tq’/ in /tq’ix/ ‘want, like’ and /ptš/ in /ptš@x/ ‘green’. Nor is this the only pidgin or creole with highly marked phonological features; both the English/Portuguese-lexifier New World creole Saramaccan and the creolized African pidgin Sango, for instance, have double-articulated stop phonemes /kp/ and /gb/ (in addition to /p/, /k/, /b/, and /g/). All these
unusual phonological features occur in at least some of the languages spoken by the contact languages’ creators, indigenous languages of the Northwest and of West Africa, respectively.

Another common claim about pidgin and creole phonological systems is that tone distinctions, even if present in some or all of the languages spoken natively by the creators of a pidgin or creole, are typically absent in the new contact language. This too is false, though it’s certainly true that many pidgens and creoles created by native speakers of tone languages have no tones. It is also true that, even when pidgens and creoles have tones, the tones often have fewer functions than in any of the input languages that have tone phonemes. But tone languages aren’t especially hard to find among contact languages. Ndyuka, for example, has at least two phonemic tones, unlike its lexifier language (English) but like the African languages spoken by its creators; Saramaccan is also a tone language, like the African languages spoken by (some of) its creators. Sango has tone phonemes, as does its lexifier language Ngbandi, though in Sango the tones serve fewer functions than in Ngbandi: they distinguish words in both languages, but they are used much more extensively in Ngbandi to make grammatical distinctions. The same is true of Kitúba, a Kikongo-lexifier creole spoken in Zaire and the Republic of the Congo; but in Kitúba only a minority of the lexical items have distinctive tones. Tày B`ôi, the French-lexifier pidgin once spoken in French-occupied Vietnam, was said to have two different phonologies that matched the systems of the two input languages; French speakers thus presumably lacked tones in their version of the pidgin, but for Vietnamese speakers Tày B`ôi was presumably a tone language.

Perhaps the best place to look for nonphonological structures that aren’t maximally simple is in pronominal systems, and one of the best places to look for elaborate pronoun systems is in islands of the Pacific Ocean, where many Austronesian languages and some others too have such pronominal categories as inclusive vs. exclusive ‘we’ and dual, trial, and occasionally paucal number as well as singular and plural (‘paucal’ means ‘a few’).
In Tok Pisin, for instance, all the pronominal morphemes are ultimately derived from English, but the system certainly is not. It lacks the English gender and case distinctions (which are unknown in the languages of the original pidgin’s Austronesian-speaking creators) but has typical local categories of person and number, with exclusive vs. inclusive ‘we’ and a system of pronominal number divided into singular, dual (two people or things), trial (three people or things), and plural (more than three people or things—as opposed to a simpler system like that of English, where plural refers to more than one). Table 2 gives the forms.

<table>
<thead>
<tr>
<th>SINGULAR</th>
<th>DUAL</th>
<th>TRIAL</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st exclusive:</td>
<td>mi</td>
<td>mitupela</td>
<td>mitripela</td>
</tr>
<tr>
<td>1st inclusive:</td>
<td>yumitupela</td>
<td>yumitripela</td>
<td>yumipela</td>
</tr>
<tr>
<td>2nd:</td>
<td>yu</td>
<td>yutupela</td>
<td>yutripela</td>
</tr>
<tr>
<td>3rd:</td>
<td>em</td>
<td>tupela</td>
<td>tripela</td>
</tr>
</tbody>
</table>

**Table 2: Tok Pisin pronouns.**

The English sources of the morphemes in the Tok Pisin pronouns are *me, you, him, two, fellow (“fella”), three, and all*, combined in very un-English ways. It’s worth noting that this system was in place before Tok Pisin expanded dramatically in its functions, i.e. while it was still an ordinary non-expanded pidgin. In spite of the absence of the English pronominal categories of case and gender, no one would want to claim that this is a simple system: it is as complex as the pronominal systems of the Austronesian input languages, though somewhat more transparent in its formation than most of them.

It’s easy to find other fairly complex grammatical subsystems and individual features in pidgin and creole languages. The African creole Kitúba, for example, is based lexically on the Bantu language Kikongo, and it was apparently created by speakers of various Bantu languages with possible input also from West Africans who accompanied Europeans in their
journeys to what later became Zaire and the Republic of the Congo. Unlike most other Bantu-based pidgins and creoles, Kitúba has a number of Bantu morphosyntactic features. The two that are most striking are a fairly elaborate system of prefixal noun classes, fewer than in Kikongo but still quite complex, and a tense/aspect system basically like that of Kikongo in categories, though not in form: the Kitúba tense/aspect system is regularized, and, though two tenses are formed with suffixes, most of the other tense/aspect categories are expressed by free particles—which, however, is also in part a feature of relevant dialects of Kikongo.

Yet another claim that is often made about pidgins and creoles is that they lack morphology, especially inflectional morphology; as noted above, these contact languages do tend to lack morphology, but the strong claim is nevertheless unjustified. The two Kitúba tense suffixes are counterexamples to this claim; other examples are the transitive suffix of Tok Pisin, the transitive suffix and causative affix of Hiri Motu, the participial suffix -ing of Pitcairnese, the reflexive suffix (from Russian) and the imperfective aspect suffix (from Mongolian) of Chinese Pidgin Russian, the diminutive and augmentative prefixes of the Portuguese-lexifier creole Fa d’Ambu, the negative prefix of Sranan, the three aspect suffixes (perfective, imperfective, and iterative, all from the substrate language Eastern Íjó) of Berbice Dutch Creole, and agreement morphology in the Lingua Franca noun phrase. These examples are only a sample; morphology is not particularly uncommon in pidgins and creoles, although no pidgins or creoles have morphological systems as elaborate as those of their most elaborate input languages.

The consonant phonemes of Chinook Jargon, the tone phonemes of Ndyuka and several other pidgins and creoles, the personal pronouns of Tok Pisin, the noun classes of Kitúba, and most of the specific examples of morphology also provide evidence against claims about similarity among all pidgins and creoles, since none of these features is common in these
contact languages worldwide. Other examples of structural diversity can be found in word order features of various pidgins and creoles. In particular, it has sometimes been claimed that SVO sentential word order is universal among creoles, even those which developed with major input from languages with other word order patterns (usually SOV); the claim has sometimes been extended to pidgins as well. Certainly this seems to be true of all well-studied languages that are generally agreed to be creoles, in the Caribbean and elsewhere: all the Caribbean creoles have SVO word order, as do the creoles of the Indian Ocean and elsewhere, e.g. Tok Pisin and the Arabic-lexifier creole Nubi, which is spoken in Kenya and Uganda. (The universal-SVO word order claim applies only to the initial word order pattern, of course. Once they are crystallized into languages, creoles, like all other languages, will undergo both internally- and externally-motivated changes, and these changes may include word order alterations; so, for instance, Sri Lanka Portuguese Creole has changed from SVO to SOV order under the influence of neighboring Dravidian languages.) Many pidgins have SVO order too, among them Russenorsk, Chinese Pidgin English, and Fanagaló.

But it isn’t true of all pidgins. The Ndyuka-Trio pidgin, for instance, is SOV in sentences with noun subjects but usually OSV in sentences with pronoun subjects—like one of its input languages (Trio) but unlike the other (Ndyuka, an SVO creole). In Hiri Motu (whose lexifier language, Motu, has rigid SOV order, as do most or all of the other input languages), the word order is SOV when both S and O are nouns; OSV or, more rarely, SVO when S is a noun and O is a pronoun; OSV when S is a pronoun and O is a noun; and either OSV or SVO when both S and O are pronouns. Pidgin Delaware had both SVO and SOV order, neither obviously predominant, and Chinook Jargon typically had VS order with adjectival predicates (in a translation of English I’m hungry, for instance) and SVO order with all other predicates. Chinese Pidgin Russian has (or had) SOV order, as does Nagamese, an Assamese-lexifier pidgin spoken in northern India (Assamese belongs to the Indic subbranch.
of the Indo-European language family). The situation is more complicated in Pidgin Yimas(-Arafundi), a trade pidgin with mainly Yimas lexicon but also with numerous Arafundi words that is spoken in the Sepik River basin in northern New Guinea (see Chapter 5 for a discussion of the Sprachbund in this region): this pidgin has verb-final sentential word order but free ordering of subject and object, so that both SOV and OSV occur frequently.

Another word order feature that varies widely among pidgins and, in this case, among creoles too is the placement of the negative element(s) in a sentence. In many creoles, the neg element appears at the beginning of the verb phrase, as in Papiamentu mi no ta bini 'I'm not coming' (lit. 'I NEG future come'), French Guiana Creole mo pa té travaille 'I hadn’t worked' (lit. 'I NEG tense work'), and Tok Pisin yu no kan go long biglain 'you can’t go to the work-group' (lit. yu NEG can go to work-group'). Some pidgins also have this ordering, as in late 19th-century Chinese Pidgin English maj no hav kači buk 'I haven’t borrowed a book' (lit. 'I NEG have caught book'). But other creoles and pidgins have different orders. In Berbice Dutch Creole, for instance, the negative element is clause-final, as in Ek suk mu lasan eni ka 'I didn’t want to leave them' (lit. 'I want go leave 3pl NEG'). In Chinese Pidgin Russian the negative element usually follows the verb. Fa d’Ambu has double negation, with one negative element before the verb and the other attached to the last word of the sentence, as in Odyai amu na be mem=bo xama-kumu-f 'I didn’t see your mother at the market today' (lit. ‘today I NEG see mother=your place-food=NEG’). In Chinook Jargon, as in all of the native Pacific Northwest languages, negation is almost always sentence-initial, as in Hilu naika kamtaks ‘I don’t understand’ (lit. ‘NEG I understand’), and in Pidgin Delaware the negator also usually precedes the subject, as in Matta ne kamuta ‘I didn’t steal it’ (lit. ‘NEG I steal’).

The purpose of this section has not been to convince readers that pidgins and creoles differ from each other in all possible respects. There are some striking similarities among
these contact languages, for instance in the famously uniform, or at least similar, placement of tense, aspect, and mood markers in many creoles. These markers typically occur in the order tense-mood-aspect, as in the following examples with anterior tense (‘before the event under discussion’), irrealis mood (‘unreal’), and nonpunctual aspect: Saramaccan Mi bi-o-tá-nján dí fisi ‘I would have been eating the fish’ (lit. ‘ANTERIOR-IRREALIS-NONPUNCTUAL-eat the fish’ and Fa d’Ambu Ineni bi ske xa tabaya ‘they would have been working’ (lit. ‘they ANTERIOR IRREALIS NONPUNCTUAL work’). At a less specific level, non-expanded pidgins in particular typically share such general linguistic features as a small lexicon, lack of extensive morphology, and lack of such (morpho)syntactic resources as clearly marked subordination and a fully elaborated tense/aspect system; as noted above, they also tend to lack a full range of social functions, being restricted to such limited functions as trade or master/slave communication.

The goal of the section, instead, has been to emphasize the fact that similarity and (in pidgins) simplicity are not the whole story, as some traditional approaches to the study of these contact languages have suggested. This emphasis is needed when we move on to a consideration of the major theories of pidgin and creole genesis, because the diversity in pidgin and creole structures reduces the appeal of some of the theories under debate: any successful pidgin/creole genesis theory must account for the diversity within this set of languages as well as for the similarities among them.

*Pidgin/creole genesis theories: where does the grammar come from?*

Again we need to begin with a definition: in this book, ‘pidgin/creole genesis’ is any process through which a pidgin or creole language comes into being. The rather clumsy phrasing of the definition has a purpose that will become evident when we examine gradualist
In any consideration of pidgin and creole genesis, the problem is to figure out the source(s) of the languages’ structures, since the lexical items are (usually) easy to trace to one or more sources. In most cases there is one main lexifier language, with a small to medium-sized set of words from one or more other languages. Some pidgins and a few creoles have more heterogeneous lexicons, but here too there’s no great difficulty in finding the sources. Saramaccan, for instance, has (by one estimate) about 50% English words and about 35% Portuguese words, together with a sizable number of words from African languages, in particular the Bantu language Kikongo and the Kwa languages Ewe, Fon, and Twi. Most of the lexicon of early Chinook Jargon, including the bulk of the basic vocabulary, comes from the Lower Chinook language, but two or three dozen words (some of them quite basic) are from Nootka, and a smaller number of words come from Salishan and other languages of the region. In addition, quite a few words entered the pidgin from French and then English, especially after the mid-19th century. There’s no mystery, therefore, about the lexical sources of pidgins and creoles. All the controversy centers on the process(es) by which the grammars emerged.

Treating this complex set of issues in a one section of one chapter risks obscuring a major topic of debate: can the question of pidgin and creole origins reasonably be considered as a single problem, or should they be considered separately? This question has no simple answer, for at least two reasons. First, it depends on what type of creole we’re talking about: the development of a creole out of a stable pidgin via nativization has no direct parallel in pidgin genesis, so any parallels that exist in processes of pidgin and creole genesis must be sought in creoles that develop without going through a stable pidgin stage. Second, and this is related to the first reason, it appears that both pidgins and creoles have emerged through a variety of developmental processes, so that a search for the genesis process for either or both sets of languages is misguided. And, as we will see, most of the processes involved do appear to
have been operative both in pidgin genesis and in creole genesis.

I believe that some, perhaps many or even most, of the controversies surrounding the topic of pidgin and creole genesis will vanish if we recognize that the common assumption of a single universal developmental route to creole genesis—which is the main locus of the controversies—is unmotivated. If specialists agree that there are several ways in which creoles (and pidgins) arise, then it may turn out that several theories are viable, but for different languages. An empirical question that arises here is whether different genesis routes lead to different linguistic results; unfortunately, and this no doubt accounts in part for the uniform-genesis assumption, they apparently don’t always do so (see below for a bit more discussion of this point).

From a historical linguist’s viewpoint, a vitally important consideration is that an origin theory that includes both pidgins and creoles is to be preferred, all other things being equal, to a theory that requires totally different routes of development for pidgins and creoles. This is a standard principle in any historical linguistic analysis, as it is in historical (and other!) sciences generally: if two theories account equally well for the data, then, all other things being equal, the simpler and more comprehensive theory wins. In the case of pidgins and creoles, an approach that comprehends both will be simpler than an approach that requires separate genesis scenarios for these two types of contact language. Of course scholars can and do disagree sharply about what counts as “all other things being equal”. But in the present context, the appeal of a theory that accounts equally well for pidgin and creole genesis is that pidgins and creoles obviously do share certain linguistic and social features, as we’ve seen; in particular, both arise under circumstances in which at least some of the groups in contact don’t learn any of the other groups’ languages—that is, under conditions where we’d expect to find linguistic effects of imperfect learning.

I should acknowledge my own bias as a historical linguist: given the things that pidgins
and creoles have in common, I prefer to try for a simpler unified approach to the genesis problem, rather than starting from the assumption that pidgins and creoles follow entirely different developmental paths. That’s why I’m treating both in a single section. As I’ll argue below, this approach works well in many cases, but it doesn’t always work, because clearly there are also some significant differences between pidgin and creole genesis processes.

With these preliminary comments out of the way, we’ll begin our examination of pidgin/creole genesis theories by surveying some of the major former and current approaches.

Monogenesis

One of the oldest theories about pidgin and creole genesis is the MONOGENESIS HYPOTHESIS. In its strong form, this hypothesis states that all pidgins and creoles are descendants of the original Lingua Franca of the Mediterranean, albeit with relexification—lexical replacement—for all pidgins and creoles that don’t have Italian lexicon, i.e. almost all known modern pidgins and creoles. Most proponents of the hypothesis posit as the first step the development (out of the Lingua Franca, or at least under the influence of the Lingua Franca) of a Portuguese-lexifier pidgin, which became well known in West Africa. The idea is that European explorers, traders, colonizers, and slavers learned West African Pidgin Portuguese and took it along in their travels around the globe, replacing its Portuguese lexicon with that of their own native language. Relexification also happened, according to the hypothesis, in the formation of non-European-lexifier pidgins and creoles. This would not be descent in the historical linguist’s framework of genetic linguistics, because a process of lexical replacement doesn’t fit the standard notion of descent with modification; but, if correct, the hypothesis would nevertheless link the Lingua Franca directly to the vast majority of the world’s pidgins and creoles. One obvious exception would be Maridi Arabic, which predates the Crusades; other exceptions would be pidgins and creoles that arose completely independently
of European expansion, such as the indigenous trade pidgins of New Guinea. Moreover, the monogenesis hypothesis can account for creoles only if all creoles arose via nativization of well-established pidgins; but, as we’ll see below, this doesn’t seem to be the case.

The appeal of the monogenesis hypothesis is that it would account nicely for the structural similarities among pidgins and creoles. This no doubt explains its popularity in the 1960s and 1970s, when most specialists viewed pidgins and creoles as grammatically very similar. But that was early in the history of pidgin/creoles studies—the field is quite new, in spite of early pioneers like Hugo Schuchardt in the late 19th and early 20th centuries and Robert A. Hall, Jr., in the middle decades of the 20th century. The last few decades have seen greatly increased intensive study of pidgin and creole languages all over the world, so that we now know much more about their structural diversity than we used to, and also much more about the fairly numerous pidgins that lie well off the beaten track of European exploration and colonization. The monogenesis hypothesis can’t account for any of the structural diversity that is now known to exist, and it also lost much of its initial appeal once its claim to (near-)universality turned out to be unjustified. Modified versions of the hypothesis are still taken seriously, for instance for all English-lexicon Caribbean creoles, which some scholars believe to be direct descendants, in the standard historical linguist’s sense, of a pidgin that arose on the coast of West Africa. But the strong form of the hypothesis is no longer considered a viable candidate in the search for a global route to pidgin and creole genesis.

_Abrupt genesis scenarios_

The major current views of pidgin/creole genesis fall into two major categories, with several subcategories in each. The focus of theoretical research has long been on creole genesis rather than on pidgin genesis, in part because Caribbean creoles and, to a lesser extent, the European-lexifier creoles in Africa still receive much more attention than pidgins.
The reason for this emphasis probably has to do ultimately with numbers of speakers: many or most of the most interesting pidgins are now extinct or nearly so, while most known creoles are flourishing as the primary languages of large speech communities. In any case, most people who propose contact-language genesis theories talk only about creole genesis.

The two categories of theories can be characterized as abrupt vs. gradual genesis scenarios. The simplest way to think of contact-language genesis seems to me to be abrupt creation, because it fits the traditional view of the onset of the appropriate kind of language contact, especially for pidgins: a new pidgin arising in a new multilingual contact situation for use in limited domains, a new creole arising in a new multilingual contact situation for use in all domains. In this scenario, the creole arises without going through a stable pidgin stage. The idea of ‘abrupt’ has to be somewhat elastic here; it isn’t meant to suggest an overnight event. Estimates of the time required for abrupt contact-language genesis range from a few years to 25 years, because time is needed for full crystallization of the language in the contact community—development of a shared vocabulary and grammar that must be learned by anyone who wishes to speak the language, including native speakers of the lexifier language.

Abrupt creolization scenarios in turn fall into three different subcategories. Derek Bickerton’s Language Bioprogram Hypothesis (LBH), one of the earliest important theories in this category, treats creole genesis as the outcome of first-language acquisition in a context of restricted linguistic input from the surrounding community. According to Bickerton, creoles—specifically plantation creoles—arise in new contact situations where a new community-wide language is needed immediately for all purposes. The adults in the community communicate with each other in a macaronic pre-pidgin—no stable shared grammar at all—and their children, growing up with only the unstable pre-pidgin as input for their language-learning task, construct a grammar derived from grammatical structures that are literally genetically
programmed, hard-wired in every newborn human infant’s brain. (He allows for some time, a generation or two, to crystallize the new creole’s grammar.) Bickerton finds support for his theory in aspects of ordinary first-language acquisition as well as in creole genesis, and he argues that children who learn languages with normal input do not display the Bioprogram features because those features are overlaid by the grammatical features of their community’s language, English or Quechua or Tagalog or whatever it is. The LBH could account for the shared features in the classic creoles, especially the tense/aspect/mood features; it cannot of course account for the diversity found in plantation creoles, and it is not meant to account for any grammatical features of non-plantation creoles like Pitcairnese and Tok Pisin, or for grammatical features of fully crystallized pidgins, which in everyone’s views are adults’ (or at least not primarily children’s) creations. The LBH also doesn’t account for shared nonsyntactic features of plantation creoles, though it could in principle be extended in that direction.

The LBH was especially popular in the 1980s, but is now considered a viable theory mainly by formally-oriented linguists who are not creolists. The reasons for most creolists’ skepticism about it are partly historical: detailed investigations of the demographics of slave populations in numerous Caribbean colonies have revealed that there were few slave children for the first fifty years or more—it was cheaper to import adult slaves, and in any case women slaves were all too often barren after suffering the rigors of travel in slave ships from Africa to the Caribbean. Presumably because of such findings, Bickerton now refers mainly to Hawaiian Creole English as an example of a creole resulting from the operation of the LBH.

Other reasons for skepticism about the LBH are social. It seems very unlikely that whoever raised slave children (when there were any) would have spoken to them exclusively in a macaronic pre-pidgin; all the adults in the community, after all, knew at least their
native languages, and they would inevitably have talked to children in those languages. So the input received by infants was surely never restricted to a pre-pidgin. In addition, slaves who were born and raised in the Caribbean were likely to have relatively privileged positions in the plantation societies, with greater access to the lexifier language.

Still another reason for skepticism has to do with the lack of generality, even for plantation creoles: if the LBH can’t account for syntactic structures of many Caribbean creoles because there were too few children on hand to manifest the LBH features there, and if Caribbean creoles nevertheless share the typical plantation creole syntactic features (like the widespread TMA patterns, which were first noticed in the Caribbean), then obviously those creoles had to get the shared features in some other way. We still need an explanation for those shared features; the LBH isn’t available for those languages; so why suppose that it was operative even in the Hawaiian Creole English context, where there apparently were enough children early enough to make its operation possible? Why not try, instead, to find a single explanation that works for the shared features in Hawaiian Creole English and also in Caribbean and other plantation creoles? Such reasoning has led most creolists to reject the LBH. But, although Bickerton’s strong claims for the LBH seem shaky on the empirical evidence, his notion of universal innate features is by no means a dead issue: this is one version of a major component of several pidgin/creole genesis theories.

Another theory in the abrupt creation category is Claire Lefebvre’s Relexification Hypothesis. Although the strong version of the monogenesis hypothesis has vanished from the creolization debates, relexification continues to be discussed; it is the basis of the theory explored by Lefebvre and her colleagues in a long-term research program, the Haitian Creole Project. According to Lefebvre’s hypothesis, ‘creoles are created by adults who develop a new lexicon by combining the phonetic shapes of one language with the semantic and syntactic information of another language’; she proposes this as ‘the central process in creolization’.
Her method of searching for evidence to support this hypothesis is to compare syntactic structures of Haitian Creole, a French-lexicon Caribbean creole, with syntactic structures of Fon, a West African language that was spoken by a significant proportion of the slaves in Haiti during the creole’s formative period. The time frame for the relexification process is not specified, but of course it must have been completed while Fon was still spoken by the slaves, which argues for an abrupt rather than a gradual process. Proponents of the relexification hypothesis have not discussed its possible relevance for pidgin genesis, but in principle, if correct for creoles, it could be extended to account for pidgins too. Like Bickerton’s LBH, Lefebvre’s Relexification Hypothesis is grounded in formal syntactic theory; also like the LBH, it has met with considerable skepticism from creolists, though not primarily on demographic grounds.

In the framework presented in this book, the creole genesis process envisioned by Lefebvre is borrowing: speakers of Fon graft French phonetic strings onto their native Fon grammar. Imperfect learning plays no role in the process, because this is incorporation of French material into Fon by native Fon speakers, not Fon interference creating a Fon-influenced variety of French. The process envisioned by Lefebvre and her colleagues doesn’t match other known instances of relexification. Those few pidgins that seem to have undergone relexification are, by hypothesis at least, cases where (for instance) native speakers of Spanish relexified a Portuguese-based pidgin that they already spoke; and the case cited by Lefebvre herself as a parallel, the Media Lengua of Ecuador, is a case of language creation by bilinguals for use as an in-group language (see Chapter 8). In other words, in other cases of relexification the relexifiers have been bilingual, and they have not needed a new medium of communication for a new contact situation. The Media Lengua, in particular, is not at all similar to Haitian Creole (or any other creoles) as a social phenomenon. Moreover, the Relexification Hypothesis takes no account of any tendency, or need, for “negotiation” in the process of creating a new
contact language: its proponents do not consider the effects of other languages in the new creolizing context, but instead assume that Fon speakers will simply use their native grammatical structures in trying to communicate with other members of the community. This assumption seems simplistic in light of what is known about second-language acquisition, where even the first efforts of learners tend not to match their native-language structures exactly, and about shift-induced interference, where interference features very often don’t match the source features in the shifting group’s language exactly.

The third major line of research on abrupt creolization, in contrast to Bickerton’s exclusively universalist/innatist approach and Lefebvre’s relexification approach, is based on the assumption that both pidgin and creole genesis are akin to second-language acquisition and thus to processes of shift-induced interference. There are several versions of this approach; I’ll describe my own, but it’s compatible with most of the other approaches. This approach encompasses both pidgin and creole genesis; the basic scenario can be constructed by combining the discussions of mechanisms #4 and #5 in Chapter 6, “negotiation” and second-language acquisition strategies. All the strategies used in pidgin/creole genesis according to this hypothesis will usually involve “negotiation”, however, because in most cases there’s probably no serious effort to learn the lexifier language as a whole: instead, the learning process is one in which the people in the new contact situation learn to communicate with each other by deploying the new vocabulary with grammatical structures they hope will be understood by their interlocutors. The idea is that people’s “right” guesses about what the others will understand become part of the emerging contact language. The hypothesis does not predict that they will simply make use of their native structures; on the contrary, it predicts that they will likely abandon those native structures that aren’t understood by speakers of the other languages. The structures they settle on will be those best understood by all the other people—primarily unmarked structures, but also marked structures that are
common in most or all of the languages in contact. The resulting pidgin or creole grammar, on this hypothesis, is a cross-language compromise among the languages of the pidgin/creole creators.

This approach can account for features that are widely shared among pidgins and creoles, to the extent that those features are universally unmarked and/or derivable from most of the native languages spoken by the creators of the contact languages. The approach has been criticized because the famously shared features (for instance the tense/aspect/mood patterns in creoles) don’t all seem to be shared by the languages of the creoles’ creators; they may be universally unmarked in some sense, but that has not been demonstrated. But caution is needed in interpreting the patterns historically. First, Caribbean creoles all have members of the ‘Standard Average European’ (SAE) linguistic area as lexifier languages (see Chapter 5), and there is considerable typological congruence among some groups of substrate languages as well, e.g. the Kwa languages that were so prominently represented at the relevant period in the history of Haitian Creole. Even Hawaiian Creole has an SAE language as its lexifier; and nobody has compared its substrate languages systematically to find out whether they share relevant features with the substrate languages underlying Caribbean and other plantation creoles.

Second, the notion of a cross-language compromise via “negotiation” does not predict that the structure of the resulting contact language will be a lowest common denominator of the structures of the input languages; instead, it is the cumulative result of different speakers perceptions of shared structures or of their interlocutors’ unshared structures. A superficial analysis of the shared features is likely to miss crucial connections. Consider, for instance, the SVO word order pattern of Caribbean creoles. This word order is present in most of the European lexifier languages (English, French, Spanish) and in many of the African substrate languages, so it requires no special explanation for those languages. But Berbice Dutch
Creole has SVO order in spite of the fact that it is based lexically on Dutch, an underlying SOV language (though one with SVO as a common surface order, perhaps more common in foreigner talk), and it had as its only important substrate language Ijo, which is also SOV. At first glance this seems to argue for a universalist, perhaps an innatist, explanation for the SVO order. But Silvia Kouwenberg argues instead for a process of “negotiation” on the basis of a detailed analysis of relevant structures in the two input languages:

‘...the development of [SVO] properties of BD [= Berbice Creole Dutch] is reconstructed from the application of compromise strategies in a situation in which speakers of Du and EI [Eastern Ijo] were involved in a process of linguistic negotiation. These strategies are aimed at the creation of a linguistic system which optimally exploits perceived similarities between Du and EI. Surface structures of different languages may be comparable, even when underlying structures are in fact irreconcilable. Also, marked strategies of one language may be comparable to unmarked strategies in another....the development of VO ordering in BD can plausibly be explained as continuity of comparable superficial orderings in Du and EI.’

Similarly, the English-lexifier pidgin Bislama, spoken in the New Hebrides, has developed a number of structures that are shared by English and the local Austronesian languages, although it also has structures that are found in the local languages but not in English.

It must be emphasized that there is no reason to assume that speakers of the lexifier language will always participate in the creation of an abrupt pidgin or creole: sometimes they do, as in the case of Berbice Creole Dutch, Bislama, Pitcairnese, and Pidgin Delaware, but sometimes they don’t, as in the case of Hawaiian Creole English (if Bickerton is right about its being an abrupt creole). If the lexifier language’s speakers do participate, then one
can reasonably expect the cross-language compromise to reflect that participation; if not, then the lexifier language should have little or no impact on the emergent pidgin or creole.

An approach based on “negotiation” can also account without difficulty for diversity in pidgin and creole structures. The elaborate consonant system of Chinook Jargon, the elaborate system of personal pronouns in Tok Pisin, and the noun classes of Kitúba are all quite reasonable as cross-language compromises, given that the crucial features are widely shared among the local languages spoken by the pidgins’ creators. The notion of cross-language compromise can’t be taken as a rigid prediction, however. For instance, the double-articulated stop phonemes /kp/ and /gb/ that appear in Saramaccan and Sango do not provide evidence that every input language had these phonemes: they indicate only that some or most of the input languages (or perhaps even just one prominent input language) had the phonemes. The idea of “negotiation” leading to a cross-linguistic compromise is indeed predictive, but it is loosely predictive. As with ordinary contact-induced language change, social factors that can’t be discovered after the fact will inevitably influence the specific outcome of a process of pidgin or creole genesis.

This completes the roster of major approaches to the explanation of pidginization and creolization that assume an abrupt genesis scenario. These approaches make good sense for some pidgins and creoles, for instance the English-lexifier language Pitcairnese, which was certainly an abrupt creation that crystallized as a language within one or at most two generations. (A quibble is possible here, as it is possible that one or more of the Bounty mutineers already spoke a pidgin or even a creole before the mutiny; a likely candidate would be the mutineer who came originally from St. Kitts in the Caribbean. Still, Pitcairnese itself surely arose on Pitcairn Island; it certainly has some striking Tahitian structural features, and it has not been reported to share features with Caribbean creoles that would be surprising in a two-language English-Tahitian creole.) There is also no reason to suppose that functionally
restricted trade pidgins like Pidgin Delaware and Chinook Jargon arose gradually; such pidgins are likely to have begun forming as soon as contact was established and to have crystallized with their eventual (rudimentary) grammars soon afterward. But it is becoming increasingly obvious that abrupt creation is not the only way in which pidgins and creoles arise, and maybe not even the major way, at least for creoles.

*Gradual genesis scenarios*

Of course one type of non-abrupt creole has always been recognized in the literature, namely creolized pidgins—creoles that develop from fully crystallized pidgins, often after the pidgin has undergone considerable lexical and grammatical expansion as a result of acquiring new social functions. In such cases, e.g. in Tok Pisin, the creolization process is typically gradual socially, as mixed marriages produce native speakers of the parents’ only shared language—the pidgin—and as the pidgin gradually acquires more functions even in adults’ daily lives. Linguistically, the effects of creolization in these cases may be negligible, depending on how expanded the pidgin was before it acquired native speakers. The linguistic expansion process itself is almost sure to have been gradual, of course.

But creolized pidgins are not the focus of the current boom in gradualist creolization theorizing. Instead, all these theories concern plantation creoles, mostly in the Caribbean but also in the Indian Ocean and in Africa—creoles which, in the opinion of most creolists, did not emerge from a stable pidgin. Pidgins and pidgins-turned-creoles are only rarely discussed in this literature. Several different lines of research have recently led to theories of gradual creole genesis, but these theories don’t necessarily all fit together: different authors often mean quite different things by ‘gradual creolization’, although probably all would subscribe to the characterization of creolization in at least some cases as ‘a gradual process, extending over a number of generations of speakers’.
Almost all the authors who espouse a gradualist position make extensive use of detailed demographic information about the patterns of settlement and the importation of slaves in various (eventual) plantation settings. Perhaps the most important demographic fact, as far as creole genesis theories are concerned, is the finding that there were two distinct stages in the development of the relevant agricultural societies: at first, agriculture was practiced on a small scale, and the population comprised European masters, African slaves, and (in some places) indentured Europeans who worked side by side with the slaves. Later—the time of the change varies from colony to colony, but it was sometimes 30-50 years after the founding of a colony—the economy shifted to large-scale plantations, usually sugar plantations. The workers on these plantations were overwhelmingly African slaves, who were imported in large numbers. The switchover from small farms to large plantations figures in most of the gradualist theories: with the advent of large plantations, the slave population rapidly came to outnumber their masters greatly (and indentured Europeans mostly vanished from the scene).

One prominent gradualist creole genesis hypothesis is Robert Chaudenson’s. Chaudenson’s main focus is on French-lexifier creoles in the Indian Ocean, but he argues for the general application of his theory to other French-based creoles (and indeed to all creoles). On his view, the first slaves to arrive in the colonies worked and lived with native French speakers and therefore learned French (specifically regional dialects of French)—but no doubt imperfectly, so that their variety of French incorporated shift-induced interference features from their native languages. Once the economy switched from small farms to large plantations, newly arrived slaves no longer had much (if any) contact with their French-speaking masters; they therefore learned French from the first group of slaves, so that the later arrivals’ target language was a variety of French that already showed changes from the masters’ French. Subsequent waves of slaves learned increasingly divergent varieties of French, until
at last the general language of the slaves, shared by all of them, was a creole language, not French at all. The process, on this hypothesis, was gradual divergence from the masters’ French; each new wave of imported slaves learned a TL, but the TL was in effect a moving target, with more and more interference features with each new wave of imported slaves. At each stage, the process was ordinary L2 acquisition, with perfectly ordinary contact-induced change; the only unusual aspect was the fact that it happened again and again until the resulting language could no longer be called a variety of French.

Chaudenson himself sees the resulting creoles as extensively modified French, allowing little scope for interference from the slaves’ languages; he rejects theories that propose significant structural contributions from African (or other non-French) languages. He has been sharply criticized for this insistence on French as the source of many, most, or all creole structures, given the large amount of evidence for African-language influence. He has also been criticized for his assumption that French itself (in whatever form) was the target language for successive waves of newly-arrived slaves; it seems likely that, as in many other pidgin/creole genesis contexts, there was no target language in the usual sense. But his scenario for gradual creolization does not in itself require a French-structure focus; it is compatible with theories that claim extensive structural contribution from the slaves’ languages and/or from universal structural tendencies based on markedness considerations. And in fact other gradualists do allow a greater role for substrate-language structural contribution.

The major appeal of all the current gradualist theories, including Chaudenson’s, is that they provide a reasonable way of dealing with a fact that is very awkward for any theory that assumes abrupt creolization for all plantation creoles: in colonies where slaves at first worked closely with their masters, and often with indentured Europeans as well, for thirty or more years, then they must have had adequate access to the masters’ language for language-learning purposes; so why would a creole emerge at all? The answer, for a gradualist, is
that it didn’t—not until much later, after the shift to a large plantation economy. The major appeal of Chaudenson’s particular theory is that it requires no special assumptions at all about the process of creolization: for him, creole genesis is simply a matter of repeated instances of imperfect language learning of French (and, by extension, of other European slavemasters’ languages). A possible problem for Chaudenson’s theory has to do with the size and timing of the influx of new slaves after the shift to a sugar economy: if the influx was large and fast enough, the resulting new slave society might have fit an abrupt creolization scenario better than a truly gradual genesis scenario: the new slaves might have learned their lexicon from French-speaking earlier slaves, but they might not have learned much if any of the modified French grammar of those earlier slaves. That is, the earlier contingent of slaves might not have been large enough to provide sufficient access for the newcomers to learn French. Other problems for Chaudenson’s hypothesis, especially if it is claimed to account for all French-lexifier or even all European-lexifier creoles, are demographic. In the Indian Ocean, for instance, the island Réunion may fit his scenario rather well (though there are doubters), and its creole, Réunionese, has numerous French structural features; but on the nearby island Mauritius, slaves outnumbered whites within ten years of the initial French settlement.

Other gradualist theories do not posit such a seamless progression that starts with a European language learned imperfectly from Europeans, then proceeds to an altered European language learned imperfectly from Africans, and finally results in a variety so distant linguistically from the original European target language that it qualifies as a separate creole language. Instead, these authors analyze particular linguistic features (in practice, the features are always syntactic) and conclude either that the typical (later) creole features didn’t emerge for quite a while after the founding of a colony or that the features showed a lot of variation until quite a while after the founding of a colony. The proponents of
these theories usually make extensive use of early textual data, emphasizing the structural differences between early data (e.g. the earliest Sranan text, of 1718) and later data (e.g. Sranan after 1750). Sometimes authors argue that linguistic change in a creole (or pidgin) may occur faster than linguistic change in a non-creole (non-pidgin) language. Sorting such factors out, as everyone acknowledges, is no easy matter; there is, for instance, too little cross-linguistic information available about the time required for any general kind of linguistic change to justify an argument that creoles change more rapidly than other languages. This is especially true given that ordinary contact-induced change often proceeds quite rapidly, and many changes in young creoles are sure to be contact-induced. It is even more difficult to decide when a speech form has enough structure to be called a language: as with any other linguistic phenomenon that can change into something else, the boundary between language and not-yet-language will inevitably be fuzzy. In any case, this approach raises a crucial question: how can we distinguish between a process of creole genesis and a process of ordinary language change in an existing language? In Chaudenson’s approach, of course, there is no distinction, as creole genesis is simply contact-induced language change taken to an extreme; but in other gradualist approaches the question arises and must be dealt with somehow.

Yet another variant of gradualism is John V. Singler’s. Singler, arguing from demographic evidence about (among other things) the ethnographic origins of slaves at different periods and the numbers of people of color relative to the numbers of whites at different periods, holds that ‘the principal agents of creole genesis in the sugar colonies of the Caribbean were those present in the period immediately following the sugar boom’—for instance, in Haiti, speakers of Gbe dialects (including, but not only, speakers of the Fon dialect), who comprised the dominant linguistic group during the first fifty years after the start of the sugar-plantation economy. The relevance of Gbe speakers is that Haitian Creole displays a
number of Gbe features in its structure, a fact that makes sense only if Gbe speakers played a prominent role in its formation.

In spite of the considerable appeal of gradualist theories for at least some creoles, they seem to leave important questions unanswered. Most crucially, how were African slaves communicating with each other during the transitional period between the founding of a new colony and the onset of the sugar plantation? Was everyone speaking some version of the European slavemasters’ language? If so, when each colony shifted to a sugar economy, did the sudden influx of slaves cause a sharp break in the transmission of the language, so that what occurred was in fact abrupt creolization, in spite of the delay in its onset? Chaudenson believes there was never a sharp break in the transmission of the European language; Singler suggests the possibility of an abrupt genesis after the shift to sugar. If everyone wasn’t speaking the European slavemasters’ language in the pre-sugar economy, what were they speaking? A fully crystallized pidgin developed in the colony? A fully crystallized pidgin imported from Africa? (In such cases, of course, the process of creolization would be nativization of a pidgin, not the emergence of a creole language without a stable pidgin stage.) This issue has been addressed occasionally; there is, for instance, growing evidence for the continued use of African languages in some colonies where large numbers of slaves did share a native African language. But most of these questions remain unanswered.

Another question arises here: if we agree that some creoles arose gradually (in some sense of gradual), is this route to a mixed language exclusively relevant to creoles, or does it apply sometimes to pidgins as well? The answer, interestingly, is that some pidgins seem to have arisen by a gradual process that is comparable to the genesis scenario proposed by Chaudenson for creoles. For these pidgins, the starting point is a foreigner-talk version of a lexifier language; this variety (simplified by native speakers of the lexifier language and perhaps modified also by interactions with people who don’t speak the lexifier language)
later crystallizes into new contact languages as the speaker group expands and diversifies. An example is Hiri Motu. As we saw above, the Motu people avoided using their “true” language in talking to outsiders; instead, they used foreigner talk. Their position in and around Port Moresby in New Guinea was prominent enough that newcomers had to depend on them for (among other things) food, and foreigner-talk Motu became the region’s lingua franca—probably with different variants used by and with different non-Motu groups in the community. The community itself grew rapidly after the first Europeans settled there in 1874, with an influx of various foreign groups. Foreigner-talk Motu was widespread in Port Moresby before it crystallized into Police Motu (now called Hiri Motu) as a result of its use in training the ethnically diverse police force, whose members took the language along with them in their travels throughout British New Guinea.

Although the social process in a case like the development of Hiri Motu was presumably different from the social dynamics among slaves on a sugar plantation, the linguistic processes—in which an ethnically mixed society takes a lexicon (and in this case some grammar) from a single prominent language in the community and forms a contact language—seem parallel to a considerable degree: the emerging contact language becomes less and less like the lexifier language as it is learned and changed by speakers of different languages, who “negotiate” a shared grammar in the process of communicating with each other. One difference between the two processes might be the deliberate distortion of their language by Motu speakers; but even that, for all we know, could have a parallel in plantation creole genesis. Another feature of plantation creole genesis is that it isn’t at all clear that it makes sense to speak of a target language in the usual sense of second-language acquisition (though the linguistic processes are similar, the goals differ); this too was quite likely a feature of the genesis of Hiri Motu, once the training of police became institutionalized.

What can we conclude from our survey of all these theories? It seems to me that the main
conclusion is that it’s a serious mistake to search for the route to pidgin and creole genesis. On the contrary, there’s solid evidence for more than one route. The development of many pidgins was probably abrupt, but Hiri Motu emerged gradually from foreigner-talk Motu. The development of Pitcairnese must surely have been abrupt, with no fully crystallized pidgin stage. Haitian Creole may or may not have developed abruptly; if it is an abrupt creole, the demographic evidence indicates that its abrupt genesis didn’t begin until thirty years after the settlement of Haiti by French speakers and their slaves (because the relevant Africans weren’t there until the shift to sugar after thirty years); if it is a gradually developed creole, it might or might not have arisen from a fully crystallized pidgin. Tok Pisin certainly emerged from a fully crystallized, and indeed an expanded, pidgin.


Do different genesis routes have different linguistic outcomes?

A vital question is whether different genesis processes correlate with different linguistic results. If they do, then we can infer the process retrospectively from the structure of the resulting pidgin or creole; if they don’t, we can’t tell what the process was unless we have either external information or (as we do for some creoles) internal evidence in the form of early vs. late attestations. Unfortunately, there doesn’t seem to be any strong indication that we can expect different linguistic outcomes from different genesis processes. Scholars were struck by the structural similarities among pidgins and especially creoles way back when the standard view was that all creoles developed from stable pidgins—i.e. long before the diversity in genesis routes was recognized—and those similarities remain. Pitcairnese is structurally similar to Tok Pisin in some notable respects, for instance. True, we can point to a few significant differences: pidgins, presumably because their speakers continue to use their native languages as well as the pidgin, tend to have more structural contributions from the substrate languages than creoles have; this tendency also seems to be shared to some
extent by two-language creoles. These effects can be seen in languages like Chinook Jargon, Tok Pisin, Pitcairnese, and Berbice Dutch Creole. But even this generalization doesn’t lead to firm predictions, given the large amount of substrate languages’ contribution to (for instance) Haitian Creole.

Even more unfortunately, adequate external information to support a particular claim about just how a pidgin or creole arose is all too often lacking. (But this may be an overly pessimistic assessment, given the increasingly sophisticated analysis of demographic data in creole- genesis contexts that is helping to clarify many previously obscure points.) At present, at least, we can be certain that diverse genesis processes have produced pidgins and creoles, but these processes usually cannot be distinguished retrospectively, either in principle or in fact. This means that the controversies about how pidgins and creoles arise will surely continue. As we get more detailed historical information, we will be able to refine existing theories further, and we may be forced to abandon some theories that turn out to be incompatible with the external historical evidence. Still, it is probably wise for pidgin/creole specialists to admit the possibility that the controversies may be unresolvable due to the permanent absence of adequate historical information.

The rest of the field

This chapter is already very long, but it has barely scratched the surface of the immense complexity of pidgin/creole studies. It has, I hope, given something of the flavor of the lively theoretical discussions in this corner of academia: readers who enjoy the intellectual stimulation of sharply differing scholarly opinions—not to mention the interactions in print of colorful and uninhibited scholars—will find further exploration of the field rewarding. Other intriguing and important issues must be left untouched here, among them the notion of a creole continuum, with varieties closer to and more distant from the lexifier language’s
grammar; the recurring question of whether English, with its enormous component of French and Latin lexicon, is a creole (it isn’t); the issue of fuzzy boundaries and their relevance to proposed semi-creoles like Afrikaans, which is spoken by descendants of Dutch immigrants in South Africa; and the roots of African-American English (is it a creole in origin?). The social functions of pidgins and creoles have only been touched on in passing in this chapter. This is a pity, because many (but not all) of these languages have traditionally been viewed with contempt even by their own speakers, and especially by speakers of the lexifier language—partly because they seem to laymen to be bastardized versions of the lexifier language, and partly because most speakers of most creoles are descended from slaves, so the languages are seen as fit only for slaves. These attitudes have been changing dramatically in recent decades, however; examples of the trend are Haitian Creole, which became one of Haiti’s official languages in 1987, and Seselwa (or Seychellois, the French-lexifier creole of the Seychelles in the Indian Ocean), which has been the country’s first official language, with English and French, since 1981. There is also the question of whether, or how, pidgins and creoles are connected with bilingual mixed languages; but this topic, at least, will be addressed in the next chapter.

**Sources and further reading**

There is an enormous scholarly literature on pidgin and creole languages. An excellent general introduction to the subject is the 1995 book *Pidgins and Creoles: An Introduction*, edited by Jacques Arends, Pieter Muysken, and Norval Smith. Besides the usual general sections on pidgins and creoles (including genesis theories), the book covers such topics as oral and written creole literature and five specific grammatical features, and it also has sketches of
one pidgin and seven creoles. Overall, as the mix of pidgin and creole sketches suggests, the book’s coverage is better for creoles than for pidgins. Another especially valuable source is the two-volume work *Pidgins and Creoles* by John Holm (1988, 1989), which covers general topics and surveys a large number of pidgin and creole languages. Both of these works contain lists, as complete as the authors could make them, of pidgins and creoles around the world. The *Journal of Pidgin and Creole Languages* (published by John Benjamins Publishing Company, Amsterdam & Philadelphia) is an excellent source of both general articles and specific case studies on pidgins and creoles.

Parts of this chapter draw on material in Thomason & Kaufman 1988 and on my 1997 article ‘A typology of contact languages’.

The observation from De Goeje 1908 is taken from p. 102 of George Huttar & Frank Velantie’s 1997 article ‘Ndyuka-Trio Pidgin’.

The definition of contact languages given at the beginning of the chapter is controversial: it isn’t accepted by all language contact specialists. One alternative view is that creoles, at least, can be defined by a list of characteristic features that they contain; this view was much more popular in earlier decades than it is now, but for a recent example, see John McWhorter’s 1998 article ‘Identifying the creole prototype: vindicating a typological class’. Another is that creoles, at least, can’t be defined at all in contradistinction from non-creole languages—that is, that they don’t differ significantly in either synchronic or diachronic respects from other languages; see especially various writings by Robert Chaudenson (e.g. his 1992 book *Des îles, des hommes, des langues* [‘Islands, men, languages’] and by Salikoko Mufwene (e.g. his 1996 article ‘The founder principle in creole genesis’). On this view, creoles (at least) do not comprise a coherent class of languages, either synchronically or diachronically; they may be identified by (for instance) political criteria, but not by any set of linguistic criteria. Most specialists would probably agree with the pioneer creolist Robert
A. Hall, Jr., who wrote in 1966 that ‘there are no structural criteria which, in themselves, will identify a creole as such, in the absence of historical evidence’ (p. 122); and most specialists would probably also agree that pidgins and creoles do form a class (or perhaps two separate classes) of languages, even if some languages seem to fall on the borderline between contact language and “ordinary” language.

Many pidgins and creoles are mentioned in this chapter, and I will not provide references to detailed discussions of all of them. Instead, references are given here only to the languages mentioned most prominently in the chapter. Sources for information about other pidgins and creoles can be found in general works, especially Arends et al., eds., 1995 and Holm 1988, 1989.

The structure and history of Ndyuka-Trio Pidgin are described by George Huttar & Frank Velantie in their 1997 article ‘Ndyuka-Trio Pidgin’, and Ndyuka itself is described in the 1994 book Ndyuka, by George L. Huttar and Mary L. Huttar. Berbice Dutch Creole is analyzed most fully in Silvia Kouwenberg’s 1994 book A Grammar of Berbice Dutch Creole See also her article ‘Berbice Dutch’ on pp. 233–243 of the 1995 book edited by Arends et al.; the example of a negative Berbice Dutch Creole sentence is from p. 237 of Kouwenberg’s 1995 article. The Ijo component of Berbice Dutch Creole is discussed in a 1987 article by Norval Smith, Ian Robertson, and Kay Williamson, ‘The Ijo element in Berbice Dutch’. Umberto Ansaldo and Stephen Matthews argue, in their paper ‘The Minnan substrate and creolization in Baba Malay’, for the status of Baba Malay as a two-language creole; the estimate of its number of speakers is from p. 358 of Norval Smith’s annotated list of contact languages in Arends et al., eds., 1995. The scholarly literature on Tok Pisin is very large; one of the most prominent specialists is Gillian Sankoff, who has published numerous articles on the language, among them ‘The genesis of a language’ (1979) and, with Suzanne Laberge, ‘On the acquisition of native speakers by a language’ (1974, reprinted in Sankoff’s 1980 book The
The claim that Chinese speakers created Chinese Pidgin English because they were unwilling, not unable, to learn English is from p. 8 of Robert A. Hall, Jr.’s, classic 1966 book *Pidgin and Creole Languages*.

The only information I have about the Nez Percé Jargon is from the source quoted at the beginning of the chapter, but there is some literature on Pidgin Delaware: the most comprehensive structural and historical analysis is in Ives Goddard’s 1997 article ‘Pidgin Delaware’. The quotations about Pidgin Hamer and Pidgin Motu are, respectively, from p. 397 of Jean Lydall’s 1976 article ‘Hamer’ and pp. 95–96 of Percy Chatterton’s 1970 article ‘The origin and development of Police Motu’, as cited by Tom Dutton and H.A. Brown on pp. 760–761 of their 1977 article ‘Hiri Motu: the language itself’. The best source for information about Hiri Motu is Tom Dutton’s 1985 book *Police Motu: iena sivarai*; the first half of the book’s title refers to the spread of the pidgin by policemen who traveled into the interior of Papua New Guinea to carry out their duties, and the second half is a Hiri Motu phrase meaning ‘its story’. For a summary of his research results on the pidgin, see Dutton’s 1997 article ‘Hiri Motu’.

There have been various claims over the years about which Romance language was the main lexifier of the medieval Lingua Franca. Robert A. Hall, Jr., argued for a Provençal lexifier, for instance, in his classic 1966 book *Pidgin and creole languages*. Convincing evidence for Italian as the lexifier for the Lingua Franca has been presented by Guido Cifoletti in his 1989 book *La Lingua Franca Mediterranea*; Jacques Arends (personal communication, 2000) observes that western Mediterranean varieties of the language acquired much Spanish lexicon as well. Maridi Arabic is discussed in a 1986 article by Sarah G. Thomason and Alaa Elgibali, ‘Before the Lingua Franca: Pidginized Arabic in the Eleventh Century A.D.’. For a broader view of Arabic-lexifier pidgins and creoles, see Jonathan Owens’ 1997 article
'Arabic-Based Pidgins and Creoles'. The location of Maridi is disputed, because the textual and linguistic indications conflict: Thomason & Elgibali argue for a location in modern Mauritania on the basis of internal evidence in al-Bakri’s text, and Owens argues for a location in Upper Egypt or northern Sudan on the basis of linguistic features that point to an eastern dialect source of the Arabic material in the pidgin.

For a good recent list of pidgins and creoles around the world, see Norval Smith’s 1995 article ‘An annotated list of creoles, pidgins, and mixed languages’ in the book edited by Arends et al. (For Smith and the other authors in that book, the term ‘mixed languages’ is used only for what I call ‘bilingual mixed languages’; in my usage, and therefore throughout this book, the term encompasses pidgins and creoles as well as bilingual mixed languages.) A recent survey of lingua francas, with good coverage of pidgins (more than of creoles), is the *Atlas of Languages of Intercultural Communication* edited by Stephen A. Wurm, Peter Mühlhäusler, and Darrell T. Tryon (1996).

For information on Pitcairnese, see Alan S.C. Ross & A.W. Moverley’s 1964 book *The Pitcairnese language*. A good recent source on Pidgin Eskimo is Hein van der Voort’s 1996 article ‘Eskimo pidgin in West Greenland’. The history and structure of Sango were described recently by Helma Pasch in her 1997 article ‘Sango’. Pasch describes the creole variety of Sango; this is presumably different from the pidgin variety, for which, she says, data are not readily available. William J. Samarin has also discussed Sango in numerous articles, most prominently in his 1967 book *A grammar of Sango*. Chinook Jargon is the subject of a large scholarly and popular literature; for a survey of its structure and history, together with references to data sources and other analyses, see my 1983 article ‘Chinook Jargon in historical and areal context’. A good recent sketch of Saramaccan is the article ‘Saramaccan’, by Peter Bakker, Norval Smith, and Tonjes Veenstra, in Arends et al., eds.

Calling Hawaiian Creole English a creolized pidgin is a controversial statement; the op-
posing view is that Hawaiian Pidgin English was never a fully-crystallized pidgin, and that the creole arose largely independently, without significant grammatical contribution from the pidgin. The most prominent advocate of this position is Derek Bickerton, who has argued for lack of grammatical input from a fully crystallized pidgin in numerous writings, e.g. his 1999 article ‘How to acquire language without positive evidence: what acquisitionists can learn from creoles’. Bickerton discusses the issue from a more general perspective in his 1981 book *Roots of Language*; this book is also a very good source of information on features shared widely among creole languages, especially Caribbean creoles and in particular in the tense/aspect/mood system in verb phrases (see the discussion later in this chapter).

Tây B`ôî was discussed in John Reinecke’s 1971 article ‘Tây B`ôî: notes on the Pidgin French spoken in Vietnam’. A good recent source on the structure and history of Kitúba is Salikoko Mufwene’s 1997 article ‘Kitúba’, which also contains much useful bibliography on the language. Two useful sources on Chinese Pidgin Russian Gunter Neumann’s 1966 article ‘Zur chinesisch-russischen Behelfssprache von Kjachta’ (in German) and Johanna Nichols’ 1980 article ‘Pidginization and foreigner talk: Chinese Pidgin Russian’. I’m grateful to Jacques Arends (personal communication, 2000) for the news that Lingua Franca noun phrases had agreement inflection. Among the language sketches in the volume edited by Arends et al. are ‘Fa d’Ambu’, by Marike Post, and ‘Sranan’, by Lilian Adamson and Norval Smith; the example of a negative Fa d’Ambu sentence is from p. 197 of Post’s article. I owe my information about Nagamese word order to Robbins Burling (personal communication, 1999). Information about Pidgin Yimas(-Arafundi) comes from an undated handout, ‘On a non-European based pidgin’, from a talk given by William A. Foley; the pidgin is discussed in print in Foley’s 1986 book *The Papuan languages of New Guinea* and his 1988 article ‘Language birth: the processes of pidginization and creolization’.

The examples illustrating the ordering of tense, aspect, and mood markers is from the
1995 article ‘TMA particles and auxiliaries’, by Peter Bakker, Marike Post, and Hein van der Voort, in Arends et al., eds. Their account in turn relies in part on Derek Bickerton’s well-known work, in particular his 1981 book *Roots of language*.

Here are a few references to theories of pidgin and creole genesis. A good starting point is the section on ‘Theories of genesis’ in Arends et al., eds.; the four articles in this section are ‘Theories focusing on the European input’, by Hans den Besten, Pieter Muysken, and Norval Smith, pp. 87–98; ‘Theories focusing on the non-European input’, by Jacques Arends, Silvia Kouwenberg, and Norval Smith, pp. 99–109; ‘Gradualist and developmental hypotheses’, by Jacques Arends and Adrienne Bruyn, pp. 111–120; and ‘Universalist approaches’, by Pieter Muysken and Tonjes Veenstra, pp. 121–134. The monogenesis hypothesis was proposed by R.W. Thompson in his 1961 article ‘A note on some possible affinities between the creole dialects of the Old World and those of the New’ and espoused in 1973 by Jan Voorhoeve in ‘Historical and linguistic evidence in favor of the relexification theory in the formation of creoles’; see also Keith Whinnom’s proposals for the Lingua Franca as a source for at least some creoles, e.g. his 1977 article ‘Lingua Franca: historical problems’ (and compare his 1956 book *Spanish contact vernaculars in the Philippine Islands*).

Derek Bickerton’s Language Bioprogram Hypothesis is discussed in numerous works, of which the most prominent is his 1981 book that is mentioned above. For extended discussion of his views, see his 1984 article ‘The language bioprogram hypothesis’, which was published together with cricial responses by several other scholars and then Bickerton’s response to the responses. An excellent recent critical article that focuses on Bickerton’s LBH and Claire Lefebvre’s relexification hypothesis is John V. Singler’s 1996 article ‘Theories of creole genesis, sociohistorical considerations, and the evaluation of evidence: the case of Haitian Creole and the Relexification Hypothesis’. The point about plantation-born slaves often having greater access to the lexifier language is from p. 196 of Singler’s article.
The relexification hypothesis is discussed in various writings by Claire Lefebvre and her colleagues, e.g. in Lefebvre’s 1993 article ‘The role of relexification and syntactic analysis in Haitian Creole: methodological aspects of a research program’ and in her 1999 book *Creole genesis and the acquisition of grammar: the case of Haitian Creole*. The quotations stating the hypothesis are from Lefebvre’s and John S. Lumsden’s announcement of a 1994 conference, The MIT Symposium on the Role of Relexification in Creole Genesis: The Case of Haitian Creole; they are cited John V. Singler on p. 186 of his 1996 article (cited above). In this article Singler also offers various criticisms of the hypothesis.

Pidgin/creole genesis as a kind of second-language learning—also called the substratum theory of pidginization and creolization—is proposed by such specialists as John Holm (see his 1988/1989 work), Norbert Boretzky (in his 1983 book *Kreolsprachen, Substrate und Sprachwandel*) and Thomason & Kaufman (in our 1988 book, as well as in other writings of mine). The quotation about the origins of SVO word order in Berbice Dutch Creole is from p. 296 of Silvia Kouwenberg’s 1992 article ‘From OV to VO: linguistic negotiation in the development of Berbice Dutch Creole’. The information about Bislama comes from W.G. Camden’s 1975 paper ‘Parallels in structure of lexicon and syntax between New Hebrides Bislama and the South Santo language spoken at Tangoa’.

The quotation characterizing ‘gradual creolization’ is from p. 111 of Jacques Arends & Adrienne Bruyn’s 1995 article ‘Gradualist and developmental hypotheses’ in Arends et al., eds.; Arends & Bruyn also discuss a gradualist model of creole genesis, based on a combination of demographic information and the time of emergence of particular syntactic features, and apply it to Sranan. Chaudenson’s writings are almost entirely in French; for a clear statement of his position (on both gradual creolization and proposed French sources for structural features of French-lexifier creoles), see his 1992 book (cited above). One of Chaudenson’s most devoted critics is Philip Baker; see, for instance, his 1996 review article
of Chaudenson’s 1992 book. See Salikoko Mufwene’s 1996 article ‘The founder principle in creole genesis’ for a good summary of a gradualist view that is closely related to Chaudenson’s. Derek Bickerton has addressed the problem of distinguishing between creole genesis and change in an existing creole in his 1991 article ‘On the supposed “gradualness” of creole development’. The quotation about the agents of creolization is from p. 226 of John Singler’s 1996 article (cited above). The gradual emergence of Hiri Motu is discussed by Tom Dutton in his 1985 book *Police Motu: iena sivarai* and his 1997 article ‘Hiri Motu’.

See Thomason & Kaufman’s 1988 book (cited above) for a lengthy discussion of the English-as-a-possible-creole issue. Thomason & Kaufman also has a case study on Afrikaans, concluding that it is a semi-creole; but since the publication of that book I’ve decided that the category ‘semi-creole’ doesn’t really exist (see my 1997 article ‘A typology of contact languages’ for discussion). An excellent recent study of African-American English is John R. Rickford’s 1999 book *African American Vernacular English: Features, Evolution, Educational Implications*. 
