CHAPTER PREVIEW

The previous chapter introduced the topic of language change and language family. In this chapter, we will explore one type of language change in more detail, the change that occurs when speakers of different languages interact with each other, in other words, when the languages are in contact. This chapter will introduce linguistic borrowing, the incorporation of forms or even meanings from one language into another. It will examine the motivations for borrowing, the range of borrowed elements, factors which promote or impede borrowing, and what happens as borrowed items are incorporated into another language.

The chapter will also discuss correlations between the sociopolitical relationships of language communities and the degree and nature of change from language contact. It will introduce special contact languages, pidgins, and creoles, which come about when groups of people with no language in common must work together or engage in trade. It will conclude by presenting linguistic areas, geographic regions that result from long-term contact between neighboring languages, which results in the sharing of features and structural convergence. This chapter demonstrates that languages are strongly influenced by the larger societal, political, and geographical features that define the speech communities.
LIST OF AIMS
After reading this chapter, students should be able to:

- discuss linguistic borrowings and their functions;
- identify linguistic borrowings in English;
- present arguments for language change conditioned by language contact;
- give examples of bilingual and multilingual societies;
- identify features of creoles and distinguish them from pidgins;
- understand the features of koinés;
- understand and define the term “linguistic area”;
- discuss sociocultural prerequisites for the formation of linguistic areas.

13.1 Language contact and areal linguistics

When you start learning a new language, you may wonder why a word in one language is similar to a word in a language you already know. This situation can result from pure chance. In Latin, ‘two’ is duo/dua. We know this from English words like dual or dualism. In Malay, the national language of Malaysia, dua also means ‘two.’ This is a coincidence, a curious fact that tells us nothing about the history of these languages, or those who speak them.

Now look at Table 13.1. Color terms in English and in German are very similar to each other. This is no coincidence. The forms and meanings are so similar because these languages come from a single common ancestor, the Germanic subgroup of Indo-European. Their similarities are due to genetic inheritance, as explained in Chapter 12.

Forms and meanings across languages can be similar for yet another reason. Languages and dialects do not exist in a vacuum. Speakers of different languages come into contact with each other: they may trade, intermarry, meet for ceremonies, and so on. The languages are then in contact, with many speakers of one having some knowledge of the other. Speakers cannot help borrowing linguistic features back and forth: habits of pronunciation, significant sounds (phonemes), grammatical categories, vocabulary items, and even some grammatical

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>white</td>
<td>weiss</td>
</tr>
<tr>
<td>red</td>
<td>rot</td>
</tr>
<tr>
<td>green</td>
<td>grün</td>
</tr>
<tr>
<td>blue</td>
<td>blau</td>
</tr>
</tbody>
</table>
forms. Thus, contact is another source of similarity between languages, whose speakers adjust their speech habits in line with those of their neighbors.

Every language has been influenced by another, at least to some extent. In English, we find many words that are not native vocabulary. For example, the word umbrella comes from Italian ombrella, which literally means ‘little shade,’ and the word cherries comes from French cérise. The impact of language contact is substantial and easy to identify in some languages. In others, it is not so easy. In contrast to English, Hungarian has very few words taken from other languages; speakers prefer to coin their own words rather than borrow. For example, in English the word cosmonaut, literally ‘world navigator,’ comes from Greek. In Hungarian the word ur-hajó (‘world navigator’) means the same, but all the morphemes are native to Hungarian. Similarly, most languages of Europe use a variant of the word president for the head of a company or a republic. Not so in Hungarian: the word for president is elnök, which literally means ‘first one.’

This shows that some speech communities purposely reject foreign imports. Their language ideologies place negative values on foreign words as unacceptable tokens of language mixing. Sometimes government bodies are even set up to ensure that people speak a “pure” language. Recently, the Academy of the French Language became alarmed at how many people are slipping English words into their French (see Textbox 13.1). They established hefty fines for those who fail to use a French word and use an English word instead.

**TEXTBOX 13.1 L’ACADÉMIE FRANÇAISE**

Established in 1635 and located in Paris, the French Academy is a body of forty elected officials who are experts on the French language. Elected officials hold office for life and are in charge of compiling, editing, and publishing France’s official dictionary, *Dictionnaire de l’Académie française*, which includes approving or denying foreign loanwords for use in French and making occasional revisions to the standard orthography. The Academy also awards prizes to French artists and scholars in literature, painting, poetry, theater, cinema, history, and translation. In an effort to keep French “pure,” the Academy has taken a stance against English loanwords like *email* and *software* and against local minority languages including Basque, Catalan, and Occitan.

13.2 **Borrowings**

When speakers of different languages interact, they borrow forms and meanings. How much they borrow depends on cultural and social factors, including the degree of knowledge of each other’s languages, speakers’ sense of purism, and also on the structure of the languages in contact. Complex morphological patterns in a language can make the incorporation of a foreign word difficult, and thus serve as a natural obstacle to foreign intruders.
Borrowing is one of the primary effects of language contact. A borrowed item is called a loan. Some borrowed words are easy to recognize. A person can be said to experience angst or schadenfreude; a language may have ablaut, and a country can be said to engage in realpolitik. What we have here, in italics, are lexical borrowing: they are words from one language (German) adopted into another (English).

If something is really “cool,” an English-speaking youth could refer to it as über-cool. The root cool is English. But the prefix über is German; it means ‘super.’ This word contains a grammatical borrowing, the borrowing of a derivational prefix. Lexical and grammatical forms can be borrowed directly. Or they can come via an intermediate language (SeeTextbox 13.2).

**Figure 13.1** L’Académie française, the Academy of the French Language, in Paris

**Textbox 13.2**

The preposition via ‘by way of’ comes from the Latin word via (the ablative singular form of the Latin noun via ‘road, channel, course,’ literally meaning ‘from the road, by the road.’) This came into English in about 1779 (Barnhart 2008: 1202).

After South America was colonized by the Spanish and the Portuguese, many words from local languages made their way into English through the intermediaries of these two languages. The word jaguar is a legacy of the now extinct Tupinambá language from South America; it came into English through the intermediary of Portuguese.
Table 13.2 lists some examples of words from Taino, an Arawak language spoken in the region of what is now Cuba, which were borrowed into Spanish in the sixteenth and seventeenth centuries. These are now common English words.

These words were borrowed as labels for novel items (for instance, a hanging bed, or hammock) or for area-specific phenomena (an unusually strong wind, a hurricane), for which Spanish (and English) had no word. For more on another type of loanword, see Textbox 13.3.

Sometimes borrowing involves lexical calquing: the creation of a new word by translating morpheme-by-morpheme from a source language. For instance, the German Ein-druck ‘impression’ (lit. ‘in-press’) has been calqued from Latin im-pressio, where each morpheme is translated from Latin into German. The term for ‘roof’ in Nigerian Arabic translates literally as ‘the head of the house’ (Owens 1996): this is how speakers of the surrounding Chadic languages refer to a roof.

In addition to lexical calques, some calques can also be grammatical. As an example, consider Pennsylvania German (also referred to as Pennsylvania Dutch in the United States), a language spoken by the Mennonite Anabaptists of Swiss-German

Table 13.2 Loanwords in English, borrowed from an Arawak language Taino via Spanish.

<table>
<thead>
<tr>
<th>Original form</th>
<th>Original meaning</th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>barnacoa</td>
<td>‘raised frame of sticks, a table used for sleeping or cooking’</td>
<td>barbacoa (first noted 1655–1665)</td>
<td>barbecue (first noted 1697)</td>
</tr>
<tr>
<td>hamaca</td>
<td>‘hammock’</td>
<td>hamaca (first noted in 1545–1555)</td>
<td>hammock (first noted c. 1555)</td>
</tr>
<tr>
<td>hurakán</td>
<td>‘strong wind’</td>
<td>huracán (first noted in 1510–1515)</td>
<td>hurricane (first noted c. 1650)</td>
</tr>
</tbody>
</table>

(Sources: Corominas (1961: 83); The Oxford English Dictionary (1989); Gastambide Arrillaga (1990: 15); Barnhart (2008: 1202.).)
origin who left Pennsylvania for Ontario, Canada, after the American Revolutionary War. Most speakers of Pennsylvania German are bilingual in English. Their variety of German is replete with calques. One is the immediate future tense, which is expressed by geh ‘go,’ a development inspired by English gonna (Burridge 2006: 183).

**Meanings can also be borrowed.** This can be seen when the meaning of a word is extended to match the range of meanings found in a neighboring, and often dominant, language. The Lakota verb *inyan* originally meant ‘to run,’ that is, to describe ‘the activity of moving fast on one’s legs.’ In contemporary Lakota, due to extensive contact with English, this same verb is commonly used to mean ‘to run for election, to function, operate, work as a device, machine, system, concept’ (Ullrich 2008: 775). Semantic extensions that arise through borrowing are referred to as **semantic loans**.

**Phonemes can also be borrowed.** Imagine a language spoken in a geographical area surrounded by languages of a different genetic affiliation. Such a language is likely to develop features atypical for the family it belongs to but shared with its neighbors. This is how Armenian, an Indo-European language, developed glottalized consonants, a feature rarely found in Indo-European; Armenian speakers were in intensive contact with speakers of surrounding Caucasian languages, in which glottalized consonants are ubiquitous (see Chirikba 2008).

Another example is the English phoneme /v/. As noted in Chapter 12, this was originally an allophone of the phoneme /f/ with a predictable distribution, as in *wife* versus *wives*. Then many French words with *v* found their way into English, among them *very*, *valley*, *vain*, and *ravine*. These new words had *v* in a wide range of environments, disrupting the old English pattern of complementary distribution between the allophones [v] and [f]. This caused the two to become different phonemes.

**In addition to words, many languages borrow grammatical forms,** including derivational and inflectional affixes, conjunctions, and discourse markers. For example, speakers of Pennsylvania German freely use English *well* and *I see* in their speech (see Burridge 2006: 189), discourse markers that are commonly used in English. **Bound morphemes – prefixes and suffixes – are not immune to borrowing.** The prefix *über*, borrowed from German into English, was exemplified above. Another example is the Spanish plural marker *-s*, which has made its way into Mexicano, a Uto-Aztecan language from central Mexico. The Mexicano word for ‘basket’ is *chiquihuitl*; to pluralize it, one adds the Spanish suffix *-s*, obtaining *chiquihuite-s* (Hill and Hill 1986: 165). See Textbox 13.4 on the difference between borrowing and code-switching.

13.2.1 How do words change through borrowing?

**Once a foreign form is borrowed, it often assimilates to the phonological patterns of the recipient language.** The words *spaghetti* and *gelato* have both been borrowed into English from Italian. In the American English pronunciation of both words, the final syllable begins with an alveolar flap [ɾ], a quick
ballistic movement of the tongue against the alveolar ridge. This follows the normal phonological pattern in most American English words – one finds the same sound in the final syllables of *potato* and *convoluted*. But in Italian, these words are pronounced differently: the final syllable of the word *gelato* has a voiceless aspirated stop, and the final syllable of *spaghetti* has a geminate consonant, or a lengthened stop, represented in writing by the doubling of the letter *t*. We can say that as these words were borrowed, they were assimilated to the phonological patterns of American English. This is known as **adaptation**: *when a foreign sound in a borrowed word does not exist in the recipient language, it is replaced by the nearest phonetic equivalent*. Words that have been perfectly integrated into the phonological and morphological systems of the language can be difficult to identify as borrowings. It takes a linguist to detect that the English words *cherries*, *very*, and *beauty* are in fact loans, from Old Norman French *cherise*, Old French *verai*, varai, vrai, and Old French *bealte*, beaute, biaute* respectively.

In contrast, unassimilated loans can stand apart from native words in their phonological make-up. In Mazateco, an Oto-Manguean language from Mexico, all voiceless stops become voiced after nasals; thus we never find the sequence [nt], only [nd]. However, in some Spanish loans, among them *siento* (from Spanish *ciento*), a frequently used word for ‘one hundred,’ one does find the native Spanish sequence [nt]. Mazateco has been described as having two coexistent phonemic systems – one native, and one for loans (see Fries and Pike 1949).

If enough loans are unassimilated, new sounds can enter the phoneme inventory. The case of English /v/ was noted above. Another example is /ʒ/, which was not a phoneme in English prior to contact with French. Like /v/, this sound developed into a separate phoneme in English as a result of numerous loans from French, such as *rouge*. Its adoption reinforced an independent sound change: the palatalization of /zj/ to /ʒ/, as in the word *Asia* (see Campbell 1999: 62).

Some loanwords are borrowed with their meanings intact, while others undergo semantic shift; the meanings of loanwords are never fully predictable. We can see this when a single word is borrowed more than once in the history of a language; the different instances of borrowing can develop different meanings in the recipient language, for example, English *chief* and *chef* were both borrowed from the same French word, but *chief* was borrowed
into Middle English in the fourteenth century, while chef came into Modern English in the nineteenth century.

Some loans retain morphological features of the source language. For example, some Latin borrowings into English require the Latin plural, e.g., the plural of colloquium is colloquia, whereas others allow either the Latin plural or the English plural, for example, syllabus can be pluralized either as syllabuses or as syllabi. Other borrowed nouns are fully assimilated into the English morphological system and take only the English plural, e.g., diplomata or pastas instead of the Italian plural pasti.

Some loans retain syntactic features of loanwords, resulting in the development of new syntactic distinctions. An example is Tetun Dili, an Austronesian language spoken in East Timor. This language is currently borrowing many words from Portuguese, a national language. Portuguese nouns, like those of other Romance languages, fall into masculine and feminine gender classes. Articles, adjectives, and other dependent elements in the noun phrase agree in gender with the head noun. Tetun Dili, on the other hand, has no native gender system. However, the Portuguese loanwords are now being borrowed in such great numbers that the genders are being borrowed as well; these nouns require agreement within the noun phrase, as in Portuguese. The result is that Tetun Dili has two systems of nouns, one gendered and one not, with only the former requiring agreement. A similar principle applies to adjectives. For example, the borrowed Portuguese adjective meaning 'pretty, handsome' will have two forms: masculine bonitu, from Portuguese bonito, and bonita, from Portuguese bonita. But native adjectives, such as di’ak ‘good,’ will have just one form for both genders (Hajek 2006: 202-203).

13.2.2 Are some forms or some meanings easier to borrow than others?

One of the hardest tasks in comparative linguistics is to tease apart similarities due to genetic inheritance and those due to borrowing and contact. This task would be easier if certain categories of forms or structures were never borrowed, since then similarities in these categories across languages could easily be attributed to shared genetic inheritance. However, while some types of forms are more resistant to borrowing than others, no linguistic feature is entirely "borrowing-proof." Words of all lexical classes – including possibly surprising categories, such as numerals, personal pronouns, conjunctions, and discourse markers – can be borrowed.

As stated above, affixes and grammatical elements can also be borrowed. We saw above how the German prefix über and the Latin preposition via have made their way into English. We also find English suffixes of French origin, including -ment, as in develop-ment, and -age, as in out-age. In Bolivian Quechua, the Spanish diminutive -itu is used with native words, e.g., rumi-tu ‘little stone.’ And the Spanish plural suffix -s can appear on native nouns together with the native plural marker -kuna: runas-kuna means ‘men,’ where runa is the native Quechua word for ‘man’ (see Appel and Muysken 2005: 172-173).
Some grammatical and features are particularly amenable to borrowing. About one-quarter of the world’s languages have grammatical systems that obligatorily mark how the speaker has come to know the information being expressed. Such systems are said to mark the speaker’s information source, or his or her evidence for the facts being conveyed; this is known as evidentiality. Getting your information source right is a prerequisite for successful communication in such languages. This is the case in numerous North American Indian languages, including Southern Paiute and Verde Valley Yavapai (Bunte and Kendall 1981). Because the expression of this category is central to their communicative practice, speakers of these languages will sometimes use native evidential markers when they speak English. For example, if a Paiute speaker knows that a woman named Minnie is pregnant because someone told him, he will pass the news onto another by saying Minnie is pregnant /ikm: the particle /ikm indicates that the speaker learned this information through a third source, as opposed to learning it through direct observation or by some other means. The fact that evidential meanings - and forms - are prone to borrowing reflects their importance to speakers who have these as part of their linguistic systems; their prevalence in the world’s languages shows how important they are in human communication.

13.2.3 Why borrow?
There are many reasons why people borrow features from other languages. With loanwords, the most common reason is to fill a gap for an item or a concept that the language has no word for. This explains borrowings such as hammock (see above), umlaut, from German, or karma, from Sanskrit. However, there are also more subtle reasons for borrowing. Borrowing may be motivated by prestige, language attitudes, or civic institutions.

Borrowing may be motivated by prestige. Consider the English forms for hooved animals in Table 13.3. The forms in the left-hand column are native, and the ones in the right-hand column are borrowed from French. They belong to an era when French culture dominated the English after the Norman Conquest of 1066, putting emphasis on cuisine and elegant dining (Hock 1991: 385). The terms in the left-hand column refer to animals themselves, the area relegated to peasants who had no

<table>
<thead>
<tr>
<th>English native forms</th>
<th>Forms borrowed from French</th>
</tr>
</thead>
<tbody>
<tr>
<td>cow, bull, ox</td>
<td>beef</td>
</tr>
<tr>
<td>calf</td>
<td>veal</td>
</tr>
<tr>
<td>pig/hog/swine</td>
<td>pork</td>
</tr>
<tr>
<td>sheep</td>
<td>mutton</td>
</tr>
</tbody>
</table>
exposure to French, or to elegant life. The terms in the right-hand column belong to the sphere of cuisine, where French influence was strong: this explains why they are of French origin.

The cultural dominance of the French-speaking court in administration and warfare is also reflected in the numerous terms from these lexical fields, such as justice, legal, court, curfew, and grand jury. In the current era, we see a different direction of borrowing, with many terms related to information technology being borrowed from English into other languages. For an example of this in Spanish, see Textbox 13.5.

Language attitudes of speakers, sometimes institutionalized through governmental or civic organizations, can determine whether loanwords are acceptable or not. In many languages, “foreign” importations are rejected tokens of unacceptable “language-mixing.” A cultural inhibition against recognizably foreign items and ensuing linguistic purism provides a mechanism for stopping an influx of borrowed forms. Once speakers become conscious of the foreign material in their lexicon – or grammar – they can try to get rid of it. This has happened in the history of various literary languages, including Hungarian, Finnish, and Estonian (see Fodor 1984, Tauli 1984). Such moves towards linguistic purism reflect the association of a group’s identity with its own language, in contrast to surrounding, often more dominant, groups (Thurston 1987: 93).

As we have seen, the nature and degree of language contact are reflections of the history, language attitudes, and social interactions of speech communities. They can also reflect the relative political status of each group. A minor language is likelier to borrow forms from a dominant one than the other way around. This is why we find numerous loans from Spanish into indigenous Meso-American and South American languages, and just a handful of loans from indigenous languages into Spanish.

**TEXTBOX 13.5 HOW TO EMAIL IN SPANISH?**

The following discussion was recently spied on a blog about the Spanish language. What does this exchange tell us about the role of English in the domain of electronic media?

**Question:** I saw in a sample Spanish sentence recently that you used the word *emails*. First of all, why didn’t you use a real Spanish word for email? Second, even if that were a Spanish word, why isn’t the plural *emailes* instead of *emails*?

**Answer:** These days, believe it or not, *email* (or *e-mail*) is a Spanish word, very commonly used, although it has not been recognized yet by the Spanish Royal Academy and is considered by many to be an Anglicism. It even has a verb form, *emailear*, that is sometimes used. It is one of those English words that has been adopted into Spanish even though some perfectly good “real” Spanish alternatives exist.
13.3 Bilingual communities, diglossia, and language shift

Any individual is a member of a linguistic community. If a community is composed of a number of groups each speaking a different language, an individual may be bilingual or multilingual. Bilingualism can be individual: for instance, some descendants of Swedes in America may still speak some Swedish. Or it can be societal: to be able to communicate within your group, you need to know more than one language. Examples of societal bilingualism include Slovak and Hungarian in the border areas between Slovakia and Hungary, or French and English in French Canada.

The use of different languages may correlate with different social settings. For example, for many years Latin was the language of the Church, and local European languages were used in day-to-day communication. Using two varieties of the same language under clearly specified conditions is known as diglossia. This is when a more prestigious, and usually more archaic, form of a language is used in “high” functions (such as church, school, literature), and a less prestigious colloquial variety is used in “low” functions (day-to-day oral interaction). For instance, in Arabic diglossia, the “high” language of the Koran coexists with local vernaculars. And in Switzerland, Swiss German (Schwyzerdütsch) varieties are the language of day-to-day interaction in the German-speaking cantons, with Standard German being the “high” variety.

In a multilingual situation, the “low” and the “high” varieties are often represented by different languages. For many generations, Spanish was a “high” language for Quechua and Aymara speakers, so was restricted to social domains such as religion and government. What may help a minority language survive is its defined role in the society and its special value for the speakers as a repository of heritage and tradition, and a symbol of identity. This is part of what is keeping many Quechua varieties alive.

In a diglossic or multiglossic situation, languages typically influence each other. This usually works in the direction from high to low. For example, three languages are spoken in Sauris, a German linguistic enclave in northeastern Italy (Denison 1971, Lehiste 1988: 53-54): an archaic form of Southern Bavarian German, Italian (the national language), and Friulian (a Romance language closely related to Italian but distinct from it). German is the “low” language: it is used within the family. Italian, the language of organized schooling and religion, is the high language. Friulian is in between high and low; it is the language of communication with the surrounding Friulian speakers, and also serves as a symbol of in-group solidarity among young males who did their secondary schooling in a regional center. Denison noticed that Italian and Friulian elements are acceptable in German (the low language). However, German elements are not introduced into either Italian or Friulian. And Italian – the high language – is also immune from intrusions of the “lower” Friulian. The social
relationships between languages and their domains of use correlate with the degree of influence of one onto the other.

In certain types of social situations, the speakers of one language may shift to speaking another, which is typically the language of a socially dominant group. When the majority of a community shifts, it can lead to the endagerment, impending obsolescence, and death of minor languages. When a community loses its language as it adopts a dominant one, it is said to have undergone language shift. The original language may leave its traces on the newly acquired language. For example, the variety of English spoken in Ireland has a distinctive prosodic pattern (colloquially referred to as a “lilt”), which was brought into the language when Gaelic speakers learned English. Like most speakers who acquire a new language as adults, they did not adopt the prosody of the language they are acquiring, but spoke it with their native patterns of rhythm and pitch. Owing to the numerical predominance of Gaelic people in the region, these Gaelic prosodic patterns were adopted by English children, making it a regular feature of the dialect (also see Filippula 2003).

13.4 Pidgins and creoles

In the process of European colonization, enslaved or subjugated people from many different linguistic groups were often forced to work closely together. They frequently spoke different languages from their fellow workers and from their employers or overseers; however, they still needed a common idiom for the purpose of business communication. Hence the name “pidgin” which comes from the English word business. Pidgins are contact languages that are not native to any of those who speak it. Pidgins are used for fairly limited communication between speakers of different languages who have no other language in common. French, Spanish, Portuguese, English, and Dutch – the major languages of the colonizers - provided the foundation for European-based pidgins.

Not all pidgins resulted from colonization. Trade pidgins also developed outside the world of European colonization. Among them is Chinook Jargon in the Pacific Northwest of North America, Iatmul pidgin in the Sepik area of New Guinea, and various pidgins in the Arctic area (see Jahr and Broch 1996). Pidgins are typically limited in their lexicon, and simplified in their structure, compared to the languages they are based on.

As speakers of a pidgin start marrying each other, and the language becomes their only way of communication, it begins to evolve a more complex structure and richer vocabulary. And for children born of such couples, the pidgin becomes their first language. A former pidgin which has acquired native speakers is called a creole. An example is Tok Pisin, a national language of Papua New Guinea.
Pidgins and creoles are generally agreed not to belong to any single language family; they arise as a makeshift means for filling communicative needs, rather than evolving from natural transmission in the way of regular, non-contact languages. (For an explanation of yet another way a new language can arise, see Textbox 13.6 on Koinés.)

The language of the socially dominant population in a situation of creolization is referred to as the superstrate language. As it typically contributes most of a creole’s vocabulary, it is also called the lexifier language. Minority languages that contribute to the formation of a pidgin or a creole are called substrate languages (see Sidebar 13.1). Some vocabulary from the substrate language typically makes it into a creole. In addition, meanings are often influenced by the indigenous languages that provide the substrate. Tok Pisin, which is spoken in new Guinea, is an English-based creole. The vocabulary includes the noun gras, from English grass, but in Tok Pisin it refers to grass, hair, and fur. This polysemy is the effect of the substrate from Oceanic languages, especially Tolai, that has a noun with these three meanings.

A typical creole looks similar to the lexifier language in terms of its vocabulary, but the grammar is very much unlike it. Consider example (1) from Tok Pisin. You will notice that many - but not all - words are English-based.

(1) 

Mi no save yet sapos bai mi stap long hia o bai mi go long Brisbane long disela taim
‘I don’t know yet if I will be here or if I go to Brisbane at that time’

Example (2) comes from Kristang, a Portuguese-based creole from Malaysia (Baxter 1988: 213). The forms are Portuguese, but the way they are put together (the grammar) is completely different.
13.5 Substratum interference

If a particular population needs to learn the language of a smaller yet politically dominant group (as is sometimes the case following foreign invasions), they often acquire the target language in an incomplete fashion, and so speak it in an altered form. The children of the dominant group might then adopt the altered forms of the target-language speakers and so change the target language itself. Such processes of substratum interference have occurred in the history of many languages. Examples include English varieties spoken in Singapore, Papua New Guinea, and other places, or the Afrikaans variety spoken in Zimbabwe.

Majority languages that have dominated indigenous languages in many countries can bear the substrate impact of the minority variety. The English spoken in Ireland has been influenced by the Celtic substrate (see Hilbert 2008: Filppula 2003: 167). The Irish prosodic patterns were mentioned above. Another example is the use of be instead of have in Irish English, e.g., They're gone mad instead of They've gone mad (Filppula 2003: 166–167). This structure is parallel to how be is used in Gaelic (Irish) and could be attributed to substrate influence.

13.6 Linguistic areas

Borrowings and structural similarities may extend over all or most of the languages in a geographical region, whether they are related or not. This results from large-scale linguistic diffusion, which defines the region as a linguistic area or a Sprachbund. Although languages in these situations are from different families, they become typologically similar: languages may remain different in many of their forms, but their structures will converge towards a similar prototype.

A linguistic area is defined as a geographic region including languages from at least two language families, or different subgroups of the same family, sharing significant traits. Within a linguistic area, diffusion, the spread of linguistic features from one language (or person) to the next, can be unilateral (when it primarily proceeds from one language to the others) or
Language contact and areal linguistics

Multilateral (when multiple languages exert influence on the others). Languages within an area often share a fair number of features. However, it can be difficult to determine whether those shared features are due to diffusion, as opposed to accidental coincidence, language universals, or genetic factors (if the languages in contact are related).

In an important study of Meso-America as a linguistic area, Campbell, Kaufman, and Smith-Stark demonstrated that not all shared features have the same “weight” in determining whether or not a region is a linguistic area. They state that “highly ‘marked’, exotic, or unique shared traits count more than does material that is more easily developed independently, or found widely in other languages” (1986: 535–536). Thus, the best types of evidence for a linguistic area are not traits that are inherited from a common ancestor or due to chance or universals, but those that can be shown to be diffused.

Thus, a highly frequent phenomenon – such as verb-final constituent order, the existence of nasalized vowels, or the presence of a past versus non-past opposition in the tense system – would not be assigned as much weight in determining a linguistic area as would a more exotic, unusual characteristic. An example of such an exotic feature is the marking of evidentiality in the Vaupés area of Brazil; we will return to this shortly.

While a single typologically common trait cannot by itself define a linguistic area, the clustering of traits can be area-specific. When we look at different linguistic areas throughout the world, we see that each is defined by a different cluster of properties. Here are four examples of linguistic areas and the features which define them:

1. Meso-America (Campbell et al. 1986). The languages of this region belong to several different language families, including Mayan, Mixe-Zoquean, Totonacan, and Otomanguean, among others.
   (i) Nominal possession of the type *his-dog the man* (‘the man’s dog’);
   (ii) Body-part nouns being used as markers of spatial relationships (e.g., ‘head’ for ‘on top of’)
   (iii) Vigesimal numeral systems (i.e., systems of counting based on twenty as opposed to ten);
   (iv) Basic constituent orders that are not verb final;
   (v) Numerous lexical collocations, or ways of expressing particular concepts that are common throughout the Mesoamerican area, e.g., ‘knee’ as ‘head of the leg’ or ‘boa-constrictor’ as ‘deer-snake.’

2. The Balkan Peninsula. All the languages belonging to the Balkan linguistic area are Indo-European, but from different subgroups. The Slavic languages include Serbian, Croatian, Bulgarian, and Macedonian; there is also Romanian (a Romance language), Greek, and Albanian. Some scholars add to this Romani (the language of the Gypsies, from the Indo-Aryan branch of Indo-European) and Turkish, an unrelated Turkic language.
   (i) A central vowel /i/ or /a/ (absent from Greek and Macedonian);
   (ii) A single affix that is used for both the dative and genitive cases;
   (iii) Articles that follow the noun (absent from Greek);
(iv) A future tense marked by an auxiliary corresponding to ‘want’ or ‘have’ (absent from Bulgarian and Macedonian);
(v) The use of the auxiliary verb corresponding to ‘have’ in constructions of the type *I have gone* in English (i.e., perfect aspect constructions);
(vi) The absence of infinitives in complement clauses;
(vii) The use of a pronoun in addition to a full noun phrase to refer to an animate object, so that the object is marked twice (Friedman 2006).

3. The Vaupés River Basin in Brazil and Colombia. Languages from this region belong to the genetically unrelated Tucanoan and Arawak families. The area is characterized by obligatory societal multilingualism, based on the principle of linguistic exogamy: one can only marry someone who speaks a different language.
   (i) Nasalization that occurs on most or all segments in a word;
   (ii) Four to five evidential morphemes, which mark the way in which the speaker has acquired the information (whether seen, heard, inferred, assumed, or learned from someone else);
   (iii) Numerous classifiers used with demonstratives, numerals, and in possessive constructions;
   (iv) Small systems of genders in verbal agreement;
   (v) Nominative-accusative case-marking;
   (vi) A single locative case-marker that indicates direction (‘to’), location (‘in, at’), and source (‘from’);
   (vii) Numerous lexical collocations, e.g., ‘father of goods’ = ‘rich man.’

4. South Asia (Emeneau 1956; Masica 1976, 2001). This linguistic area is composed of languages from the Indo-Aryan, Dravidian, Munda, and Tibeto-Burman families.
   (i) Retroflex consonants, especially stops;
   (ii) Dative-subject constructions (where some subject arguments are marked with the dative case);
   (iii) A compound verb construction whereby a special auxiliary combines with the immediately preceding main verb and the two verbs referring to a single event;
   (iv) The use of an affix to mark causation;
   (v) Verb-final constituent order.

It is important to reiterate that no single one of the features is found only in the particular linguistic area, as you can find similar features throughout the world. However, the clustering of the properties is area-specific; it is only in the particular area that the whole set of properties recur in language after language.

In a situation of intensive language contact within a linguistic area, the gradual convergence of languages may result in structural isomorphism. In such a case, the grammar and the semantics of one language are almost fully replicated in another. A classic example in the literature is the village of Kupwar in India, where long-standing language contact between Indo-Aryan and Dravidian languages has resulted in extreme convergence (see Gumperz and Wilson (1971); and also Nadkarni 1975).
Another example of a striking **structural isomorphism** comes from two languages spoken within the Vaupés River Basin linguistic area. Consider the two sentences below, both taken from traditional stories involving a female cannibal. Example (3) is from Tariana, an Arawak language, and (4) is from Tucano, a Tucanoan language, which is in constant contact with Tariana. Compare the content and order of the morphemes in these two examples; you will see that they are almost identical.

(3) **Tariana**

\[
\text{nese \ pama \ di-na} \\
\text{then one.NUM.CLF.F 3SG.NF-OBJ} \\
\text{du\-yana\-sita\-pidana} \\
\text{3SG.F-cook-ALREADY-REM.PST.REPORT} \\
\text{‘She had reportedly cooked him already’ (reportedly, a long time ago)}
\]

(4) **Tucano**

\[
\text{tiîta \ ni’kol \ kû-re} \\
\text{then one.NUM.CLF.F 3SG.M-OBJ} \\
\text{do’á-toha-po’} \\
\text{cook-ALREADY-REM.PST.REPORT.3SG.F} \\
\text{‘She had reportedly cooked him already’}
\]

**Figure 13.2** Tariana men from Santa Rosa, who have Tariana as their first language and who also speak Wanano, Desano, and Piratapuya. The women speak Piratapuya, Wanano, Siriano (Tucanoan), and Baniwa (Arawak), but hardly any Tariana. Everyone also knows Tucano, and most people know Portuguese.
Linguistic convergence does not always result in the creation of identical grammars. It is also not the case that categories in language contact always match. Languages in contact often maintain some distinct typological traits. This can be seen from Examples (3) and (4): Tariana maintains its prefixes, and Tucano its suffixes.

Examples of convergence within linguistic areas show that languages may come to be structurally similar due to areal diffusion, notwithstanding the original differences between them.

CHAPTER SUMMARY
Languages and their speakers do not exist in a vacuum. Different linguistic communities frequently come into contact with each other. Their languages are then in contact, with many speakers of one language having some knowledge of the other. Speakers naturally borrow linguistic features back and forth; habits of pronunciation, significant sounds (phonemes), grammatical categories, vocabulary items, and even some grammatical forms are taken from one language and applied in another.

Here are some of the general principles that have been presented:

- Languages reflect the sociolinguistic history of their speakers. If one language community dominates the other, we expect the language of the dominant group to have more impact than the language of the other group.
- Language contact may result in direct diffusion (borrowed lexical and grammatical forms) and indirect diffusion (borrowed or calqued patterns).
- Language contact may result in the creation of pidgins and creoles. These are non-genetic in their origins, and do not belong to any linguistic family.
- Languages may share features, or combinations of features, as members of an extensive linguistic area.
- Convergence within linguistic areas and contact situations may result in one language adopting the structure of the other. Or it may result in one language adopting new patterns and forms, while at the same time preserving some of its own properties. The outcomes of convergent development depend on the degrees of dominance within each area.

TEXTBOX 13.7 GLOSSING CONVENTIONS USED IN THIS CHAPTER

<table>
<thead>
<tr>
<th>Convention</th>
<th>Meaning</th>
<th>Convention</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
<td>NF</td>
<td>non-feminine</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
<td>NUM.CLF</td>
<td>numeral classifier</td>
</tr>
<tr>
<td>ALREADY</td>
<td>already</td>
<td>OBJ</td>
<td>objective</td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
<td>PRF</td>
<td>perfect aspect</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
<td>REM.PST</td>
<td>remote past</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
<td>REPORT</td>
<td>non-personal knowledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>singular</td>
</tr>
</tbody>
</table>
SUGGESTIONS FOR FURTHER READING

A concise introduction to language contact, with a focus on bilingualism and language acquisition.

A brief general introduction to language contact, full of insights and generalizations, with an incisive characterization of concepts such as linguistic area and creolization.

An overview of language contact and contact-induced change, with an emphasis on sociopolitical factors rather than linguistic features in language contact.

A classic work on language contact, with a focus on bilingual communities, and a must-read for any student of language contact.

An up-to-date introduction to issues related to language contact, with a special focus on the formation of creoles and pidgins.

EXERCISES

1. When words are borrowed into a language, they can retain their basic meanings but take on different stylistic connotations. This is often the case with French and Latin loans in English.

   A prominent linguist was asked to write a grant proposal for a high-status grant-giving agency. He then was asked to write a popular article for a students' magazine on the same topic as his grant proposal. Two extracts from the grant application and two from the popular article are given below.

   Extract 1
   *In similar fashion, scholars have demonstrated the genetic unity of Australian languages.*

   Extract 2
   *In exactly the same way, linguists have shown that almost all the languages of Australia belong to one language family.*

   Extract 3
   *The documentation of endangered languages is, for a number of reasons, one of the highest priorities facing mankind at the turn of the millennium.*

   Extract 4
   *Writing grammars and descriptions of languages that are about to die is, for all sorts of reasons, one of the most important things to be done in the world as we near the year 2000.*
a. Identify which two extracts come from the grant application, and which two come from the popular article. What specific features of each extract served as clues as you considered this?
b. Using a dictionary that provides reliable etymologies (such as the Merriam-Webster Online), look up the source language of each of the following words and state whether the word was originally Germanic (i.e., evolved from Old English and is related to Old High German, Old Norse, etc.) or whether it was borrowed into the language from a Romance language (i.e., Latin or French) or Greek.

Extract 1
similar
fashion
scholar
demonstrate
genesis (attested in this extract in the etymologically related word genetic)
unity
Austr- (attested in the word Australian)
language

Extract 2
exact (attested in the word exactly)
same
way
linguist
show (attested in the word shown)
almost
all
belong
one
family

Extract 3
document (attested in the word documentation)
danger (attested in the word endangered)
number
reason
high (attested in the word highest)
prior (attested in the word priority)
face (attested in the word facing)
man
kind
turn
millennium

Extract 4
write
grammar
scribe (attested in the word description)
about
die
sort
reason
most
important
Language contact and areal linguistics

c. Based on your answers to tasks (a) and (b), what can you say about the relationship of the historical source of vocabulary to academic and popular genres of written English?

2. Identify the origin of each of the following loanwords in English and explain why you think they were borrowed:
   a. Karma
   b. Svastika
   c. Cumquat
   d. Pajamas
   e. Croissant
   f. Cockatoo
   g. Kangaroo
   h. Mutton
   i. Canoe
   j. Torso

   a. The following data presents a list of loanwords borrowed into Zulu from English.
      i. Determine the original English source for each loanword and write it in the space provided.
         (Examples (e), (q), and (aa) are proper nouns.)
      ii. The initial vowel is a separate Zulu morpheme that is added based on the semantics of the loan. What semantic class of noun takes u- instead of i-?
      iii. State the phonological rules by which English words are assimilated into Zulu phonology.
         (Note: It is always better to state a general rule, such as “voiceless stops become voiced,” than to state multiple specific rules, such as p->b, t->d, and k->g.)

   A. Ipuulatifomu
   B. Isigina
   C. IpiKingi
   D. Ipalagilafu
   E. Ifulansi
   F. Ikilasi
   G. Igilamafoni
   H. Iminiti
   I. Igilamu
   J. Ipuulani
   K. Ikiliniki
   L. Ibizinisi
   M. Ujamu
   N. Uwisiki
   O. Itenisi
   P. Ibesiboli
   Q. Ifuntshi
   R. Ipuleti

SIDEBAR 13.3
ORTHOGRAPHIC NOTE

<table>
<thead>
<tr>
<th>Zulu standard orthography</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsh</td>
<td>[tʃʰ]</td>
</tr>
<tr>
<td>J</td>
<td>[dʒ]</td>
</tr>
<tr>
<td>Y</td>
<td>[j]</td>
</tr>
</tbody>
</table>
b. How would the following English words sound if borrowed into Zulu?
A. flat
B. cap
C. cricket
D. globe
E. film
F. gin
G. ginger
H. brake
I. inch
J. juice

c. What does the proper noun *ingilandi* mean in Zulu?


a. The words given below are Kikuyu words that were borrowed from English into Kikuyu, as well as the English source word. Make a list of the phonological changes that these words underwent when they were adapted to Kikuyu phonology. Again, generalize where possible.

<table>
<thead>
<tr>
<th>Kikuyu</th>
<th>English source</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. thukuru</td>
<td>school</td>
</tr>
<tr>
<td>ii. haithukuru</td>
<td>high school</td>
</tr>
<tr>
<td>iii. anderethi</td>
<td>address</td>
</tr>
<tr>
<td>iv. rainbarari</td>
<td>library</td>
</tr>
<tr>
<td>v. miriti</td>
<td>military</td>
</tr>
<tr>
<td>vi. korenji</td>
<td>college</td>
</tr>
<tr>
<td>vii. thateraiti</td>
<td>satellite</td>
</tr>
<tr>
<td>viii. mbuku</td>
<td>book</td>
</tr>
<tr>
<td>ix. thonda</td>
<td>soda</td>
</tr>
<tr>
<td>x. mboi</td>
<td>boy</td>
</tr>
<tr>
<td>xi. rumu</td>
<td>room</td>
</tr>
</tbody>
</table>

b. How would the following English words sound if borrowed into Kikuyu?
   i. dresser
   ii. colour-bar
   iii. glue
   iv. agenda

5. Tok Pisin is a creole language spoken in Papua New Guinea; it was discussed briefly above. The creole is derived from English, but is clearly distinct from it. The excerpts below are taken from a cartoon in a newspaper. They both feature Isuzu trucks; in the second one, the truck is being used
Language contact and areal linguistics

as PMV (Public Motor Vehicle) – a type of public transport. Each excerpt is given in Tok Pisin, with an English translation provided.

Make a glossary containing each Tok Pisin word and its English counterpart. Wherever possible, identify the original English word or words that were initially adopted to express each meaning. Finally, note any grammatical features that differentiate Tok Pisin from English.

Example I (Picture shows men squeezing into a pickup truck)

Man A: Mipela i kam nau long Los Angeles lukim ol profesenel basketbol ... Ol mai hia i bikpela liklik ...
Tasol mipela inap putim beksait long Isuzu na karim ol i go bek long hotel.
‘We are now coming to Los Angeles to see professional basketball ... Men here are a bit big. But we can put our backsides into an Isuzu truck and take all back to the hotel.’

Man B: Tru tumas, Isuzu em inap ... ‘Very true, Izuzu can do that/is sufficient …’

Example II (Picture shows a mother dragging a complaining child into a truck, which is being used like a taxi)

Child to mother: Mama, mama, mi no laik ...
‘Mama, mama, I don’t want …’

Mother to child: yu kam, yu blary hambag ...
‘You come, you naughty humbug.’

Driver: Mi amamas tru ranim dispela Isuzu PMV ... Olgeta de mi bungim planti ol naispela pipel ... Em i gutela wok tru ...
‘I am happy to run this Isuzu PMV ... Every day I meet many nice people ... This is a really good job …’