Though Weinreich focuses specifically on the phenomenon of bilingualism, his statement can, *mutatis mutandis*, apply equally well to the study of all contact situations. Moreover, the field of contact linguistics is not limited to just the study of "interference," but covers all the linguistic consequences of contact, including phenomena such as simplification and various other kinds of restructuring that characterize the outcomes of contact. Weinreich's goal of "prediction" is perhaps ambitious, but he himself is well aware of the complexity of the problem. In particular, he emphasizes that the components of an explanatory framework must include "purely structural considerations . . . psychological reasons. . . . and socio-cultural factors" (1953: 44). The need to explore the latter two types of factor arises from the fact that, first, contact situations which appear quite similar in terms of the linguistic inputs present can and do result in quite different linguistic outcomes. Moreover, for any given contact situation, predictions of contact-induced changes based solely on structural factors fail miserably. This point will be discussed in later chapters, when we consider the various linguistic constraints on such changes. Weinreich's outline of the main concerns of "interference" studies is worth quoting in full. He notes:

> In linguistic interference, the problem of major interest is the interplay of structural and non-structural factors that promote or impede such interference. The structural factors are those which stem from the organization of linguistic forms into a definite system, different for every language and to a considerable degree independent of non-linguistic experience and behavior. The non-structural factors are derived from the contact of the system with the outer world, from given individuals' familiarity with the system, and from the symbolic value which the system as a whole is capable of acquiring and the emotions it can evoke. (1953: 5)

It follows, first, that we need to distinguish among the various social contexts of language contact if we are to understand the nature and direction of contact-induced change. Second, it is necessary to examine, where possible, the actual speech behavior of persons in each contact situation in order to uncover the factors that motivate them to change their language in one way or another.

Scholars have long been aware that differences in the social setting lead to differences in the outcomes of contact. For instance, Wackernagel (1904) distinguished three kinds of contact situation – when a conquered group adopts the language of its conquerors, when the reverse occurs, and when there is mutual influence leading to a "mixed language." Every outcome of language contact has associated with it a particular kind of social setting and circumstances that shape its unique character. The goal of contact linguistics is to uncover the

various factors, both linguistic and sociocultural, that contribute to the linguistic consequences of contact between speakers of different language varieties. Toward that end, we need a framework of analysis that includes a variety of components. In the rest of this chapter, we provide a broad overview of types of contact situation, their outcomes, and the social settings in which they emerge. We will consider each of these situations in more detail in subsequent chapters. There too we will explore the mechanisms and types of change involved as well as the factors, both linguistic and non-linguistic, which influence the patterns of cross-linguistic influence.

1.4 Types of Contact Situation

We can in general distinguish three broad kinds of contact situation: those involving language maintenance, those involving language shift, and those that lead to the creation of new contact languages. Most cases of language contact can be assigned clearly to one or another of these categories. However, as we will see, there are many situations that cannot be classified so readily. Some are characterized by interplay between maintenance and shift, like the "fuzzy" cases found in Sprachbünde or linguistic areas such as the Balkans, discussed in chapter 3. Others involve types of interaction and mutual accommodation which make it difficult to place them in a single category, for instance the kinds of extreme structural convergence found in Northwest New Britain, where languages of the Austronesian and non-Austronesian families have become structurally isomorphic (see chapter 3). Similar difficulties arise in the case of the so-called "new" contact languages, pidgins (chapter 8), creoles (chapter 9), and bilingual mixed languages (chapter 6). These are cases neither of maintenance nor of shift in the strict sense, though they share characteristics with the latter situations. Each of them presents its own problems of definition and classification.

1.4.1 Language maintenance

1.4.1.1 Borrowing situations

Language maintenance refers simply to the preservation by a speech community of its native language from generation to generation. Preservation implies that the language changes only by small degrees in the short run owing to internal

developments and/or (limited) contact with other languages. Hence the various subsystems of the language – the phonology, morphology, syntax, semantics, and core lexicon – remain relatively intact.

Cases of maintenance may involve varying degrees of influence on the lexicon and structure of a group's native language from the external language with which it is in contact. This kind of influence is referred to as "borrowing." Since this term has been used in a variety of senses, it is necessary to emphasize that it is used here, following Thomason and Kaufman (1988: 37), to refer to "the incorporation of foreign features into a group's native language by speakers of that language." This makes it clear, first, that the borrowing language is maintained, though changed in various ways by the borrowed features, and that the agents of change are its native speakers. As van Coetsem (1988: 3) points out, borrowing involves recipient language agentivity, and this crucially distinguishes it from the other major type of cross-linguistic influence that involves source language agentivity in cases of second language learning (see section 1.4.2 below). The borrowing language may be referred to as the recipient language, and the foreign language as the source language. Both of these terms may also be used in a wider sense, to refer respectively to (a) any language that incorporates features from another and (b) any language that provides the relevant input.

Borrowing is also sometimes referred to as "borrowing interference" (as opposed to "interference via shift"), reflecting a tendency within the field to use the term "interference" as a cover term for all kinds of contact-induced change (Thomason and Kaufman 1988). Since the term "interference" has been used in a variety of conflicting senses, some general, some rather narrow (for instance, Weinreich 1953: 1 defines it as "deviations from the norm of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language"), the term will be avoided as far as possible here. Instead, we will use terms like "contact-induced changes" and "cross-linguistic influence" as general labels to cover all kinds of influence by one language on another.

Borrowing may vary in degree and kind from casual to heavy lexical borrowing, and from slight to more or less significant incorporation of structural features as well. As already noted, situations involving primarily lexical borrowing, that is, borrowing of content morphemes like nouns, verbs, etc., are extremely common, and most, if not all, languages have been subject to this kind of influence at some time or another. Sometimes, as we shall see later, significant lexical borrowing may have effects on the lexical semantics as well as other aspects of a language's structure. Situations involving structural borrowing, that is, borrowing of features in phonology, morphology, syntax, and semantics, are somewhat rarer, though examples can be found. Borrowing situations will be discussed in chapter 2.

1.4.1.2 Situations of structural convergence

Structural diffusion often occurs where languages are spoken in close geographical proximity, for example in border areas, or in communities characterized by a high degree of multilingualism. Examples of the former type of situation are Sprachbünde or linguistic areas. Perhaps the best-known of these is the Balkan Sprachbund, where long-standing contact between languages like Albanian, Bulgarian, Greek, Macedonian, and others led to significant diffusion of structural features. In cases involving bi- or multi-lingualism within the same speech community, the results of language contact are often manifested in increasing structural convergence between the languages involved. A well-known case in point is the village of Kupwar in India. Here, a long history of interaction between speakers of Marathi, Kannada, and Hindi-Urdu led to a surprising degree of isomorphism in structure, to the point where it has been claimed that simple replacement of lexical items from each language within the same structural frame is often possible. Long-term pressure on the language of a minority group surrounded by a larger dominant group can sometimes lead to significant structural and lexical diffusion from the latter to the former. This can in some cases lead to a radically altered version of the recipient language. Cases in point include Asia Minor Greek, which incorporated many features from Turkish, and Wutun, a Chinese language heavily influenced by Tibetan.

Sometimes, diffusion of features across languages may be so widespread that the boundaries between the languages become blurred, even for the speakers themselves. Thurston (1987, 1994) describes situations like this in Northwest New Britain, an island that forms part of Papua New Guinea. Here, as in Kupwar, convergence has led to structural isomorphism among the languages involved, with lexicon serving as the primary means of distinguishing one from the other. Thus, though they belong to quite distinct language families (Austronesian versus non-Austronesian), or to different subgroups within these families, all languages use practically the same syntactic strategies. For example, requests for items follow the same pattern: first the requested item is named, followed by a third person form of the verb come; then there is a first person verb expressing what the speaker will do with the desired item. The following examples illustrate. Anêm is non-Austronesian. Mouk and Lusi belong to the Bibling and Bariai subgroups of Austronesian respectively. Amara is an Austronesian isolate:

(5)

Anêm:	uas	gox	o-mên	da-t
Mouk:	uas	silaŋ	max	ŋa-ŋan
Lusi:	uasi	eta	i-nama	ŋa-ani

Amara: aguas kapso i-me e-kenen
tobacco some 3s-come 1s-eat
"Hand me some tobacco to smoke" (Thurston 1987: 69)

In cases like these, it is often difficult to identify the agents of change, whether they may be native speakers of language A who maintain it while borrowing, or speakers of language B who shift to A and introduce features of B which native speakers of A eventually adopt. These situations will be discussed more fully in chapter 3.

1.4.1.3 Code-switching situations

Language maintenance situations also include more or less stable bilingual speech communities in which bilingual mixture of various types is usual, leading to the phenomena known collectively as code switching. This involves the alternate use of two languages (or dialects) within the same stretch of speech, often within the same sentence. For example, Puerto Ricans in New York city switch between Spanish and English with great facility, as illustrated in the following example from Blanca, a 9-year-old girl living in Spanish Harlem, New York city. Spanish items are italicized:

(6) Hey Lolita, but the Skylab, the Skylab *no se cayó pa(-ra) que se acabe el mundo*. It falls in pieces. *Si se cae completo*, yeah. The Skylab *es una cosa que (e-)stá rodeando el* moon taking pictures of it. *Tiene tubos en el medio. Tiene tubos en el medio*. It's like a rocket. It's like a rocket.

(Hey Lolita, but the Skylab, the Skylab ("didn't fall for the world to end"). It falls in pieces. ("If it falls whole"), yeah. The Skylab ("is something that's going around the") moon taking pictures of it. ("It has tubes in the middle") [repeated]. It's like a rocket [repeated]. (Zentella 1997: 117)

Notice how Blanca switches languages from clause to clause, but also mixes items from the two languages within the same clause. These are examples of inter- and intra-sentential switching, which reflect somewhat different kinds of bilingual competence, as we shall see.

In many bi- or multi-lingual communities, the choice of one code or another is dependent on the situation or domain of use, so that the codes tend to be used in mutually exclusive functions. Such situations are referred to as cases of diglossia, or (where more than two languages are involved) polyglossia. An example of the former is Spanish/Guaraní bilingualism in Paraguay, while the latter is exemplified by the situations in Singapore and Malaysia, where speakers alternate between English, Malay, and other ethnic languages like Mandarin depending on the interlocutor and the situation (Platt 1977). Situations like these, of course, also allow for a certain degree of code alternation and code mixture within a single interaction. The social and linguistic aspects of code switching will be discussed in chapters 4 and 5 respectively.

1.4.2 *Language shift*

In other situations, contact between different linguistic groups can lead to language shift, the partial or total abandonment of a group's native language in favor of another. In some cases, the shift results in successful acquisition of the target language (TL), with little or no influence from the native language (L1) of the shifting group. For instance, by the third generation, most immigrant groups in the United States succeed in achieving native proficiency in American English. In many cases, however, shift is accompanied by varying degrees of influence from the group's L1 on the TL. Such situations fall into two broad categories. First, there are cases involving immigrant or other minority groups that shift either partially or completely to the language of the dominant majority, but carry over features of their L1 into their version of the TL. Sometimes, the shifting group is eventually absorbed into the TL community and the innovations that they introduced are imitated by the TL community as a whole, thus becoming permanently established in the language. This happened, for instance, when speakers of Norman French shifted to English in the late Middle English period, leading to significant lexical and some structural (especially phonological) influence from French on English. In other cases, a minority group may preserve its L1 for certain functions, while acquiring the dominant language for other uses. Such situations typically result in significant L1 influence on the TL, as for example in the second language varieties of German used by "guestworkers" in Germany from the late 1950s on. Such influence tends to be confined to the minority group and does not usually spread into the language of the host community as a whole.

The second category of situation where shift leads to L1 influence on a recipient language involves languages that become targets of shift after being introduced into new communities by invaders or colonizers. The indigenous community then adopts the foreign language either as a replacement for its original native language(s), or as a second language to be used in addition to the latter. Such "indigenized" varieties of a foreign language are especially common in areas that were formerly colonized by external powers. Indian English and Irish (Hiberno-) English are two examples. Second language versions of target

languages such as these, which result from untutored learning in "natural" community settings, are clearly similar in certain ways to the varieties of second or foreign languages acquired in formal settings such as the classroom. "Interlanguage" phenomena in classroom second language acquisition (SLA) often arise from the same kinds of L1 influence that characterize "untutored" SLA, that is, targeted language shift. Moreover, both types of learning may be subject to other principles and constraints, such as the universal tendency toward simplification of target structures, at least in the early stages of learning. There is therefore much to be gained from a close comparison of all these types of language acquisition.

Language shift obviously implies the gradual or complete abandonment of a previous native language in favor of the TL. Such situations provide interesting insight into the phenomenon of *language death*, the slow attrition and decay of the language previously used by the shifting group.

As noted above, many of the changes in a TL which accompany shift are the result of influence from the shifting group's L1. Such changes have been referred to by various names, including "interference through shift," "transfer," "substratum influence," and "imposition." Some of these labels are problematic in one way or another. We've already seen that "interference" is used in several conflicting senses. The same is true of "transfer," which is used by some as a cover term for all kinds of contact-induced change (hence "borrowing transfer" versus "substratum transfer"), and by others to refer only to L1 influence on an L2. Most SLA researchers use the term "transfer" to refer only to L1 influence on (learner versions of) a target language. Van Coetsem (1988: 3) introduced the term "imposition" to refer to this kind of contact-induced change. Though this term has failed to gain currency, his description of the change itself is quite insightful. As he notes, it involves the agentivity of source language speakers who "impose" their L1 habits on the recipient or target language.

The term "substratum influence" is popular among creolists, who use it to refer to much the same phenomena that SLA researchers describe as (L1) transfer – hence the growing rapport between these fields, as we shall see in chapter 9. Creolists use the term in a somewhat different sense from historical linguists. The latter generally use it to refer to influence from the language of a subordinate group, distinguishing it from "superstratum" and "adstratum" influence from the languages of dominant and equal groups respectively. Creolists on the other hand use it to refer specifically to influence from a subordinate group's language on pidgin and creole formation. Henceforth, we will use the term "L1 influence" or "substratum influence" to refer to the influence from a speaker or group's L1 on an outcome of language contact. It is immaterial whether the outcome is a second language variety of a TL or a new creation

such as a creole, or what the relative statuses of the languages (groups) in contact may be.

Thomason and Kaufman seem to have this sense in mind when they define substratum influence as the result of "imperfect group learning during a process of language shift" in the course of which the learning group commits "errors" that may spread to the TL as a whole. This definition may not be precise enough. In the first place, the results of "imperfect learning" may include strategies ("errors") other than substratum influence, such as simplification of TL structures. Second, not all cases of substratum influence result in spread of such influence to the TL as a whole. There are indeed such cases, usually when the shifting group is absorbed by the TL community. However, there are also cases where the shifting group constitutes a separate community in its own right, and the changes they introduce remain restricted to their version of the TL (e.g., Hiberno-English and other "indigenized" Englishes). In addition, we may want to distinguish between individual and group shifts. Thomason and Kaufman are right to note that group shifts promote substratum influence in a TL. But we can gain much insight into this type of cross-linguistic influence by investigating the strategies employed by individual learners in both "natural" and "tutored" contexts. As Mufwene (1990: 2) notes, "interference" from an L1 at the individual level is the first stage in the establishment of substrate influence in the language of the group. When the same types of change are replicated by various individuals and are adopted by many others, they become conventionalized as part of the community's linguistic system and at this point they can be described as substratum features.

Substratum or L1 influence, like borrowing, may be found at all levels of linguistic structure. But, in general, borrowing begins with vocabulary, and the incorporation of structural features into a maintained language comes only after substantial importation of loanwords. By contrast, substratum influence begins with sounds and syntactic patterns and sometimes also morphology, and is therefore characterized by more structural than lexical influence from the L1 on the TL. Thomason and Kaufman offer a sketch of the difference between borrowing and shift as illustrated by Rayfield's (1970: 85) description of mutual influence between English and Yiddish as spoken by a group of bilinguals in the United States (see table 1.1).

As table 1.1 shows, the process of borrowing from English into the Yiddish of these immigrants involves the lexicon much more than either phonology or morphosyntax. On the other hand, structural influence from Yiddish on the English of this group is much more pronounced than lexical influence.

These differences in the patterns of contact-induced change in borrowing as opposed to shift situations appear to be quite common, perhaps even predictable, and the distinction is therefore crucial to our understanding of what goes

Table 1.1 Degrees of "interference" in bilinguals' languages

	English → Yiddish (borrowing)	Yiddish → English (substratum influence)
Lexicon	Very strong	Moderate
Phonology	Weak	Strong
Morphosyntax	Moderate	Strong

Source: Thomason and Kaufman 1988: 40, table 2

on in different contact situations. It has important implications for both our methodology and our theories of contact-induced change. Methodologically, it means that we must understand the precise nature of the contact situation to determine the directionality of change and its agents. As far as theory is concerned, it means that explanations or predictions of the results of contact will vary depending on which of the two major vehicles of change is involved.

Exercise 2
Rayfield (1970) predicts that in situations of second language learning, lexical borrowing from the L2 will be much more frequent than structural borrowing in the L1 of the learners, while structural changes due to L1 influence will be more frequent in the learner's version of the L2. Investigate the use of English or any other language as a second language by international students at your university. Does Rayfield's prediction hold true as far as their usage is concerned?

1.4.3 Language creation: new contact languages

In addition to maintenance and shift situations, there are other kinds of contact setting which have yielded rather special outcomes: the contact languages referred to as pidgins, creoles, and bilingual mixed languages. These outcomes involve such extreme restructuring and/or such pervasive mixture of elements from more than one language that they cannot be considered cases of either maintenance or shift in the strict senses of those terms. It is also difficult at times to decide which outcomes of contact should be included in each of the above categories of contact language. The labels "pidgin" and "creole," for instance, have each been applied to a very heterogeneous group of languages,

which differ both in the circumstances of their creation and in their structural characteristics. For this reason, it is necessary to refer to "prototypical" examples of each category, and attempt as far as possible to relate other potential members of the class to the prototype (Thomason 1997c).

1.4.3.1 Bilingual mixed languages

Bilingual mixed or intertwined languages arose in settings involving long-term contact between two ethnic groups leading to bilingualism and increasing mixture of the languages. In these cases, that mixture became conventionalized as a community norm, resulting in the creation of hybrid languages whose components could clearly be traced to one or the other source language. We saw one example of a bilingual mixed language, Anglo-Romani, earlier in this chapter. Another example is the Media Lengua of Ecuador, a language which incorporates Spanish lexicon into a virtually unchanged Quechua grammatical framework. The latter preserves intact not just the syntactic rules of Quechua, but also its highly complex morphology. Here is a brief example, in which a Media Lengua speaker explains how the language is made up. Items derived from Spanish are in italics:

(7) *Media Lengua*-ga *asi* Ingichu-munda *Castallanu*-da *abla*-na
Media Lengua-TOP thus Quechua-from Spanish-ACC talk-NOM
kiri-xu-sha, *no abla*-naku-ndu-mi *asi*, chaupi-ga *Castellanu* laya,
want-PROG-SUB not talk-pl-SUB-AFF thus, half-TOP Spanish like,
i chaupi-ga Ingichi laya *abla*-ri-na ga-n.
and half-TOP Quechua like talk-REFL-NOM-be-3.

"Media Lengua is thus if you want to talk Spanish from Quechua, but you can't, then you talk half like Spanish, and half like Quechua." (Muysken 1997a: 377)

Other somewhat similar examples are Michif, a language in which Cree VP structure is wedded to French NP structure, and Mednyj Aleut, in which Russian finite verb morphology and other structural features have been fused with Aleut grammatical systems. In general, it is fair to say that these vernaculars fuse the grammar of one source with the lexicon (at least the phonological representations of the lexical items) of another. However, this picture is simplistic, since it ignores many respects in which a bilingual mixed language may differ from either of its source languages. Moreover, no single formula can be applied to describe or predict the mixture, even though there are many similarities in design among them. These and other aspects of the genesis and structure of bilingual mixed languages will be discussed further in chapter 6.

1.4.3.2 Pidgins

Trading contacts between groups speaking different languages have often led to various types of linguistic compromise to facilitate communication. Such compromises often result in pidgins, highly reduced languages with minimal vocabulary and grammar whose functions are restricted primarily to barter and exchange. An example of the pidgin English used for trading between English speakers and Pacific islanders in the nineteenth century was provided earlier in this chapter. Pidgins are a rather mixed bag of languages. Some involve more lexical mixture than others. For instance, Russenorsk, used in trade between Russians and Norwegians up to the nineteenth century, employed vocabulary from both groups' languages. Other pidgins, like Eskimo Trade Pidgin and Chinese Pidgin English, derive their vocabulary primarily from one source, Eskimo in the former, English in the latter. The primary source language in these cases tends to be the language of the group that has control of the trade or its location. Pidgins have also arisen in contexts other than trade, for instance in cases of military occupation (Pidgin English in Japan during the post-war period) or in domestic settings for communication between employers and servants of different language backgrounds (Indian Butler English) or on plantations (Hawai'i Pidgin English).

The cases mentioned so far are all examples of prototypical pidgins. The label is necessary because there is in fact a great deal of controversy over the scope of reference of the term "pidgin." The reason is that the degree of reduction in structure as well as range of functions may differ significantly from one case to another. Prototypical pidgins are severely restricted in terms of their social functions, and clearly reduced in form and structure, containing a minimal lexicon and a rudimentary grammar. Bickerton (1981) describes them as lacking inflectional morphology, tense/mood/aspect sytems, movement rules, embedding strategies, and other structural characteristics associated with fully developed natural languages. The sociohistorical and structural criteria by which such pidgins are defined will be outlined further in chapter 8.

By contrast, other languages to which the term "pidgin" has been applied, for example, Tok Pisin, Nigerian Pidgin, etc., are far more elaborate in terms of social function and structure, and hardly meet the criteria for inclusion in this class. These more elaborate contact languages may be placed in two broad categories: extended pidgins and simplified languages, though once more, the boundaries between these two are not always clear.

So-called extended pidgins apparently began as highly reduced (prototypical) pidgins which then underwent varying degrees of elaboration in both vocabulary and grammar when their range of functions extended beyond the confines of their original contexts of use. In such cases, there is usually incorporation of

features from both the lexifier (superstrate) language and the native (substrate) languages of indigenous groups. Contact vernaculars like these can achieve such a degree of elaboration in this way that they become indistinguishable from other fully developed natural languages. Examples include Tok Pisin and Bislama, official languages of Papua New Guinea and Vanuatu respectively, both descended from an earlier plantation pidgin, in turn rooted in early Pacific Trade Pidgin. Other examples include varietes of West African Pidgin English, such as Nigerian Pidgin English, that are used as lingua francas in various parts of West Africa. These contact languages have much more in common, both functionally and structurally, with creoles than with prototypical pidgins.

There are other contact vernaculars to which the label "pidgin" has been applied which do not appear to involve the degree of structural reduction characteristic of prototypical pidgins. For instance, languages like Trade Motu or Pidgin Yimas appear to be somewhat simplified forms of Motu and Yimas respectively, only partially reduced so as to facilitate their use by non-native speakers in trading and other contacts with native speakers. Their degree of reduction is not nearly as extensive as that found in, say, Russenorsk. Hence they should arguably be referred to as simplified languages, rather than pidgins. All of these cases and others like them will be discussed more fully in chapter 8.

1.4.3.3 Creoles

European colonial expansion during the fifteenth to nineteenth centuries led in many cases to the creation of new communities peopled primarily by groups transplanted from distant regions of the world. In the plantations of the New World, where huge numbers of slaves were transplanted from West Africa, contact between the latter and European settlers led to the emergence of creole languages, so called because they were used by the creole or locally born descendants of slaves (as well as Europeans and other freemen) in the colonies. A typical example is Sranan Tongo, a brief sample of which was provided earlier in this chapter. Other well-known Caribbean creoles include Jamaican and Guyanese creole (English lexicon); Haitian creole (French lexicon); Papiamentu, a creole used in the former Dutch islands of Aruba, Bonaire, and Curacao (Spanish/Portuguese lexicon) and Berbice Dutch, once spoken in the interior of modern Guyana (Dutch lexicon).

Similar languages emerged in the Indian Ocean and other areas where European colonies were established. For instance, there is Isle de France creole, a French-lexicon creole with varieties spoken in Mauritius and the Seychelles. In South East Asia, we find creoles such as Daman Creole Portuguese, spoken in India, and Papia Kristang, spoken in Malaysia and Singapore. There are also

several other creole languages spoken in West Africa, including Krio (English-lexicon), spoken in Sierra Leone, and Guinea Kriyol (Portuguese-lexicon), spoken in Guinea-Bissau. Some of the earliest creoles known arose on plantation settings on islands off the West African coast. Well-known examples include Cape Verde Crioulo and other Portuguese-lexicon creoles spoken on São Tomé, Principe, and other islands in the Gulf of Guinea.

The formation of these languages involved varying degrees of input from the superstrate languages of the colonizers and the native languages of the subjected peoples. Creoles, like other contact vernaculars, differ significantly in the nature and extent of the respective inputs. Just about every aspect of these languages, their origins and sources, their typological characteristics, their classification, etc., remains a matter of controversy. These issues will be discussed more fully in chapter 9.

As with "pidgins," there are substantial differences among so-called "creoles" in terms of both their processes of formation and their structural make-up. Essentially, such differences have to do with the nature and extent of the substratum contribution to the creole's formation. On the one hand, there are radical creoles like Sranan and its Surinamese relative Saramaccan, and varieties of the Eastern Maroon Creole, a substantial part of whose grammar can be traced to West African (especially Gbe) sources. For this reason, it is difficult to accept Thomason and Kaufman's characterization of them as cases of shift "whose structure can be accounted for under a hypothesis of extreme unsuccessful acquisition of a TL" (1988: 48). One might just as well argue that they are akin to cases of maintenance, though, as usual, the truth lies somewhere between these two extremes.

By sharp contrast, the so-called intermediate creoles of the Caribbean, such as Bajan, urban Guyanese, or Trinidadian creole, are arguably cases of shift and far more akin to products of "unsuccessful" acquisition of a TL such as Hiberno-English, Singapore English, Taiwanese Mandarin, etc. than they are to radical creoles. Once more, between these poles lie many other points on a continuum that includes contact vernaculars in the Caribbean, Pacific, Indian Ocean, and elsewhere to which the label "creole" has traditionally been applied.

1.5 Overview of Contact Situations and their Outcomes

At this point, it may be useful to provide a brief taxonomy of contact situations and the types of cross-linguistic influence they involve. Table 1.2, based partly on Thomason and Kaufman (1988: 50), illustrates the major outcomes of language contact. The table distinguishes three general categories of outcome, those

Table 1.2 Major outcomes of language contact

(A) Language maintenance

I Borrowing situations

Degree of contact	Linguistic results	Examples
Casual	Lexical borrowing only	Modern, English borrowings from French, e.g., *ballet*
Moderate	Lexical and slight structural borrowing	Latin influence on Early Modern English; Sanskrit influence on Dravidian languages
Intense	Moderate structural borrowing	German influence on Romansh

II Convergence situations

Type of contact	Linguistic results	Examples
Contiguous geographical location	Moderate structural diffusion	Sprachbünde, e.g., the Balkans
Intra-community multilingualism	Heavy structural diffusion	Marathi/Kannada influence on Kupwar Urdu
Intense pressure on a minority goup	Heavy structural diffusion	Tibetan influence on Wutun; Turkish influence on Asia Minor Greek
Intense inter-community contact (trade, exogamy)	Heavy lexical and/or structural diffusion	The languages of Northwest New Britain; the languages of Arnhem Land, Australia

(B) Language shift

Type of shift	Linguistic results (substratum)	Examples
Rapid and complete (by minority group)	Little or no substratum interference in TL	Urban immigrant groups shifting to English in the US
Rapid shift by larger or prestigious minority	Slight to moderate substratum interference in TL	Norman French shift to English in England

Table 1.2 (*cont'd*)

Type of shift	Linguistic results (substratum)	Examples
Shift by indigenous community to imported language	Moderate to heavy substratum interference	Shift to English by Irish speakers in Ireland (Hiberno-English); shift to English dialects in seventeenth-century Barbados (intermediate "creole")

(C) Language creation (new contact languages)

Type	Characteristics
Bilingual mixed languages	Akin to cases of maintenance, involving incorporation of large portions of an external vocabulary into a maintained grammatical frame
Pidgins	Highly reduced lingua francas that involve mutual accomodation and simplification; employed in restricted functions such as trade
Creoles	Akin to cases of both maintenance and shift, with grammars shaped by varying degrees of superstrate and substrate influence, and vocabulary drawn mostly from the superstrate source

pertaining to language maintenance situations (here subdivided into borrowing and convergence situations), those relating to language shift, and those involving the creation of new contact vernaculars, viz., pidgins, creoles and bilingual mixed languages.

1.6 The Social Contexts of Language Contact

Precisely what factors determine the varied outcomes of the contact situations we have just surveyed? We have already emphasized the complementary roles of external and internal factors in shaping such outcomes. Early scholars such as Müller (1875) and Jakobson (1938) argued that structural (linguistic) constraints were the primary determinants of contact-induced change. But the wide body

of evidence available to us now shows that practically any linguistic feature can be transferred from one language to another, if the circumstances are right. The reason is that extralinguistic factors – the social ecology of the contact situation itself – can override any purely structural resistance to change. Moreover, it is such factors that explain one of the key problems of language contact studies – why all potential forms of contact-induced change may not actually materialize in a given situation. This does not mean, of course, that explanations in terms of purely linguistic constraints are not possible or relevant. It is of prime importance for us to seek explanations as far as possible in linguistic structure, But ultimately, as Weinreich (1953: 3) so aptly stated: "A full account of interference in a language contact situation, including the diffusion, persistence and evanescence of a particular interference phenomenon, is possible only if the extra-linguistic factors are considered."

We will consider the various linguistic constraints on contact-induced change in some detail in our discussions of specific contact situations and their outcomes in later chapters. For the present, let us survey briefly the sociocultural factors that play so important a role in regulating these outcomes.

1.6.1 Language contact in its social settings

It bears repeating that the broad distinctions we have made between situations involving language maintenance, language shift, and the creation of new contact languages are crucial to explaining the linguistic outcomes of contact. Without a clear understanding of the history and social dynamics of the contact situation, we are in no position to explain anything. Not just the mechanisms of change but also its directionality and agentivity vary according to the type of situation involved. It follows that the constraints on the changes that can occur will vary from one case to another as well. In general, however, the same set of sociocultural factors is present in every contact situation, though the particular mix varies from case to case, with consequent variation in the results. These sociocultural factors include the types of community settings, the demographics of the populations in contact, the codes and patterns of social interaction among them, and the ideologies and attitudes that govern their linguistic choices. Other factors that play a role include the degree of bilingualism among the individuals and groups in contact, the history and length of contact, the power relationships between the groups, and so on. Obviously, it is no easy task to integrate all the relevant factors into a comprehensive and coherent picture of the social ecology of a given contact situation. In the following chapters, we will try to examine the social setting of each type of contact in more detail, and show, as far as possible, how it contributes to the particular outcome in

of lexical borrowing, and determine which items become permanently installed in the recipient language, as distinct from being employed as nonce switches in bilingual code switching. Further discussion of the similarities and differences in code-switching patterns in bilingual contact situations will be found in chapter 4.

2.6 The Processes and Products of Lexical Borrowing

What exactly is a lexical borrowing? We've proceeded so far as though the answer to this question was clear. However, the phenomena that have been referred to by this label are quite varied. Some are close imitations of foreign items (e.g., *rendezvous* borrowed from French into English). Others are items that have been thoroughly transformed in shape (e.g., Costan Rican Spanish *chinchibí* from English *gingerbeer*), while still others are inventions that employ only recipient language materials in imitation of some foreign pattern (e.g., Spanish *rascacielos* modeled on English *skyscraper*). In fact, many so-called "borrowings" are not the result of a direct or complete adoption of a foreign item with both its form and meaning intact. The process of borrowing can be very selective, adopting a foreign form but assigning it a new meaning (e.g., Japanese *sumato* "slim, slender" < Eng. *smart*), or adopting a foreign meaning or concept and assigning it to a native form (e.g., Japanese *sara*, extended to include Western-style "plate"). Also, many of the outcomes of lexical borrowing involve innovations or creations that have no counterpart in the donor language. Some of these innovations may be created out of donor materials (e.g., Japanese *wan-man-ka* "bus without a conductor" < English *one+man+car*). Others may be created out of native materials, for example Zapotec *éxxuwí* "fig" < *exxu* "avocado" + *wí* "guava" and (older) coinages in the Pima language such as "wrinkled buttocks" for "elephant" and "dog person" for "monkey" (Herzog 1941: 68). Still other creations are blends of native and foreign items (e.g., Yaqui *lios-nóoka* "pray" < Span. *Dios* "God" + Yaqui *nóoka* "speak"). It would appear that the composition of lexical entries can be manipulated and rearranged in a variety of ways to produce these outcomes of contact. [1]

Attempts to establish a coherent framework for dealing with contact-induced changes in the lexicon began as early as the nineteenth century with Paul (1886) and others, and continued in the first decades of the twentieth century with scholars like Seiler (1907–13) and later Eugen Kaufman (1939). Perhaps the most comprehensive of the early frameworks was that of Betz (1949), whose basic distinction between *Lehnwort* (loanword) and *Lehnprägung* (loancoinage) still forms the basis for current descriptions.

Haugen (1950a, 1950b, 1953) added a new dimension to existing classifications with his distinction between importation and substitution – a dichotomy based on the presence or absence of foreignness markers (1950b). Importation refers to the adoption of a foreign form and/or its meaning, and may involve complete or only partial imitation. Substitution refers to the process by which native sounds or morphemes are substituted for those in the donor model. For example, in producing *rendezvous*, English speakers generally fail to reproduce the uvular [R] of French, using their own continuant [ɹ] instead. This is a case of phonemic substitution. Cases where a meaning or concept is borrowed but expressed by a native form are instances of morphemic substitution. An example is Spanish *rascacielos*, discussed above. In short, for Haugen, "every loan [is] part importation and part substitution" (1953: 388).

Following Haugen (1953), we can classify lexical contact phenomena into two broad categories – *lexical borrowings*, which involve imitation of some aspect of the donor model, and *creations*, which are entirely native and have no counterpart in the donor language. Lexical borrowings can be further subdivided into two categories. First, there are *loanwords*, in which all or part of the morphemic composition of the loan derives from the external source language. Second, there are *loanshifts*, in which the morphemic composition of the item is entirely native, though its meaning derives at least in part from the donor language. Each of these categories can be further subdivided, according to the types of importation and substitution involved.

Loanwords may be divided into two categories; "pure loanwords" and "loanblends." Pure loanwords may consist either of single words like *rendezvous* or compounds like *chincibiri*. Sometimes these undergo semantic modification of some sort. For instance, the English word *corner* is borrowed into Dutch only in its football (soccer) sense, to refer to a corner kick. As usual, borrowed compounds or phrases may also be adjusted both phonologically and syntactically, like the phrase *objetores conscientes* borrowed into Florida Spanish from English *conscientious objectors* (Ortoz 1949, cited in Weinreich 1953: 50).

"Loan blends" involve the transfer of part of the foreign model and the reproduction of the rest (importation of a foreign morpheme combined with substitution of a native one). Examples of such "hybrids" include Pennsylvania German (PG) *esix-jug* "vinegar jug" and *home-plato* in Tampa Spanish (Weinreich 1953).

Loanshifts or loan meanings fall into the following subtypes. In some cases, a native word may undergo extension of its meaning on the model of a foreign counterpart. These are cases of "extensions" or "semantic loans." For example, Yakut *tahym*, originally "water level," was extended to mean all kinds of level, both concrete and abstract (e.g., of water, of skill, of development, etc.) on the model of Russian *uroven'* (Mordinov and Sanžejev 1951: 41, cited by Weinreich

1953: 48). In other cases, native morphemes are employed to express new meanings imported from foreign sources. For instance, American Portuguese (Am. Port.) *humoroso*, originally "capricious," acquired the meaning "humorous" on the model of its English counterpart.

Loanshifts or coinages involving compounds allow for a wider variety of results combining direct transfer with "substitutions" of various types. For example, loanshifts may take the form of "pure loan translations" or calques in which the foreign model is replicated exactly by native words, for example American Portuguese *estar direito* "to be right" modeled on English. Sometimes the model is adjusted somewhat to fit native patterns of order, like Spanish *rascacielos* modeled on English *skyscraper*. Similar to this are "loan renditions" in which the model compound provides a general hint for the native imitation, like German *Wolkenkratzer*, lit. "cloud scraper," also based on English *skyscraper*. We also find "loan creations," that is, new coinages based on a foreign model, for example Yiddish *mitkind*, lit. "fellow-child," modeled on English *sibling*, German *Geschwister*, and the like (Weinreich 1953: 51). For the sake of simplicity, we will refer to all of these compound loan formations as "loan translations."

Table 2.3 presents a brief summary of types of lexical contact phenomena, based on Haugen's (1953) classification. I have modified his terminology and description somewhat. I have also expanded his category of "native creations" to include a third subcategory ("creations using only foreign morphemes," e.g., Japanese *wan-man-ka*), which was not included in Haugen's classification.

Many of the lexical phenomena included in table 2.3 are not direct results of the process of borrowing itself, but rather due to additional processes applied to borrowed items. Most "loanblends," for instance, arise when native (recipient language – RL) derivational processes are applied to previously imported words, for example, PG *bassig*. Others may result from applying a foreign derivational process to a native item, such as Japanese *ichigo-edo* "strawberry" + *-ade*. Loanblends of the first type are really due to the more general process of integrating loan items into the morphology of the recipient language.[2]

In the process of morphological adaptation, loanwords can be subjected to various other processes, such as clipping. This can affect single words, for example, Japanese *baito* "part-time job" < German *Arbeit*, as well as imported compounds, such as Japanese *wa-pro* < *wo[rd] pro[cessor]*. As Loveday (1996: 79) notes, these processes help to simplify the pronunciation of the loans and facilitate their integration.

Creative word formation involving imported items is another interesting by-product of lexical borrowing, which Haugen includes in his category of "native creations." New compounds may be built entirely out of native materials to express new concepts, for example Pima "wrinkled buttocks." As we saw earlier, some of these "native creations" may also be blends of foreign and native

Table 2.3 A classification of lexical contact phenomena

Types	Processes involved	Examples
I Borrowings (modeled on the donor language)		
A Loanwords:		
1 "Pure" loanwords	Total morphemic importation of single or compound words	*rendezvous*
	Varying degrees of phonemic substitution[a]	*chinchibiri*
	Possible semantic change	Dutch *corner*
2 Loanblends[b]	Combination of native and imported morphemes	
2a Derivational blend	Imported stem + native affix	PG *bassig* Eng. *boss* + Germ. *-ig*
	Native stem + imported affix	Jap. *ichigo-edo* "strawberry" + *-ade*
2b Compound blend	Imported stem + native stem	PG *blaumepie* "plum" + *pie*
B Loanshifts (loan meanings):		
1 "Extensions" (semantic loans)	Shifts in the semantics of a native word under influence from a foreign word	
	a Phonological resemblance	Am. Port. *humoroso* "humorous"
	b Partial semantic resemblance	Am. Port. *frio* "cold infection"
2 Loan translations (calques)	Combination of native morphemes in imitation of foreign pattern	Germ. *Wolkenkratzer* cf. Eng. *skyscraper*
II Native creations		
1 Purely native creations	Innovative use of native words to express foreign concepts	Pima "wrinkled buttocks" for "elephant"
2 Hybrid creations	Blends of native and foreign morphemes to express foreign concepts	Yaqui *lios-nóoka* "pray"
3 Creations using only foreign morphemes	Combinations of foreign morphemes for new concepts	Jap. *wan-man-ka*

Notes:

a Some cases that appear to belong in this category involve the phonological adjustment of a native word on the model of a foreign one, without change in content. Weinreich (1953: 50) cites the example of Tampa Spanish *europa* becoming *uropa* on the model of English *Europe*, and American Yiddish *vakátsje* becoming *vekejʃn* on the model of American English *vacation*. It's difficult to say whether these are really cases of phonological adjustment of the native word as distinct from importation (imitation) of the foreign counterpart.

b Haugen (1953: 399) includes what he calls "blended stems" under the category of "loanblends." He gives the example of American Norwegian *kârna*, which appears to be a blend of English *corner* and Norwegian *hyrna*. Such cases seem to be rare, and might well be treated as cases of morphemic importation with phonemic substitution. Hence I omit them from the present classification.

words, like Yaqui *lios-nóoka* "pray." But we also find new compounds being created entirely out of foreign materials. Examples include Japanese *wan-man-ka*, *gaadoman* "guard" < *guard* + *man*, and, most interestingly, *sukinshippu* "intimate, physical closeness" < Eng. *skin* + *-ship*. All of these are innovations based on native patterns, or creative extensions of a foreign pattern, which have no counterparts in the source language. We therefore need to distinguish the process of lexical borrowing from other processes that may apply to imported items. Moreover, we should distinguish all such processes (borrowing, adaptation, integration) from the products they create.

Exercise 4

The following are a number of Japanese lexical items that are products of lexical borrowing and attendant processes of integration. First, try to assign each item to one of the categories in Haugen's classification (table 2.3). What difficulties do you have with this, and what changes in classification would you suggest to resolve them? Second, describe the processes of adaptation and change that the borrowings have undergone. The examples are from Loveday (1996) and Ishiwata (1986):

apaato "apartment" < English *apartment*
dai-sutoraiku < Jap. *dai* "big" + Eng. *strike*
dansu paati "dance" < Eng. *dance* + *party*
dorai "unsentimental" < Eng. *dry*
goo-sutoppu "traffic signal" < Eng. *go* + *stop*
han-suto "hunger strike" < Eng. *hunger strike*
kaa "car, bus or truck" < Eng. *car*
ofisu redi "office girl" < Eng. *office* + *lady*
poteto furai "fried potatoes" < Eng. *potato* + *fry*
raisu "rice served on a plate with a Western-style dish" < Eng. *rice*
(Compare *gohan* "cooked rice served in a bowl and eaten with chopsticks" and *kome* "uncooked rice.")

2.7 The Integration of Loanwords

2.7.1 Phonological integration of loanwords

In cases of relatively light to moderate contact, lexical borrowings tend to be adapted in terms of the phonology and morphology of the recipient language,

and become essentially indistinguishable from native items. It is quite easy to find examples of this kind of integration (or, in Haugen's words, importation with phonological and morphological "substitution"). For instance, English loanwords in Japanese tend to be adapted to Japanese pronunciation as well as its preferred CV syllable structure. Japanese accomplishes this adaptation by various means, including epenthesis (e.g., *baseball* > *besuboru*), cluster simplification (*sweater* > *seta*), and syllabification of glides (*quizz* > *kuizu*). Another excellent illustration of the processes of integration is provided by the integration of borrowings from English and other languages into Hindi (and other Indic languages). Part of this involves a complex pattern of substitution of foreign stops and fricatives by perceived equivalents in Hindi. Hock (1991: 393) calls this an example of a system-based pattern of substitution. The pattern is summarized as follows:

Foreign	/p^h, t^h, k^h/	→	Indic	/p, t̪, ʈ, k/
	/f, θ, x/	→		/p^h, t^h, k^h/

English →	*Hindi*	*Example*
p^h	p	proof → $prup^h$
t^h	ʈ	tin → ʈin
k^h	k	concrete → kaŋkrit
f	p^h	phone → p^ho:n
θ	$t̪^h$	thermos → $t̪^h$armas
x (Arabic)	k^h	xatam → k^hatam

To find reasons for these substitutions, we need first to compare the very different structures of the Hindi and English obstruent systems, as in the following chart:

English				*Hindi*				
p	t	č	k	p	t̪	ʈ	č	k
				p^h	$t̪^h$	$ʈ^h$	$č^h$	k^h
f	θ	s	š	s			š	

Hock (1991: 394) suggests that foreign non-sibilant fricatives (f, θ, x) are "nativized as the corresponding Hindi voiceless aspirated stops because the friction noise of these aspirates approximates the acoustic impression of the foreign fricatives." Hence: English /θ/, a genuine dental, is rendered as Hindi dental /$t̪^h$/.

English /t/ is really alveolar, post-dental, and it is this "post-dentality" that is captured by the post-dental retroflex /ʈ/ of Hindi. Also, English aspirated stops like [p^h, t^h, k^h] may not be perceived as turbulent enough to be considered

instances of "true" aspiration by Hindi speakers, where aspiration has a much higher level of turbulence.

Hock's explanation of the substitutions recalls Weinreich's notion of "interlingual identifications" at the phonetic and other levels.

Exercise 5
Revisit exercise 4 in this chapter. Find out about Japanese phonemes and phonotactics and try to explain how Japanese rules of phonology might have influenced these changes.

2.7.2 *Morphological integration of loanwords*

In general, loanwords pose little problem for syntactic adaptation, simply behaving like their counterparts of different syntactic categories in the recipient language. However, morphological adaptation can prove more difficult, especially if the recipient language has complex rules involving case, number, gender, and the like. In many cases, borrowed words are treated like native stems of equivalent categorial status, and take the bound morphology and other properties appropriate to the class they are assigned to. But class assignment itself may be problematic. In (standard) Swahili, for example, nouns fall into 15 morphologically defined subclasses, each with its own pair of singular and plural suffixes, some of which are covert (Mkude 1986: 519). Differences in class membership are signaled by agreement markers appearing on demonstratives and other word classes which have to agree with the noun in question. Some examples are sufficient to illustrate:[3]

Class	*Sing. prefix*	*Plur. prefix*	*Examples*
Class 1–2	m(u)-	wa-	*mtu*, *watu* "person(s)"
Class 3–4	m(u)-	mi-	*mkia*, *mikia* "tail(s)"
Class 5–6	ø	ma-	*harage*, *maharage* "bean(s)"
Class 7–8	ki-	vi-	*kiti*, *viti* "chair(s)"
Class 9–10[4]	ø	ø	*nyama* "meat"
Class 11	u	n/a	*uhuru* "freedom"

In some cases, foreign loans are assigned to a noun class simply on the basis of a formal similarity to native stems. Thus, Arabic *kitab* "book" has been reanalyzed as *ki-tabu* and assigned to class 7–8, with the plural *vi-tabu*. A similar example is *ki-biriti* "match" (< Arabic *kibrit*), with the plural *vi-biriti*. Whiteley (1967) reported on some interesting cases of adaptation among the

speakers he observed. For example, some speakers assigned English loans like *madigadi* (< *mudguard*) and *machingoda* (< *marching-order*) to the *o-/ma-* (5–6) class on the basis of their initial CV sequence. Even more interestingly, they created singular forms like *digadi* and *chingoda* from these loans (via "back formation") by analogy with native singular items. These kinds of adaptation are not found in today's Swahili. It may be that they are a function of the degree of bilingualism and proficiency in the foreign language among borrowing speakers. The more usual strategy is to place foreign loans into classes 5–6 and especially 9–10, which lack overt class prefixes. Some estimates suggest that loanwords now make up more than 50 percent of the total words in these classes (Mkude 1986: 520).

Another interesting aspect of morphological adaptation involves the treatment of borrowed nouns and adjectives in languages like Dutch, French, German, etc., which assign grammatical gender to such items. The conventional wisdom is that the rules of the recipient language determine the assignment. But this may depend on various factors, including formal criteria (similarity in phonological shape), meaning, and analogy. For instance, English loanword *stress* is assigned masculine gender in German by analogy with nouns like *kampf* "struggle" which are semantically similar.

Poplack et al. (1988) investigated the role of five factors in gender assignment to English nouns borrowed into Montréal French: sex of (animate) referent; phonological shape; (semantic) analogy; homophony; and shape of suffix. They found that only the first of these was significant, though other factors played some role as well. This seems to be the case as well with French borrowings into Dutch. French nouns which refer to males (*agent* "agent," *facteur* "postman," etc.) receive masculine gender, while nouns referring to females (*danseuse* "female dancer," *madame* "madam") are assigned feminine gender (Treffers-Daller 1994: 130). As in Montréal, analogy plays a small role, but only in the case of disyllabic nouns with stress on the second syllable (*canon* "cannon," *palais* "palace," *prison* "prison"). These tend to be assigned neuter gender in Dutch by analogy with native deverbal nouns, whether or not they are masculine or feminine in French. In many cases too, Dutch assigns gender to foreign nouns on the basis of formal criteria. For instance, French nouns ending in *-ment* (*gouvernment* "government," *appartement* "apartment," etc.) generally receive neuter gender, while loans ending in *-iteit* (e.g., *variabiliteit* "variability") are assigned feminine gender on the basis of the suffix (Treffers-Daller 1994: 124).

There is also evidence that social factors such as degree of bilingualism and proficiency in the foreign language may influence the gender assignment of borrowed nouns. This seems to be the case with French nouns borrowed into Brussels Dutch, which tend to keep their original gender (either masculine or

feminine). This might be explained by the fact that the Brussels Dutch speakers are mostly bilingual in French, and familiar with gender assignment in the latter language. Moreover, the gender systems of the two languages match to a large extent, since Brussels Dutch distinguishes masculine, feminine, and neuter genders. Treffers-Daller (1994) finds that approximately 80 percent of the borrowed French nouns keep their gender in Brussels Dutch. This contrasts with French nouns borrowed into Standard Dutch, many of which are assigned neuter gender despite being either masculine (e.g., *bureau* "office," *numero* "number") or feminine (e.g., *station* "station," *terasse* "terrace") in French. Part of the explanation for this is that the gender systems of the two languages do not match, French nouns being classified as either masculine or feminine, while Standard Dutch nouns are either neuter or non-neuter. The fact that many French-origin nouns are assigned neuter gender in Standard Dutch may be due in part to borrowing speakers' unfamiliarity with French gender assignment (Treffers-Daller 1994: 125).

On the whole, it is clear that no single general rule applies to the way gender is assigned to borrowed nouns from one contact situation to another. The interplay of linguistic and social factors may vary significantly from one case to another, yielding different results.

The integration of loan items into the morphological structure of the recipient language can also involve creative processes of adaptation that yield additional lexical entries. In Japanese, for example, English loans are treated as uninflected nouns or stems which can be converted to other classes by the addition of suffixes or a helping verb (Loveday 1996: 118). For example, borrowed nouns may be converted into adjectives (or adjectival nouns) by attaching the suffix *-na* (e.g., *romanchikku-na* "romantic") or into adverbs via affixation of *-ni* (e.g., *romanchikku-ni* "romantically"). Borrowed nouns may also be converted for use as verbs by adding the dummy verb *suru* "do, make," for example *sain suru* "sign," *enjoi suru* "enjoy," etc. These strategies conform fully to Japanese patterns of derivation. Even the "clipping" of loan items common in Japanese (e.g., *han-suto* < *hanga-sutoraiki* < *hunger strike*) is a way of making such importations conform more closely to native Japanese morpho-phonology (Loveday 1996: 118).

The various types of integration we have examined here demonstrate that so-called "borrowing" involves complex patterns of lexical change that create new lexical entries or modify existing ones in response to culture contact. In all cases, borrowed items are manipulated so that they conform to the structural and semantic rules of the recipient language. This is what distinguishes the mechanisms of change associated with borrowing from those that characterize other vehicles of cross-linguistic influence, such as substratum influence. The kinds of adaptation and integration found in borrowing are also quite common in code switching and other outcomes of bilingual contact such as bilingual mixed languages, as we will see in chapters 5 and 6.

2.8 Linguistic Constraints on Lexical Borrowing

In addition to social factors, there are structural (linguistic) constraints which condition the degree and type of lexical borrowing. The most general constraint involves the well-known "hierarchy of borrowability," according to which open-class content items like nouns and adjectives lend themselves most easily to borrowing, while closed-class function items like pronouns and conjunctions are least likely to be adopted. Hierarchies of borrowing were proposed as early as the nineteenth century by Whitney (1881), and later by Haugen (1950b) and Muysken (1981b). The most comprehensive of these is the following, from Muysken:

> nouns > adjectives > verbs > prepositions > co-ordinating conjunctions > quantifiers > determiners > free pronouns > clitic pronouns > subordinating conjunctions

Part of the reason for the greater accessibility of nouns and adjectives lies in the fact that they form less tightly knit subsystems of the grammar than functional morphemes do. Moreover, they occur frequently in contexts where they can be isolated and extracted as loans. At the same time, the open-ended nature of these categories in the recipient language makes them more receptive to new additions. By contrast, the structuredness of classes such as pronouns, prepositions, etc. makes them highly resistant to borrowing. This reflects the more general hierarchical constraints on lexical versus structural borrowing to be discussed in chapter 3. Muysken supports his hierarchy with evidence from Spanish borrowings in Quechua. But the borrowing hierarchy in this case may not be fully representative of all situations. Appel and Muysken (1987: 171) emphasize the need to distinguish counts of tokens as distinct from types to ensure a more accurate picture of the hierarchy of borrowing. Still, the general outlines of the borrowing hierarchy are supported by other research, such as Poplack et al.'s study of English loans in Ottawa/Hull French, and Treffers-Daller's study of French loans in Brussels Dutch.

Syntagmatic constraints relating to the morphological and syntactic properties of lexical classes may also operate to favor or inhibit borrowing. This may explain why categories like verbs or prepositions, which govern other categories and assign case, tend not to be as heavily borrowed as nouns and

adjectives. Moreover, the greater the degree of morphological complexity in the paradigms of a lexical class, the more resistant it is to borrowing. Again, this may be why verbs, which tend to be morphologically complex as well as central to the syntax of the sentence, tend to be borrowed less than other open-class categories. The borrowing of verbs tends to be facilitated in cases where there is close typological similarity in verbal structure between the languages in contact, or where the borrowed item can be fitted easily into the morphology of the recipient language. Thus, most French verbs borrowed into Brussels Dutch tend to be from the *-er* class, since these lend themselves most readily to incorporation into the class of regular Dutch verbs whose infinitival suffix is *-en*. So we find *blesseren* "hurt" (< Fr. *blesser*); *rappeleren* "remember" (< Fr. *rappeler*), and so on (Treffers-Daller 1994: 110). In fact, so many of these French verbs have been borrowed that the French suffix *-er* has become somewhat productive in Brussels Dutch as a means of incorporating French verbs which do not even belong to the *-er* class. Thus we get BD *ofreren* "offer" (< Fr. *offrir*) and *finisseren* "finish" (< Fr. *finir*). Indeed, *-er* is often combined with French- or Latin-derived nouns to form verbs that aren't found in French, such as *fantaseren* "to fantasize" (ibid.: 111). We will see other examples of typologically favored borrowing of verbs in the discussion of convergence in Arnhem Land, Australia, in the following chapter.

Constraints having to do with degree of structural complexity may also explain the preference for morphologically simple lexical items over more complex ones in bilingual borrowing (Poplack et al. 1988: 60). In some cases, borrowing speakers may resort to strategies of simplification to facilitate the borrowing of verbs. A well-known example is provided by languages like Mayan, whose speakers borrow Spanish infinitives and use a Mayan verb meaning "do" as an auxiliary to which Mayan inflections can be added to convey tense/aspect meanings. A similar strategy is also found in Persian borrowings from Arabic, Japanese borrowings from English, and many other cases throughout the world. The same strategy is used in code switching (see chapter 5). Other languages follow the Brussels Dutch strategy of borrowing an infinitive and attaching a verb-forming suffix to it, for example, German borrowings from French and Russian borrowings from various languages (Thomason and Kaufman 1988: 349). Cases like these led Weinreich to suggest that the reasons why nouns tend to be borrowed more frequently than verbs has less to do with structural constraints than with "lexical-semantic" motivation. Still, structure does seem to have much to do with it.

Weinreich (1953: 61) also notes that typological differences in word structure may inhibit direct borrowing and promote the use of strategies like loanshifts or loan translations instead, when contact is sufficiently intense. He cites as an example the different types of borrowing from Sanskrit and Chinese into

Tibetan. Tibetan has borrowed directly from Chinese because of the similarity in word structure between the two languages, but has resorted to loan translations in borrowing from Sanskrit because of the mismatch between their word structures. Loan translations are particularly common when compounds are involved. Thus we find new formations in Brussels Dutch such as *ijzerweg* "railway" modeled on French *chemin de fer*, and *schoonbroer* "brother-in-law" (modeled on French *beau-frere* (Treffers-Daller 1994: 98). Some scholars, such as Heath (1984: 367), prefer to view such cases of "pattern transfer" or "calquing" as instances of structural convergence rather than lexical borrowing per se.

While there is much evidence for structural constraints on lexical borrowing, there are nevertheless many exceptions that do not follow the predicted patterns. Weinreich (1953: 62) provides several examples of the borrowing of words whose structure is typologically different from that of words in the recipient language. As always, structural constraints may not apply when the right social conditions prevail. As Weinreich (ibid.) puts it: "The unequal degrees of resistance to transfers and the preference for loan translations over transfers are a result of complex sociocultural factors which are not describable in linguistic terms alone."

We might note, finally, that there are strong constraints on the borrowing of basic or core as opposed to peripheral vocabulary. Indeed, the assumption that basic vocabulary is almost immune to replacement via borrowing is vital to assessments of language relatedness via the comparative-historical method. Some scholars use this criterion to establish whether contact-induced change is due to borrowing under language maintenance or to changes induced by shift. Thomason and Kaufman (1988), for example, argue that Ma'a, a bilingual mixed language, is a result of massive grammatical borrowing from Bantu languages into a previous Cushitic language, since most of the core vocabulary of Ma'a is of Cushitic origin. It's not clear that this conclusion rests on solid ground. In the following chapter, we will see cases in which a great deal of core vocabulary diffuses across language boundaries, suggesting that the constraints on such diffusion may not be as absolute as the traditional wisdom holds. However, it is difficult to say precisely what factors – structural or sociocultural – facilitate or impede changes in core vocabulary.

2.9 Structural Consequences of Lexical Borrowing

In cases of relatively intense contact, heavy lexical borrowing can be accompanied by the introduction of new sounds as well as morphemes which can affect the phonology and morphology of the recipient language. In fact, it has been claimed