I
The study of code-mixing

This book is about intra-sentential code-mixing and how it can help us understand language interaction as the result of contact, yielding a new perspective on central aspects of the human linguistic capacity. The question discussed here is: how can a bilingual speaker combine elements from two languages when processing mixed sentences? I am using the term code-mixing to refer to all cases where lexical items and grammatical features from two languages appear in one sentence. The more commonly used term code-switching will be reserved for the rapid succession of several languages in a single speech event, for reasons which will be made clear. However, sometimes the terms switch, switch point, or switching will be used informally while referring to the cooccurrence of fragments from different languages in a sentence. Of course, it will also be necessary to separate cases of code-mixing from lexical borrowing. The term language interaction will be used occasionally as a very general cover term for different, frequently highly innovative, results of language contact, both involving lexical items (as in code-mixing) and otherwise (e.g. phonological or syntactic interference).

In most models portraying the functioning of the speaker/listener, pictures we carry in our minds or see portrayed in a textbook, a single grammar and a single lexicon are embedded in the network of relations that constitutes the model. This is so commonplace that the essential enrichment of having several grammars and lexicons participate in it at the same time is often seen as a threat, a disruption, a malady. This is particularly the case in the structuralist tradition in linguistics. Ronjat (1913) and Leopold (1939–1949) formed the basis for the single parent/single language approach to bilingual child rearing – bilingualism in the family is ok, but it should remain tidy. Weinreich (1953: 73) thought that intra-sentential code-mixing was a sign of lack of bilingual proficiency and interference. An ideal bilingual ‘switches from one language according to appropriate changes in the speech situation (interlocutor, topic, etc.) but not in an unchanged speech situation and certainly not within a single sentence.’ A growing number of studies have
Bilingual speech

demonstrated, however, that many bilinguals will produce mixed sentences in ordinary conversations. What is interesting to me, as it has been to many others in recent years, is that such sentences are produced with great ease and complete fluidity. Indeed, for some speakers it is the unmarked code in certain circumstances (Myers-Scotton 1993a). Neither does it reflect limited proficiency in either of the languages involved. Rather, speakers who code-mix fluently and easily tend to be quite proficient bilinguals (Poplack 1980, Norrier 1990). Finally, we cannot assume either that it is word-finding difficulties or specific cultural pressures that lead to the mixture (even if language contact itself is culturally conditioned). Often, the element introduced corresponds to a household word.

In the last fifteen years, a large number of studies have appeared in which specific cases of intra-sentential code-mixing are analysed from a grammatical perspective. These cases involve a variety of language pairs, social settings, and speaker types. It is found that intra-sentential code-mixes are not distributed randomly in the sentence, but rather occur at specific points. Where much less agreement has been reached is with respect to general properties of the process.

This book is an attempt to present a general account of the very complex intra-sentential code-mixing phenomena that have been discovered. By now the amount of material collected for different language pairs is both diverse and substantial, and it is time to attempt a first synthesis. Rather than introducing one single data set, I will try to integrate the results of a great many different studies, some still unpublished. The present work is grounded both in structural linguistics and in sociolinguistics. Many of the characteristics of the mixing patterns are determined by the structural features of language; I will adopt the general tools and concepts of generative grammar in accounting for these (while trying to stay clear of highly specific formalisms and analyses). Structural analysis along generative lines will be combined with quantitative analysis as in the work of Labov and Sankoff, and comparative typological work. Occasionally, I will try to relate my interpretations to notions from psycholinguistics such as activation and processing. I will only infrequently have recourse to pragmatic and conversational analysis, partly because of my own lack of expertise, and partly because the wide-ranging comparative approach I am adopting here necessarily relies on data gathered by others less suited for detailed textual analysis, and often taking the form of isolated mixed sentences and tables.

The work reported on here could be considered to represent a taxonomic phase in the discipline, an attempt to tie together a set of intermediary results rather than giving a conclusive account. I feel the results from current studies
are so diverse that some tidying up is called for. Although the focus of the present work is grammar, it does not mean attention will not be given to the crucial role of psycholinguistic and sociolinguistic factors influencing code-mixing, such as degree of bilingual proficiency, mode of bilingual processing, political balance between the languages, language attitudes, and type of interactive setting. However, these factors are considered in so far as they are related to or manifest themselves in the grammatical patterns of code-mixing encountered. Indeed, any synthesis at present must depart from the enormous variation in code-mixing patterns encountered, variation due to language typological factors in addition to sociolinguistic and psycholinguistic factors such as those mentioned.

I do not propose a single ‘model’ of code-mixing, since I do not think there is such a model, apart from the general models provided by grammatical theory and language processing. The challenge is to account for the patterns found in terms of general properties of grammar. Notice that only in this way can the phenomena of code-mixing help refine our perspective on general grammatical theory. If there were a special and separate theory of code-mixing, it might well be less relevant to general theoretical concerns.

Different processes
The patterns of intra-sentential code-mixing found are often rather different from one another. Much of the confusion in the field appears to arise from the fact that several distinct processes are at work:

- **insertion** of material (lexical items or entire constituents) from one language into a structure from the other language.
- **alternation** between structures from languages
- **congruent lexicalization** of material from different lexical inventories into a shared grammatical structure.

These three basic processes are constrained by different structural conditions, and are operant to a different extent and in different ways in specific bilingual settings. This produces much of the variation in mixing patterns encountered. The three processes correspond to dominant models for code mixing that have been proposed.

Approaches that depart from the notion of **insertion** (associated with Myers-Scotton 1993b) view the constraints in terms of the structural properties of some base or matrix structure. Here the process of code-mixing is conceived as something akin to borrowing: the insertion of an alien lexical or phrasal category into a given structure. The difference would simply be the size and type of element inserted, e.g. noun versus noun phrase.
Bilingual speech

Approaches departing from alternation (associated with Poplack 1980) view the constraints on mixing in terms of the compatibility or equivalence of the languages involved at the switch point. In this perspective code-mixing is akin to the switching of codes between turns or utterances. This is the reason I avoid using the term code-switching for the general process of mixing. Switching is only an appropriate term for the alternational type of mixing. The term code-switching is less neutral in two ways: as a term it already suggests something like alternation (as opposed to insertion), and it separates code-mixing too strongly from phenomena of borrowing and interference.

The distinction I make here between alternation and insertion corresponds to Auer’s distinction between code-switching and transfer (1995: 126). Some authors have used the term ‘switching’ for language interaction between clauses, and ‘mixing’ for intra-clausal phenomena. This distinction parallels my distinction between alternation and insertion, but does not coincide with it, since in my framework alternation often takes place within the clause as well.

The notion of congruent lexicalization underlies the study of style shifting and dialect/standard variation, as in the work of Labov (1972) and Trudgill (1986), rather than bilingual language use proper. The exception is the bilingual research by Michael Clyne (1967) on German and Dutch immigrants in Australia. This comes closest to an approach to bilingual language use from the perspective of congruent lexicalization.

In this book I am claiming that these different models or approaches in fact correspond to different phenomena: there is alternation between languages, insertion into a matrix or base language, and congruent lexicalization, in the code-mixing data reported in the literature. In chapters 3, 4, and 5 criteria are proposed, both structural and quantitative, for giving substance to the three-way distinction.

In some cases, a single constituent is inserted into a frame provided for by the matrix language:

(1) kalau dong tuakan bikin dong tuakan bikin
    when they always make they always make
    voor acht personen dek orang suka nganga dong makan
    for eight persons and then people only look they eat
    ‘When they [cook], it is always for eight people, and then they only look
    at it, they eat . . .’

   (Moluccan Malay/Dutch, Huwai 1992)

While in (1) this is an entire Dutch prepositional phrase inserted into a Moluccan Malay sentence, in (2) it is a single English verb stem used in a complex Navaho verbal structure:

4
The study of code-mixing

(2) na’ish-crash là 1sg:pass out-crash EMPH
‘I am about to pass out.’ (Navaho/English; Canfield 1980: 219)

In (3) the temporal expression *por dos días* is clearly related to the verb *anduve*, encapsulating the inserted *in a state of shock*:

(3)  Yo *anduve in a state of shock* por dos días.
‘I walked in a state of shock for two days.’ (Spanish/English; Pfaff 1979: 296)

With insertion, there is embedding. The English prepositional phrase is inserted into an overall Spanish structure. Insertion is akin to (spontaneous) lexical borrowing, which is limited to one lexical unit. There is considerable variation in what is or can be inserted: in some languages this consists mostly of adverbial phrases, in others mostly single nouns, and in yet others again determiner + noun combinations. Insertion and the distinction between code-mixing and borrowing are taken up again in chapter 3.

In other cases, it seems that halfway through the sentence, one language is replaced by the other. The two languages alternate:

(4) maar ‘t hoeft niet  li’anna ida left ana . . .
but it need not for when I see I . . .
‘but it need not be, for when I see, I . . .’
(Moroccan Arabic/Dutch; Noort 1990: 126)

(5) Les femmes et le vin, *ne pomimayu*.  ‘Women and wine, I don’t understand.’ (French/Russian; Tinn 1978: 312)

(6) Andale pues *and do come again*.
‘That’s all right then, and do come again.’
(Spanish/English; Gumperz and Hernández-Chavez 1971: 118)

In the case of alternation, there is a true switch from one language to the other, involving both grammar and lexicon. Thus in (6) there is no reason to assume that the Spanish first segment is embedded in the English second segment or vice versa. Alternation is just a special case of code-switching, as it takes place between utterances in a turn or between turns. In chapter 4 alternation is studied in more detail.

In a third set of cases, it appears that there is a largely (but not necessarily completely) shared structure, lexicalized by elements from either language, congruent lexicalization. Consider the following examples:

(7) Weet *jij [whaar]* Jenny is?
‘Do you know where Jenny is?’ (Dutch: war Jenny is)
(English/Dutch; Crama and van Gelderen 1984)

The sequence *where Jenny is* could as easily be English in structure as Dutch. Furthermore *where* is close to Dutch *waar* (particularly when pronounced by bilinguals), *Jenny* is a name in both languages, and *is* is homophonous.
Bilingual speech

A similar example is:

(8) En de partij dy’t hy derby blassde, (Frison)
en de partij die hij erbij bleef
And the part that he thereby blew
is net [foar herhaling vatbaar]. (Frison)
is niet voor herhaling vatbaar. (Dutch)

‘And the song he sang then is not fit to be repeated.’

Here, Frisonian foar ‘for’ is sufficiently similar to Dutch voor ‘for’ to be an ambiguous switchpoint; Dutch herhaling vatbaar is not a constituent, but two words that form an idiom together with voor.

While English/Dutch and Frisonian/Dutch are two closely related language pairs with many cognates, we may find something similar to these examples in the English/Spanish material analysed by Poplack (1980) as well:

(9) (A) Why make Carol sentarse atrás (B) pa’que everybody
    sit at the back so that

    has to move (C) pa’que se salga. (Spanish/English, Poplack 1980: 589)

Here sentence fragment (B) is a complement to (A), and (C) is a complement to (B). Notice that the first Spanish fragment here contains both a verb phrase, sentarse atrás and a purposive complementizer, pa’que. There is no particular grammatical relation between the two English fragments nor between the Spanish ones. The example could perhaps be analysed as backfire insertions within insertions. However, this is rather counter-intuitive, both because the switched fragments are not unique constituents and because they do not appear to obey rules specific to the supposed matrix constituent, but rather rules common to both languages.

Consider a similar example:

(10) Bueno, in other words, el flight [que sale de Chicago around three o’clock]. (Spanish/English; Pfaff 1976: 250)

In (10) que sale de Chicago ‘that leaves Chicago’ or even el flight que sale de Chicago (assuming Chicago to be part of the Spanish stretch for the sake of the argument – in fact it may be the trigger for the subsequent switch to English) is a constituent, but not a unique one, since it also includes the English fragment around three o’clock.

The term congruent lexicalization refers to a situation where the two languages share a grammatical structure which can be filled lexically with elements from either language. The mixing of English and Spanish could be
The study of code-mixing

interpreted as a combination of alternations and insertions, but the going back and forth suggests that there may be more going on, and that the elements from the two languages are inserted, as constituents or as words, into a shared structure. In this perspective, congruent lexicalization is akin to style or register shifting and monolingual linguistic variation. The latter would be the limiting case of congruent lexicalization.

I want to explore these three separate patterns of intra-sentential code-mixing and study them through the systematic exploration of bilingual corpora, in addition to the detailed structural analysis of individual examples (van Hout and Muysken 1995).

The structural interpretation of these notions is as follows. Consider the following trees, in which $A$, $B$ are language labels for non-terminal nodes (i.e. fictitious markers identifying entire constituents as belonging to one language), and $a$, $b$ are labels for terminal, i.e. lexical, nodes, indicating that the words chosen are from a particular language.

(11) insertion

\[
\begin{array}{c}
A \\
\quad \triangleleft a \quad \triangleleft b \quad \triangleleft a \\
\end{array}
\]

In this situation, a single constituent $B$ (with words $b$ from the same language) is inserted into a structure defined by language $A$, with words $a$ from that language.

(12) alternation

\[
\begin{array}{c}
A \\
\quad \triangleleft a \\
\end{array}
\quad
\begin{array}{c}
B \\
\quad \triangleleft b \\
\end{array}
\]

In this situation, a constituent from language $A$ (with words from the same language) is followed by a constituent from language $B$ (with words from
Bilingual speech

that language). The language of the constituent dominating $A$ and $B$ is unspecified.

(13) congruent lexicalization

Finally, in (13) the grammatical structure is shared by languages $A$ and $B$, and words from both languages $a$ and $b$ are inserted more or less randomly.

Having presented the three processes, insertion, alternation, and congruent lexicalization, I will try to suggest a number of diagnostic features which may be used to distinguish these three patterns in chapters 3, 4, and 5. There are a number of criteria I would like to consider, and I will illustrate their application with concrete cases. The criteria are rarely knock-down criteria by themselves, but should be used conjointly to characterize a sentence or a bilingual speech sample as a case of alternation, insertion, or congruent lexicalization.

In addition to the structural interpretation of the three patterns, in terms of labels in tree configurations, there can also be a psycholinguistic and a sociolinguistic one. The psycholinguistic interpretation of the three-way distinction made here could be in terms of different degrees of activation of components of both languages in speech production. In the case of alternation, activation would shift from one language to another, and in the case of insertion, activation in one language would be temporarily diminished. For congruent lexicalization, the two languages partially share their processing systems. Psycholinguistic factors determining the choice between these different processes include bilingual proficiency, level of monitoring in the two languages, the triggering of a particular language by specific items and the degree of separateness of storage and access systems.

The interpretation of the three patterns can also be sociolinguistic, in terms of bilingual strategies (an example would be Sankoff, Poplack, and Vanniarajan’s (1990) distinction between equivalence and insertion). The sociolinguistic embedding of these three patterns, i.e. their use as bilingual strategies, can be described as follows. The process of alternation is particularly frequent in stable bilingual communities with a tradition of language separation, but occurs in many other communities as well. It is a frequent
The study of code-mixing

Figure 1.1 Schematic representation of the three main styles of code-mixing and transitions between them

and structurally little intrusive type of code-mixing. Insertion is frequent in colonial settings and recent migrant communities, where there is a considerable asymmetry in the speakers’ proficiency in the two languages. A language dominance shift, e.g. between the first and third generation in an immigrant setting, may be reflected in a shift in directionality of the insertion of elements: from insertion into the language of the country of origin to the presence of originally native items in the language of the host country. Congruent lexicalization may be particularly associated with second generation migrant groups, dialect/standard and post-creole continua, and bilingual speakers of closely related languages with roughly equal prestige and no tradition of overt language separation.

The three types of code-mixing can be conveniently viewed in terms of a triangle, an image to which I return in chapter 8 when I discuss the various factors that help determine which type of mixing occurs in a specific setting. The differences between the three types are gradual rather than absolute, as can be seen in figure 1.1.

Between insertion and alternation there is a transition zone since insertion
Bilingual speech

of longer fragments leads to increasingly more complete activation of the second grammar. There are different modules potentially involved in the insertion: phonetic shapes, lexical meaning, morphosyntax, which involve activation of the second grammar to different degrees.

Since only some higher-level structures may be shared between languages, there is a gradual transition as well between alternation and congruent lexicalization: alternation when only the top node (the sentence node) is shared, and congruent lexicalization when all or most nodes are shared between the two languages.

Finally, there can be a gradual shift from one base language to shared structure and on to the other base language, possibly varying with individual bilingual proficiency and over time. This implies particularly that in many immigrant communities, insertion of new items and expressions into the home language can evolve into congruent lexicalization and then possibly into alternation (with set phrases and expressions from the ethnic language interspersed in the new language).

The literature abounds both with proposals for various specific constraints on code-mixing, and with claims that general constraints do not hold. Romaine (1989, 1995) takes an intermediate position, namely that none of the constraints covers all cases. A similar position is taken by Clyne (1987). I want to argue against too much relativism, however, both because it is overly pessimistic of the relevance of linguistic structure, and because I believe it only portrays a limited picture of the often quite regular array of code-mixing patterns to be found. Rather, a particular constraint can only be assumed to hold for a specific type of mixing.

The role of a dominant research experience in shaping one’s vision: the development of code-mixing research

It has been said somewhat maliciously that an Indo-Europeanist will tend to make reconstructed Proto-Indo-European look like the language that she or he thoroughly studied early on. If the scholar starts out with Albanian, her or his version of the proto-language will have many features of Albanian, and so on. This subjectivism (if the observation is at all valid) may seem reprehensible, but I think there are many ways in which research that one does early on shapes one’s vision of a certain domain. This does not mean one is condemned to the one perspective, but rather that this perspective is the frame of reference onto which later ideas are often grafted. In this way, research on code-mixing has been shaped by the language pairs encountered by researchers, and the key notions were introduced one by one on the basis of the data encountered.