The bilingual individual

FRANÇOIS GROSJEAN

This article presents a general overview of the adult bilingual individual. First, the bilingual is defined and discussed in terms of the complementary principle, i.e. the fact that bilinguals acquire and use their languages for different purposes, in different domains of life, with different people. Next, the various language modes bilinguals find themselves in during their everyday interactions are examined. These range from the monolingual mode when they are communicating with monolinguals (and they have to deactivate all but one language) to the bilingual mode when they are interacting with other bilinguals who share their two (or more) languages and with whom they can mix languages if they so wish (i.e. code-switch and borrow). The article ends with a rapid survey of the psycholinguistics of bilingualism and, in particular, of how bilinguals access their lexicon when perceiving mixed speech. The regular bilingual is compared to the interpreter bilingual whenever possible.

A number of myths surround bilingualism and bilinguals. For many people, bilingualism is a rare phenomenon that is found only in bilingual or multilingual countries such as Canada, Switzerland and Belgium. Supposedly bilinguals grow up speaking two (or more) languages, they have equal speaking and writing fluency in their languages, they have accentless speech, and they are excellent translators and interpreters. The reality is in fact quite different. There are more bilinguals in the world than monolinguals, they are found in every country of the world, in all classes of society and in all age groups, and they usually acquire their languages at various times during their lives and not just in early childhood. In addition, they are rarely equally fluent in all their languages (many cannot read or write one of their languages), they often have an accent in the language(s) they acquired late, and few bilinguals are proficient translators and interpreters.

Several aspects of the bilingual individual will be discussed in this paper. The adult bilingual, as opposed to the child bilingual, will be the object of study as will the stable bilingual, that is the person who is no longer in the process of acquiring a second or third language. First the bilingual person will be described in terms of language knowledge and language use. Then the language modes bilinguals find themselves in when interacting with monolinguals and with other bilinguals in their everyday life will be examined. Finally, a rapid survey of the psycholinguistics of bilingualism and in particular of how bilinguals undertake lexical access when perceiving mixed speech will be presented. Throughout the paper an attempt will be made to compare the "regular bilingual" with the distinct bilingual that the interpreter has become (the latter will be called the "interpreter bilingual"). Student interpreters start off as regular bilinguals and slowly, through extensive training and practice, become interpreter bilinguals. Emphasis here will be put on regular bilinguals, not only because they have been studied in more depth but also because they share many more traits with interpreter bilinguals than one might think at first.

1. Describing the bilingual

1.1 *Definition*

Although a few researchers have defined bilinguals as those who have native-like control of two or more languages (Bloomfield, 1933; Thiery, 1978), most others agree that this position is not realistic. If one were to count as bilingual only those people who pass as monolinguals in each of their languages, one would be left with no label for the vast majority of people who use two or more languages regularly but who do not have native-like fluency in each. This has led researchers to propose other definitions of bilingualism, such as: the ability to produce meaningful utterances in two (or more) languages, the command of at least one language skill (reading, writing, speaking, listening) in another language, the alternate use of several languages, etc. (Beatens-Beardsmore, 1986; Hakuta, 1986; Haugen, 1969; Romaine, 1995). In what follows, bilinguals will be defined as those people who use two (or more) languages (or dialects) in their everyday lives. (For a discussion of notions such as dominance, fluency, balance, etc., as applied to bilinguals, see Grosjean, 1982, 1985b).

Bilinguals differ from one another on a number of variables. In addition to such factors as age, sex, socio-economic and educational status, one finds the following: number and type of languages known and global competence in these languages; language history (when and how the languages were acquired and used); language stability (are one or several languages still being acquired (or restructured) or has a certain language stability been reached)?; competence in each of the four skills (reading, writing, speaking, listening) in each language; function of the languages (which languages are used, when, with whom and for what reason?); language modes (how often and for how long do the subjects find themselves at the various points along the language mode continuum?); amount of code-switching and borrowing normally done, etc. Despite the great diversity that exists between bilinguals, they all have one thing in common: they lead their lives with two (or more) languages. (Bilinguals who no longer use their different languages but who have retained knowledge of them are termed "dormant bilinguals.") It is clear that the many variables listed above need to be taken into account when assessing future interpreters, and they have to be kept in mind during actual interpreter training. They can explain in part the great variability found in the bilingualism of student interpreters and, to some extent at least, they can account for differential behavior when interpreters are working under various conditions (from optimal to less than optimal).

1.2 *The complementary principle*

There are many reasons that bring languages into contact and hence foster bilingualism: migrations of various kinds (economic, educational, political, religious), nationalism and federalism, education and culture, trade and commerce, intermarriage, etc. These factors create various linguistic needs in people who are in contact with two or more languages and who develop competencies in their languages to the extent required by these needs. In contact situations it is rare that all facets of life require the same language (people would not be bilingual if that were so) or that they always demand two languages (language A *and* B at work, at home, with friends, etc.). This leads to what is called the complementary principle:

> Bilinguals usually acquire and use their languages for different purposes, in different domains of life, with different people. Different aspects of life require different languages.

It is precisely because the needs and uses of the languages are usually quite different that bilinguals rarely develop equal and total fluency in their languages. The level of fluency attained in a language (more precisely, in a language skill) will depend on the need for that language and will be domain specific. If reading and writing skills are not needed in a language, they will not be developed. If a language is spoken with a limited number of people in a reduced number of domains, it may be less fluent and more restricted than a language used extensively. If a language is never used for a particular purpose, it will not develop the linguistic properties needed for that purpose (specialized vocabulary, stylistic variety, some linguistic rules, etc.).

In general, the failure to understand the complementary principle has been a major obstacle to obtaining a clear picture of bilinguals and has had many negative consequences (Grosjean, 1985b): bilinguals have been described and evaluated in terms of the fluency and balance they have in their two languages (when in fact they are rarely balanced); language skills in bilinguals have almost always been appraised in terms of monolingual standards (but monolinguals use only one language for all domains or life whereas bilinguals use two or more); research on bilingualism has in large part been conducted in terms of the bilingual's individual and separate languages (the use of language A or of language B when in fact both languages are often used simultaneously); and, finally, many bilinguals evaluate their language competencies as inadequate. Some criticize their mastery of language skills, others strive their hardest to reach monolingual norms, others hide their knowledge of their "weaker" language, and most do not perceive themselves as being bilingual even though they use two (or more) languages in their everyday lives.

The complementary principle can help us understand a number of phenomena. First, it reflects the configuration of the bilingual's language repertoire: what languages are known and to what extent, what they are used for, with whom and when, why one language is less developed than another, etc. Second, it helps to explain why the bilingual's language repertoire may change over time: as the environment changes and the needs for particular language skills also change, so will the bilingual's competence in his or her various language skills. New situations, new interlocutors and new language functions will involve new linguistic needs and will therefore change the language configuration of the person involved. Extreme cases of restructuring are language forgetting and a return to functional monolingualism, be it in the

person's first, second or third language. On this point, it should be noted that interpreters must constantly strive to maintain their knowledge of, and their fluency in, their various languages despite changes in their own personal lives which may result in a particular language being used very little outside of work. Third, an increasing understanding of the complementary principle has changed researchers' view of bilinguals over the last few years. Bilinguals are now seen not so much as the sum of two (or more) complete or incomplete monolinguals but rather as specific and fully competent speakers-hearers who have developed a communicative competence that is equal, but different in nature, to that of monolinguals. This competence makes use of one language, of the other, or of the two together (in the form of mixed speech, see below) depending on the situation, the topic, the interlocutor, etc. Thus, it is perfectly normal to find bilinguals who can only read and write one of their languages, who have reduced speaking fluency in a language they use only with a limited number of people, or who can speak only about a particular subject in one of their languages. This in turn is leading to a redefinition of the procedure used to evaluate the bilingual's competencies. Bilinguals are now being studied in terms of their total language repertoire, and the domains of use and the functions of the bilingual's various languages are now being taken into account (see e.g. Romaine, 1995).

Finally, the complementary principle accounts for why regular bilinguals are not usually very good translators and interpreters. Some may not know the translation equivalents in the other language (words, phrases, set expressions, etc.) which in turn will lead to perception and production problems. At the level of words, for example, some of the bilingual's domains of life are covered by the lexicon of only one language, others by the lexicon of the other language only, and some by the two. Unless bilinguals acquired their second language in a manner which involves learning translation equivalents, many will find themselves lacking vocabulary in various domains (work, religion, politics, sports, etc.) even though they may appear to be perfectly fluent in their two languages. Another reason that accounts for less than perfect translation and interpreting (prior to training, of course) is that some bilinguals may not have the stylistic varieties needed in their two languages (they simply do not need them in one or the other language). Yet another reason is that some bilinguals may not have the cultural knowledge (pragmatic competence) required to understand an utterance in one of their languages (how many are totally bicultural?). And a final reason is that most bilinguals have

not developed the necessary transfer skills needed for translation and inter-pretation, not to mention the specific on-line processing and memory mecha-nisms needed for the latter behavior. Taking into account the complementary principle is a crucial prerequisite in interpreter training: interpreter bilinguals, unlike regular bilinguals, will have to learn to use their languages (and the underlying skills they have in them) for similar purposes, in similar domains of life, with similar people. This is something regular bilinguals do not often need to do.

2.   The bilingual's language modes

In their everyday lives, bilinguals find themselves in various language modes that correspond to points on a monolingual-bilingual mode continuum (Grosjean, 1985b, 1997). A mode is a state of activation of the bilingual's languages and language processing mechanisms. At one end of the con-tinuum, bilinguals are in a totally monolingual language mode in that they are interacting with monolinguals of one — or the other — of the languages they know. One language is active and the other is deactivated. At the other end of the continuum, bilinguals find themselves in a bilingual language mode in that they are communicating with bilinguals who share their two (or more) lan-guages and with whom they can mix languages (i.e. code-switch and borrow). In this case, both languages are active but the one that is used as the main language of communication (the base language) is more active than the other. These are end points but bilinguals also find themselves at intermediary points depending on such factors as interlocutor, situation, content of dis-course and function of the interaction.

Figure 1 is a visual representation of the continuum. The languages (A and B) are represented by a square located in the top and bottom parts of the figure and their level of activation is depicted by the degree of darkness of the square: black for a highly active language and white for a deactivated lan-guage. In the figure, three hypothetical positions for the same speaker are presented (discontinuous lines numbered from 1 to 3). In all positions, the speaker is using language A as the main language of communication (the base language) and it is therefore the most active language (black squares). In position 1, the speaker is in a monolingual mode: language A is totally active whereas language B is deactivated (Green, 1986, would even say that it is
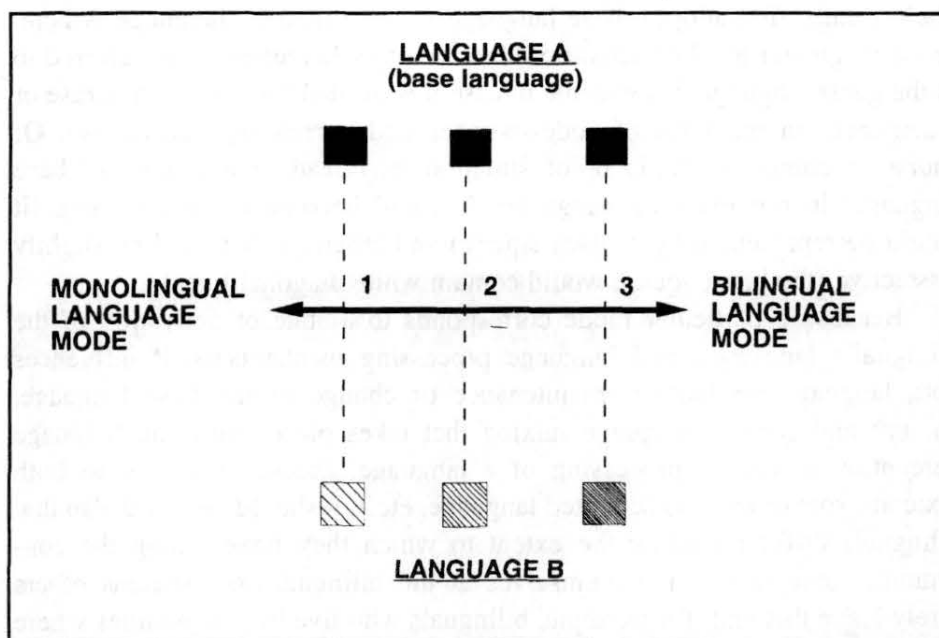
Figure 1. Visual representation of the language mode continuum. The speaker's positions on the continuum are represented by the discontinuous vertical lines and the level of language activation by the degree of darkness of the squares (black is active and white is inactive).

inhibited). This mode arises when the person being spoken to (the interlocutor) is monolingual (in this case, in language A), and/or the topic, the situation and the purpose of interaction require that only one language be spoken to the exclusion of the other(s). In position 2, the speaker is in an intermediary mode. Language A is still the most active language (it is the language of communication) but language B is also partly activated. This kind of mode arises, for example, when a bilingual is speaking to another bilingual who does not wish to use the other language (in this case, language B) or when a bilingual is interacting with a person who has limited knowledge of the other language. Any number of combinations of the factors listed above (interlocutor, topic, situation, etc.) can lead to this intermediary position. In position 3, the speaker is at the bilingual end of the continuum. Both languages are active but language B is slightly less active than language A as it is not currently the language of communication. This is the kind of mode bilinguals find themselves in when they are interacting with other bilinguals who share their two (or more) languages and with whom they feel comfortable mixing languages.

They usually first adopt a base language to use together (language A here, hence its greater level of activation) but the other language, often referred to as the guest language, is available in case it is needed for a word, a phrase or a sentence (in the form of code-switches and borrowing; see below). Of course, a change of topic or of situation may lead to a change of base language. In our example, language B would become the most active (it would be represented by a black square) and language A would be slightly less active (the black square would contain white diagonal lines).

Because a particular mode corresponds to a state of activation of the bilingual's languages and language processing mechanisms, it influences both language production (maintenance or change of the base language, amount and type of language mixing that takes place, etc.) and language perception (speed of processing of a language, access to one or to both lexicons, role of the less activated language, etc.). It should be noted also that bilinguals differ regarding the extent to which they travel along the continuum; some rarely find themselves at the bilingual end whereas others rarely leave this end (for example, bilinguals who live in communities where the language norm is mixed language). The two end points of the continuum will now be examined as will the bilingual's language behavior in the mono-lingual and bilingual language modes. This will be followed by a discussion of the manner in which the language mode continuum model has to be adapted to take into account simultaneous interpreting.

2.1 The monolingual language mode

In this mode (discussed in more detail in Grosjean, 1982), bilinguals adopt the language of the monolingual interlocutor(s) and deactivate their other language(s) as completely as possible. Bilinguals who manage to do this totally and, in addition, who speak the other language fluently and have no foreign accent in it, will often "pass" as monolinguals. Although such cases are relatively rare (many interpreters are among them), it is precisely these cases that have led people to think that bilinguals are (or should be) two monolinguals in one person. In fact, deactivation of the other language is rarely total as is clearly seen in the interferences bilinguals produce (these are also known as between-language deviations). An interference is a speaker-specific deviation from the language being spoken due to the influence of the other deactivated language. Interferences can occur at all levels of language

(phonological, lexical, syntactic, semantic, pragmatic) and in all modalities (spoken, written or sign). They are of two kinds: static interferences which reflect permanent traces of one language on the other (such as a permanent accent, the meaning extensions of particular words, specific syntactic structures, etc.) and dynamic interferences (the ones interpreters know well) which are the ephemeral intrusions of the other language as in the case of the accidental slip on the stress pattern of a word due to the stress rules of the other language, the momentary use of a syntactic structure taken from the language not being spoken, etc. Following are some examples of interferences produced by a French person speaking English: At the phonetic level, pronouncing *Thank 'eaven for this* instead of *Thank heaven for this*; at the lexical level, using *corns* (from French *cornes*) instead of *horns* in *Look at the corns on that animal!* ; at the syntactic level, saying *I saw this on the page five* (instead of *on page five*), and in writing, misspelling *adress* or *appartment* (based on the French *adresse* and *appartement*).

In addition, if one of the bilingual's languages is mastered only to a certain level of proficiency (this does not concern interpreters), deviations due to the person's interlanguage (also known as within-language deviations) will also occur. These include overgeneralizations (for example, taking irregular verbs and treating them as if they were regular), simplifications (dropping pluralization and tense markers, omitting functions words, simplifying the syntax, etc.) as well as hypercorrections and the avoidance of certain words and expressions. Between — and within — language deviations are clearly observable when bilinguals are in a monolingual language mode but they also occur in the bilingual language mode. In this latter mode, however, deviations such as interferences are often difficult to distinguish from perfectly normal borrowings. Unfortunately, this has resulted in grouping very different manifestations of language contact under one heading such as interference or code-switching. It should be noted finally that both types of deviations, although sometimes quite apparent (such as a foreign accent), usually do not interfere with communication. This is because bilinguals generally develop their languages to the level of fluency required by the environment.

2.2 *The bilingual language mode*

In this mode, bilinguals are usually interacting with one another. First they adopt a language to use together, that is a base language (also known as the

"host" or "matrix" language). This process is called language choice and is governed by a number of factors that are similar to those that play a role in positioning the bilingual on the monolingual-bilingual mode continuum: the interlocutors involved (the usual language of interaction, language proficiency, language preference, socioeconomic status, age, sex, occupation, education, kinship relation, attitude toward the languages, etc.); the situation of the interaction (location, presence of monolinguals, degree of formality and intimacy); the content of the discourse (topic, type of vocabulary needed) and the function of the interaction (to communicate information, to create a social distance between the speakers, to raise the status of one of the interlocutors, to exclude someone, to request something, etc.). Language choice is a well-learned behavior (a bilingual rarely asks the conscious question, "Which language should I be using with this person?") but it is also a very complex phenomenon which only becomes apparent when it breaks down. Usually, bilinguals go through their daily interactions with other bilinguals quite unaware of the many psychological and sociolinguistic factors that interact to help choose one language over another. It should be noted again that the base language can change several times during a single conversation if the situation, topic, interlocutor, etc. require it.

Once a base language has been chosen, bilinguals can bring in the other language (the "guest" or "embedded" language) in various ways. One of these ways is to code-switch, that is to shift completely to the other language for a word, a phrase, a sentence. For example, *Va chercher Marc and bribe him avec un chocolat chaud with cream on top* (Go get Marc and bribe him with a hot chocolate with cream on top). Code-switching has long been stigmatized (Baetens Beardsmore, 1986; Grosjean, 1982), and has been given a number of pejorative names such as Franglais (the switching between French and English) or Tex-Mex (the switching between English and Spanish in the southwestern part of the United States). The consequence of this has been that some bilinguals never switch while others restrict it to situations in which they will not be stigmatized for doing so. Recently, code-switching has received considerable attention from researchers (for an overview, see Milroy & Muysken, 1995). For example, sociolinguists have concentrated on when and why switching takes place in the social context. Reasons that have been put forward are: to fill a linguistic need, to continue the last language used, to quote someone, to specify the addressee, to exclude someone from the conversation, to qualify a message, to specify speaker involvement, to mark

group identity, to convey emotion, to change the role of the speaker, etc. Linguists, on the other hand, have sought to study the types of code-switches that occur (single words, phrases, clauses, sentences, etc.) as well as the linguistic constraints that govern their appearance. Although there is still considerable controversy over this latter aspect (are constraints universal or language specific? how broad can a constraint be?), it is now clear that switching is not simply a haphazard behavior due to some form of "semi-lingualism" but that it is a well governed process used as a communicative strategy to convey linguistic and social information.

The other way bilinguals can bring in the other, less activated, language is to borrow a word or short expression from that language and to adapt it morphologically (and often phonologically) into the base language. Thus, unlike code-switching, which is the juxtaposition of two languages, borrowing is the integration of one language into another. Most often both the form and the content of a word are borrowed (to produce what has been called a loanword or more simply a borrowing) as in the following examples taken from French-English bilinguals: "*Ca m'étonnerait qu'on ait code-switché autant que ça*" (I can't believe we code-switched as often as that) and "*Maman, tu peux me tier /taje/ mes chaussures*" (Mummy, can you tie my shoes?). In these examples, the English words "code-switch" and "tie" have been brought in and integrated into the French sentence. A second type of borrowing, called a loanshift, consists in either taking a word in the base language and extending its meaning to correspond to that of a word in the other language, or rearranging words in the base language along a pattern provided by the other language and thus creating a new meaning. An example of the first kind of loanshift would be the use of *humoroso* by Portuguese-Americans to mean 'humorous' when the original meaning is 'capricious'. An example of the second kind is the use of idiomatic expressions that are translated literally from the other language, such as "*I put myself to think about it*" said by a Spanish-English bilingual, based on "Me puse a pensarlo". It is important to distinguish idiosyncratic loans (also called "speech borrowings" or "nonce borrowings") from words which have become part of a language community's vocabulary and which monolinguals also use (called "language borrowings" or "established loans"). Thus, in the following text, every third or fourth word is an established loan from French which has now become part of the English language: "*The poet lived in the duke's manor. That day, he painted, played music and wrote poems with his companions.*"

Current research is examining, among other things, the differences and similarities that exist between code-switches and borrowings (and within the latter, between idiosyncratic borrowings and established borrowings), as well as the impact of the two on language itself, such as first- and second-language restructuring.

2.3 *The case of simultaneous interpreting*

What language mode are interpreters in when they are doing simultaneous interpretation? Figure 2 is an attempt to answer this question. First, as can be seen, the interpreter is in a bilingual mode and both languages are active. However, one language is not more active than the other as is normally the case in the bilingual mode. Both the source language and the target language are active to the same extent as both are needed, for perception and production respectively. Second, input and output components have been added to each language and it is their level of activation that varies. The addition of
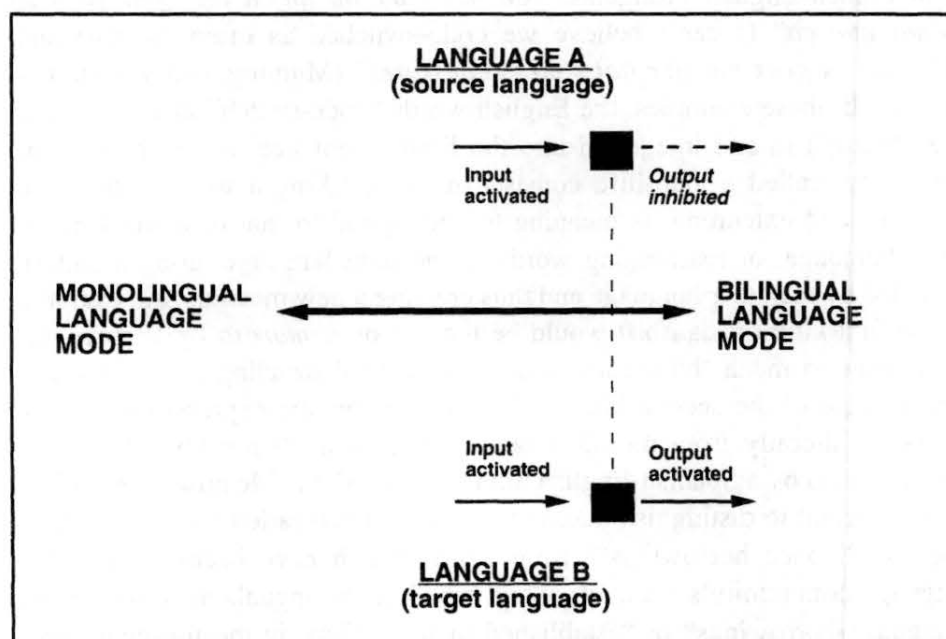

Figure 2. Visual representation of the interpreter's position on the language mode continuum when doing simultaneous interpreting. Both languages are active (black squares) but they differ as to the level of activation of their input and output mechanisms.

these components allows us to make a distinction between language (grammar) and language processing mechanisms. The languages are equally active but the processing mechanisms are not. In this way, the interpreter will be able to input the source language (and to a lesser extent the target language, see below) and to output the target language only. Third, the input component of the source and of the target language are both activated. No comments are needed concerning the source language input except that it is probably more strongly activated than its target language counterpart. As for the latter (the target language input), at least three reasons require that it also be activated: the interpreter must be able to monitor his/her overt speech (Levelt, 1989), the client's occasional use of the target language must be processed, and a fellow interpreter's cues must be heard or read. Fourth, the target language output mechanism is activated and the source language output mechanism inhibited (both for obvious reasons). In sum, the output mechanisms are in a monolingual mode (only one language should normally be output) whereas the input mechanisms are in a bilingual mode (input will take place in both the source and the target language).

Of course, this is a very tentative view of interpreting in terms of the monolingual-bilingual mode continuum model and it needs to be refined. One concern is how strongly the source language output mechanism should be inhibited. For example, it is well known that interpreters sometimes naturalize a source-language term (Gile, 1995) by adapting it to the morphological and phonological rules of the target language. (This is called a nonce borrowing in the bilingualism literature). It is also common knowledge that interpreters sometimes transcode a word or expression into the target language (called a loanshift or loan translation in the field of bilingualism). Although the processing mechanisms that account for these two phenomena are not clear, they do not probably require the source language output mechanism to be activated. However, there remains the problem of what Gile (1995) refers to as reproducing the sounds heard in the source-language speech (with or without an explanation in the target language) which is simply code-switching in the bilingualism terminology. Thus, although it is the target language that is being output, the interpreter sometimes code-switches to the source language for a word or phrase. How this takes place in processing terms remains an unanswered question as the source language output mechanism is normally inhibited. A momentary release of inhibition probably has to occur for a source language utterance to be output. Future research should examine this issue as well as how bilingual interpreters shift from being regular bilinguals in a

bilingual language mode when interacting with other bilinguals (they adopt a base language and code-switch or borrow from time to time, etc.) to being distinct bilinguals when interpreting; although they remain in the bilingual mode, they can no longer mix their languages in this particular way.

3. The psycholinguistics of bilingualism

The psycholinguistics of bilingualism is aimed at studying the processes involved in the production, perception, comprehension and memorization of the bilingual's languages (spoken, written or signed) when used in a monolingual or a bilingual language mode. Until recently the emphasis has been put on the independence of the bilingual's languages (How does the bilingual keep the two languages separate? Does the bilingual have one or two lexicons?) to the detriment of issues such as the on-line processing of language, be it in a monolingual or in a bilingual language mode. Much research was conducted in earlier years, for example, on the coordinate-compound-subordinate distinction (see Grosjean, 1982, for a review). According to it, there are three types of bilinguals: coordinate bilinguals who have two sets of meaning units and two modes of expression, one for each language (this means that the words of the two languages are totally separate entities); compound bilinguals who have one set of meaning units and two modes of expression (equivalent words in different languages have the same meaning); and subordinate bilinguals who have the meaning units of the first language and two modes of expression: that of the first language and that of the second, learned by means of the first (here the bilingual processes words of the weaker language through the words of the stronger language). Despite the inherent appeal of this distinction, no amount of experimentation has brought conclusive evidence for this trichotomy.

Another area of considerable investigation has been whether bilinguals possess one or two internal lexicons. Proponents of the one-lexicon view state that linguistic information is stored in a single semantic system. Words from both languages are organized in one large lexicon, but each word is "tagged" to indicate the language it belongs to (see, for example, Kolers, 1966; Schwanenflugel & Rey, 1986). Other researchers have claimed that bilinguals have two lexicons (e.g. Tulving & Colotla, 1970; Taylor, 1971), and that the information acquired in one language is available in the other only through a

translation process. Despite a large number of studies, no clear-cut results have been found. In fact, it has since been proposed that bilinguals have three stores, one conceptual store corresponding to the bilingual's knowledge of the world and two language stores, one for each language (Paradis, 1980). In recent years, the question of lexical representation has been studied with different paradigms and new models have been proposed (the word association model, the concept mediation model, the asymmetrical model, etc.) but it is still too early to say which one accounts for most of the data obtained (Schreuder & Weltens, 1993; de Groot & Kroll, 1997). In fact, de Groot (1995) has recently proposed that the lexical memory of every bilingual contains structures of various types (each one corresponding to a different kind of model) and that these structures occur in different proportions across bilinguals.

A third issue of interest has been the ability of bilinguals to keep their two languages separate when in the monolingual mode. Researchers have postulated the existence of a language switch which allows bilinguals to gate out the other language (e.g. Macnamara, 1967; Macnamara & Kushnir, 1971), and experimental studies have been conducted to find evidence for this proposal. The results obtained have been inconclusive or, at the very least, questionable (see Grosjean, 1982, for a review), and currently it is felt that no switch, be it psycholinguistic or neurolinguistic, exits in bilinguals. Rather, it has been proposed that bilinguals are probably using various activation and deactivation procedures to maintain their languages separate in the monolingual mode and to make them interact in the bilingual mode (Green, 1986; Paradis, 1989; Grosjean, 1985b, 1997).

Now that it is more generally accepted that the bilingual is not two monolinguals in one person, but a unique speaker-hearer using one language, the other language, or both together depending on the interlocutor, situation, topic, etc. (see above), current psycholinguistic research is trying to understand the processing of language in the bilingual's different language modes. Researchers are studying how bilinguals in the monolingual mode differ from monolinguals in terms of perception and production processes, and they are investigating the actual interaction of the two languages during processing in the bilingual mode. Following is a review of some recent work on spoken word recognition in the bilingual mode as it has some relevance to the type of lexical access the interpreter has to engage in as he or she is interpreting simultaneously.

3.1 *Lexical access in the bilingual language mode*

A great deal of research has been conducted on lexical access in monolinguals. It is well known, for example, that certain properties of words affect their recognition: their frequency of use, their length, their phonotactic configuration, their uniqueness point, their lexical neighborhood. It is also known that when words are presented in context, their lexical properties interact with various sources of knowledge (linguistic rules, knowledge of the world, discourse, etc.) to speed up or slow down the recognition process (see Frauenfelder & Tyler, 1987, for a review). The exact nature of the "interaction" between the properties of the words and the sources of knowledge remains to be described adequately, and the controversy concerning the moment at which "top-down" information enters the lexical access process has yet to be resolved (Forster, 1976; Marslen-Wilson, 1987; McClelland & Elman, 1986; Swinney, 1982). But one conclusion that emerges from this research is that recognizing a word may not be a simple mapping between its acoustic-phonetic properties and its entry in the mental lexicon (although see Klatt, 1979). Instead, it may well be a rather complex process that involves various narrowing-in and monitoring stages, correcting strategies, post-access decision stages, and even look-ahead and look-back operations (Grosjean, 1985a; Grosjean & Gee, 1987; McLelland & Elman, 1986).

Word recognition in bilinguals has received much less attention, especially as it concerns the access of code-switches and borrowings (often grouped under the label "guest words"). The literature most closely related to this latter question dates back a number of years and examines the perception and production of language mixtures, most of them ungrammatical (Kolers, 1966; Macnamara & Kushnir, 1971; Neufeld, 1973; etc.). In a first exploratory study, Soares & Grosjean (1984) investigated the lexical access of base language words and code-switched words by means of the Phoneme Triggered Lexical Decision task (Blank, 1980). English-Portuguese bilingual subjects were presented with sentences and were asked to listen for a word or a nonword within them which began with a prespecified phoneme. Once this word (or nonword) was found, the subjects had to indicate as quickly as possible whether the item was a real word or not. English monolingual subjects were run on English sentences only, whereas bilingual subjects were tested on three separate sets of sentences (English, Portuguese and code-switched). Before each set, every effort was made to induce the appropriate

language set: English, Portuguese or bilingual.

Two main findings emerged from this study. The first was that although bilinguals accessed real words in English as quickly as English monolinguals, they were substantially slower at responding to nonwords. This finding provided additional evidence for the residual activation of the other language when the bilingual is processing only one language (Altenberg & Cairns, 1983; Obler & Albert, 1978). It was hypothesized that a nonword triggers a complete search of the base language lexicon (or an activation of the lexicon, depending on the access theory one espouses), which is then immediately followed by at least a partial search (or activation) of the other, less active, lexicon. This occurs before the stimulus is classified as a nonword, hence the longer reaction times. The second finding was that bilinguals took longer to access code-switched words in the bilingual language mode than they did base language words in the monolingual language mode. Although at first this was accounted for by suggesting that bilinguals always search the base language lexicon before the less activated "guest" lexicon, in a later paper (Grosjean & Soares, 1986) it was suggested that a number of factors could account for the delay, irrespective of the access strategy. These have been grouped recently into four categories (Grosjean, 1997):

a. The listener. This category concerns the listener's fluency in the guest language, the language mode the listener is in (one expects slower recognition of guest words when the listener is not totally in the bilingual language mode), the listener's attitude towards code-switching and borrowing (a negative attitude will usually have an inhibitory effect on the guest lexicon) and the listener's expectations for code-switches and borrowings based on the topic, the speaker, the situation, etc. The higher the expectation, the easier will be the guest word recognition. Although these various factors are probably important, empirical evidence for the role they play during guest word recognition is still lacking.

b. Base and guest language activation. This category is divided into two parts: base language activation and guest language activation. As concerns base language activation, more than a quarter of a century ago, Macnamara & Kushnir (1971) showed that the bilingual listener has certain "expectations" for strings of words and that one such expectation is that all words should be in a single language. Although the terminology has changed a bit today (one would speak in terms of activation and/or inhibition), there is now consider-

able evidence that the base language being spoken (which normally makes up some 80% to 90% of a mixed utterance) has a strong effect on language processing. It is more strongly activated and hence base language units (phonemes, syllables, words) are favored over guest language units, at least momentarily (for evidence, see Grosjean & Soares, 1986; Bürki-Cohen, Grosjean & Miller, 1989; Grosjean, 1988). As concerns guest language activation, evidence is emerging that the density of the code-switches (that is, the number of code-switched words in a sentence) also influences their recognition (Soares & Grosjean, 1984; Leuenberger, 1994).

c. Code-switching constraints. This category covers the higher order constraints (syntactic, semantic and pragmatic) that govern code-switching. These have been studied extensively by linguists and sociolinguists (Romaine, 1995; Myers-Scotton, 1993; Muysken, 1995; di Sciullo, Muysken, & Singh, 1986; Poplack, Sankoff & Miller, 1988) but have not been the object of many word recognition studies (see however, Leuenberger, 1994; Li, 1996).

d. Guest word properties. This final category concerns the properties of guest words that affect their recognition. The first property is the word's phonotactics, that is the sequential arrangements (or groupings) of the word's units such as consonant sequences, syllables, etc. It has been hypothesized (e.g. Grosjean, 1988) that the more phonotactic cues there are for the word to belong to the guest language, the easier it should be to recognize. The second property concerns the actual phonetics of the word. If it contains sounds that are specific to the guest language, if it is said clearly and fully in the phonetics of the guest language (and not in that of the base language) and if it is said with the prosody of the guest language, then all this should speed up its recognition as the appropriate word will be activated more easily in the less activated lexicon. One difference between code-switches and borrowings is precisely their degree of phonetic integration in the base language; a code-switch is not usually integrated in the base language (unless the speaker has an accent in that language) whereas a borrowing is. One can therefore expect differences in the recognition of code-switches and borrowings. Finally, the presence of interlanguage neighbors (i.e. words that are phonologically similar in the base language) should affect the recognition of guest words. If the guest word has a close homophone in the base language and, furthermore, if this homophone is more frequent than the guest word, then the latter should be recognized with more difficulty. Evidence for the role of some of these

factors (phonotactics, language phonetics, interlanguage neighbor proximity) has been provided by Grosjean (1988), Li (1996), Bürki, Grosjean and Miller (1989), among others.

In summary, a number of studies show that the recognition of guest words in bilingual mixed speech is a highly complex process governed by a number of factors pertaining to the listener, the degree of activation of the two languages, the linguistic constraints governing code-switching and borrowing, and the properties of the guest words. A model that can account for some of these factors will now be described.

3.2 *BIMOLA: A bilingual model of lexical access*

Grosjean (1988) proposed an interactive activation model of word recognition in bilinguals, which has since been named BIMOLA (Bilingual model of lexical access; Léwy & Grosjean, in preparation). It is strongly inspired by TRACE (McClelland & Elman, 1986) and is governed by two basic assumptions. First, it is assumed that bilinguals have two language networks (features, phonemes, words, etc.) which are independent yet interconnected. They are independent in the sense that they allow a bilingual to speak just one language but they are also interconnected in that the monolingual speech of bilinguals often shows the active interference of the other language, and in that bilinguals can code-switch and borrow quite readily when they speak to other bilinguals. This view has long been defended by Paradis (1986, 1989), who proposes that both languages are stored in identical ways in a single extended system, though elements of each language, because they normally appear only in different contexts, form separate networks of connections, and thus a subsystem within a larger system. According to this "subset hypothesis", bilinguals have two subsets of neural connections, one for each language (each can be activated or inhibited independently because of the strong associations between elements), while at the same time they possess one larger set from which they are able to draw elements of either language at any time. The second assumption is that in the monolingual language mode, one language network is strongly activated while the other is only very weakly activated (the resting activation level of the units of this other network is therefore very low) whereas in the bilingual language mode, both language networks are activated but one more than the other (see Green, 1986; Myers-Scotton, 1993, for a similar assumption).
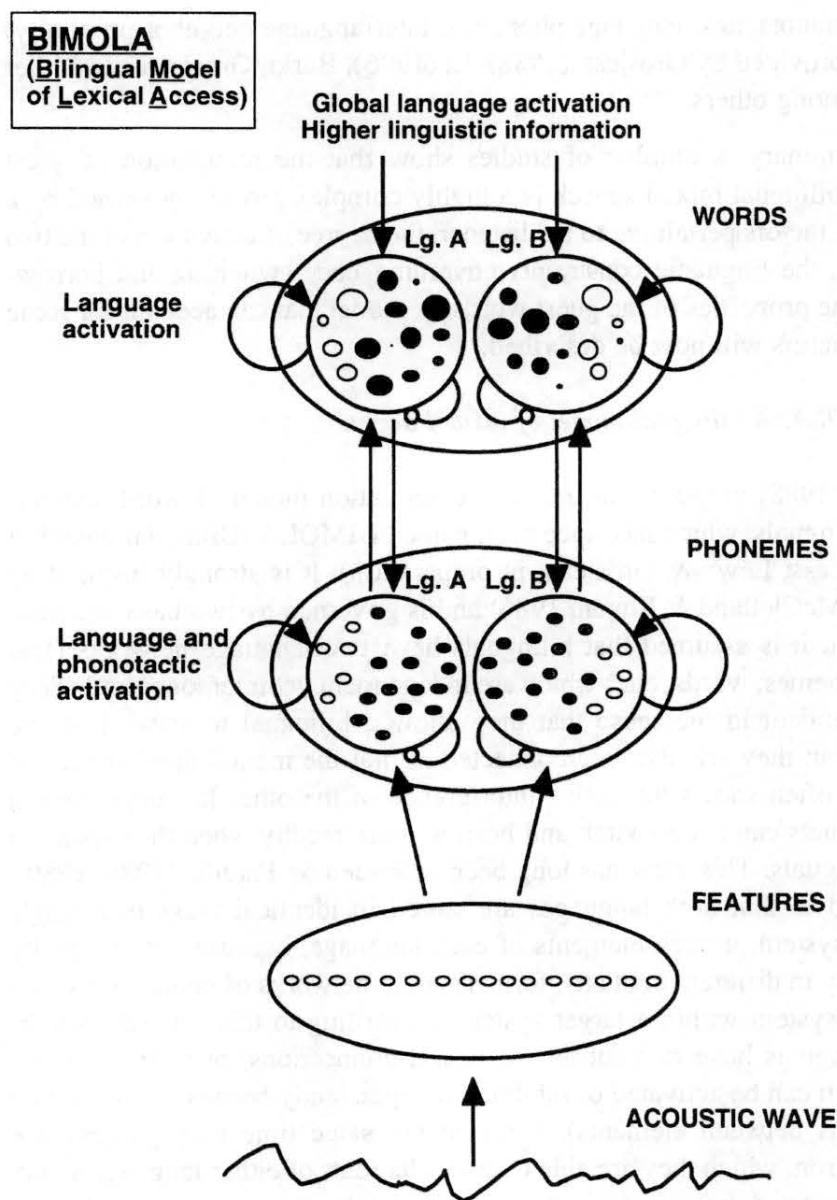
Figure 3. Visual representation of the BIMOLA model of lexical access in bilinguals. The
         proximity of neighboring units (phonemes, words) is depicted by the degree of
         darkness; darkly shaded units have close neighbors in the other language
         whereas lightly shaded units do not. At the word level, word frequency is
         represented by the size of the units; the larger the unit, the more frequent the
         word.

Figure 3 presents a visual representation of the BIMOLA model as it stands today. As can be seen, the feature level is common to both languages but the next two levels — phonemes and words — are organized according to the subset hypothesis, that is, both independently (each language — Lg. A and Lg. B — are represented by a subset of units) but also interdependently (both subsets are enclosed in a larger set). At both the word and phoneme levels, units can have close or distant form neighbors, both within a language and between languages. This is depicted by the degree of darkness of the units; darkly shaded units have close neighbors in the other language whereas lightly shaded units do not. At the word level, word frequency is represented by the size of the units — the larger the unit, the more frequent the word. Connections (mainly excitatory) are unidirectional between features and phonemes and bidirectional between phonemes and words. Features activate phonemes which in turn activate words. Descending connections bearing information about the listener's base language and language mode, and information from the higher linguistic levels (semantic, syntactic), serve to activate words which in turn can activate phonemes. Language activation (reflected by the overall activation of one language system over the other) takes place through these descending connections but also through within language connections at the phoneme and word levels. Finally, at the phoneme level, between phoneme connections within a language can allow for phonotactic activation.

The model is currently being refined and implemented on computer (Lévy & Grosjean, in preparation) and it appears to be able to account for a number of effects found experimentally (see the previous section). Considerable time has been spent developing a combined French and English feature system which will be used to activate phonemes of one or of both languages depending on their similarity. Work is now being carried out on the higher levels, interconnecting phonemes and words in the languages. Once the model has been implemented, it will be assessed by comparing its behavior to the data obtained experimentally and to that of the BIA model (Bilingual Interactive Activation) model of visual word recognition (Grainger & Dijkstra, 1992; van Heuven, Dijkstra, & Grainger, 1996). Although also an interactive activation model, BIA has been developed for visual word recognition and is different in a number of ways (e.g. presence of a language node, many inhibitory connections, etc.). Finally, the BIMOLA model will need to be adapted at some point to explain what takes place when the interpreter

doing simultaneous interpretation accesses the words of the source language (and sometimes of the target language) without letting the words of the target language interfere too much with the process. The latter are highly active (they are being used by the output mechanism) and yet they must not be too involved in the recognition process. How the interpreter keeps the input and output processes separate, at this level of processing but also at all other levels, remains an intriguing question.

4.   Conclusion

Despite what is already known about the bilingual individual, much more research needs to be conducted on the topic. The emergence of a wholistic view of bilingualism is encouraging researchers to move away from the monolingual yardstick and develop a true linguistics and cognition of bilingualism. However, many issues require further study: the structure and organization of the bilingual's different languages; the various processing operations involved in the perception, production and memorization of language when the bilingual is in a monolingual language mode or in a bilingual language mode; the linguistic and psycholinguistic differences (and similarities) between code-switches, borrowings and interferences; the organization of the bilingual brain; and finally, cognitive processes in the bilingual individual.

Université de Neuchâtel

References

Altenberg, E., & Cairns, H. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior, 22,* 174-188.

Baetens Beardsmore, H. (1986). *Bilingualism: Basic principles.* Clevedon: Multilingual Matters.

Blank, M. (1980). Measuring lexical access during sentence processing. *Perception and Psychophysics, 28,* 1-8.

Bloomfield, L. (1933). *Language.* New York: Holt, Rinehart and Winston.

Bürki-Cohen, J., Grosjean, F., & Miller, J. (1989). Base language effects on word identification in bilingual speech: Evidence from categorical perception experiments. *Language and Speech, 32(4),* 355-371.

de Groot, A. (1995). Determinants of bilingual lexicosemantic organisation. *Computer Assisted Language Learning, 8(2-3),* 151-180.

de Groot, A., & Kroll, J. (1997). *Tutorials in bilingualism: Psycholinguistic perspectives.* Majwah, NJ: LEA.

di Sciullo, A., Muysken, P., & Singh, R. (1986). Government and code-mixing. *Journal of Linguistics, 22,* 1-24.

Forster, K. (1976). Accessing the mental lexicon. In R. Wales & E. Walker (Eds.), *New approaches to language mechanism* (pp. 257-287). Amsterdam: North-Holland.

Frauenfelder, U., & Tyler, L. (1987). *Spoken word recognition.* Cambridge, Mass.: MIT Press.

Gile, D. (1995). *Basic concepts and models for interpreter and translator training.* Amsterdam: John Benjamins.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207-220). New York: North-Holland.

Green, D. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27,* 210-223.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism.* Cambridge, Mass: Harvard University Press.

Grosjean, F. (1985a). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38,* 299-310.

Grosjean, F. (1985b). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development, 6(6),* 467-477.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3(3),* 233-274.

Grosjean, F. (1997). Processing mixed language: Issues, findings and models. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225-254). Majwah, NJ: LEA.

Grosjean, F., & Gee, J. (1987). Prosodic structure and spoken word recognition. *Cognition, 25,* 135-155.

Grosjean, F., & Soares, C. (1986). Processing mixed language: Some preliminary findings. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and*

neuropsychological perspectives (pp. 145-179). Hillsdale, NJ: Lawrence Erlbaum.

Hakuta, K. (1986). *Mirror of language: The debate on bilingualism.* New York: Basic Books.

Haugen, E. (1969). *The Norwegian language in America: A study in bilingual behavior.* Bloomington, IN: University of Indiana Press.

Klatt, D. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics, 7,* 279-312.

Kolers, P. (1966). Reading and talking bilingually. *American Journal of Psychology, 3,* 357-376.

Leuenberger, M. (1994). L'accès au lexique de code-switchs chez le bilingue: Effets de la densité et du contexte. *Travaux neuchâtelois de linguistique (TRANEL), 21,* 61-72.

Levelt, W. (1989). *Speaking: From intention to articulation.* Cambridge, Mass: MIT Press.

Léwy, N. & Grosjean, F. A computational model of bilingual lexical access. In preparation.

Li, P. (1996). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language, 35,* 757-774.

Macnamara, J. (1967). The bilingual's linguistic performance: A psychological overview. *Journal of Social Issues, 23,* 59-77.

Macnamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior, 10,* 480-487.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71-102.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1-86.

Milroy, L., & Muysken, P. (1995). *One speaker, two languages: Cross-disciplinary perspectives on code-switching.* Cambridge: Cambridge University Press.

Muysken, P. (1995). Code-switching and grammatical theory. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages* (pp. 177-198). Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching.* Oxford: Clarendon Press.

Neufeld, G. (1973). The bilingual's lexical store. *Working Papers on Bilingualism, 1,* 35-65.

Obler, L., & Albert, M. (1978). A monitor system for bilingual language processing. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 105-113). Columbia, SC: Hornbeam Press.

Paradis, M. (1980). The language switch in bilinguals: Psycholinguistic and neurolinguistic perspectives. In P. Nelde (Ed.), *Languages in contact and conflict* (pp. 501-506). Weisbaden: Franz Steiner Verlag.

Paradis, M. (1986). Bilingualism. In *International encyclopedia of education* (pp. 489-493). Oxford: Pergamon Press.

Paradis, M. (1989). Bilingual and polyglot aphasia. In F. Boller & J. Grafman Eds), *Handbook of neuropsychology,* vol. 2 (pp. 117-140). Amsterdam: Elsevier.

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics, 26*, 47-104.

Romaine, S. (1995). *Bilingualism.* London: Blackwell.

Schreuder, R., & Weltens, B. (1993). *The bilingual lexicon.* Amsterdam & Philadelphia: John Benjamins.

Schwanenflugel, P., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language, 25*, 605-618.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition, 12(4)*, 380-386.

Swinney, D. (1982). The structure and time-course of information interaction during speech comprehension: Lexical segmentation, access and interpretation. In J. Mehler, E. Walker, & M. Garrett (Eds.), *Perspectives on mental representations* (pp. 151-167). Hillsdale, NJ: Lawrence Erlbaum Associates.

Taylor, I. (1971). How are words from two languages organized in bilinguals' memory? *Canadian Journal of Psychology, 25*, 228-240.

Thiery, C. (1978). True bilingualism and second language learning. In D. Gerver and H. Sinaiko (Eds.), *Language, interpretation and communication* (pp. 145-154). New York: Plenum Press.

Tulving, E., & Colotla, V. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1*, 86-98.

van Heuven, W., Dijkstra, T., & Grainger, J. (1996). *Orthographic neighborhood effects in bilingual word recognition.* Unpublished manuscript, NICI, University of Nijmegen, The Netherlands.