10

The Role of Guest-Word Properties*

This chapter is best introduced by an anecdote. Olivier, a 5-year-old French-English bilingual boy, comes up to his mother and is overheard by a bilingual onlooker as saying, "Maman, tu peux me tailler mes chaussures?" (Mummy, can you sharpen my shoes?). With no apparent hesitation, the mother kneels down and starts to *tie* his shoelaces, while the onlooker strives to understand what Olivier said: "tailler des chaussures (sharpen shoes)? No, that doesn't make sense . . . ah, he's asking to have his shoes tied." Any reader who knows both English and French will have understood the predicament the onlooker was in: by inserting the English "tie" into his French sentence and adapting it morphologically and phonetically, Olivier unwittingly brought the English guest word ("tier") into conflict with an existing word in French ("tailler") and led the onlooker down a word recognition garden path. The mother, used to hearing Olivier employ "tie" in French, accessed the English meaning with no problem and went about the job of lacing Olivier's shoes. In the present study we will explore the underlying processes that take place when bilinguals have to recognize different types of guest words such as the word "tie" in the above example.

Bilinguals, that is those who use two languages (or dialects) in their everyday lives, move in and out of different speech modes depending on the interlocutor they are facing and the situation they are in. They are in a monolingual speech mode when speaking to monolinguals who speak only one of their two languages, and they are in a bilingual speech mode when they are speaking to other bilinguals who share the same two languages, and with whom they normally mix languages.

* This chapter first appeared as an article: Grosjean, F. (1988). "Exploring the recognition of guest words in bilingual speech", *Language and Cognitive Processes* 3: 233–74. The author wishes to thank Taylor and Francis (www.informaworld.com) for permission to reprint it here.

(For a discussion of speech modes and the continuum they belong to, see Grosjean (1982, 1985c); Grosjean and Soares (1986); and Chapter 4 of this book.) In the monolingual speech mode, bilinguals adopt the language of the monolingual interlocutor and reduce the activation of the other language. Some researchers have proposed various mechanisms, such as a switch or a monitor, that allow this reduction in activation (Macnamara 1967; Obler and Albert 1978; Penfield 1959); but others, notably Paradis (1980), have argued that such mediating devices are not necessary. All agree, however, that bilinguals rarely manage to deactivate totally the language not being spoken. This can be seen in various types of production and perception interferences, that is, the involuntary influence of one language on the other. In production, one notes pronunciation "errors", accidental lexical borrowings, "odd" syntactic constructions, etc., and in perception the residual activation of the other language can be observed in cross-language Stroop tests (Obler and Albert 1978; Preston and Lambert 1969), word–nonword judgments (Altenberg and Cairns 1983), and comprehension tasks using the phoneme monitoring paradigm (Blair and Harris 1981).

In the bilingual speech mode, the mode that is of interest to us in this study, both languages are activated, and bilinguals often use elements of one language when speaking the other. One language usually serves as the base language (the main language of communication) and the other language—we will call it the "guest" language—is brought in at various points during the interaction when the need occurs. Note that simply speaking to another bilingual does not automatically entail the use of the other language; a number of factors account for the presence of language mixing and for how much takes place (Grosjean 1982). Bringing the other language (the "guest" language) into the base language is usually done in two different ways: by code-switching or by borrowing. In code-switching, the bilingual usually shifts completely to the other language for a word, a phrase, or a sentence. For example:

(1) C'était des *wild guys* à cheval
"Those were wild guys on horseback"

(2) J'ai l'impression d'être *back in the country*
"I've got the feeling I'm back in the country"

Code-switching is a phenomenon that has received considerable attention from researchers: linguists have studied the syntactic constraints that govern the alternation between languages within a sentence (Joshi 1985; Lipski 1978, 1982; Pfaff 1979; Poplack 1980; Timm 1975; Woolford

1983); sociolinguists have studied the factors that account for code-switching (Gal 1979; Gumperz 1970; Scotton and Ury 1977; Valdes Fallis 1976); and developmental psycholinguists have studied the development of code-switching abilities in children (e.g. McClure 1977).

The other way of bringing the guest language into the base language is by borrowing a word from that language and integrating it phonetically and morphologically into the base language. For example:

(3) On peut SWITCHER les places?
"Can we switch the seats?"

(4) Il a SLASHÉ le rideau
"He slashed the curtain"

Here the English words "switch" and "slash" are pronounced in French and are integrated morphologically into the sentence. Note that these borrowings (which are also called "speech" or "nonce" borrowings) are different from "language borrowings" (or "loan words") which are borrowings only in a historical sense (Haugen 1969). The latter are now an integral part of the base language and are used by monolinguals and bilinguals alike (e.g. "weekend", "jazz", "transistor" in French; "fiancé", "croissant" in English). The borrowings we will be concerned with here are the "speech" or "nonce" borrowings produced by bilinguals when speaking with other bilinguals in a mixed language speaking mode. These kinds of borrowings, along with code-switched words, only belong to the lexicon of the other (or guest) language.

The aim of the present study is to explore how guest words, produced as borrowings or code-switches, are processed by bilingual listeners. Although much research has been undertaken to understand the processes underlying the recognition of spoken words in monolinguals (see e.g. Cole and Jakimik 1978; Foss and Blank 1980; Grosjean 1980, 1985a; McClelland and Elman 1986; Marslen-Wilson and Welsh 1978; Morton 1969), much less is known about how guest words are processed by bilinguals in a mixed-language interaction.

Recently, though, Soares and Grosjean (1984) investigated the recognition of base language words and of code-switched words in monolingual and bilingual sentences. They used Blank's (1980) "Phoneme Triggered Lexical Decision" task and obtained two interesting results. The first was that although bilinguals, in a monolingual speech mode, accessed base-language words as quickly as 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 in a monolingual mode. The second result of interest was that bilinguals took longer to access code-switched words than base-language words. It seemed that such factors as the phonetic and phonotactic characteristics of the guest word, the base-language context, the amount of code-switching that has occurred up to that point, etc., can account for the delay in processing.

The object of the present study is not to study further whether code-switches take more time to process than base-language words, or to investigate how the delay is made up during the ensuing speech. Such questions are important and are currently being studied. Rather, the aim here is to explore the underlying processes that are involved in the lexical access of guest words (that is, code-switches and borrowings) when they are produced and perceived in a bilingual speech mode. We will assume that the bilingual has two lexicons, which are interconnected in some way, and that guest words are stored, and therefore have to be accessed, in the other, less activated, lexicon.[1] Our exploration will revolve around the roles of two variables in the recognition of guest words—a structural or "word type" variable, and an output or "language phonetics" variable.

Concerning word type, we wish to ask the following questions. First, will guest words that are marked phonotactically as belonging to the guest-language lexicon only (because of the initial consonant cluster, for example) be recognized sooner and with more ease than words not marked in this way? Thus, will words like "snap", "blot", and "quit", which have initial consonant clusters that are more frequent in English than in French, be accessed more easily than words that do not have such language-specific cues? Second, will guest words that belong solely to the guest lexicon be identified sooner and with more ease than words that do not belong to just one lexicon? In other words, will the access of guest words like "lead" (/lid/), "tag", and "tease" be facilitated because they are nonwords in French (although possible words)? Third, will words in the guest-language lexicon that have close homophones in the base language be processed with more difficulty than other guest-language words? Thus, will "pick", "cool", and "knot", which have base-language counterparts with different meanings—"piquer"

[1] Although this assumption appears to be quite categorical, it is not meant to be a defense of the independence position in the debate on the organization of the bilingual's two lexicons (see Grosjean 1982 for a review of the controversy). Our use of the word "lexicon" refers to the set of lexical items that belong to one language; we make no claim, at this point, about the independence or interdependence of the two lexicons.

(to sting, puncture, steal), "couler" (to sink), and "noter" (to note down, mark)—be accessed with more difficulty than guest words with no counterparts in the base language?

The second variable we will study, the language phonetics of a word (also called "word phonetics" in this chapter), pertains to whether the guest word is pronounced in the guest language (as a code-switch) or in the base language (as a borrowing). The question of interest is whether code-switches, which normally retain a phonetic cue as to which lexicon they belong to, are easier to process than borrowings which are integrated phonetically and morphologically into the base language and thus have lost some of the cues that can help the listener access the correct lexicon. Will the language of pronunciation of a guest word affect its recognition, especially when the word is pronounced quite differently in the two languages? And what happens to guest words, such as "pick" and "cool", that have a base-language counterpart? Will they be accessed more easily when produced as code-switches than as borrowings? Although two quite distinct versions of the same word will be compared in this study (they will be produced in unaccented French and English), we should keep in mind that the borrowing and code-switching versions of a word are not always so distinct. When a bilingual has an accent in the guest language, for example, what is meant to be a code-switch will often resemble a borrowing (at the phonological level at least). These more hybrid cases will be the object of a later study.

Because the aim of the study is to gain some insight into the underlying processes involved in the identification and recognition of guest words, and not simply to study the role of word type and word phonetics, we will use an experimental paradigm that allows us to uncover some of the underlying operations involved in word recognition, namely the gating paradigm (Grosjean 1980; see Ohman 1966 and Pollack and Pickett 1963 for earlier versions).

In this task, a spoken word is presented from left to right, in segments of increasing duration. At the first presentation, only the first 40 ms of the word are presented; at the second presentation, the first 80 ms are presented; and so on, until at the last presentation, the whole of the word is presented. The subject's task, after each presentation, is to guess the word being presented and to give a confidence rating based on the guess. The gating paradigm presents a number of advantages which make it a useful tool in the study of the word recognition process. First, it allows one to assess how much of a word is needed to be identified or "isolated" correctly. This is done by determining a word's isolation point, that point in the presentation sequence at which the listener has correctly guessed the word and does not subsequently change his or her guess. It has been proposed (Grosjean 1985a) that the "isolation point" reflects the moment, in the left to right recognition process, at which the listener has a strong candidate in mind but has not yet decided to use it in the construction of the interpretative representation of the ongoing message. This point is quite close to the word's uniqueness point as defined in Marslen-Wilson and Welsh's (1978) cohort model—some 20–80 ms according to Tyler and Wessels (1983, 1985)—and corresponds quite closely to what Bradley and Forster (1987) mean when they say that a word has been accessed.

A second advantage of the paradigm is that one can examine the confidence ratings proposed by listeners at various points in time: at the isolation point, at the end of the word, and at the end of the sentence if gating continues after the word (as in Grosjean 1985a). One can also examine where, in the left to right sequence, a perfect confidence rating is given to the word. This "total acceptance point" may be the moment in time at which the word starts being used in the construction of the interpretative representation (Grosjean 1985a). This point occurs later than the uniqueness point and corresponds quite closely to what Bradley and Forster (1987) mean by word recognition, that is, the listener's fixation of belief that he or she has indeed heard word X.

A third advantage of the paradigm is that the word candidates proposed before the isolation point give some insight into the word isolation process itself. By examining responses across subjects we can infer the path followed by the individual listener when he or she is narrowing-in on a word. Thus, in this chapter, we will study the early preference bilinguals have for the base-language lexicon (as shown by Grosjean and Soares 1986), how and when they shift their preference to the guest lexicon, the conflict that arises when both a base language word and guest-language word are possible, and how that conflict is resolved.

In addition to employing all the information provided by the gating paradigm in the exploration of the underlying processes involved in the recognition of guest words in bilingual mixed speech, we will also undertake side analyses. We will study, for example, the relationship that exists between the acoustic information given to listeners (as defined by spectrographic analysis) and the moment at which a word is guessed correctly. We will also examine the impact of the "frequency pull" of words which come into conflict in the recognition process, that is guest words and their base-language homophones. We will end the chapter with a proposal for how an interactive activation model of word recognition can be modified to take into account not only the effects found in monolingual word recognition research, but also the effects that are specific to bilingual language processing. A complete model of how spoken words are recognized in bilingual speech is still far off, but we hope that our proposal can be a first step in that direction.

10.1 Method

10.1.1 Participants

A total of twelve French-English bilingual adults, with no reported speech or hearing defects, served individually in a session lasting 90 minutes. All participants had the following common characteristics. They were native speakers of French and had only started learning English in secondary school; they had moved to the United States as adults and had lived in the Boston region for at least four years (it is on their arrival in the United States that English became a language of communication for them and stopped being a language known only formally); they used their two languages on a daily basis (French in the family and with friends; English at work, in the community, and with American friends); they code-switched and borrowed when speaking French to bilingual friends and family members; and they had served previously as participants in the bilingual research project based in the Psychology Department of Northeastern University. Note that no effort was made to test the bilingual's proficiency in English and French or to use "balanced" bilinguals (see Grosjean (1982, 1985b), for a discussion of the problems linked with proficiency tests and with the use of "balanced" bilinguals). Membership in a bilingual community (the European French speaking community in Boston) and daily use of English and French for at least four years were the critical variables in choosing the participants.

10.1.2 Materials

In total, twenty-four monosyllabic English verbs and eight French filler verbs were chosen for the study. The English verbs all had the same uniqueness point, that is, that point in the left to right sequence of phonemes at which the word distinguishes itself from every other word (Marslen-Wilson 1984). The English items belonged to one of three

TABLE 10.1 A description of the two variables used in the study. Three types of words (structural variable) were pronounced either in English as a code-switch or in French as a borrowing (output variable: language phonetics of word)

Structural variable	Type of word		
	1	2	3
Initial CC or CV favors	English	French	French
French homophone counterpart?	No	No	Yes
Examples	Slash	Feed	Pick
	Blot	Lean	Knot
	Drop	Tease	Sit

Output variable: Language phonetics of word

English (code-switch)	French (borrowing)
Il faudrait qu'on /slæʃ/ . . .	Il faudrait qu'on /slaʃ/ . . .
Il faudrait qu'on /blɔt/ . . .	Il faudrait qu'on /blɔt/ . . .
Il faudrait qu'on /fid/ . . .	Il faudrait qu'on /fid/ . . .
Il faudrait qu'on /lin/ . . .	Il faudrait qu'on /lin/ . . .
Il faudrait qu'on /pIk/ . . .	Il faudrait qu'on /pik/ . . .

groups, each group containing eight words (see Table 10.1). Type 1 words, like "slash", "blot", and "drop", contained initial consonant clusters (/sl/, /bl/, /dr/, etc.) that are infrequent in French but quite frequent in English. A general comparison of French and English words which have the clusters in question (accomplished by examining the *Micro Robert Dictionary* and *Webster's II New Riverside Dictionary*) showed that the English/French ratio for words with this initial consonant cluster was always in favor of English (the values ranged from 42: 1 to 2.3: 1 with a mean of 22.6). For example, twenty-six English words were found with initial /dr/ as compared to six French words, and twenty-nine English words were found with initial /sl/ but only one French word, etc. Type 2 words, such as "feed", "lean", and "tease", contained an initial CV which occurs more frequently in French than in English. The English/French ratios ranged from 0.83: 1 to 0.40: 1 with a mean of 0.59. Thus, for example, eleven French words started with French /fi/

but only five English words began with English /fi/, and sixteen French words started with French /li/ but only nine words in English began with English /li/, etc. Type 3 words, such as "pick", "knot", and "sit", were similar to Type 2 words in phonotactic configuration (the mean English/French ratio was 0.45, a non-significant difference with Type 2 words), but all had a French homophone counterpart. Thus when words like "pick", "knot", and "sit" are pronounced in French, they cannot be distinguished from their French counterparts with different meanings—"pique", "note", and "cite".

In sum, as can be seen in Table 10.1, Type 1 words were marked phonotactically as belonging to English (because of the initial cluster) and had no French counterparts; Type 2 words were not marked phonotactically as belonging to English (in fact, their phonotactics favored French) but, like Type 1 words, they had no French counterparts; and Type 3 words were phonotactically similar to Type 2 words but, unlike the first two types, they had French counterparts. The uniqueness point of all twenty-four words (with the exception of one word in each group) fell on the last consonant. Finally, the mean frequency of occurrence of the words in the three groups, as measured by Kučera and Francis (1967), was similar: 6.89, 5.10, and 8.57 respectively—$F(2,21) = 0.21$, N.S. The eight French filler verbs were one or two syllables long and began with CCs and CVs (e.g. "soulève", "pratique", "stipule", "grignote", "brosse", etc.).

Each word was embedded in a sentence that began with "Il faudrait qu'on" ("We should") and ended with a three-word NP in which each word was a monosyllable. The initial part of the sentence was chosen so that the morphological integration of the guest word, when said as a borrowing, did not necessitate an inflection and hence increase the number of syllables of the borrowing as compared to the code-switch. As for the final NP, care was taken to make sure that its last word added semantic context to the sentence. Examples of the complete sentences containing the stimulus verbs (in capitals) are:

(5) Il faudrait qu'on SLASH tous les prix
"We should slash all the prices"

(6) If faudrait qu'on LEAN contre le mur
"We should lean against the wall"

(7) Il faudrait qu'on KNOT ces deux cordes
"We should knot these two ropes"

(8) Il faudrait qu'on PICK les bons chiffres
"We should pick the right numbers"

Two type-written versions of the twenty-four experimental sentences were then prepared for the recording of the English verbs as code-switches or borrowings. In the first, the verb was typed normally in the sentence; in the second, the verb was spelled in French. Thus, "slash" was written "slache"; "feed" was written "fide"; "knot" was written "notte"; "fool" was written "foule". The filler sentences were added to this French version. A bilingual French-English female speaker, with no apparent accent in either language, was then asked to read the two versions of the sentences. This person was chosen because she code-switches and borrows naturally when speaking to other bilinguals and has been used repeatedly to prepare experimental tapes with natural sounding code-switches. For the code-switching set, she was asked to switch naturally to English for the word in question, and for the borrowing set she was requested to read the whole sentence in French.

A waveform analysis of the code-switching and borrowing versions confirmed that all sentences were read naturally (there were no pauses before or after the stimulus words) and that the reader did in fact code-switch when requested to do so. To verify the latter, two acoustic analyses were undertaken on a subgroup of words. In the first, we measured the stop-initial voice onset time (VOT) of the code-switching and borrowing versions of the nine words that began with a stop consonant (four voiced and five unvoiced). The mean VOT value of the code-switches (English) was, as expected, longer than that of the borrowing (French): 46 and 27 ms, respectively ($t = 1.99$, $p < 0.05$). In the second analysis, we measured the duration of the high amplitude periodic portion of the waveform corresponding to the /i/ vowel in the code-switching and borrowing versions of the eight words containing that phoneme (English /i/ and French /i/). Again, as expected, the periodicity lasted longer in English than in French: 154 ms as compared to 110 ms ($t = 6.32$, $p < 0.01$). We concluded from this that the reader had indeed produced two different versions of the experimental words—a code-switching version and a borrowing version (see Table 10.1, bottom).

The recordings of the 56 sentences (24 with code-switches, 24 with borrowings, and 8 monolingual French filler sentences) were digitized at a sampling rate of 20 kHz and gated using a waveform editing program on a PDP 11/44 (see Grosjean 1980, 1985a for a general description of the procedure).[2] For each sentence, the "onset" of the stimulus word and of each of the next three words was located as best as possible by inspecting the speech waveform and by using auditory feedback. Most stimulus words began with a fricative or a stop consonant, and their "onsets" corresponded respectively to the start of the frication in the speech wave and to the end of the silence preceding the release burst.

The presentation set of each gated sentence was prepared in the following way. The first gate contained "Il faudrait qu'on" up to, but not including, the onset of the stimulus word. The second gate contained the same information plus the first 40 ms of the word. From then on, gates were incremented by 40 ms until the end of the word was reached. When the duration of the stimulus word was not an exact multiple of 40, the gate containing the full stimulus word was incremented by the amount remaining. Once the full word had been presented, three "after offset" gates were added to the presentation set. Unlike the stimulus word gates, which were incremented by 40 ms, these three gates were incremented by a whole word. The first "after offset" gate contained the carrier sentence, the stimulus word, and the first word of the following NP; the second gate contained all the previous information plus the second word of the NP; and the third gate (which was also the final presentation gate) presented the whole sentence, including the final NP.

Two experimental tapes were made from these presentation sets. Each tape contained thirty-two sets (eight for each type of word and eight fillers). The order of the fillers and of the word type exemplars was randomized. The only difference between the two tapes was that one tape presented the code-switched version of a particular word and the other tape presented, in exactly the same position, the borrowing version of the same word. Each tape contained four borrowing and four code-switch exemplars of each word type.

10.1.3 Procedure

The twelve participants were split into two groups of six, and were run individually on one of the two experimental tapes. This meant that subjects heard each of the twenty-four stimulus words either as a code-switch or as a borrowing. (As indicated above, they heard as

[2] The overall software package for speech processing was developed at Northeastern University by Thomas Erb and Ashish Tungare, and is based in part on the BLISS system developed by John Mertus at Brown University.

many code-switch exemplars as borrowing exemplars for each word type—four in each case.) The sessions were conducted in French (the usual language of communication between the experimenter, a bilingual himself, and the participants) and the instructions were written in French. The participants were told that they would be hearing English or French verbs, presented in segments of increasing duration, followed by a short three-word phrase, presented one word at a time after the stimulus word. They were also told that in the case of English verbs, the word could be pronounced in English or in French. They were asked to listen to the presentations and, after each presentation, to do three things: (1) write down the word they thought was being presented after "Il faudrait qu'on"; (2) indicate how confident they were about their guess by circling a number on a scale of 1–10 (anchored with "Incertain" (Unsure) and "Certain" (Sure)); and (3) indicate whether they thought the word was French or English (that is, belonged to the French or English language) by circling "F" (français) or "A" (anglais) on the right of the confidence rating scale. The answer sheet was arranged in such a way that the sequence of events was first to write down a word, second to give a confidence rating, and third to indicate the language of origin of the word. The participants were given 8 seconds between each presentation to accomplish these three tasks. They were asked to give a response after every presentation, however unsure they might feel about the stimulus word, and they were asked to write the English words with English orthography, even if these words were pronounced in French. A break of 15 minutes was given to the participants halfway through the 90-minute sessions.

10.1.4 Data analysis

Response sheets provided three kinds of information. The first concerned the isolation point of the word, that is, that point at which the subject correctly guessed the stimulus word and did not subsequently change his or her guess. A first analysis indicated whether this point occurred before the offset of the word, after the offset but before the end of the sentence, or never occurred at all (the participant never guessed the word). A second analysis indicated, when appropriate, at what point within the stimulus word the word was isolated (this was expressed as a percentage of the way through the word), and a third analysis indicated, again when appropriate, where during the final NP the stimulus word was isolated correctly. Note that the subjects' orthography of the words and their "A" or "F" indications always permitted one to determine whether the responses were French or English words.

The second type of information concerned confidence ratings. The total acceptance point of a word (that point at which a perfect confidence rating, i.e. 10, was given to the correct response) was located in the same way as the isolation point (before, after, never; within the stimulus word; after offset, etc.).[3] The third type of information obtained was the erroneous candidates proposed prior to the isolation point. These candidates were analyzed in terms of their language of origin and the error type they belonged to (homophone error, segmentation error, etc.). Precise indications of the measures used, and of the tests of significance that were conducted on them, are given below with the results and discussion.

10.2 Results and discussion

10.2.1 The isolation point

In this section, we will first compare the three types of words used in the study: Type 1 words which are marked phonotactically as belonging to English; Type 2 words which are not marked in this way but which only belong to English; and Type 3 words which have French homophone counterparts. We will then separate Type 1 and Type 2 words from Type 3 words, and examine the first two types of words with a measure more appropriate to them. We will end by studying two variables that account for the isolation point of Type 3 words.

The majority of words (76 percent in all) were isolated before their acoustic offset; the remainder were isolated after their offset, but before the end of the sentence (16 percent), or were never isolated at all (8 percent). Figure 10.1 presents the percentages of words that fell into each of these three categories (before, after, never) as a function of word type and language phonetics of the word. As can be seen, practically all Type 1 and Type 2 words were isolated before their ending (97 and 90 percent respectively), whereas less than half of Type 3 words (43 percent) were isolated by then. The remaining Type 1 and Type 2 words were isolated before the end of the sentence (with the exception of 4 percent of Type 2 words), but a full 20 percent of the Type 3 words were not isolated by that point. This clearly indicates that the properties of guest words, such

[3] Because of lack of space, the results pertaining to the confidence ratings will not be presented in this chapter. They simply confirm the isolation point data, and can be obtained from the author upon request.

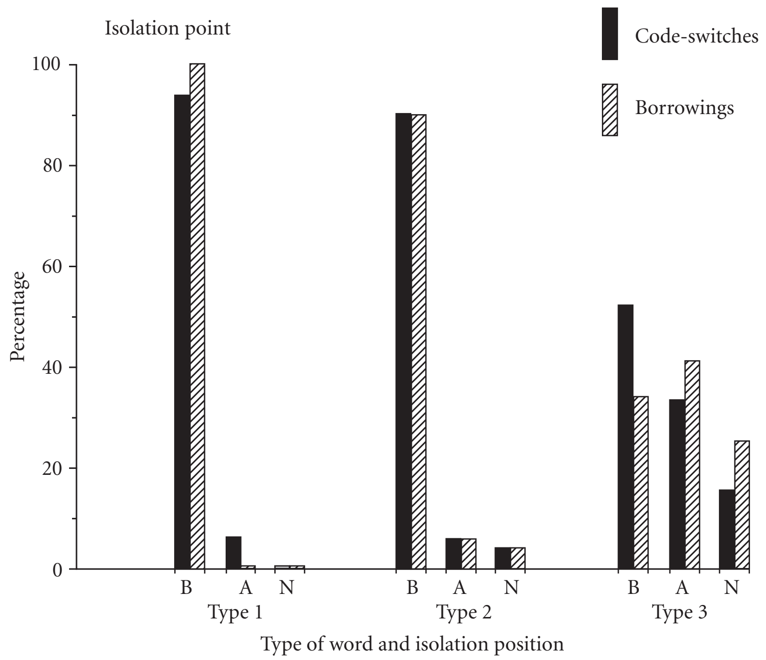

FIGURE 10.1 Percentage of words that are isolated as a function of position of isolation (before their offset (B), after their offset (A), or never isolated (N)), type of word (Types 1–3) and language phonetics of word (pronounced as a code-switch (English) or as a borrowing (French))

as their phonotactic configuration and their single lexicon membership (the English lexicon in this case), can facilitate their isolation. The presence of an initial consonant cluster in Type 1 words ("sn" in "snap", "bl" in "blot", etc.) and the presence of Type 1 and Type 2 words (such as "feed", "tag", "lean", etc.) in the English lexicon, but their absence from the French lexicon, facilitated the isolation of these words before their acoustic offset. Type 3 words, on the other hand, were difficult to isolate, not only because they are phonotactically possible in both languages, but also because they exist in both English and French ("knot" and "note", "cool" and "coule", etc.). The problem caused by these words (57 percent were either isolated late or never isolated at all) is a clear indication that in everyday interactions, bilingual listeners will have problems with such words. If the context is not constraining enough, they will mistakenly use the lexical meaning of the base-language homophone in the construction of the internal representation (the meaning of "couler" and not of "cool", for instance), and will have to backtrack later when enough contradictory information becomes available.

The second important point that emerges from Figure 10.1 is that the language phonetics of a word, that is, whether it is said as a borrowing or a code-switch, appears to play more of a role in the isolation of Type 3 words than in the identification of Type 1 and Type 2 words (at least with the present isolation measure). We note, for example, only a 6 percent difference between the number of code-switches and borrowings that were isolated before offset in Type 1 words (94 and 100 percent, respectively), no difference between them in Type 2 words (90 percent in both cases), but a 19 percent difference between code-switches and borrowings in Type 3 words (52 percent of the code-switches were isolated before offset as compared to 33 percent of the borrowings). Here the difference is in the direction predicted: code-switches, which are marked phonetically for the lexicon they belong to, were isolated sooner than borrowings. Although there were slightly more borrowings than code-switches isolated after offset (41 percent as compared to 33 percent), the deficit accrued by borrowings in the first category was not overcome before the end of the sentence, and a full 25 percent of the borrowings were never isolated, as compared to 15 percent of the code-switches. We conclude from this that pronouncing a guest word in the appropriate language may help in its identification, especially when it is not already "tagged" for the language phonotactically or lexically (as were Type 1 and 2 words).

In order to test the results obtained so far, every word response was given a position rating: a score of 1 if the word was isolated before its acoustic–phonetic offset; a score of 2 if the word was isolated after its offset but before the end of the sentence; and a score of 3 if the word was never isolated. Two analyses of variance were then conducted on these ratings, one over subjects and the other over items. A main effect was found for word type in both analyses—over items: $F\,(2,22) = 49.26$, $p < 0.01$; over subjects: $F\,(2,21) = 23.19$, $p < 0.01$. An *a posteriori* test in the over items analysis (Tukey HSD: Kirk 1967) showed that Type 1 and Type 2 words were not different from one another, but that each was different from Type 3 words ($p < 0.01$). No main effect was found for word phonetics in either analysis, but there was a significant Type × Phonetics interaction in the analysis over items: $F\,(2,22) = 4.32$, $p < 0.05$. An *a posteriori* test showed that only the difference between borrowings and code-switches in Type 3 words was significant ($p < 0.05$). Although the difference between Type 3 borrowings and code-switches is weakened somewhat by the fact that it was found in the over-items analysis only, further evidence will be presented throughout the chapter to show that the language phonetics of a word does indeed appear to play a role in its recognition, especially when that word has a base-language homophone.

Two findings emerge from the analysis so far, therefore. The first is that Type 1 and Type 2 words behave quite differently from Type 3 words; the former are mostly isolated before their offset, whereas the latter are difficult to isolate, and many are never isolated at all. (Note that the broad isolation measure used so far does not allow us to make any claims about the difference that may exist between Type 1 and Type 2 words.) The second finding is that the language phonetics of a word appears to play a role, especially when that word has a base-language homophone; if that word is said as a code-switch, then it will be isolated sooner than if it is said as a borrowing.

We will now examine Type 1 and Type 2 words separately from Type 3 words so as to better understand the isolation process involved in the two subsets of words.

Type 1 and Type 2 words

A within word isolation point, defined as the percentage of the way through the word needed for isolation, was computed for every response. When a particular value was missing, as when the word had been isolated after its offset, or never isolated at all, it was replaced by the mean value for the word calculated over subjects (3 percent of Type 1 values and 10 percent of Type 2 values were replaced in this way). Figure 10.2 represents the amount of a word needed to isolate it as a function of word type and word phonetics (language phonetics of the word). Two findings are apparent. The first is that Type 1 words are isolated earlier than Type 2 words: participants needed, on average, 66 percent of a Type 1 word to isolate it as compared to 78 percent of a Type 2 word. The fact that words like "blot", "slash", "snap", and "quit" are isolated sooner than words like "tag", "feed", "sip", and "beep" appears to indicate that the language specificity of the initial consonant cluster of Type 1 words helps listeners narrow-in more rapidly on the appropriate lexicon and, therefore, on the specific item within it. Note that such factors as word frequency and uniqueness point are not involved here: Type 1 and Type 2 words have similar frequencies and identical uniqueness points.

The second point of interest is that the language phonetics of Type 1 and Type 2 words appears to play little role in the time it takes to isolate them: listeners needed, on average, 70 percent of the borrowings

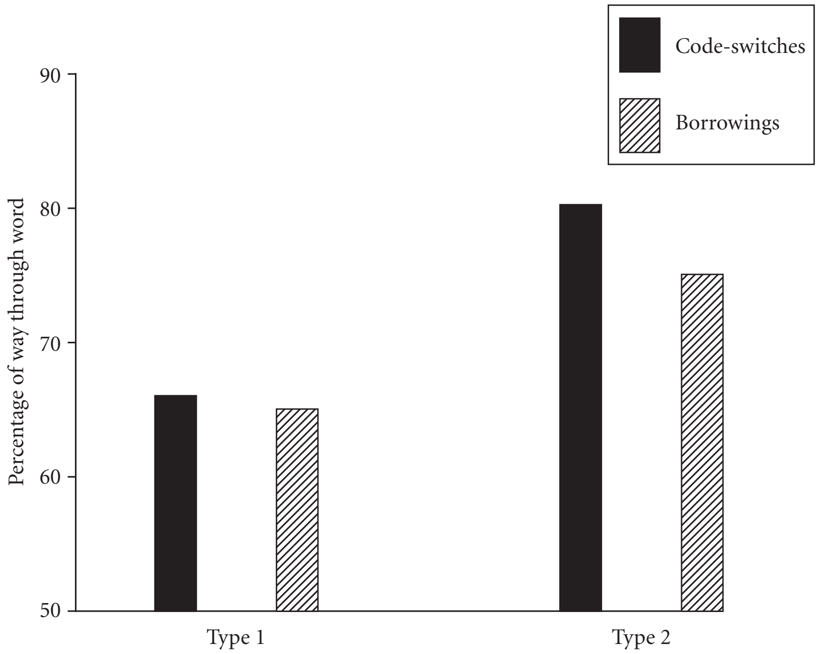

FIGURE 10.2 Amount of a word needed to isolate it (expressed as the percentage of the way through the word) as a function of type of word (Types 1 and 2) and language phonetics of word (pronounced as a code-switch (English) or as a borrowing (French)). Each point is based on 48 observations

(over both types of words) to isolate them and 73 percent of the code-switches. Thus, in this case at least, pronouncing an English word in English or in French has no effect on its identification, unless, as we saw above, it can be mistaken for a word that already exists in the base language. Two analyses of variance, one over subjects and one over items, confirm the pattern obtained. The only main effect obtained was for word type—over items: $F(1{,}11) = 33.60$, $p < 0.01$; over subjects: $F(1{,}14) = 5.09$, $p < 0.05$. No main effect was found for word phonetics, and there was no interaction.

Two aspects of the results pertaining to the language phonetics of a word are surprising. The first concerns the absence of an effect in Type 2 words which, unlike Type 1 words, do not have an initial language specific consonant cluster to indicate the appropriate lexicon. The second concerns the fact that the isolation point of borrowings (b) sometimes occurs before that of code-switches (cs). Thus, for example, participants needed 82 percent of "beep"(cs) on average to isolate it, but only 50 percent of "beep"(b); they needed 91 percent of "tag"(cs) to isolate it, but only 69 percent of "tag"(b), etc. We decided to investigate these two aspects by conducting an acoustic analysis of Type 2 word pairs (borrowings and code-switches).

We located on the speech wave form the vowel "offset" of each pair, and calculated for these items the time that elapsed between the beginning of the word and the "end" of the vowel (see Repp 1981 for a discussion of linguistic categories and their physical correlates in the speech wave). We reasoned that, by the end of the vowel, the listener would have received enough consonantal information (through co-articulation) to be able to isolate the word correctly. (The uniqueness point of these CVC words was on the final consonant.) The vowel "offset" was located in different ways depending on the words in question: (1) for the two words that ended with a voiceless stop consonant ("beep" and "sip"), the offset was defined as that point where the high amplitude periodicity associated with the vowel ended and the closure silence began; (2) for the four words that ended with a voiced stop consonant ("feed", "tag", "lead", "dab"), the offset was that point where the high amplitude periodicity ended and the closure periodicity began; (3) for the word "tease", vowel offset was that point where the periodicity ended and the aperiodic energy associated with the fricative began; and (4) for the word "lean", the offset was that point where the periodicity associated with the vowel ended and the nasal periodicity began. In each case, auditory feedback and an examination of the spectrogram was used to confirm the measurement decisions.

The vowel offset values for the sixteen Type 2 words (eight code-switches and eight borrowings) were transformed into a percentage of the way through the word and were correlated with the corresponding isolation points (also expressed as a percentage of the way through the word). The Pearson product–moment correlation obtained was 0.55 ($p < 0.05$), indicating a rather strong relationship between vowel offset and isolation point—the earlier the vowel offset, the earlier the isolation point, and vice versa. This relationship is illustrated quite clearly in Figure 10.3, where we present the spectrogram of "tag" pronounced as a borrowing (top) and as a code-switch (bottom). (For presentation purposes, the two versions are displayed without their preceding and following context.) As can be seen, the linear arrangement of the acoustic characteristics is quite different in the two languages (the initial consonant burst is longer in English, the final consonant release is longer in French, etc.) and,

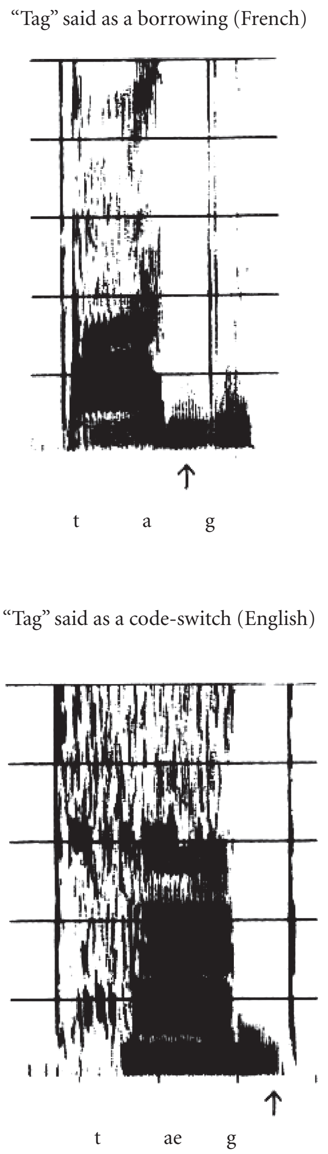

FIGURE 10.3 The spectrograms of “tag” when said as a borrowing (French: top) and as a code-switch (English: bottom). The Y-axis goes up to 5 kHz. The vertical arrows indicate the isolation points of the two versions

therefore, the vowel offset occurs at different points in the borrowing and in the code-switch—54 and 75 percent of the way through the word, respectively. It is no surprise, therefore, that the isolation points are different for the two words—69 percent for the borrowing and

91 percent for the code-switch. What is critical, however, is that these isolation points occur at very similar informational points—during the closure preceding the release of the /g/ (see the black arrows below the spectrograms).

In order to control for other variables that may have played a role in the isolation of the words (frequency, familiarity, number of candidates after each phoneme, etc.) and which, *de facto*, would be keeping the correlation coefficient at its 0.55 level, we calculated for each word pair (borrowing, code-switch) a vowel offset difference (the difference between the end of the vowel in the code-switch and in the borrowing) and an isolation point difference (the difference between the isolation point of the code-switch and the borrowing). We then correlated these two sets of differences, and obtained a much higher coefficient of correlation: 0.82 ($p < 0.01$). Thus the greater the difference between the vowel offset of a code-switch and of a borrowing, the greater the difference between the isolation point of the two words (and vice versa). We conclude from this that the important factor in the isolation of a guest word with no base-language homophonic counterpart is whether the critical acoustic–phonetic information has been received, and not whether the word has been pronounced in the guest language or in the base language. Of course, this generalization does not include words with base-language homophones (Type 3 words) or words whose phonetic configuration changes quite drastically when pronounced in the other language—either because specific phonemes are absent in that language and close neighbors have to be used (for example, when the French /f/ replaces the English /ð/) or because the speaker has a strong accent in the guest language.

We can summarize the word isolation results so far by stating that word type (which in our case includes the phonotactic configuration of a word and its presence or absence in the base-language lexicon) is an important variable in the recognition process of guest words. The language phonetics of a word, on the other hand, plays less of a role. (Its role is more important during the narrowing-in stage, prior to word isolation, as we will see below.) Code-switches are not isolated sooner than borrowings when the words are marked phonotactically (Type 1 words) or when they belong to only one lexicon (Type 1 and Type 2 words), but only when the guest word comes into conflict with a base-language homophone. In this case, the fact that the code-switch retains some phonetic cues regarding its lexicon of origin helps in its identification.

Type 3 words

We saw in the first part of this section that Type 3 words (those with cross-language homophones) behave quite differently from Type 1 and Type 2 words: not only are they isolated later (and quite often never isolated at all), but they also appear to be the only kinds of words in which the language phonetics—whether they are pronounced in English or in French—plays a role in their isolation. In what follows, we will examine the role played by the post-offset syntactic and semantic context on the isolation of these words, and we will study how two variables—the frequency of the stimulus word and of its base-language homophone, and the specific phonetic characteristics of the guest word—can speed up or slow down the isolation of these words.

When one examines the exact isolation position of Type 3 words identified after their acoustic offset (37 percent of them in all), an interesting pattern emerges. Figure 10.4 presents the percentage of words isolated as a function of post-offset position: during the next word (+1), during the word after (+2) or during the last word of the sentence (+3). The post-offset percentages obtained for Type 1 and Type 2 words have been included for comparison. As can be seen, the narrowing-in pattern for Type 3 words is quite distinct. Whereas the few Type 1 and

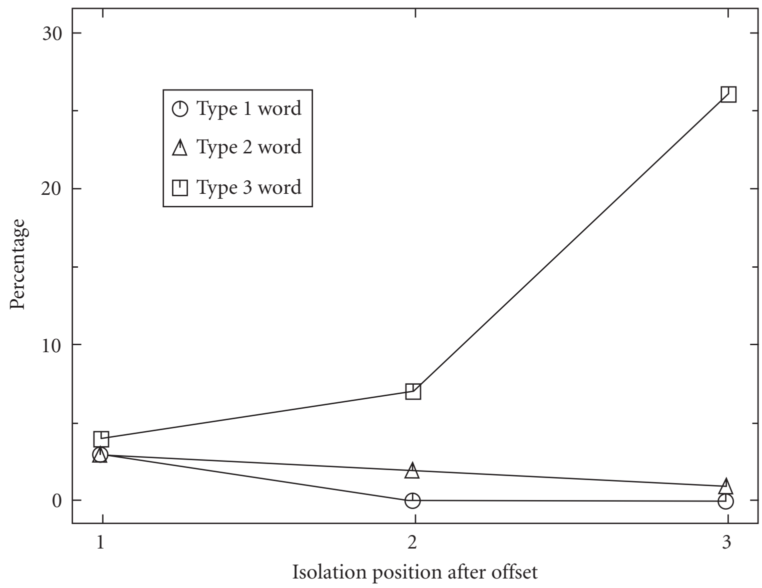

FIGURE 10.4 Percentage of words isolated as a function of word type (Types 1–3) and post-offset position: during the next word (+1), during the word after (+2), or during the last word of the sentence (+3)

Type 2 words that remain to be identified are rapidly isolated in positions +1 and +2 (where the phonetic, lexical, syntactic, and semantic information all help in the choice of the appropriate word), the isolation pattern of Type 3 words is quite different. Instead of being characterized by a slightly negative function, as in the case of Type 1 and 2 words, the isolation function is positive and rather steep between positions +2 and +3: 4 percent of the words are isolated in the first position, 8 percent are isolated in the second position, and as many as 26 percent are identified in position +3. This is a clear indication that in the case of words with base-language homophones, the isolation of the words will need to await the appropriate semantic information. This information is presented primarily in the last word of the sentence (position +3). For example, "cordes" in the sentence "knot ces deux cordes" (knot these two ropes) tells the listener that the word is not "note" (to note) but "knot". Another example involves "bières" in the sentence "cool ces deux bières" (cool these two beers); before hearing "bières" many listeners thought they were dealing with "coule" (to sink) and not "cool". It was only on hearing the last word of the sentence that a number of subjects modified their proposal. We should note that more borrowings were isolated in position +3 than code-switches (31 percent as compared to 21 percent); this is because more borrowings remained to be isolated after their acoustic offset (67 percent as compared to 48 percent) and the semantic information carried by the noun in position +3 allowed some of them to be "caught" before the end of the sentence.

A second point of interest concerning Type 3 words is the apparent role played by two variables in the isolation process: the frequency of the stimulus word and of its base-language homophone, and the specific phonetic characteristics of the word. As regards the first variable, we were struck by the rather large variability in the isolation results of Type 3 words (see Figure 10.1)—some were isolated before word offset, others were isolated during the next word or words, and some were never isolated at all. For example, when we combined the code-switching and borrowing results, we found that "peel" was isolated eleven times out of a possible twelve before its acoustic offset, and "sit" was isolated eight times before its offset. On the other hand, "knot" was always isolated after its offset (ten times during the last noun of the sentence), and "cool" failed to be isolated on five occasions. We hypothesized that this large variability in the isolation results could perhaps be explained by the frequency "pull" of the English words and of their French counterparts. If an English word is more frequent than its French homophone, then the guest word (pronounced as a borrowing or a code-switch) should be identified quite quickly. If, on the other hand, the English word is less frequent than its French homophone, then the listener should be "pulled" towards the French item, and the stimulus word should be isolated later (or maybe even never).

To test this hypothesis, we obtained subjective frequency ratings for the Type 3 English words in their infinitive form (e.g. "to peel", "to fool", "to knot", etc.) and for their homophonic counterparts (e.g. "piler", "fouler", "noter", etc.). In all, eleven French-English bilinguals were asked to rate the sixteen words on a scale of 1–10, where 1 corresponded to very infrequent words and 10 corresponded to very frequent words. This subjective estimation approach was used because there are no published frequency lists for the bilingual population we used and because Segui *et al.* (1982) have reported a very high correlation (in the order of 0.85–0.90) between subjective and observed word frequency. The ratings were averaged over participants and an "English pull index" was calculated for each word pair by subtracting the rating of the French word from the rating of the English word. The eight indices obtained in this way ranged from positive values, indicating a higher frequency for the English item, to negative values, indicating a higher frequency for the French item. Thus, for example, the pull index for "pick" was 1.18 because the estimated frequency for "pick" was 8.64 and that for "piquer" was 7.46; the index for "knot" was −3.46 because the estimated frequency for "knot" was 4.27 and that for "noter" was 7.73, etc. These eight indices were then correlated with the corresponding mean position indices of the words averaged over code-switches and borrowings (as we indicated in the first part of this section, individual indices ranged from 1 for words isolated before their acoustic–phonetic offset to 3 for words that were never isolated).

The resulting Pearson correlation coefficient was a surprisingly high −0.77 ($p < 0.05$): the stronger the pull towards the English word, the earlier the isolation point and, conversely, the stronger the pull towards the French homophone, the later the isolation. Two examples will illustrate this relationship. The pull index for "peel" was a rather high 2.45 ("to peel" is more frequent than "piler") and the position index for the word was therefore quite low (1.17; 11 of the 12 tokens of the word were isolated before offset). On the other hand, the pull index for "knot" was −3.46 ("to knot" is less frequent than "noter") and the position index for the word was therefore quite high (2.09; 11

of the 12 cases were isolated after offset). We conclude from this that the ease with which a Type 3 word is identified depends, in part, on the "frequency pull" of that word. If the word is more frequent than its base-language homophone, then it will be identified quite early on. If, on the other hand, the base-language homophone is more frequent, then the identification of the guest word will be delayed.

A second factor which appears to affect the isolation of Type 3 words, but for which we only have a small amount of evidence, is the specific phonetic characteristics of the words. A side analysis showed that code-switches that are "flagged" phonetically as being English—such as "pick"(cs) or "wrap"(cs)—are isolated relatively early, whereas their borrowing counterparts—"pick"(b) and "wrap"(b)—are isolated late. On the other hand, code-switches that are not as strongly marked phonetically, such as "knot"(cs), are isolated in the same amount of time as their borrowing counterparts, such as "knot"(b).

We can conclude this first part by stating that the two variables under examination—word type and the language phonetics of a word—both play a role in the identification of guest words. They do so, however, to different degrees. Word type is a strong variable that accounts for the different isolation points of three types of guest words: words that are marked phonotactically as belonging to the guest-language lexicon only (Type 1 words); words that are not marked in this way, but that only exist in the guest language (Type 2 words); and words that have base-language homophones (Type 3 words). The language phonetics of a word, on the other hand, is a variable that appears to take on some importance mainly when there is an ambiguity concerning the origin of the lexical item, that is, whether it belongs to the base-language or to the guest-language lexicon. And even then, other variables, such as the frequency of occurrence of the guest-language word and of its homophonic counterpart, as well as the phonetic specificity of individual sounds in the guest-language pronunciation, will intervene to increase or decrease the effect of the language phonetics variable.

10.2.2 The word isolation process

The experimental paradigm we have used in this study allows us not only to determine how much of the stimulus word is needed to isolate it, but also to better understand the word-isolation process itself. This is done by analyzing the candidates proposed prior to the isolation point. As in earlier research (Grosjean 1980, 1985a), we will assume that by examining responses across subjects we can infer the path followed by the individual listener. We will also assume that the candidates proposed by the subjects on the basis of gated information are similar, in part at least, to those that would be available were we able to tap into the word-isolation process as it takes place during online processing of language. In what follows, we will first illustrate the narrowing-in process of three exemplars of the guest words used in the study, one for each of the three types of words. We will then examine in more detail three aspects of the word isolation process.

Figure 10.5 presents the candidates proposed for the word "snap" (vertical axis) as the length of the gate increased in duration (horizontal axis). The top part of the figure presents the candidates proposed for the code-switched version of the word, and the bottom part the candidates for the borrowed version. The word offset is marked by a horizontal dashed line, and the three gates beyond that represent the post-offset presentations where the stimulus word is presented along with the next word ("tous"), the next two words ("tous en"), or the next three words ("tous en rythme"). Candidates proposed at only one gate duration are depicted by a dot; those that are proposed over two or more gates are depicted by a line. The number of subjects proposing a particular candidate is represented by the thickness of the line—the more subjects, the thicker the line. English candidates are written in capitals on the vertical axis and are represented by continuous black lines or bars; French candidates are written in lower case and are depicted by discontinuous lines or bars.

A number of interesting points emerge from the figure. The first is that the two versions of "snap" (which have slightly different total durations) are isolated at very similar points, and much before their offsets. This simply illustrates what has been stressed so far about Type 1 words: their initial consonant cluster and their single lexicon membership allows them to be isolated very early on. A second point that emerges is that the candidates at the very early gates are more often French words than English words: of the twenty-four candidates proposed at the first two gates of the two versions, thirteen are actual French words and four, for which no actual lexical item is written (we have marked this with a Ø), are thought to be of French origin. This is further evidence of the base-language (assimilation) effect studied by Grosjean and Soares (1986): when listening to a base language, the listener "expects" (or is primed for) the next item to be in the base language, unless "warned" otherwise. The third point is that the phonotactic (consonant cluster)

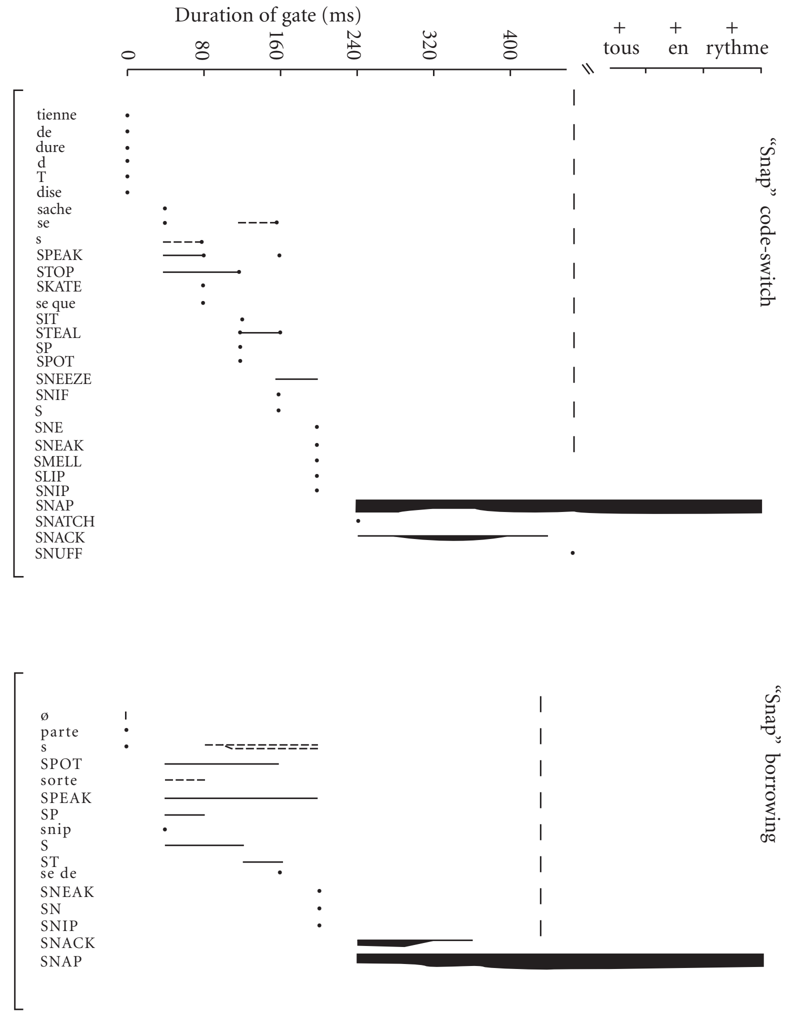

FIGURE 10.5 Candidates proposed for "snap" when produced as a code-switch (top) and as a borrowing (bottom)

Note: The candidates are listed on the vertical axis and the duration of the gates are marked along the horizontal axis (gates were incremented by 40 ms). The dashed vertical line marks the offset of the words; after that point the gates were incremented by a full word. Candidates proposed at only one presentation are depicted by a dot; those proposed over two or more presentations are depicted by a continuous line. The number of subjects proposing a particular candidate is represented by the thickness of the line—the more participants, the thicker the line. English candidates are written in capitals and are represented by continuous black lines; French candidates are written in lower case and are depicted by discontinuous white lines or bars.

information that is given in the early gates leads to a rapid decline of French candidates and a rapid increase of English candidates that begin with /s/ plus a consonant. The fourth point is that the decline of French candidates is not quite as rapid for the borrowing as for the

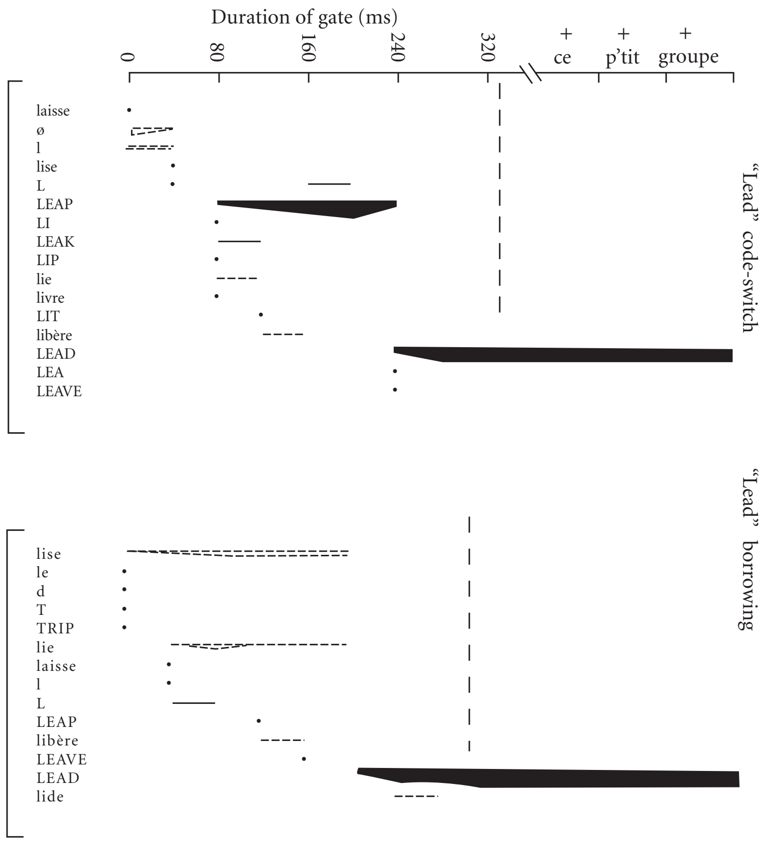

FIGURE 10.6 Candidates proposed for "lead" when produced as a code-switch (top) and as a borrowing (bottom)

Note: See the caption of Figure 10.5 for details.

code-switch: 40 percent of all erroneous candidates (tokens) are French words when the borrowing is presented, as compared to 28 percent for the code-switch. This word phonetics effect is only momentary, however, and as soon as enough acoustic–phonetic information has specified the word "snap" (whether it is pronounced in English or French), it is isolated by the majority of subjects.

Figure 10.6 presents the narrowing-in patterns for the two versions of the Type 2 word "lead". Like "snap", both versions of the word are isolated before their offset, albeit later than the Type 1 word. This early isolation is due to the fact that "lead" belongs unambiguously to the English lexicon and that, by the time the offset is reached, the listener has received enough information to isolate it correctly. The actual isolation points of the two versions are very similar (82 and 77 percent of the way through the word for the code-switch and the borrowing, respectively), but the narrowing-in patterns are quite different. What is especially striking is the rather rapid selection of English candidates in the case of the code-switch (at the third gate, four of the six candidates are already English words) as compared to the maintenance of French candidates when the borrowing is being heard. Overall, 60 percent of the erroneous candidates (tokens) for the code-switch are English words, whereas only 17 percent of the candidates for the borrowing belong to that language.

This difference between the two versions of "lead" is probably due to the distinct pronunciations of the word in the two languages. The initial /l/ in English is very different from its counterpart in French, and this difference is reinforced by the different articulations of English /i/ and French /i/. Thus, when listening to the code-switched version, listeners quickly opt for English words that start with English /li/ (note the "leap" garden path), whereas when listening to the borrowed version, listeners choose French /li/ words (note the "lise" garden path). The cue for French is so strong in the case of the borrowing that one subject actually wrote "lide?" (and circled "F" on the answer sheet). The error was not due to a problem in spelling as the listener switched her guess to "lead" and circled "A" (for "anglais") two gates later. It is interesting to note that despite the early choice of the English lexicon in the case of the code-switch, the isolation of the actual item takes place no earlier in time than when the borrowing is presented. This is because words are isolated when the sequence of sounds allows them to become unique, as we saw above, and this point is reached at about the same time in the two versions of "lead" (the closure of the final /d/ begins 82 percent of the way through the code-switch and 78 percent of the way through the borrowing).

Figure 10.7 presents the candidates proposed for the two versions of the Type 3 word "pick". As is clearly evident, very different narrowing-in patterns emerge from the presentation of the two versions. With the code-switch, after two gates of uncertainty (and hence of French candidates), listeners quickly propose English candidates and then narrow-in very rapidly on the stimulus word. The pattern for the borrowing is quite different. Although some subjects opt for the stimulus word quite early on (it is slightly more frequent than the base-language counterpart "pique"), the acoustic–phonetic cues just prior to word offset force almost all of them to switch over to the French homophone (at the offset, five of the six participants proposed "pique"). In the

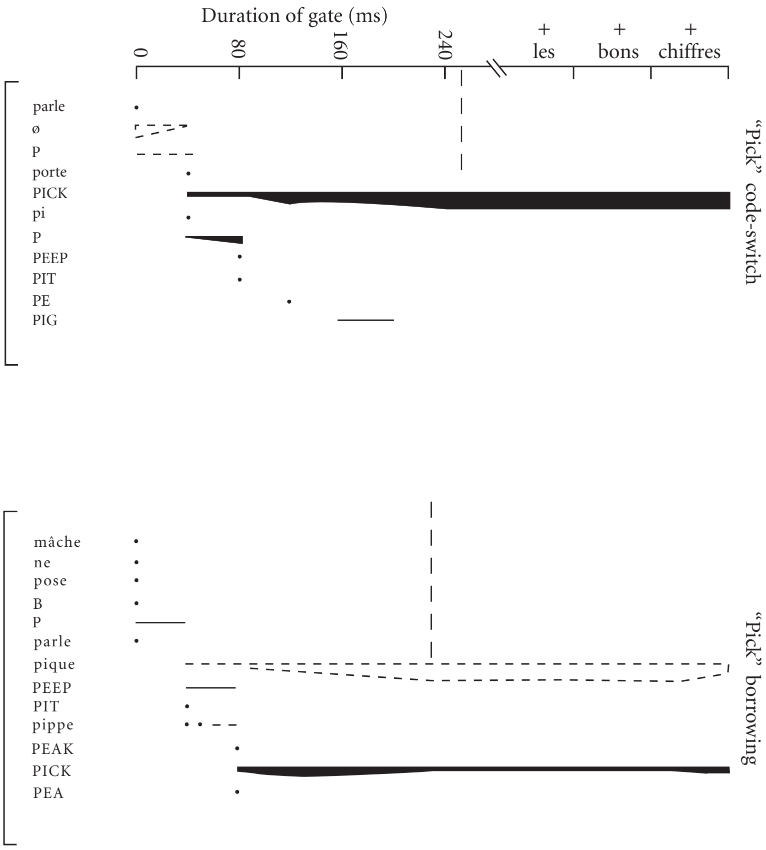

FIGURE 10.7 Candidates proposed for "pick" when produced as a code-switch (top) and as a borrowing (bottom)

Note: See the caption of Figure 10.5 for details.

post-offset presentations, the rather constraining context created by the last word of the sentence—"chiffres" in "les bons chiffres" ("the right numbers")—only made one of the five participants change over to "pick" at the last gate; the remaining four ended the gating sequence with the erroneous candidate "pique". This is a clear example of how Type 3 words can be affected by the language they are said in: if they are pronounced in the base language and neither the context nor the frequency pull is in favor of the guest language, then there is every chance that the base-language homophone will be accessed; if, on the other hand, they are pronounced in the guest language, then the listener will often opt for the correct word. This difference in access strategies will be even greater when the phonetics of the two versions of the word are clearly those of the respective languages, as was the case for "pick"

(French and English /pi/ are pronounced quite differently). Of course, not all guest words are cued so strongly for one or the other lexicon and, in that case, both the stimulus word and the base-language homophone may be proposed as candidates.

In what follows we will examine in more depth three aspects of the word isolation process: the candidates proposed at the early gates, the language of erroneous candidates before word isolation, and the erroneous candidates of Type 3 words in the post-offset presentations.

The early candidates

Figure 10.8 presents, for the three types of words used in the study, the average number of English candidates proposed during the first five gates (the results are averaged over borrowings and code-switches). Two clear patterns emerge from Figure 10.8. The first is that subjects show a strong base language effect at the beginning of the word. It is rare that they propose an English candidate at the first gate (where very little, if any, information concerning the word is presented) and it is only over the next two or three gates, as the phonetic, phonotactic, and lexical information starts arriving, that they begin to propose words from the other lexicon. This finding is very similar to the ones obtained

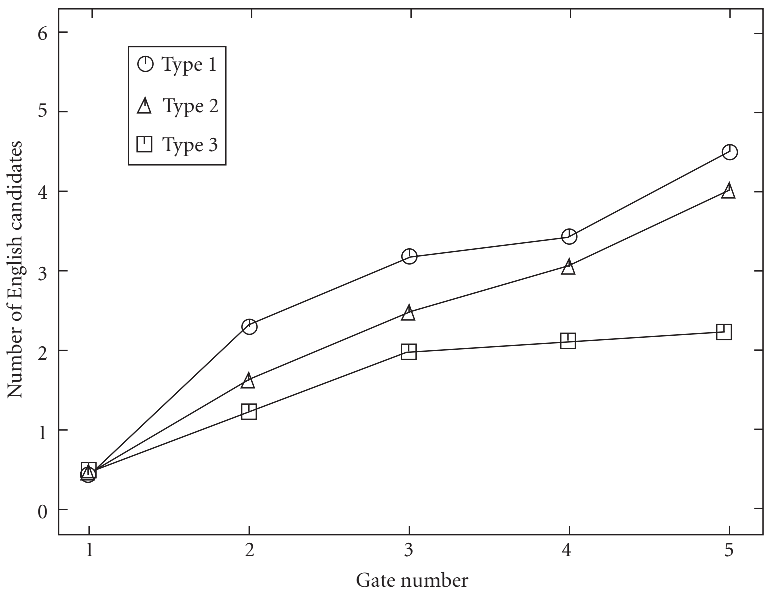

Figure 10.8 The average number of English candidates presented as a function of the number of the gate (from the first to the fifth) and the type of word (Types 1–3)

Note: Each point is based on 96 observations.

by Grosjean and Soares (1986), where subjects were given gated nonsense code-switched words in context and had to guess the language of these words. Invariably, during the first two or three gates, subjects thought the language of the word was the context (base) language. This base-language effect (studied in a different manner by Macnamara and Kushnir 1971, and discussed in Chapter 6) could explain why many studies have found some delay in the recognition of code-switched words as opposed to base-language words (see e.g. Soares and Grosjean 1984). In the case of code-switches, listeners may at first search the wrong lexicon, whereas in the case of base-language words they immediately search the correct lexicon.

The second pattern to emerge from Figure 10.8 is that the functions of Type 1 and Type 2 words rise more rapidly than that of Type 3 words. Whereas the early information in the first two types of words indicates that the items are probably English words, the early information of Type 3 words points towards French words (the base-language homophones). Thus, after a short rise in the number of English candidates, the Type 3 function stabilizes and remains low until further information, usually in the post-offset position, "shocks" the participant to switch over to the guest-language lexicon.

An analysis of variance over participants confirms the pattern just described. A main effect was found for position—$F(4,35) = 17.47$, $p < 0.01$—and for type—$F(2,70) = 9.23$, $p < 0.01$—and there was no significant interaction. An *a posteriori* test (Tukey HSD) showed that Type 1 and Type 2 words were not different from one another, but that each was different from Type 3 words ($p < 0.05$). In addition, a significant difference was found between both Type 1 and Type 2 words and Type 3 words at gate number 5 ($p < 0.05$).

In order to determine whether there was an effect of language phonetics over the early gates, separate analyses of variance were conducted (over subjects) for each word type. All three showed a significant main effect for position, but none produced a language phonetics main effect (all three were in the expected direction and the Type 3 analysis was close to showing an effect). We reasoned that the language phonetics effect was probably reduced by the fact that not all words were marked by "strong" phonetic cues when pronounced in English or in French. In order to assess whether this explanation was correct, we conducted a separate analysis of Type 2 words. We took the four words with "strong" phonetic cues, that is those which started with consonants that are very different in English and French (in our case, the consonants /t/ and /l/

in “tease”, “lead”, “lean”, “tag”), and examined the candidates proposed for these words when said as code-switches and borrowings.

We found that the number of English candidates proposed increases as more of the word is given, but this increase is different for the two versions. For the borrowings, the acoustic–phonetic cues at the beginning of the words (the short VOT for the /t/ or the “clear” /l/, for example) clearly indicate a French word and thus French candidates are proposed. It is only at gate 5 that listeners realize that, despite the unambiguous acoustic–phonetic cues, no French word corresponds to the phonetic sequence; they then start proposing English candidates. The pattern for the code-switches is quite different. The early acoustic–phonetic cues (long VOTs, “dark” /l/) all point to English words and the number of English candidates, therefore, increases rapidly at each gate. The lexicon membership information that comes in later only confirms the correct choice of lexicon made at the beginning. An analysis of variance conducted over participants confirms these findings. There is a main effect for position—$F(4,30) = 4.76$, $p < 0.01$—and a main effect for word phonetics—$F(1,30) = 6.57$, $p < 0.05$—but no interaction. We conclude from this that if a guest word is pronounced very differently as a code-switch and as a borrowing, then the choice of the early candidates will be greatly affected. In the case of the code-switch, listeners will propose candidates from the guest lexicon and will then narrow-in on the appropriate candidate in that lexicon. In the case of the borrowing, however, listeners will first start with candidates from the base language lexicon and then revert to candidates from the guest lexicon when they realize that no base-language words correspond to the sequence of sounds being heard.

The language of erroneous candidates

In our examination of erroneous candidates, it appeared to us that the language in which a guest word was pronounced affected not only the language of the first candidates proposed in the gating sequence (especially when the word was pronounced very differently in the two languages), but also the language of the erroneous candidates all the way up to the isolation point. If this proved to be correct, the language phonetics of a word would join the phonotactic configuration of the word and its lexicon membership in accounting for the language of the erroneous candidates. In order to test this, we calculated for every word the percentage of erroneous candidates that were English. To do this we counted the number of candidate tokens (minus the correct stimulus

TABLE 10.2 Percentage of erroneous candidates that are English when each word type is presented as a code-switch (English) or as a borrowing (French)

Language phonetics of word	Type of word		
	1	2	3
English (code-switch)	51.75	45.38	27.75
French (borrowing)	44.63	31.25	14.75

word candidate tokens) that were proposed up to the last gate of the word (the gate just prior to the first post-offset gate), and calculated the percentage that were English words. (When we had doubts regarding the language identity of a word, as when it is spelled similarly in the two languages, we consulted the language identification answers given by the participants.)

Table 10.2 presents the results averaged over word exemplars. As can be seen, both the type and the language phonetics of the word appear to play a role in the proportion of guest-language candidates proposed during the isolation process. The larger percentage of English candidates is obtained with Type 1 words where both the phonotactics and the single lexicon membership are clear indications that an English word is being presented; an intermediate percentage is obtained with Type 2 words where only the absence of the words in the French lexicon is an indication that they are English words; and the lowest percentage is obtained with Type 3 words where the presence of a base-language homophone leads the listener down a base-language garden path. In addition, in each of the three cases, the percentage of English candidates is larger for the code-switches than for the borrowings, clearly indicating that the language phonetics of the word plays a role in the candidates proposed. An analysis of variance over participants confirms these results. A main effect was found for word type—$F(2,21) = 5.39$, $p < 0.01$—and for word phonetics—$F(1,21) = 4.34$, $p < 0.05$—and there was no interaction.

We conclude from this that, in the early stages of word isolation, the two variables we have been studying (word type and language phonetics) both play a role. It is only in the later stages—the actual isolation of the stimulus word and the total acceptance of the word—that word phonetics loses some of its impact (although not for Type 3 words); as for word type, it continues to play a role throughout the isolation/recognition process.

Type 3 erroneous candidates in the post-offset presentations

Unlike Type 1 and Type 2 words, which were practically all isolated before their acoustic–phonetic offset, as many as 37 percent of Type 3 words were isolated after offset. Given this result, it is interesting to examine the narrowing-in process of Type 3 words when the next word or words were given along with the stimulus word. In what follows, we will examine the erroneous candidates proposed in place of Type 3 stimulus words when the latter were presented in offset positions +1, +2, and +3 words. We classified the 114 erroneous post-offset candidates into one of three categories:

1. Phoneme error candidates, that is candidates which differed from the stimulus words by the addition, omission, or substitution of one or more phonemes (from the same or the other language). For example: "sit"(cs) → "set"; "fool"(cs) → "fourre"; "cool"(cs) → "coure".
2. Segmentation error candidates, that is bisyllabic candidates that blend information from the stimulus word and the word following it. For example: "lease(cs) ce" → "listen"; "knot(cs) ces" → "answer".
3. Homophone candidates, that is candidates that are base language homophones. For example: "knot"(cs) → "note"; "lease"(cs) → "lisse"; "pick"(b) → "pique"; "sit"(b) → "cite".

Table 10.3 presents the percentage of borrowing and code-switching candidates that fall into each of the three categories (note that these results are based on 112 errors—not 114—because two erroneous

Table 10.3 Percentage of types of errors made for Type 3 words in the three offset presentations as a function of the language phonetics of the guest word (English or French)

Type of error	Language phonetics of word	
	French (borrowing)	English (code-switch)
Phoneme	6	71
Segmentation	0	13
Homophone	94	16

candidates could not be classified). As can be seen, the types of errors made when listening to borrowings and code-switches in the post-offset presentations are very different. This is confirmed by a highly significant Pearson's Chi Square for proportion computed on raw frequencies—$\chi^2 = 71.7$, $p < 0.01$. As expected, the candidates proposed for a borrowing are primarily base-language homophones. A full 94 percent are the French counterpart of English words: "pile" for "peel"(b), "coule" for "cool"(b), "lisse" for "lease"(b), etc. As for the remaining candidates, 6 percent fall into the phoneme error category and none are in the segmentation error category. From this we infer that if an English cross-language homophone is borrowed into French, and thereby becomes identical to an existing French counterpart, then the listener will assume it is the French word (and will feel quite confident about it). It is only when contradictory semantic information is heard (the last noun in the sentence, in our case) that the listener will be forced to backtrack and to access the English counterpart.

The candidates proposed for Type 3 code-switches reflect a more complex narrowing-in process. Whereas listeners go down the homophone garden path quite systematically with borrowings (and are only "shocked" out of it with later occurring top-down information), such is not the case with code-switches. Only 16 percent of the Type 3 code-switch candidates are base-language homophones; the remaining 84 percent fall into the phoneme error category (71 percent) and the segmentation error category (13 percent). We hypothesize that fewer homophones are proposed with code-switches because the English phonetics of the code-switch enters into (momentary) conflict with the internal representation of the French homophone. (On hearing a Type 3 code-switch, a subject commented that the speaker had an "odd" pronunciation in French!) Given that the homophone garden path is partly closed (but not the French lexicon path—89 percent of all post-offset code-switch candidates are French words), the listener proposes candidates that are phonetically similar (phoneme error candidates) or even candidates that combine information from the stimulus word and the following word (segmentation error candidates). Phoneme substitutions in two different words account for practically all of the phoneme error candidates: in the first, subjects proposed "coure" and not "coule", when presented with "cool"(cs); in the second, they proposed "fourre" and not "foule", when given "fool" (cs). The acoustic–phonetic information of the code-switch and the internal phonological representations of several possible candidates ("cool", "coule", "coure"; "fool", "foule", "fourre") interacted in such a way that neither the stimulus word, nor the French homophone were proposed; it was, instead, a third "compromise" candidate that emerged. (Note that both words were perceived as homophones when presented as borrowings.)

We should stress, finally, that Type 3 code-switches are the only items that fall into the segmentation error category ["knot(cs) ces" → answer; "lease(cs) ce" → "listen"]. These errors, although few in number, reflect once again the difficulties listeners had with words where the language phonetics signaled one lexicon but the base-language context and the presence of a base-language homophone signaled the other lexicon.

In summary, the analysis of the word isolation process has confirmed the importance of the two variables under study. The type of word that is presented affects how early guest-language candidates are proposed and how many there are as compared to base language candidates. As for the language phonetics of the word, it too plays an important role in the early stages of the recognition process: there are proportionally more English candidates when the word is said as a code-switch than as a borrowing, and the candidates of Type 3 words differ greatly in nature depending on their language phonetics. The variable takes on even more importance when guest words are pronounced very differently in the two languages, as we saw with Type 2 initial /t/ and /l/ words. As we noted, the importance of the language phonetics variable is at its highest during the narrowing-in stage; its importance diminishes (at least for Type 1 and Type 2 words) in the isolation and total acceptance stages. Finally, additional evidence was found for the base-language effect: when listeners hear the very beginning of a guest word presented in context, they propose base-language candidates in preference to guest language candidates.

10.3 Elements of a model of guest-word recognition

The results of numerous experiments, and the outcome of much theorizing in current psycholinguistic research, have substantially increased our knowledge of how spoken words are recognized during the online processing of speech. However, this work has been conducted mainly with monolinguals, and thus little is known about how bilinguals recognize spoken words in real time, especially when they are in a bilingual speech mode. A model that can account for mixed language word recognition has yet to be developed, but in what follows we will point to a few of the main features it could have.

The model will have to account for the general effects that have been found in studies of word recognition in monolinguals as well as the effects that are specific to bilingual language processing. Among the general effects we find the following:

1. Low-frequency words take more time to recognize than high-frequency words (Foss 1969; Howes 1957; Rubenstein and Pollack 1963).
2. Words are not always recognized from left to right, from onset to offset (Nooteboom 1981; Salasoo and Pisoni 1985).
3. When words *are* recognized from onset to offset, recognition occurs close to the word's uniqueness point, that is the point in the left to right phonotactic configuration at which the word diverges from other words (Marslen-Wilson 1984).
4. Words in continuous speech are not always recognized one word at a time, that is two words can be recognized simultaneously, or a later occurring word can be recognized before an earlier occurring word (Grosjean 1985a; McClelland and Elman 1986).
5. The syntactic, semantic, and pragmatic contexts of the sentence in which a word occurs affect its recognition (Grosjean 1980; Miller and Isard 1963; Morton and Long 1976; Tyler and Wessels 1983).
6. Various sources of knowledge, such as the listener's knowledge of the world and the rules of the language, also affect the word's recognition (Cole and Jakimik 1978; Marslen-Wilson and Welsh 1978).

It should be noted that existing models of word recognition, such as those of Forster (1976), Marslen-Wilson and Welsh (1978), Morton (1969), and McClelland and Elman (1986) account for a number of these effects, but none accounts for all of them.

In addition to these general effects, the model for bilinguals will need to capture a number of effects found in this study:

1. *The base-language effect.* When a guest word is presented in a base-language context, and only its very beginning has been heard, the candidates proposed are invariably members of the base-language lexicon.

2. *The phonotactic effect*. Words marked phonotactically as belonging to the guest language only (Type 1 words) are recognized sooner and with more ease than words not marked in this way.
3. *The single lexicon effect*. Words that belong solely to the guest lexicon (Type 1 and Type 2 words) are recognized sooner and with more ease than words that do not belong to just one lexicon.
4. *The base-language homophone effect*. Words in the guest-language lexicon that have close homophones in the base language (Type 3 words) are processed with more difficulty than other guest-language words.
5. *The language phonetics effect.* (a) During the narrowing-in stage preceding the isolation of a word, the proportion of guest-language candidates is affected by the language phonetics of the word (i.e. the language it is pronounced in). (b) Strong language phonetic cues will activate the lexicon that contains words characterized by these cues, and thus affect the language of the candidates proposed and, at times, the final isolation point of the appropriate candidates (as with those with cross-language homophones). (c) Cross-language homophones pronounced in the base language (Type 3 borrowings) are isolated later than when they are pronounced in the guest language (Type 3 code-switches), and the nature of the candidates prior to isolation are quite different for the two versions of the words.
6. *The frequency effect for cross-language homophones*. The ease with which a guest language homophone is identified depends on the "frequency pull" of that word as compared to that of its base-language homophone.

Although most existing models of monolingual spoken word recognition could be extended to account for word recognition during mixed speech processing, the type of model that may have the most promise is an interactive activation model, such as the TRACE model proposed by McClelland and Elman (1986). According to this model, language processing takes place through the excitatory and inhibitory interactions of a large number of processing units, each working continuously to update its own activation on the basis of the activations of other units to which it is connected. In TRACE, the units are organized into three levels: features, phonemes, and words. Throughout the course of processing, each unit is continually receiving input from other units, continually updating its activation on the basis of these inputs and, if it is over threshold, it is continually sending excitatory and inhibitory signals to other units. Connections between levels are bidirectional and there is no between-level inhibition (inhibition only exists within one level, between units that are inconsistent with one another). Although neither word frequency nor context effects are at present accounted for by the model, these can be built in quite easily, according to the authors: word frequency can be accommodated in terms of variation in the resting activation level of word units, and contextual influences can be thought of as supplying activation to word units from even higher levels of processing.

How could an interactive activation view of word recognition be modified in order to accommodate word processing in bilinguals, be it in a monolingual or a bilingual speech mode? First, we will assume that bilinguals have two language networks (features, phonemes, syllables, words, etc.) which are both independent and interconnected. They are independent in the sense that they allow a bilingual to speak just one language; they are interconnected in the sense that the monolingual speech of bilinguals often shows the active interference of the other language and that, when bilinguals speak to other bilinguals, they can code-switch and borrow quite readily. This view has long been defended by Paradis (1981, 1986), 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 Paradis, 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.[4]

[4] It should be noted that this proposal does not address head-on the question of whether the bilingual has one or two lexicons. The reason is that there probably exist as many experimental studies that find evidence for the one-lexicon view as studies that defend the two-lexicons hypothesis (Grosjean 1982). Unfortunately, however, these studies have often confounded the basic question (one versus two lexicons) with the task employed to examine the question; thus, many of the results obtained have reflected the experimental paradigm and not the underlying reality. In addition, the types of bilinguals used as subjects have varied from one study to another, making any definite statement problematic. The mixed model proposed by Paradis (1981, 1986) is thus not only intellectually appealing but also a nice compromise.

Other assumptions that can be made are the following. In the monolingual speech mode, one language network is strongly activated while the other is activated very weakly; the resting activation level of the units of this other network is therefore very low. In the bilingual speech mode:

1. Both networks are activated but the base-language network is more strongly activated (this accounts for the base-language effect). The resting activation level of the language not being used as the base language (the guest-language network) can be increased or decreased depending on the amount of mixed language (code-switching, borrowing) that occurs during the interaction.
2. The activation of a unit in one network and of its "counterpart" in the other depends on their degree of similarity. Thus, for example, if English /b/ is activated, French /b/ will also be activated (to some extent, at least) as the two consonants are quite similar. On the other hand, the activation of English word initial /p/ will lead to a much lower level of activation of French word initial /p/, as the two consonants are quite different. And when English /r/ is activated, its French counterpart should receive very little activation (apart from some possible top-down lexicon activation due to the fact that the two sounds have the same orthography). Cross-language activation of "counterpart" units concerns phonemes (as we have just seen), but also all other types of units (features, words, etc.).
3. The activation of units (or of a combination of units, such as consonant clusters) that are specific to one language increases the overall activation of that language network and thus speeds up the recognition of the words of that language (this accounts for the phonotactic effect and the language phonetics effect).
4. The activation of a word that is specific to just one language increases the overall activation of that network and thus speeds up the recognition of the words of that language (this accounts for the single lexicon effect).
5. The activation of words that are similar in the two lexicons will normally slow down the recognition of the guest-language word (this explains the cross-language homophone effect). But the frequency pull of the two homophones (reflected in their different resting activation levels), and the language phonetics of

the input, will interact with the recognition process of the guest word to speed up or slow down the access of that word (this accounts for the Type 3 word frequency and language phonetics effects).[5]

Much work needs to be done to refine this interactive activation view of the recognition of words in bilinguals. In particular, we need to think about which connections—between and within language networks—are inhibitory and which are excitatory. As we learn more about such models in general, and as more experiments on bilingual mixed speech are conducted, changes will be brought to the model. What is encouraging at this point though is that such a view does away with the switch or monitor mechanism that has been proposed by a number of researchers (Macnamara 1967; Obler and Albert 1978; Penfield 1959) and discussed by others (Grosjean and Soares 1986; Paradis 1980). According to proponents of the switch or monitor mechanism, its role is to tell the processing system which language is being spoken so as to direct the incoming signal to the processors of the appropriate language. The evidence for this mechanism is mainly based on studies which have shown that it takes bilinguals more time to process mixed speech than monolingual speech. But this evidence is both insubstantial and indirect. It is not because bilinguals may process code-switches more slowly than base-language words that one can conclude that there is a language switch/monitor involved in the processing; the delay could be due to numerous other factors (see Grosjean and Soares 1986). In addition, the proponents of the mechanism do not address pertinent questions such as: Is the switch/monitor an essential part of language

[5] An anonymous reviewer asks whether the results reported in the study cannot help distinguish between two models of bilingual lexical organization and access—on the one hand, two distinct lexicons, one of which is searched before the other; on the other hand, only one lexicon which contains a number of different acoustic features detectors which are abstract enough to discriminate between many of the different allophones that are distinct in the two languages. According to the reviewer, an interaction between the structural variable (word type) and the output variable (language phonetics) could be interpreted as evidence for a single system. Unfortunately, the data cannot help choose between these two views because a significant interaction was found in the isolation point results (see Figure 10.1 and the statistics that pertain to it), but none was observed in the confidence rating data (results that were obtained in the study but that are not reported here for lack of space). The reviewer asks to what extent we need to postulate truly distinct networks. We should point out that the view that we have adopted (based on Paradis 1981, 1986) does not defend the independence position; rather it proposes a mixed model—“separate” networks of connections which belong to a single extended system.

processing or does it "fall out" of the processing? If the former, at what stage does it come in—during the acoustic to phonetic mapping of the speech sounds or after this mapping? The data and the model we have presented do not prove the absence of a language switch or monitor, they simply show that the processing system can do without it, and that language decisions (e.g. was that word English or French?) can simply emerge from the process. Having heard a particular sound, syllable, or word, we can then make the metalinguistic statement that Language X or Language Y is being spoken. That the system needs to make this decision in order to process the incoming signal is highly unlikely.

We have shown in this study that the recognition of guest words is a highly complex process. Only further research using different paradigms, materials, and bilinguals with different pairs of languages will allow us to assess the validity of what we have proposed. The challenge for the psycholinguist interested in bilingual language processing will remain, for many years to come, to understand how processing in mixed language takes place so rapidly and so efficiently despite, as we have seen, many intricate underlying operations.