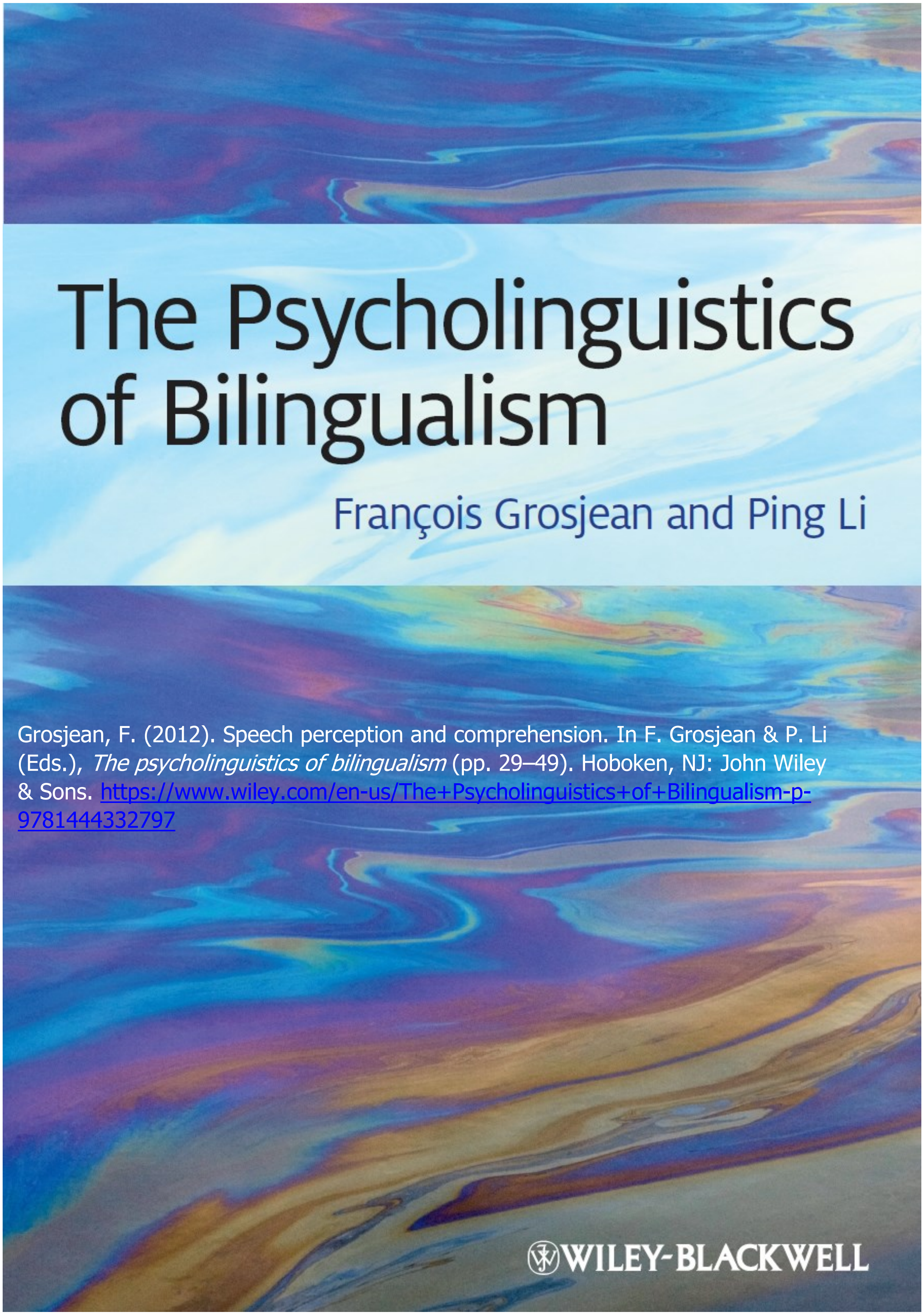

Grosjean, F. (2012). Speech perception and comprehension. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 29–49). Hoboken, NJ: John Wiley & Sons. https://www.wiley.com/en-us/The+Psycholinguistics+of+Bilingualism-p-9781444332797

Chapter 2

Speech Perception and Comprehension

François Grosjean

Even though there are many more bilinguals in the world who listen to two or more languages than read them (recall that many bilinguals may not know how to read and write one of their languages), the studies pertaining to bilingual speech perception and comprehension have been less numerous than those pertaining to reading (see Chapter 4). The situation is changing, however, and it is now possible to give an overview of how bilinguals perceive and comprehend their different languages.

This chapter has several aims. The first is to describe how bilinguals process just one language when there are no elements of the other language(s) in it – we call this "monolingual speech." Bilinguals have to deal with it in their everyday lives when they are in a monolingual mode (see Chapter 1), and it is important that we understand how they do so.

The second aim is to describe how bilinguals process bilingual speech, that is, speech that contains code-switches and/or borrowings. As we saw in Chapter 1, bilinguals in a bilingual mode often bring in the other language, and it is worth asking how listeners process both the base language – the main language being heard – as well as the guest elements of the other language.

A third aim will be to introduce readers to the methodology used to study oral language perception and comprehension in bilinguals. Unlike with speech production, where we can learn a lot about the underlying production mechanisms involved from the speech produced, with speech perception and comprehension this is not possible. We therefore have to resort to experimental paradigms to "open a window" into the mind of the bilingual listener.

This chapter contains three sections. In the first, we offer a rapid overview of how speech perception and comprehension takes place. We will end the section with a short discussion of the ways in which bilinguals are invariably different from monolinguals when doing this type of processing. In the second section, we will examine two aspects related to how bilinguals process monolingual speech. The first relates to the activation of the other language, that is, the language not being heard. In

The Psycholinguistics of Bilingualism, First Edition. François Grosjean and Ping Li.

essence, does that other language play a role and, if so, when and how? We will see that the selective/nonselective processing issue already discussed in Chapter 1, Section 1.4 is far more subtle than was thought at first and that the answer is less categorical than many have believed. The second aspect will concern the more permanent influence of the other language on processing. This pertains primarily to how the dominant language may influence the processing of the nondominant (weaker) language. Finally, the third section of this chapter will examine how speech that contains code-switches and/or borrowings is processed by bilinguals. Much less is known about this but what we do know is fascinating and well worth a description.

2.1 From the Speech Wave to the Mental Representation

The processing of spoken language – something that we do every minute of the day without being aware of it – is a highly complex, but also very rapid, process that we now understand reasonably well thanks to the research of psycholinguists over the years. For it to take place smoothly, one needs a number of components to be present, as shown in Figure 2.1. First, we have to have a speech wave (called the "speech input" in the diagram) produced by a speaker. It is referred to as "bottom-up" information, hence the upward arrow above the speech wave in the figure. We also need to know the language in question (we have called this "linguistic knowledge," on the right of the main box) and we have to have available a certain number

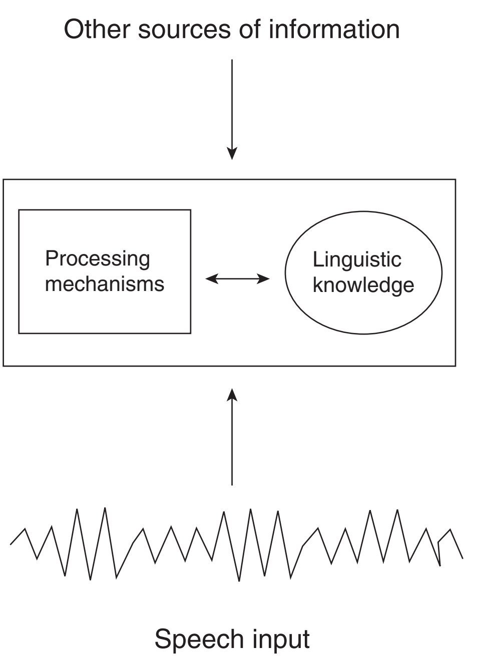

Figure 2.1: The basic components involved in speech perception and comprehension.

of processing mechanisms (see the left area of the box). Finally, speech perception and comprehension is helped greatly if we have other sources of information such as the context in which the speech situation is taking place, knowledge about the world we live in, information about what has been said so far, and so on. Note that these sources of information are often called "top-down" information (see the downward arrow in the figure). The final outcome of the work done by the processing mechanisms, with the contribution of linguistic knowledge and the other sources of information present, is called the mental (or interpretative) representation, that is, the enriched meaning of what has been said.

Although all the components of Figure 2.1 play a vital role during comprehension, we will concentrate on the processing mechanisms. To illustrate how processing takes place, let's follow a very simple spoken utterance, "The window is closed," from the moment it is said to the moment it is fully understood. As soon as the speech wave that corresponds to the beginning of the utterance is heard, phonetic processing starts to take place. That is, the phonetic and prosodic units that comprise it are identified. Thus, we perceive the two phonemes that correspond to the first word "the," the five phonemes that correspond to "window," and so on. We also perceive that the intonation is falling, that the speech rate is normal, and that there is a slight emphasis on the word "closed."

As soon as the first sounds have started to be perceived, lexical processing begins, that is, the recognition of the words "the," "window," "is," and "closed." They are activated in our internal lexicon (also known as our mental dictionary), along with other words that have similar beginnings, and little by little we isolate them and finally "accept" them. By doing so, we gain access to the information about these words that is contained in our lexicon – their meaning or meanings (if there are several), their grammatical categories, the syntactic and semantic structures that they can occur in, and so on. Along with this, we undertake morphological processing on the words that need it, for example computing that "is" refers to the third person singular of the verb "to be." With the information obtained at this processing level, we can undertake syntactic processing, that is, identifying the constituents of the utterance (e.g., "the window" is a noun phrase, "is closed" is a verb phrase) as well as the way in which the constituents are structured together.

While this is taking place, the next levels of processing are already at work, using the information that has been obtained so far. In semantic processing, the literal mental representation is computed, that is, the basic meaning is obtained. Some believe it takes the form of propositions that contain predicates and arguments. In this case, the mental representation of our utterance would look something like "closed (window)," meaning that the window is closed. Finally, pragmatic processing will use the context in which the utterance was said, as well as our knowledge about the world and of the rules of communication, to help reach a conclusion about meaning. The literal meaning may then be modified so as to produce an enriched final representation. Let us assume here that the utterance was said by someone who entered a room that was overheated (this is part of the context of the utterance). Let us also assume that it was cold outside, colder than in the room. We all know through our knowledge of the world that a way of cooling a room is to open a window IF it is colder outside than inside. Finally, we know through our internalized rules of communication that we can ask for something with an indirect speech act such as a statement. The final enriched mental representation will therefore be of the type, "She's asking me to open the window."

Two remarks need to be made at this point. First, we have greatly simplified the working of the processing stages that take place; each one is highly complex and the object of specific research in the cognitive and brain sciences. Second, even though we have talked of the processes taking place one after the other, it has been known since the pioneering work of Marslen-Wilson (1975) that processing is "parallel" in that all processes occur quasi-simultaneously. In addition, they take place "on-line" (as the utterance is being uttered, from its beginning to its end), and many researchers agree that they are also "interactive," that is, information is exchanged between processes to accomplish the task that is theirs.

A few general comments can now be made about processing in bilinguals before we examine specific issues in the rest of this chapter.

Whatever the type of speech being processed (monolingual or bilingual), the basic processing components and levels presented above are found in bilinguals as well as monolinguals. Thus bilinguals analyze the speech input with the same processing mechanisms (phonetic, lexical, syntactic, semantic, and pragmatic), their linguistic knowledge is called upon during processing, and other sources of information such as context, knowledge of the world, and information about what has been said so far play an important role in building the enriched mental representation. In addition, speech processing in bilinguals takes place in parallel, on-line, and most probably in an interactive fashion.

Despite these similarities with monolinguals, however, there are also important differences. First, bilinguals process not just one language but two or more. Thus the central components depicted in Figure 2.1 will be multiplied, at least in part, by the number of languages involved (see Sections 2.2 and 2.3). Second, since bilinguals are rarely equally fluent in all of their languages (recall that they use them for different purposes, in different domains of life, and with different people), the linguistic knowledge they have of their languages will be different, and this will have an impact on speech perception and comprehension. Some languages may be less well processed than others and, within a language, missing linguistic elements such as vocabulary items may affect perception and comprehension. This is clearly exemplified when bilinguals are asked to perceive speech in adverse conditions such as when there is background noise; they practically always do less well than their monolingual counterparts (see Lecumberri, Cooke, & Cutler, 2010, for a review).

Third, depending on the linguistic make-up of each language, the processing mechanisms called upon may function differently, at least in part. For example, at the level of phonetic processing, the perception of tones may be crucial for one language but not for the other; during lexical processing, there may be complex morphological analyses in one language but simpler analyses in the other; and at the level of syntax, specific syntactic processes may occur in one language but not in the other, and so on.

Fourth, as has been mentioned already, the utterance that has to be processed may be "monolingual" in that all the elements in it originate in just one language or be "bilingual" in that the interlocutor is in a bilingual mode and is using code-switches and borrowings. These will be contained in the speech stream and will have to be processed by the bilingual listener.

Finally, even when only one language is being processed (the signal is "monolingual"), the processing mechanisms called upon may be influenced, momentarily or in a more permanent way, by those of the other language(s). This can be due to simple coactivation of the other language(s) or by processing mechanisms and strategies that impose themselves on the language being processed.

In the following sections, we will concentrate primarily on these last two points. By doing so, we have simplified things slightly and so we should keep in mind that the overall picture that will one day emerge concerning speech perception and comprehension in bilinguals will be more complex than what is proposed here.

2.2 Processing Monolingual Speech

Figure 2.2 presents the processing components involved when the listener is bilingual. To simplify things, we represent just two languages, La and Lb, but we should

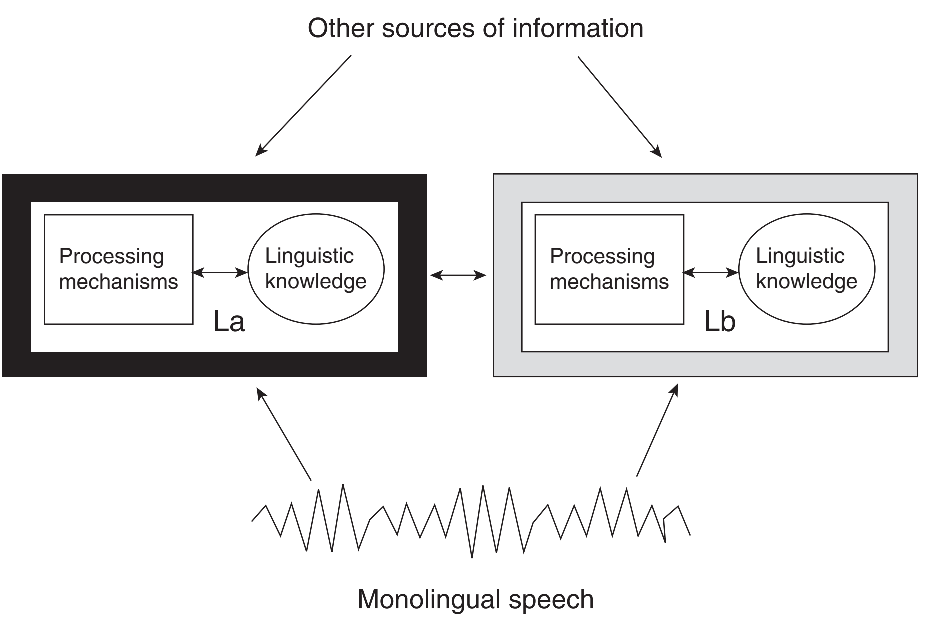

Figure 2.2: The components involved in speech perception and comprehension in bilinguals when the input is monolingual.

keep in mind that more languages may be known. In the figure, the person is listening to speech input that is monolingual (the language being heard is La) and hence only the linguistic knowledge and the processing mechanisms of that language are active (see the black rim around the left-hand box). The other language (Lb) is also present but it is deactivated, hence the grey rim. It should be noted that the monolingual speech input is transmitted to both language systems and that the other sources of information also feed into both systems.

2.2.1 The activation of the other language

The first issue we will discuss relates to the activation of the other language (Lb in our diagram), that is, the language not being heard. Even though we show it deactivated in Figure 2.2, many researchers believe that it is also active and that, in essence, it plays a role when the input is monolingual and in the other language. Hence, to the question of whether processing in bilinguals is selective (when one language is heard, only that language is active and is processed) or nonselective (the bilingual's different languages intervene while processing takes place), many researchers have opted for the latter answer in recent years. However, we will show that things are more complex than that.

Evidence for the simultaneous activation of the two languages, and hence nonselective processing, even in a monolingual mode, has come from studies using written language (see Chapter 4), and in more recent years from speech studies that have used an eye-tracking technique. The first one of its kind with bilinguals was conducted by Spivey and Marian (1999) with Russian-English bilinguals. They used a head-mounted eye tracker which allows the experimenter to see where the participant is looking while speech comprehension is taking place. We will describe the Russian part of their study here.

Their bilingual participants were asked to look at a board situated in front of them containing a number of objects. The example in Figure 2.3 is based on the description the authors give. Note that the target object, a stamp, which has to be moved by the participant, is in the bottom right-hand square. In the top left-hand square there can be one of two objects: a competitor object (in this case, a marker) or a control object (a ruler). There are also filler objects in the other two corner squares.

In the Spivey and Marian study, the participants heard pre-recorded instructions which asked them to displace the target object on the board, such as, "Poloji marku nije krestika" (Put the stamp below the cross). In the interlingual competitor condition, some of the objects on the board had English names that shared initial phonetic characteristics with the onset of the name of the Russian target object. Thus, in the example presented on the board, when the target object was a stamp ("marku" in the above sentence) the interlingual competitor object was a marker, an object whose English name shares the same word beginning as "marku."

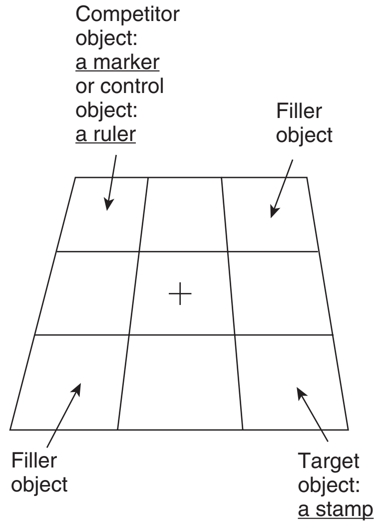

Figure 2.3: A visual representation of the board used by Spivey and Marian (1999) based on the description they give.

The researchers examined the eye movements made to this interlingual competitor object as compared to a control object, in exactly the same position, such as a ruler (see the top left-hand corner of the figure). In this second condition, the object's name bore no phonetic similarity with the name of the target object ("marku"). The results obtained showed that the participants made significantly more eye movements to the interlingual competitor object (32%) than to the control object (7%). Why is that? It would seem that the word onset of the target object (e.g., "marku") not only activated Russian words in the Russian lexicon but also English words in the English lexicon that began in a similar way ("marker" is very similar to "marku"). This happened through bottom-up processing, that is, the processing of the speech input (see the upward arrows leading from the speech wave to both languages in Figure 2.2). Based on this, the authors concluded that processing is nonselective.

Later studies were conducted by other researchers which also used the eye tracking approach but where the board was on a computer screen. Many reported similar findings. Thus, for example, Ju and Luce (2004), using Spanish-English bilinguals, showed that even a subtle phonetic cue in the other language was enough to activate both lexicons. To do this, they changed a critical aspect of the first consonant of Spanish words to its English counterpart by means of sound editing. More precisely, they manipulated the Voice Onset Time (VOT) of the consonant, that is, the brief delay between the release burst and glottal pulsing. Thus, for example, the Spanish /p/ of the word "playa" (beach) was in essence replaced with the English /p/ sound. This was enough to attract eye movements to the interlingual competitor object (a picture of "pliers") when the participants were asked in Spanish to click on the picture that corresponded to the target word ("playa" said with the English /p/ sound).

Bottom-up information can therefore sometimes activate both the language being spoken and the bilingual's other language. But can the activation of the nonused language, at other times, also be reduced and even neutralized? The answer is "yes," as we will see from two studies. In the first, Chambers and Cooke (2009), also using the eye-tracking technique but working with English-French bilinguals this time, preceded the target words (e.g., "poule" (chicken)) with nonrestrictive and restrictive sentences. In the former case, such as in "Marie va décrire la poule" (Marie will describe the chicken), there was very little prior semantic constraint on the target word (here "poule") since the verb "décrire (describe)" can be followed by any number of nouns (one can describe many things in life!). But in the restrictive sentence case (e.g., "Marie va nourrir la poule" (Marie will feed the chicken)), the predicate constrained the noun (in this example, to animate objects that can be fed). The competitor object was the picture of an interlingual homophone (the picture of a "pool" in our example). The participants listened to the French sentences and dragged the target object on the screen to the middle square. What the researchers found was that consideration of the interlingual competitor object was greatly reduced when the context sentence was restrictive: there were on average 0.28 eye movements (saccades) in this context compared to 0.48 saccades in the nonrestrictive context. Why was the number not reduced to zero? Quite simply because homophones were used in the study and they were activating both the French lexicon and the English lexicon in a bottom-up manner.

Finally, Marian and Spivey (2003) conducted a further study that shows that the activation of the other language cannot just be reduced but neutralized in certain situations, resulting in selective processing in essence. They realized that the contextual factors in their first study had pushed their participants toward a bilingual mode of processing, thereby activating both languages, and hence encouraging nonselective processing. Among the factors they mentioned was that participants knew they were taking part in an experiment on bilingualism, they were tested by bilingual experimenters who were fluent in both languages, and the two languages were tested in adjacent experimental sessions. (To these factors we can add several others presented in Chapter 1, Section 1.4). So as to put their participants in as close to a monolingual mode as possible, Marian and Spivey used different experimenters for the Russian and the English sessions who posed as monolingual speakers. (Note that what follows concerns the Russian session once again). During testing, they used only the language of the session, and participants only took part in the one or the other session. The results they obtained were quite convincing. The participants looked only at interlingual English competitor objects in 8% of the trials as opposed to 5% for the control object, a nonsignificant difference. (Recall that in their first study, the percentages had been 32% and 7%, respectively.) Hence, in this case, the other language had been totally "closed out" and processing had now become selective.

What can we conclude, therefore, about bilingual language processing and, more specifically, about whether processing is selective or nonselective? The answer is quite simply that it depends. The bottom-up, phonetic, information that is heard is processed by the language(s) that contain(s) elements of that input and this can lead to nonselective processing, such as when words are used that have similar word beginnings in the other language, or when homophones, homographs, and cognates are involved (as studies have shown repeatedly). Of course, if the input only contains elements of one language, then only one language will process it. The activation of just one language may be reinforced, in a spoken utterance, by aspects that are specific to that language: certain sounds, word forms (see Vitevitch, 2012, who found that English and Spanish share very few phonological neighbors), as well as prosodic information such as intonation, stress, rhythm, and so on.

To this should be added top-down factors ("other sources of information" in Figure 2.2) such as the interlocutor and the context which, in the case of bilinguals, influence the language mode the interlocutors are in (see Chapter 1, Section 1.4). Sometimes this top-down information may even "contradict" the bottom-up information, as when, for example, the listener is "shocked" upon hearing the speaker say something in a language that is not expected. We should add that things are further complicated by the fluency bilinguals have in their different languages, among other things. It would appear that if the stronger language is being processed, the weaker language will not be activated as much (or at all), and hence will not "interfere" with processing. However, if it is the weaker language that is being processed, then the stronger language may be active and may influence the processing that is taking place (see, for example, Weber and Cutler, 2004, who have shown this to be the case). Extreme care must therefore be taken when talking about whether processing of monolingual speech in bilinguals is selective or nonselective, since both outcomes are possible.

2.2.2 More permanent influences of the other language

In addition to the possibility that the language not being processed can be activated from time to time and hence can intervene in the processing of monolingual speech, as we have just seen, there may be more permanent influences of one language on the other (depicted by the horizontal arrows between the two languages in Figure 2.2). This pertains primarily to how the dominant language may influence the processing of the nondominant (weaker) language. We will examine two studies in particular, among many others, that show this. The first will deal with the perception of sounds and the second with the processing of grammatical gender.

It has been known for some time that the sound categories of a stronger (dominant) language can influence the sound categories of a weaker (nondominant) language. Thus, if the stronger language only has one category and the weaker language has two, these two categories may be assimilated to just one category in the weaker language. For example, the two English categories /æ/, as in "sat," and /ɛ/, as in “set,” are often assimilated by Dutch-English bilinguals to one Dutch category /ɛ/, and so one often hears Dutch speakers of English not being able to differentiate between the pronunciation of “sat” and “set.”

Pallier, Colomé, and Sebastián-Gallés (2001) showed how this phenomenon can play a role in speech perception by examining how Spanish-Catalan bilinguals perceive certain Catalan words. Spanish has only five vowels whereas Catalan has eight, among which we find the vowel /ɛ/, which does not exist in Spanish. Thus, Catalan has minimal pairs such as /netə/, (granddaughter) and /nɛtə/, (clean, feminine) whereas Spanish does not; for a Spanish speaker, these two words are homophones. The experimenters used two groups of highly skilled bilinguals fluent in Spanish and Catalan who had comparable, if not fully equivalent, command of the two languages, at least at levels beyond the phonetic level. All had been born in Barcelona and all had received the same kind of bilingual education. The difference was that the Spanish-dominant bilinguals had been raised in a monolingual Spanish environment before attending kindergarten, whereas the Catalan-dominant bilinguals had been raised in a monolingual Catalan environment before kindergarten.

The experimenters used a repetition priming task in which participants were asked to perform a lexical decision (i.e., decide whether an item is a word or a nonword) on lists of stimuli, some of which appeared twice. They cleverly used the well-known repetition effect to see how their bilinguals perceived certain sounds. The effect is quite simply that the reaction time to make a lexical decision is faster when an identical item is presented a second time in an experiment. Would there be a repetition effect, in this case, when one item of a minimal pair in Catalan was heard (e.g., /netə/ (granddaughter)), and the other item of the pair (/nɛtə/ (clean, feminine)) was heard later? The results they obtained were quite clear. Catalan-dominant bilinguals did not show a repetition effect: they processed these words as being quite distinct and showed no facilitation for the “repetition,” that is, the other element of the pair. Spanish-dominant bilinguals, on the other hand, behaved as if the words were real homophones; they showed a repetition effect that was of the same amplitude as that observed for a real repetition. The authors concluded that despite being very fluent bilinguals, their Spanish-dominant bilinguals had not created separate perceptual categories for certain Catalan sounds due to the fact that their first language (Spanish) did not have them.

We have just seen how a language can have a permanent influence on the processing of another language. In the example given, we showed that certain categories are not developed in the weaker language and that processing is affected by this. It can also be the case that, in addition to categories not existing, some processing mechanisms and strategies are not acquired because the first language did not have them and/or the second language was acquired later. Our example comes from the domain of word recognition. Guillelmon and Grosjean (2001) were interested in how gender marking is processed by bilinguals. Depending on the language, words may carry any number of genders, from two in French and Italian all the way to six in Swahili. Other word classes may not have gender but they can reflect, in their morphology, the gender of the words that do. Thus a gender agreement marking can appear before or after a noun on a determiner, adjective, pronoun, and so on. For example, in the French phrase, "la petite voiture" (the small car), both the article (la) and the adjective (petite) agree with the feminine noun (voiture) and carry a feminine ending.

We have known for some time that a congruent gender marking (as in the example just given) on the words preceding the noun will speed up the noun's recognition whereas an incongruent marking (e.g., "*le petit fille") will slow it down. Guillelmon and Grosjean asked whether bilingual listeners were also sensitive to gender marking. They tested early and late English-French bilinguals. The early bilinguals started using their two languages in everyday life as early as 5;4 years on average whereas the late bilinguals had only become regular users of their two languages at 24;8 years. The latter had learned French at school but it was only as adults, mostly because of immigration, that they had become regular users of it.

Both groups of participants heard phrases of the type, "le/la/leur joli(e) + noun" (the$_{\text{masculine}}$ / the$_{\text{feminine}}$ / their nice + noun) and were asked to repeat the noun. Depending on the determiner, the gender marking was congruent (e.g., "le joli bateau" (the nice boat), incongruent (e.g., "*la jolie bateau") or not present (e.g., "leur joli bateau"). The time needed to repeat the noun was used to understand the underlying processes involved. The early bilinguals demonstrated strong congruency and incongruency effects. They had become sensitive to gender early in life and they appeared to use gender marking in perception the way monolinguals do.

The real surprise came when the results of the late bilinguals were analyzed. They were quite simply insensitive to both gender congruency and gender incongruency. It appeared that they simply couldn't use the masculine "le" cue or the feminine "la" cue during the processing of the phrase even though they themselves made very few gender errors in production. Thus the gender processing mechanism, which can help to speed up word recognition and which is acquired by native French speakers and early bilinguals, was never acquired by late bilinguals. There is probably, therefore, a sensitive period to acquire such a mechanism (see Chapter 7) and the late bilinguals, whose first language (English) does not have gender marking of this type, missed it and could not master it later on. Of course, they still recognize words (the late bilinguals reported having very good French oral comprehension) but their lexical access is not speeded up by a congruent gender marking on the preceding word(s). In sum, one language can definitely have a permanent influence on the language processing of the other.

2.3 Processing Bilingual Speech

As was described in Chapter 1, bilinguals may find themselves in a bilingual mode when they are speaking to people who share their languages and where code-switching and borrowing can take place. Thus, the speech input bilinguals have to process may well be bilingual. In Figure 2.4 we show how the two languages are configured when this happens. One language is the base language and is the most active (La in the figure); the other language is also active but less so as it is only

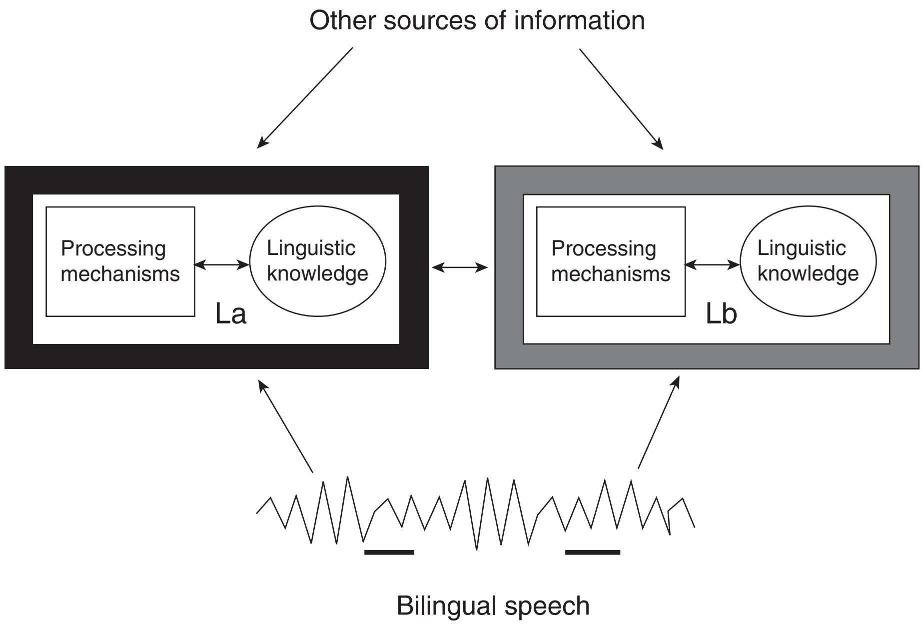

Figure 2.4: The components involved in speech perception and comprehension in bilinguals when the input is bilingual.

called upon from time to time. We have depicted this by making the border around the Lb box slightly less dark than that around La.

In what follows we will first look at the impact that the base language can have when guest words (code-switches and borrowings) are heard. Note that we have depicted them with short horizontal lines below the speech wave in the figure. We will then report on how code-switches and borrowings themselves are processed by bilingual listeners, and we will end with the description of a model which simulates many of the findings that have been obtained by researchers.

2.3.1 The base-language effect

The base-language effect concerns the impact that the base language (La in Figure 2.4) can have on the processing of guest words. Does it, for example, slow down their perception and recognition momentarily? As we will see in this section, there is now good evidence that this is the case. Since in normal bilingual discourse the base language normally makes up some 80–90% of the utterance, it is more strongly activated and hence base-language units (phonemes, syllables, words) are favored over guest-language units, at least for a short period of time.

Grosjean (2008; Chapter 6) describes four studies that all point toward the existence of a base-language effect. We will describe two of them here. Soares and

Grosjean (1984) asked English-Portuguese bilinguals to listen to sentences and to detect words or nonwords within them that began with a pre-specified phoneme. Once they had found the items, they had to make a lexical decision, that is, indicate as quickly as possible whether they were dealing with a word or a nonword. The bilinguals heard three types of sentences – English, Portuguese, and Portuguese with code-switches. Here are examples (the critical words on which a lexical decision had to be made are emboldened and the code-switches are in italics):

English: "After lunch, the children asked for a piece of **cake** for dessert."
Portuguese: "Depois do almoço os miudos pediram uma fatia de **bolo** para sobremesa."
Code-switched: "Depois do *lunch* os miudos pediram uma fatia de ***cake*** para *dessert*."

The results obtained were interesting. First, in the English condition, the bilinguals were as rapid as the English control participants. Second, there were no differences between their reaction times in the English and the Portuguese condition, indicating thereby that they were equally fluent in their two languages. Third, their mean reaction times to code-switches was significantly slower than to base language words, in the order of about 150 ms. Thus a base-language effect was clearly present in this study.

In another study, conducted by Domenighetti and Caldognetto (1999; described by Grosjean, 2008), the aims were, first, to confirm that the base language delays, however slightly, the recognition of code-switches in a neutral context and, second, to ascertain whether the word following the code-switch is also delayed. This second aim was important for our understanding of the code-switching process. If the delay is carried over to the next word, then the bilingual listener will gradually fall behind the speaker, which seems quite counterintuitive. One possibility is that the delay is momentary and is "caught up" before the code-switched word is finished and, hence, before the next word arrives.

The authors asked French-Italian bilinguals to listen to a short sentence which ended with a short list of words. For example, "J'ai entendu les mots aéroport, grenouille, sapin, collier" (I heard the words airport, frog, fir tree, collar). The participants' task was to repeat the word in the second position of the list, that is, "grenouille" in the example. In the code-switching version, the second word was replaced by an Italian word which took the same amount of time to repeat as the French word in isolation. Thus, in our example, "grenouille" was replaced with "cena" (dinner). The results showed a base-language effect once again; code-switches took on average 50 ms more to repeat than base language words. However, when a second group of participants were asked to repeat "sapin," the word in third position in both types of sentences, the repetition times were similar. Thus, the base-language effect appears to be short lived, and by the time the following word arrives in the sentence, any time that may have been lost will have been caught up. This may explain why certain studies that measure processing later on in sentences containing guest words, or at the end of the sentences, show no base-language effect (see, for example, Cheng & Howard, 2008).

2.3.2 The recognition of code-switches and borrowings

A topic that has been of interest for a number of years pertains to the factors that play a role in how guest words (code-switches and borrowings) are recognized in bilingual speech. Grosjean (2008) presents a list of these factors which he organized into four categories: those that pertain to the listener, such as how fluent the person is in the guest language; those that concern the level of activation of both the base language (La in Figure 2.4) and the guest language (Lb), and which include the base-language effect we have just discussed; those factors that involve various code-switching constraints (it is well known that code-switching is a rule-governed behavior which will probably have an impact on processing); and, finally, the factors that concern the properties of the guest word being heard. We will concentrate on this latter category in what follows since several studies have been conducted on this topic.

A first study Grosjean (1988) examined the role of a number of guest-word properties during word recognition. He presented French-English bilinguals with English guest words preceded by a French neutral context, "Il faudrait qu'on . . ." (We should) and followed by a final phrase in French. For example, "Il faudrait qu'on *slash* tous les prix" (We should slash all the prices) or "Il faudrait qu'on *lean* contre le mur" (We should lean against the wall). In each example, the guest word is in italics.

Grosjean used the gating task in which a spoken word is presented in segments of increasing duration. Thus, the first gate contained "Il faudrait qu'on" up to, but not including, the onset of the guest word. The second gate contained the same information plus the first 40 ms of the guest word. And from then on, gates were incremented by 40 ms until the end of the word was reached. Once the full word had been presented, three "after offset" gates were added covering the final phrase so that participants could hear the following context in order to resolve any remaining ambiguity concerning the guest word.

After each presentation the participants were asked to: (1) write down the word they thought was being presented after "Il faudrait qu'on"; (2) indicate how confident they were about their guess, and (3) indicate whether they thought the word was French or English.

The three word properties that were examined in the study were the language phonetics of the word, its phonotactics, and the interlanguage neighbor proximity. As concerns language phonetics, the author asked whether guest words which retain a phonetic cue as to which lexicon they belong (by being pronounced clearly in the guest language) are easier to process than words which are integrated phonetically into the base language. In other words, to use the terminology employed in

Chapter 1, would code-switched words, which are pronounced in the guest language and thereby retain phonetic cues as to the lexicon they are a part of, be accessed more easily than borrowings, which are usually integrated into the base language and hence have lost most, if not all, of their phonetic cues.

As concerns phonotactics, the permissible combination of phonemes and syllables in a language, the author asked whether guest words that are marked phonotactically as belonging to the guest language lexicon would be recognized sooner and with more ease than words not marked in this way. Thus, would English "snap" or "slash," where the initial consonant clusters, "sn" and "sl," are more frequent in English than in French, be perceived sooner than words such as "lead" or "pick," which have a consonant-vowel-consonant pattern common to both languages. The reasoning was that the initial consonant cluster would be a strong cue to the lexicon it belonged to.

Finally, as concerns interlanguage neighbor proximity, Grosjean asked whether guest words that have near homophones in the base language would be recognized with more difficulty than other guest language words (for instance, English "pick" is homophonic with French "pique," English "knot" with French "note," etc.).

A combination of these latter two variables – phonotactics and interlanguage neighbor proximity – gave three types of words: Type 1 words which favor English phonotactically and which only exist in English; Type 2 words which favor French phonotactically but which only exist in English; and Type 3 words which also favor French phonotactically but which have a close homophone in the other language.

The results confirmed the importance of the variables under study. First, words that were marked phonotactically as belonging to the guest language only (e.g., "slash," "blot") had a mean identification point that occurred sooner than for words not marked in this way. The identification point was expressed as a percentage of the word needed to identify it correctly. Percentages could range from 0% (no part of the word was needed) to 100% (the whole word was needed). Participants needed 66% of Type 1 words to identify them as opposed to 78% for Type 2 words. Second, words that belonged solely to the guest lexicon (Type 1 and 2 words) were recognized sooner than words that did not belong to just one lexicon (Type 3 words): 97% of Type 1 words and 92% of Type 2 words were identified before their ending whereas only 43% of Type 3 words fell into this category. Third, words in the guest language lexicon that had close homophones in the base language (Type 3 words) were processed with more difficulty than other guest language words: 37% of Type 3 words were isolated after their acoustic offset, but before the end of the sentence, and a full 20% were never identified at all.

As concerns the language phonetics variable, it appeared that the way a guest word was said (i.e., as a code-switch or as a borrowing) affected more the narrowing-in process that led to word identification than the actual point in time at which the word was identified. Grosjean found that during the selection phase which preceded word recognition the proportion of guest language candidates was greater for code-switches than for borrowings. He also noted an interesting interaction between the language phonetics variable and the interlanguage homophone variable. The candidates proposed for Type 3 words were quite different depending on whether they were said as borrowings (i.e., in French) or as code-switches (i.e., in English). In the former case, subjects invariably chose the base language homophone ("pique" for "pick," "note" for "knot") but, in the latter case, only about 20% fell into this category. The majority involved the addition, omission or substitution of one or more phonemes (e.g., "set" proposed for "sit," "fourre" for "fool," "coure" for "cool," etc.). This indicated the very real difficulties subjects had with items in which the language phonetics activated the two lexicons, primarily the English lexicon but also its French counterpart, and where the base-language effect reinforced the French candidates.

In addition to showing the importance of these three variables, the study examined two other variables – sound specificity and interlanguage neighbor frequency. As concerns the former, an analysis of the candidates that were proposed showed that strong language phonetic cues (such as those of a plosive or a lateral) clearly activated either the English or the French lexicon, depending on the phonetics of the guest word, and thus affected the language of the candidates proposed. Concerning interlanguage neighbor frequency, the author noted a great deal of variability in the results of Type 3 words (i.e., those that had close homophones in the other language). This was due to the degree of "frequency pull" of the guest words (i.e., the English items) as compared to their base language counterparts (the French words).

Later studies The Grosjean (1988) study was followed by other studies undertaken by different researchers and with different language pairs. Since each one contributed new information to our understanding of how guest words are processed, they are worth mentioning here. Li (1996) conducted two experiments with Chinese-English bilinguals: the first one was a gating study, similar to Grosjean's study, whereas the second one was a naming (word shadowing) study. Li examined the language phonetics of the English guest words as well as their phonotactics, as had Grosjean, but he also looked at the role of the prior context. He preceded the guest words with a short context (a bit like Grosjean's neutral context) and a long, more constraining, context, in order to see the role it would play in guest word recognition.

The gating results reported by Li were very interesting. First, he showed that guest words pronounced as code-switches provide phonetic cues to the listener and hence are easier to identify than when they are pronounced as borrowings. (It should be noted that this effect was far stronger in his study than in Grosjean's where its impact was situated more in the narrowing-in stage). Second, he found a strong effect for context: if the context was constraining, only 59% of the word was needed for correct identification as compared to 72% of the word if it was not (a result that has been obtained repeatedly with monolingual speech). Finally, he obtained a very interesting result concerning the phonotactics of the words. When the initial syllable was made up of a consonant and a vowel (CV), then code-switches and borrowings were identified at about the same point. However, when the initial syllable was made up of a consonant followed by another consonant (CC), then borrowings required far more information (79% of the word) than code-switches (55%). This comes from the fact that when words are borrowed into Chinese, their configuration is changed quite drastically with certain consonants softened or even dropped. Thus, for words such as "flight" where the code-switch version resembles the English pronunciation, the borrowing version becomes /faɪ/. It makes sense, therefore, that such words need much more information to be identified correctly. It should be noted, finally, that the naming (word shadowing) study provided converging evidence for the results obtained in the gating study concerning all three variables studied.

A few years later, Schulpen, Dijkstra, Schriefers, and Hasper (2003) examined the processing of homophones in Dutch-English bilinguals (the equivalent of Type 3 words in Grosjean's study), for example, Dutch "boel" and English "bull." In a first study, they presented these words in isolation (along with control words) using the gating task and the participants were asked, after each gate, to guess the word being presented, to rate how sure they were, and to rate their confidence that it was either a Dutch or an English word. They found, as could be expected based on prior studies, that homophones were more difficult to isolate than control words (52.8% of the homophones were isolated as compared to 76.1% of the control words). They also found that the language of the target word (Dutch, English) affected the candidates proposed prior to isolation – something both Grosjean (1988) and Li (1996) had previously found. In addition, the bilinguals showed much less confidence in the choices they made when the items were homophones than when they were control words.

In addition, the authors conducted a cross-modal priming study in which the participants heard a word (a prime), then saw a letter string on a computer screen, and then had to decide whether the string was an English word or not. The results again showed that interlingual homophones were more difficult to process than monolingual controls. In addition, participants were sensitive to sublexical cues: they reacted more slowly to the English homophones when they were preceded by spoken primes in the other language. Thus, the Dutch spoken word /li:f/ followed by the visual English LEAF led to longer reaction times than the English spoken word /li:f/ followed by visual English LEAF. The authors concluded that upon the presentation of an auditory input signal, lexical candidates from both languages are activated, depending on the degree of overlap between the input signal of a target word and its internal representation. The selection of the appropriate word in the one or the other lexicon can be facilitated by sublexical cues that are present in the input signal. This view is very similar to the one presented at the end of Section 2.2.1 and originally proposed by Grosjean (1988). We will discuss it in Section 2.3.3, along with the computational model that was developed a few years later.

2.3.3 A model of spoken word recognition in bilinguals

At the end of his 1988 paper, Grosjean proposed the outline of a model of spoken word recognition in bilinguals. Its main characteristics, which have been evoked indirectly at various points of this chapter, are that there are two language networks which are both independent and interconnected. In the monolingual language mode, one language network is strongly activated while the resting activation level of the other language network is very low (as in Figure 2.2). In the bilingual mode, both networks are activated but the base-language network is more strongly activated (see Figure 2.4). The resting level of the language not being used as the base language can be increased or decreased depending on the amount of input from the other language as well as from other, top-down, sources of information. The activation of a unit of one network (e.g., a phoneme) and of its counterpart in the other network depends on their degree of similarity. The activation of units that are specific to one language increases the overall activation of that language network and thus speeds up the processing in that language. Of course, if a unit (i.e., a phoneme or word) in one language has a similar unit in the other language (such as with homophones) then there will be a delay in recognizing it, all other things being equal.

With these general considerations in mind, and after having worked out various specificities, Léwy and Grosjean (in Grosjean, 2008) developed a computational model of bilingual lexical access (BIMOLA). It is similar in certain ways to the Bilingual Interactive Activation (BIA) computational model presented in Chapter 4 in that both deal with word recognition in bilinguals. In addition, both are based on interactive activation models of cognitive processes (see McClelland & Rumelhart, 1981) and both are implemented on computer. However, there are also major differences between the two in that BIA simulates the recognition of visual words whereas BIMOLA is a model of spoken word recognition (it was inspired specifically by McClelland and Elman's [1986] TRACE model). In addition, BIA deals with Dutch and English whereas BIMOLA simulates the recognition of French and English words. Other important differences concern the internal properties of each model (see Thomas & Van Heuven, 2005).

Figure 2.5 presents a simplified visual representation of BIMOLA (the reader may want to compare it to the representation of BIA in Figure 4.4). As can be seen there are three levels of nodes: features, phonemes, and words. The features level nodes are shared by the two languages whereas the two other types of nodes are organized both independently (as subsets) and as one large system. Features activate phonemes that, in turn, activate words. The activation connections between phonemes and words are bidirectional (see the upward and downward arrows between levels) whereas the activation connections between the features level and the phonemes level are simply bottom-up.

Words also receive top-down activation, allowing the language mode to be preset – only one language is active in the monolingual mode at the beginning of word

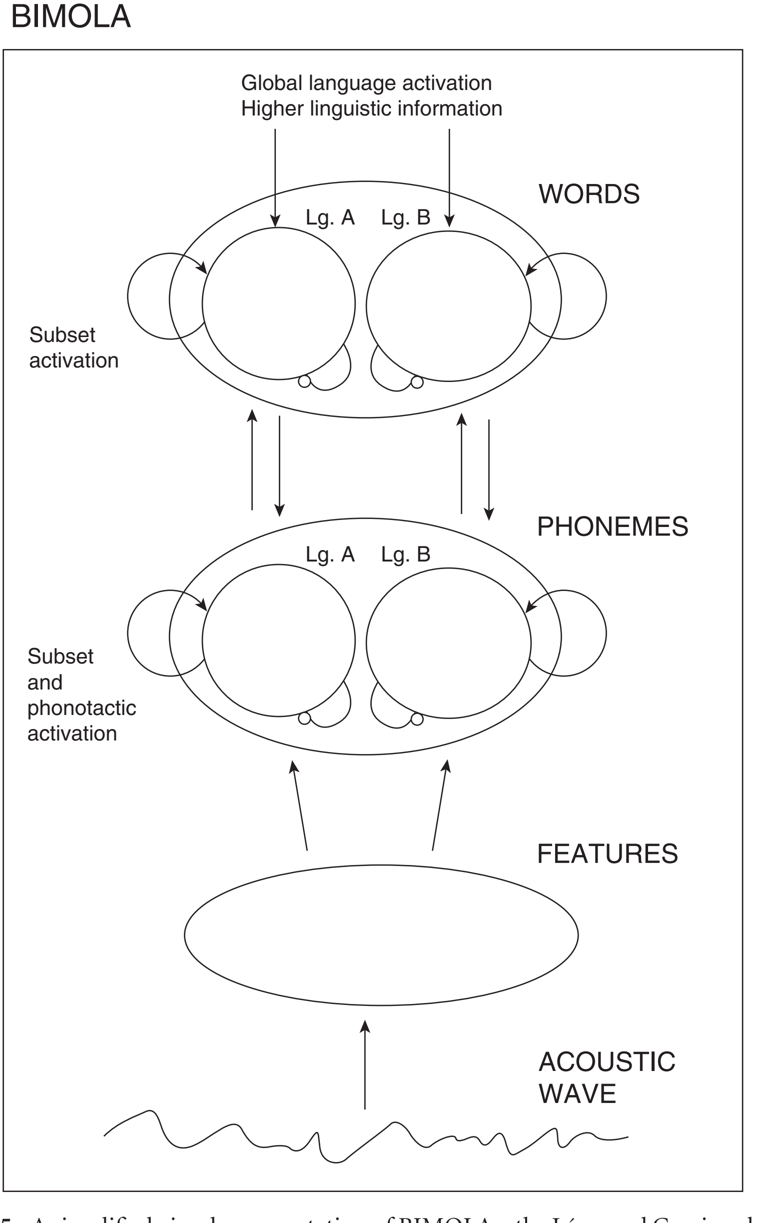

Figure 2.5: A simplified visual representation of BIMOLA – the Léwy and Grosjean bilingual model of lexical access. Adapted from Fig. 11.1, p. 204, in *Studying Bilinguals* by François Grosjean (2008). By permission of Oxford University Press.

recognition whereas the other language can also be active in the bilingual mode. There is also subset activation at the word and phoneme levels as well as phonotactic activation just at the phoneme level. Finally, units within the word and phoneme levels inhibit one another but only within a language. The model has been found to replicate the bilingual effects found in Grosjean (1988): the base-language effect, the phonotactic effect, the language phonetic effect, and the base-language homophone effect.

Research Questions

1. Why have so many researchers in psycholinguistics believed for so many years that monolingual language processing in bilinguals is nonselective?
2. What kind of study could one conduct to show the impact of *two* other languages (a second and a third language) when bilinguals are processing just one language (as depicted in Figure 2.2)?
3. Specify the role that the prosody of a language (intonation, stress, rhythm, etc.) can have on the selective/nonselective processing issue?
4. How would you test whether code-switching rules have an impact on the processing of bilingual speech?

Further Readings

Chambers, Craig & Cooke, Hilary. 2009. Lexical competition during second-language listening: sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(4). 1029–1040.

Cutler, Anne & Broersma, Mirjam. 2005. Phonetic precision in listening. In William Hardcastle & Janet Mackenzie Beck (eds), *A Figure of Speech: A Festschrift for John Laver*, pp. 63–91. Oxford: Routledge.

Grosjean, François. 2008. *Studying Bilinguals*. Oxford/New York: Oxford University Press.

Lecumberri, Maria, Cooke, Martin, & Cutler, Anne. 2010. Non-native speech perception in adverse conditions: A review. *Speech Communication* 52. 864–886.

References

Chambers, Craig & Cooke, Hilary. 2009. Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(4). 1029–1040.

Cheng, Yu-Lin & Howard, David. 2008. The time cost of mixed-language processing: An investigation. *International Journal of Bilingualism* 12(3). 209–222.

Domenighetti, Corinna & Caldognetto, Dolorès. 1999. Le rôle de la langue de base dans la reconnaissance des alternances codiques chez les bilingues. Unpublished Master's Thesis, Language Pathology Program, Neuchâtel University, Switzerland.

Grosjean, François. 1988. Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes* 3. 233–274.

Grosjean, François. 2008. *Studying Bilinguals*. Oxford/New York: Oxford University Press.

Guillelmon, Delphine & Grosjean, François. 2001. The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition* 29. 503–511.

Ju, Min & Luce, Paul. 2004. Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science* 15(5). 314–318.

Lecumberri, Maria, Cooke, Martin, & Cutler, Anne. 2010. Non-native speech perception in adverse conditions: A review. *Speech Communication* 52. 864–886.

Léwy, Nicolas & Grosjean, François. 2008. The Léwy and Grosjean BIMOLA model. Chapter 11 of Grosjean, François. 2008. *Studying Bilinguals*. Oxford/New York: Oxford University Press.

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

Marian, Viorica & Spivey, Michael. 2003. Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* 6. 97–115.

Marslen-Wilson, William. 1975. Sentence perception as an interactive parallel process. *Science* 189. 226–228.

McClelland, Jay & Elman, Jeff. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18. 1–86.

McClelland, Jay & Rumelhart, David. 1981. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88. 375–407.

Pallier, Christophe, Colomé, Angels, & Sebastián-Gallés, Núria. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12(6). 445–449.

Schulpen, Béryl, Dijkstra, Ton, Schriefers, Herbert, & Hasper, Mark. 2003. Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 29(6). 1155–1178.

Soares, Carlos & Grosjean, François. 1984. Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition* 12. 380–386.

Spivey, Michael & Marian, Viorica. 1999. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10. 281–284.

Thomas, Michael & Van Heuven, Walter. 2005. Computational models of bilingual comprehension. In Judith F. Kroll & Annette M.B. de Groot (eds), *Handbook of Bilingualism*, 202–225. Oxford/New York: Oxford University Press.

Vitevitch, Michael. 2012. What do foreign neighbors say about the mental lexicon? *Bilingualism: Language and Cognition* 15. 167–172.

Weber, Andrea & Cutler, Anne. 2004. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50(1). 1–25.