Language comprehension in bilingual speakers

Ana I Schwartz, University of Texas at El Paso
Judith F Kroll, Penn State University
Chapter 25
Language Processing in Bilingual Speakers

Ana I. Schwartz and Judith F. Kroll

We couldn’t agree where we wanted to go on vacation, y tuvimos una larga discusión
[and we had a long argument]

In the border community of El Paso, Texas, it is not at all uncommon to hear code-switched phrases like the one above. English and Spanish conversations are heard throughout all contexts of the community, whether it be on a university campus, a small café, bus, or in a formal business meeting. What an interesting feat it is then for the typical, bilingual citizen of El Paso to comprehend the mixture of language that she may hear or read. How is it that bilinguals comprehend input from their two languages? At what point in comprehension is the language of the utterance identified? At what point, if at all, does a bilingual select one language over the other?

These questions seem particularly salient when one considers that most of the world’s population is proficient in more than one language (Bhatia & Ritchie, 2004). However, it is important to point out that research on bilingualism is essential toward developing theories of language processing and cognition that extend beyond interests in bilingualism. For example, take the word “discusión” in the phrase quoted above. Although this is a Spanish word, it clearly looks like the word “discussion” in English, and this similarity makes the language membership of the word quite ambiguous. Furthermore, despite the high degree of superficial, form similarity between these two words, their meanings are actually quite distinct. In Spanish, “discusión” refers to a disagreement; whereas the meaning of the English word “discussion” does not include the same combative overtones. How do bilinguals disambiguate such words and apply the appropriate intended meaning? The issue of lexical disambiguation has been a focus in the development of psycholinguistic theories of language processing and has spurred considerable debate regarding the role of context in lexical access and the degree to which processes of language comprehension are modular vs. interactive. The study of bilingual language comprehension allows researchers to address these issues in ways that cannot be achieved through monolingual paradigms.
In this chapter, we review the recent literature on bilingual language comprehension and production. We will focus in particular on issues related to cross-language interactions that take place between a native language and a non-native language. Previous research has examined both the interactions that occur as a new language is being acquired as well as the interactions that take place during online language comprehension processes. In terms of the interactions that take place during acquisition, research has demonstrated that linguistic characteristics of the native language (L1), such as how phonologically or syntactically similar it is to the second language (L2) has an effect on its acquisition (e.g., Bosch, Costa, & Sebastian-Galles, 2000; Fernández, 1998; MacWhinney, 1997). Similarly, there is evidence that the process of acquiring an L2 also has an impact on the L1. For example, studies of syntactic processing suggest that parsing preferences that develop in L2 acquisition can transfer and modify the parsing strategies used during L1 processing (Dussias, 2003). With respect to the interactions that take place during online processing, research has provided compelling evidence that bilinguals are unable to selectively activate one of their languages during either comprehension or production. Bilingual language comprehension appears to involve the parallel activation of lexical information during both visual word recognition (Dijkstra, De Bruijn, Schriefers, & Brinke, 2000; Dijkstra, Grainger, & Van Heuven, 1999; Jared & Kroll, 2001; Jared & Szucs, 2002; Van Heuven, Dijkstra, & Grainger, 1998; Von Studnitz & Green, 2002; and see Dijkstra, 2005 for a recent review) and auditory word recognition (e.g., Spivey & Marian, 1999). Likewise, the spoken production of words in even one of the bilingual’s two languages appears to activate lexical candidates in both languages (e.g., Colomé, 2001; Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998).

Only very recently have researchers begun to examine how language non-selectivity is modified, if at all, by language or sentence context (e.g., De Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; Ju & Luce, 2004; Van Hell & De Groot, 1998). To the extent that the target language cannot be immediately selected, there are also critical consequences for the cognitive control that bilinguals must develop to negotiate the potential competition across the two languages (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004; Green, 1998). In this chapter, we will review these studies and discuss the implications for current models of the bilingual lexicon in comprehension and production.

In the review that follows, we take an inclusive approach to consider bilinguals to be anyone who actively uses two languages to some degree of proficiency. Bilinguals are rarely equally proficient or balanced in their use of the two languages, rendering one of the languages the more dominant language. Typically, the more dominant language will be the first or native language, but for bilinguals who have lived in their L2 environment for many years, the L2 may be functionally more dominant, at least for certain language-processing tasks. Among the most commonly studied bilinguals (e.g., the Dutch–English bilinguals in The Netherlands or the Catalan–Spanish bilinguals in Barcelona), many individuals are actually trilinguals. These contexts of language acquisition and use will clearly have implications for comprehension although some of the consequences, particularly in studies of lexical access, have only recently begun to be explored systematically.
We discuss them in the cases in which there are documented consequences for processing and where there is enough evidence to assess their contribution.

We begin the chapter with a review of models of the bilingual lexicon. Although the issues we consider in this chapter extend beyond the lexical level, early models of cross-language representation and processing focused extensively on the way in which words and concepts are interconnected in the bilingual’s two languages. These models provide a starting point for understanding how a learner’s cognitive system adapts to the presence of a second language. We then provide an overview of the recent research that has examined the cross-language interactions that take place during comprehension and production at the lexical and sub-lexical levels. These studies consider how single words and sounds from a bilingual’s two languages are identified and spoken. We then review the evidence for cross-language interactions at the sentence/syntactic levels. Finally, we examine a set of issues common to both the lexical and sentential level of processing that include the cognitive factors that modulate L2 processing, the consequences of cross-language similarity, the directionality of cross-language influences, and the more general implications of research on bilingualism for theories of word recognition, lexical access, and sentence processing. Although our review focuses primarily on behavioral evidence from studies of the performance of second language learners and relatively proficient bilinguals, we also consider the emerging literature on the neural basis of bilingualism.

1. MODELS OF THE BILINGUAL LEXICON

Early accounts of the bilingual’s cognitive system examined the issue of how an adult who already has a fully formed lexicon and conceptual system for their native language represents newly acquired information in a second language. Potter, So, Von Eckardt, and Feldman (1984) contrasted two alternative models for how a learner might integrate new L2 knowledge into their existing L1 language system. The models are shown in Figure 1. According to the Word Association model, new L2 words are represented by means of associations to their translation equivalents. Thus, to learn that the word *kikker* in Dutch means *frog* in English, the learner would form an association directly between kicker and frog. On this view, L2 words access meaning indirectly via the L1. In contrast, the Concept Mediation model assumes that L2 words have direct access to their respective meanings. Potter et al. tested the two models in a series of experiments in which bilinguals named pictures and translated words. They reasoned that if words in the L2 were lexically mediated via the L1, then translation into the L2 should be performed more quickly than picture naming because the translation would be directly available without conceptual processing. Picture naming would necessarily be slower than translation because to name a picture in L2 would require first accessing its meaning and name in L1, and only then could the L2 name be retrieved by its association with the L1 translation. However, if words in the L2 access their respective meanings directly, then all other things being equal, picture naming in L2 and translation into the L2 should take about the same amount of time. Potter et al. found support for the Concept Mediation predictions.
In two experiments, one with highly proficient Chinese–English bilinguals and the other with French–English second language learners, the time to name pictures and translate words was approximately the same. Contrary to the predictions of the Word Association model, the time to name pictures was, if anything, faster than the time to translate words. The results were counterintuitive because although the learners were slower in L2 than the proficient bilinguals, the pattern of results was the same, suggesting that even at early stages of L2 learning, individuals are able to access concepts directly.

The models shown in Figure 1 assume independent representations at the lexical level but a shared representation at the conceptual level. A question that we will return to is how the picture of cross-language activity suggested by the current research can be accommodated within this hierarchical framework. If even learners can access meaning directly for words in the L2 as the Potter et al. (1984) results suggest, then lexical-level interactions across languages might be expected to be quite limited. However, the immediate problem raised by the Potter et al. study was to understand why striking differences in L2 skill did not seem to influence the form of the connections between words and concepts in the bilingual’s mind. Two subsequent studies (Chen & Leung, 1989; Kroll & Curley, 1988) failed to replicate the Potter et al. results for the less-proficient L2 learners. In each of these studies, the performance of highly proficient bilinguals replicated the Concept Mediation model’s pattern, with similar latencies for naming pictures in the L2 and translating words into the L1. However, contrary to Potter et al., the performance of L2 learners followed the pattern predicted by the Word Association model, with translation into the L2 significantly faster than picture naming in the L2. The apparent discrepancy in the findings for L2 learners may be attributed to their level of proficiency. The learners in the Potter et al. study were high school students about to travel to France for a summer immersion experience. Although they were far from proficient in French, they were presumably highly motivated learners who were likely to have been beyond the very earliest stages of acquisition.
The results of Chen and Leung (1989) and Kroll and Curley (1988) suggested that learners initially associate new words in the L2 to their translation equivalents in the L1 and only later become able to access concepts directly. To account for the change in the connections between words and concepts as L2 skill develops, Kroll and Stewart (1994) proposed the revised hierarchical model (RHM). The model, shown in Figure 2, integrates the Word Association and Concept Mediation models but, unlike the earlier models, also makes assumptions about the strength of connections. At the lexical level, the RHM assumes that the connection from L2 to L1 is stronger than the connection from L1 to L2. Because learners presumably access L1 translations during early stages of acquisition for the purpose of retrieving the meaning of new L2 words, the requirement to transfer information from L1 to L2 will create a strong lexical-level connection from L2 to L1. With respect to the connections from words to concepts, the RHM proposes that the L1 is privileged with respect to accessing meaning and thus L1 connections to concepts will be stronger than those for L2 and only with increasing L2 proficiency will the L2 links to concepts begin to resemble those for L1.

The RHM can account for an asymmetry observed in the previous studies such that translation in the forward direction, from L1 to L2, is typically slower and more error-prone than translation in the backward direction, from L2 to L1. According to the model, L2 to L1 translation can be performed by retrieving lexical-level associations that provide a direct route to the translation, whereas translation from L1 to L2 is more likely to engage concept processing and subsequent lexicalization by virtue of the strong connections between L1 words and their respective meanings. The translation asymmetry is consistent with this account but can also be attributed to difficulties in accessing the phonology of the weaker L2. To provide a more critical test of the RHM, Kroll and Stewart (1994) examined the translation performance of relatively proficient Dutch–English bilinguals in contexts in which the words to be translated appeared in semantically categorized lists or in lists that were semantically mixed. If only translation from L1 to L2 engages meaning, then only translation in that direction should be influenced by the semantic organization of the list context. The results confirmed this prediction. Translation from L1 to L2 was slower in the semantically categorized lists.

Figure 2. The Revised Hierarchical Model (adapted from Kroll & Stewart, 1994).
than in the mixed lists but translation from L2 to L1 was unaffected by this manipulation. The fact that semantically categorized lists induced interference was attributed to likely competition during L2 production (see Kroll, Bobb, & Wodniecka, in press for a recent discussion of this issue).

If conceptual processing in the L2 is a function of proficiency, then why did the Dutch–English bilinguals in the Kroll and Stewart (1994) study, who are relatively skilled speakers of English, show an asymmetry for the two directions of translation? One possibility is that the words to be translated included items that were relatively low frequency, making it more likely that the L2 would revert to reliance on the L1 translation. A number of recent studies have suggested that weighting of interlanguage connections is affected not only by the proficiency of the bilingual speaker but also by the nature of the words and concepts that are translated (e.g., Francis, Tokowicz, & Kroll, 2003). When a small set of items are well learned or highly practiced, even relative novices may appear to conceptually process words in the L2 (e.g., Altarriba & Mathis, 1997). However, under other circumstances, the development of L2 proficiency appears to shift from reliance on lexical cross-language connections to reliance on conceptual interconnections with increasing skill (e.g., Kroll, Michael, Tokowicz, & Dufour, 2002; Talamas, Kroll, & Dufour, 1999). Although there is controversy with respect to the issue of whether highly proficient bilinguals ever revert to lexical mediation during translation (e.g., Duyck & Brysbaert, 2004; La Heij, Hooglander, Kerling, & Van der Velden, 1996), the evidence on L2 learners suggests that there is an early stage of acquisition in which L1 translations are highly active (Sunderman & Kroll, in press).

The RHM and the earlier Concept Mediation model share an implicit assumption that once individuals achieve sufficient expertise in the L2, the L2 can function autonomously of the L1 (but see Kroll & De Groot, 1997, for a distributed model at both the lexical and conceptual levels). As we will see in the remainder of this chapter, it is this assumption that has been challenged in the recent literature in which cross-language activity has been observed for even highly proficient bilinguals. While no one would dispute the fact that bilinguals gain automaticity in processing the L2 with increasing skill (e.g., Segalowitz & Hulstijn, 2005), recent studies suggest that proficiency does not imply an ability to use each of the languages autonomously, as if the bilingual were functionally monolingual. However, a recent study (Sunderman & Kroll, in press) demonstrated that at the lexical level, it is the activity of the translation equivalent that changes with increasing proficiency, but not the activity of lexical form relatives. As we will review below in the section on lexical-level cross-language interaction, when bilinguals read or hear words in one language, there is activation of related information in the other language. However, what is active for proficient bilinguals is the orthographic and phonological information present in words in the two languages, not the translation equivalent itself. Thus the word man might be briefly activated in English when an English–Spanish bilingual reads the word mano, meaning hand in Spanish. These words are semantically unrelated but share similar lexical form. Sunderman and Kroll showed this type of cross-language activity occurred regardless of L2 proficiency. In contrast, activation of information related to the translation equivalent (e.g., the word hambre, meaning hunger in Spanish, resembles...
hombre, the translation of man) occurred only for learners at relatively early stages of L2 acquisition. The account provided by the RHM with respect to lexical mediation thus appears characterize the performance of learners who still depend on the L1 to access meaning and more proficient bilinguals who may revert to this strategy when the words to be translated are relatively low frequency in the L2, as seen in Kroll and Stewart (1994).

A comprehensive bilingual model will require that distinctions be made between those aspects of cross-language representation and processing that change dynamically with changes in proficiency and language use, and those that reflect the way in which the linguistic structure of the bilingual’s two languages imposes constraints that are relatively stable (see Kroll & Tokowicz, 2005 for a discussion of these issues). Furthermore, it will require that a principled account of the relation between sub-lexical, lexical, semantic, and syntactic processes. In the sections that follow we report the recent evidence at each of these levels of language processing and further consider the manner in which context and skill and modulate the observed cross-language interactions.

2. CROSS-LANGUAGE INTERACTIONS AT THE LEXICAL AND SUB-LEXICAL LEVELS

2.1. The Perception of Speech in Two Languages

A focus of contemporary psycholinguistic research on both monolinguals and bilinguals concerns the mechanisms that allow individuals to perceive and uniquely identify the sounds and words of their languages. Indeed, one of the first tasks faced by all language learners is the identification of his/her native language(s) and all of its distinct sounds. In the earliest days of life, newborns are able to discriminate between unfamiliar languages, particularly if these languages belong to different rhythmic classes (e.g., Mehler et al., 1988). What about infants who are raised in multilingual environments? Recent findings suggest that bilingual infants might have enhanced language-discrimination abilities. Infants raised in bilingual environments show an early ability to discriminate languages; even those that are from the same rhythmic class (Bosch & Sebastián-Gallés, 2001).

In addition to making cross-language distinctions, young language learners must learn to identify and discriminate sounds within their native language(s). Research on monolingual language acquisition has demonstrated that within the first year of life monolingual infants are sensitive to the phonotactic patterns of their native language. For example, as early as nine months of age they are able to discriminate phonological sequences (e.g., CV structure) that conform to the regularities of the language from patterns that do not (e.g., Jusczyk, Goodman, & Baumann, 1999; Sebastián-Gallés & Bosch, 2002). How do the early sensitivities of infants exposed to multiple languages compare to those of infants raised in monolingual contexts? Recent evidence suggests that infants who are exposed simultaneously to two languages develop phonotactic sensitivities at the same rate as monolinguals. More interestingly, the degree of this sensitivity increases as
a function of the relative exposure to a given language; thus even at this early stage of life infants’ perceptual abilities can reflect a perceptual dominance for the language to which they are most often exposed (Sebastián-Gallés & Bosch, 2002). This perceptual dominance of one language over another has been shown to persist into adulthood. It has been found that even highly proficient bilinguals apply one set of segmentation rules exclusively from the language to which they were most often exposed when parsing speech (e.g., Cutler, Mehler, Norris, & Seguí, 1989).

During the early months of life, infants are also developing cognitive representations of the different sounds or phonemes that make up their native language (Kuhl, Williams, Lacerda, & Stevens, 1992; Marean, Werner, & Kuhl, 1992; Werker & Lalonde, 1988). The perceptual system is quite flexible, allowing very young infants to make non-native phonemic contrasts. Remarkably, within just 6–12 months of age, there is a significant re-organization in which perceptual processes become specifically tuned to the native language and infants’ ability to discriminate non-native phonemic contrasts diminishes (Werker & Lalonde, 1988). With increasing age and linguistic exposure, this language-specific tuning acts as a filter through which a non-native language is perceived. As a consequence, two phonemes that are contrastive within a second or less familiar language may be perceived as the same phoneme if the L1 does not distinguish between these sounds. For example, adult native Japanese speakers often have great difficulty in distinguishing between the /r/ and /l/ English phonemes, which are not contrastive in Japanese and both phonemes tend to be perceived as /l/.

This L1 filtering effect has been observed even for highly proficient bilinguals who have had long lifetime experiences and seemingly equal command of both languages (e.g., Pallier, Colomé, & Sebastián-Gallés, 2001; Pallier, Bosch, & Sebastián-Gallés, 1997; Sebastián-Gallés, Echeverría, & Bosch, 2005). For example, highly proficient Spanish–Catalan speakers performing an auditory lexical decision in their L2 showed repetition priming for Catalan words that differed only by a single vowel contrast that does not exist in Spanish. These priming effects suggested that the bilinguals were perceiving the minimal pairs as the same entity. A separate group of Catalan-dominant bilinguals did not show these priming effects. These results were particularly remarkable since the two bilingual groups were both early bilinguals who were living in the same city and had received the same bilingual education. Indeed, the Spanish-dominant group was distinguished based only on the fact that they had been raised in Spanish monolingual homes during their earliest childhood prior to entering school (Pallier et al., 2001).

However, it is important to note that the perceptual system does not ever completely lose its ability to perceive new phonemes and distinguish new contrasts. Furthermore, non-native speakers vary greatly in their ability to discriminate L2 contrasts (e.g., Strange, 1995). The relative ease with which non-native phonemes are perceived and contrasted depends critically on several factors, including the phonetic similarity of the native and non-native languages and the age at which the second language was acquired (e.g., Flege, 1995).
2.2. Recognizing Words in Two Languages when they are Read or Spoken

The studies reviewed thus far highlight the ways in which the acquisition of a native language has a fundamental and long-lasting effect on how the sub-lexical and lexical units of a new language are perceived and acquired. There is also evidence for cross-language interactions that occur during the online comprehension of language, and there is now a large body of literature demonstrating that when recognizing words bilinguals activate lexical and sub-lexical units from both of their languages in parallel (e.g., Dijkstra et al., 1999; Dijkstra et al., 2000; Jared & Kroll, 2001; Ju & Luce, 2004; Marian & Spivey, 2003; Van Heuven et al., 1998).

Many of the studies that have looked at these online interactions have focused on visual word recognition tasks such as lexical decision and naming. The general strategy has consisted of presenting bilinguals with words in one language that share some lexical property or properties with words from their other language. For example, in a lexical decision task bilinguals may be presented with non-words (blart) and two types of words: Words that share lexical form and meaning with the non-target language (e.g., the cognate piano in English and Spanish) or control words that do not share any lexical property with the non-target language (e.g., pencil). One can then compare the bilinguals’ word recognition performance, in terms of latency and/or accuracy, for these critical words relative to control words that do not share lexical properties (e.g., pencil in English and Spanish). If bilinguals are able to selectively activate a single language during word recognition, then whether the word shares lexical properties with the non-target language should be of no consequence. If, on the other hand, there is parallel activation of both languages then the processing time and/or accuracy of performance should differ from control words.

The evidence has consistently demonstrated that bilingual word recognition involves the parallel, non-language-selective activation of both languages. For example, the processing of cognates is consistently facilitated relative to non-cognate control words across a wide variety of tasks including translation (Kroll & Stewart, 1994); word association (Van Hell & De Groot, 1998) and lexical decision (Dijkstra, Van Jaarsveld, & Ten Brinke, 1998). In translation, it is perhaps not surprising to find cognate facilitation because both languages are required to be active. However, in within-language lexical decision, it should be possible to function selectively, yet even under these circumstances there are robust effects of the non-target language. Cognate facilitation is obtained when bilinguals perform visual lexical decision in their native language with no expectation that the L2 will be used, suggesting that these effects are not simply a reflection of the stronger L1 influencing the weaker L2 (e.g., Van Hell & Dijkstra, 2002). Furthermore, evidence of non-selectivity persists irrespective of the surrounding language context, task instructions, or participant expectations to process one or multiple languages (Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra & Van Hell, 2003).

It is important to note that the nature of the effects of non-selective activation depends on the relative match of the lexical codes (orthographic, phonological, and semantic)
across languages. While cognate facilitation has been observed across a variety of
different tasks, the effects of inter-lingual findings regarding interlingual homographs,
words that share form but not semantics (e.g., fin in Spanish means "end") have not been
nearly as consistent. Some studies have demonstrated inhibitory effects associated with
homograph status (Dijkstra et al., 1998; Jared & Szucs, 2002; Von Studnitz & Green,
2002), while others have failed to find any effects at all (Dijkstra et al., 1998; Gerard &
Scarborough, 1989). Furthermore, the specific nature of homograph effects, whether they
are inhibitory or facilitative in nature has varied as a consequence of differences in task
demands, the salience of the non-target language, and the relative frequency of the ho-
mographs’ lexical representations across languages (Dijkstra et al., 2000a; Dijkstra et al.,
2000b; Dijkstra et al., 1998).

In a recent study, Schwartz, Kroll, and Diaz (in press) found further evidence that
bilingual lexical processing of cognates is influenced by the relative match in ortho-
graphic and phonological codes across languages. In that study, English–Spanish bilin-
guals named cognates (e.g., piano) and non-cognate control words (e.g., lapiz/pencil)
in their L1 and L2. The cognates were classified according to the relative orthographic
and phonological similarity across English and Spanish. To illustrate, the
English–Spanish cognate base maps on to very distinct pronunciations ([bas] vs.
[báse]) and was classified as +O–P, whereas piano is pronounced much more simi-
larly ([paiˈnou] vs. [piˈaˈno]) and was classified as +O+P. The authors predicted that
in the presence of highly similar orthography (e.g., pianol/piano; base/base), the activa-
tion of the cross-language phonological representations would be particularly strong.
When these representations were more distinct (e.g., [bas] vs. [báse]) the resulting
competition would inhibit performance. The findings supported this prediction, the
+O+P cognates (e.g., piano) were named faster and more accurately than the +O–P
cognates (e.g., base) in both the L1 and L2 of the participants. This suggests that not
only is lexical access non-selective across bilinguals’ two languages, but that the sub-
tle interactions between the activated codes determine the manner in which cross-lan-
guage competition is manifest.

The studies reviewed above all focused on cross-language activation that occurs dur-
ding the recognition of visually presented words. In these studies, effects of cross-lan-
guage activation may have been particularly robust since the visual input (i.e., a string of
letters) can be completely language neutral (e.g., fin in English and Spanish), at least in
the case in which languages share the same script. It should be noted that cross-language
activation has also been observed for words that are not identical. For example, the iden-
tification of cross-language neighbors, words that share all but one letter across lan-
guages (e.g., cineldine in Spanish and English), has been shown to be slower and more
error-prone relative to control words in lexical decision (Van Heuven et al., 1998) and
naming (Jared & Kroll, 2001). There is also facilitation for recognizing cognates even
when the cognates are non-identical (Van Hell & Dijkstra, 2002). Furthermore, there is
evidence that even when the languages do not share the same script (e.g., Hebrew and
English or Chinese and English), cross-language priming effects can be observed (e.g.,
Gollan, Forster, & Frost, 1997; Jiang, 1999).
Parallel activation of information in each of the bilingual’s languages during visual word recognition has been modeled in a variant of McClelland and Rumelhart’s (1981) Interactive Activation model for monolingual word recognition called the Bilingual Interactive Activation model, or BIA (Dijkstra et al., 1998). The model is shown in Figure 3. Like the monolingual model, BIA assumes that there is parallel activation of

Figure 3. The Bilingual Interactive Activation Model (adapted from Dijkstra et al., 1998).
letter features, letters, and words, with information similar to the input string activated to some degree and producing competition across alternative candidates. However, unlike the monolingual model, BIA includes a layer of language nodes which serve to represent the top-down contextual biases and subsequently inhibit the bottom-up activation of the non-target language. In this scheme, the inhibitory effects occur relatively late in processing, once the initial components of word recognition are set in motion for all possible solutions in either of the bilingual’s two languages. BIA has been implemented as a computer model and does an excellent job of simulating bilingual word recognition performance under conditions in which the words to be recognized differ in their within and across-language orthographic properties (Van Heuven et al., 1998). At the end of the chapter, we describe an updated version of the BIA model that includes phonology and semantics as well as orthography, and also makes some different assumptions about the way in which language selection is controlled.

The observation of parallel activity across the two languages during visual word recognition does not necessarily suggest that similar cross-language activation occurs during auditory processing of the speech signal. Since languages differ in their component sounds it can be argued that within the speech signal there are language-specific cues that are not as readily available within printed text. There have been a few studies that have examined cross-language interaction during speech processing. In a seminal study, Spivey and Marian (1999) asked Russian–English bilinguals to view an array of objects as they listened to instructions in either their L1 or L2 which indicated an object that they should select (e.g., “pick up the marker”). On critical trials the instructions indicated a target object whose phonological onset was the same as that of another object in the non-target language (e.g., “stamp” in Russian is “marka”). To test whether the non-target lexical representation of the object was activated, the authors monitored the bilinguals’ eye-movements as they surveyed the array of objects and listened to the instructions. When the instructions indicated an object whose phonological onset was shared across languages, participants initially looked toward the object that shared this onset in the non-target language. This indicated that upon hearing the initial, shared phoneme, the bilinguals activated lexical candidates from both of their languages (and see Marian & Spivey, 2003).

Using a very different paradigm, Colomé (2001) found converging evidence that bilinguals activate phonemic representations from both languages in a non-selective manner. In that study highly proficient Spanish–Catalan bilinguals performed a phoneme monitoring task in their L2, in which they decided whether the name of a visually presented picture (e.g., a table) contained a target phoneme (e.g., /m/). On critical trials, the bilinguals had to reject phonemes that were not part of the Catalan name (e.g., /m/ is not present in the Catalan word taula) but were part of the contextually irrelevant Spanish translation of that object (e.g., mesa). The bilinguals took significantly longer to reject phonemes contained in the Spanish translation relative to phonemes that were not part of the picture’s name in either language (e.g., /s/).

Subsequent studies have demonstrated that aspects of the linguistic input itself may make it possible to constrain the parallel activation of the non-target language when processing spoken language. Weber and Cutler (2004), testing Dutch–English bilinguals
with an eye-tracking paradigm very similar to Marian and Spivey (2003), found significant cross-language effects from the L1 to the L2 (i.e., when bilinguals were processing the spoken targets in the non-native language) but not from the L2 to L1. Likewise, Ju and Luce (2004) replicated the basic pattern of cross-language phonological competitor effects, but then went on to demonstrate that cross-language competitor activation could be eliminated when the VOTs of the initial phonemes were spoken like the L1. That is, L2 competitors were no longer activated when the target words were perceived to be native-like speech. These results contrast with results from experiments using written stimuli and illustrate the critical role that access codes play in the activation of lexical and sub-lexical representations in bilingual language comprehension.

2.3. Cross-Language Lexical Access in Context

Two types of studies have examined cross-language influences on lexical access in context. One adapts the standard semantic priming paradigm (e.g., Meyer & Schvaneveldt, 1971) to ask whether semantically related words prime target words when the primes and targets appear in different languages. The other asks whether the parallel activation of lexical alternatives in both languages is modulated by the presence of sentence context and whether lexical access within each language is open to the semantic and syntactic influences of sentence context in the other language. We consider briefly the evidence from each of these areas of research.

2.3.1. Semantic context

Early research on the bilingual lexicon used the semantic priming paradigm as a means to test the independence of representations across the bilingual’s two languages (e.g., McCormack, 1977; Snodgrass, 1984). In semantic priming, a prime word is typically presented briefly and then followed by a target word for lexical decision. When the prime word is semantically related to the target word, lexical decision time is facilitated relative to conditions in which the prime is unrelated to the target (Meyer & Schvaneveldt, 1971). A series of studies using variants of the semantic priming paradigm (e.g., Altarriba, 1990; Chen & Ng, 1989; De Groot & Nas, 1991; Kirsner, Smith, Lockhart, King, & Jain, 1984; Meyer & Ruddy, 1974; Schwanenflugel & Rey, 1986; Tzelgov & Eben-Ezra, 1992) demonstrated that it was possible to observe semantic priming across as well as within languages. The result of these studies suggested that bilinguals access semantic representations that are shared across the two languages (and see Dufour & Kroll, 1995, for related evidence from a categorization paradigm). Although some of these studies attempted to control the methodology of the priming paradigm to ensure that any observed priming could be attributed to automatic processes, many of the early studies can be criticized on methodological grounds because they included a high proportion of related trials and a long interstimulus interval between prime and target words that may have encouraged subjects to develop expectations for the upcoming targets (see Neely, 1991). The use of long prime-to-target intervals in the bilingual experiments is particularly problematic because in some of these experiments subjects may have been encouraged to translate the prime and/or target into the same language. If primes can be translated, even on a small
proportion of trials, then the observed cross-language effects may reflect only the presence of within-language priming in the cross-language conditions.

Subsequent cross-language semantic priming studies attempted to control for these factors with the result that cross-language priming was obtained but only under some conditions. For example, using a masked priming paradigm (see Figure 4 for an illustration), in which participants cannot consciously report the prime information and are therefore not even aware of the bilingual nature of the task, De Groot and Nas (1991) found evidence for cross-language associative priming only when translation equivalents in the two languages were cognates, sharing lexical form as well as meaning. Keatley, Spinks, and De Gelder (1994) reported that even when bilinguals are highly proficient in both languages there are asymmetries in the magnitude of semantic priming, with significant facilitation only with L1 primes and L2 targets (see also Altarriba, 1990, and Tzelgov & Eben-Ezra, 1992).

Recent studies of cross-language priming have reported mixed results with respect to the conditions under which priming is observed. On one hand, a series of masked priming studies (e.g., Gollan et al., 1997; Jiang, 1999) has shown that cross-language priming is observed in lexical decision for translation equivalents even when the two languages involved do not use the same script (Hebrew–English and Chinese–English), but is only consistently reliable with L1 primes and L2 targets (and see Grainger & Frenck-Mestre, 1998, for evidence that priming is observed in semantic categorization but not in lexical decision). Finding cross-language masked priming is particularly striking because participants are unaware of the prime words. When experiments are designed so that the language of target words is blocked, the experiment can be functionally performed in “monolingual mode” (Grosjean, 2001), thus reducing the likelihood that a translation

Figure 4. Schematic illustration of the masked priming procedure. In this example, a Spanish prime word (justo meaning “just”) is preceded and followed by a forward and backward visual mask. An English target word (just) is then presented. The combination of the brief presentation of the prime word in addition to the masking stimuli minimizes conscious processing of the prime word.
strategy can account for the observed priming. However, other studies suggest that there are limits to the degree of cross-language priming that reflect constraints in the nature of bilingual language representations attributable to the age at which individuals acquired the L2 (e.g., Kotz & Elston-Güttler, 2004; Silverberg & Samuel, 2004) and the amount and kind of semantic information that is accessed for each language (e.g., Finkbeiner, Forster, Nicol, & Nakamura, 2004; Kotz & Elston-Güttler, 2004). The constraints observed in cross-language priming also appear to reflect the degree of proficiency in the L2. For example, Kotz and Elston-Güttler found that late L2 learners were able to exploit associative but not categorical relatedness, whereas early bilinguals were able to use both types of information. Although the earlier research was largely compatible with a model of the lexicon in which the two languages access the same semantic information (see Francis, 2005, for a review), the recent studies support a mixed model in which some semantic information is shared and other semantic information is distinct (see De Groot, 1993, and Van Hell, 1998, for models of partly shared cross-language semantics). It remains to be seen to what extent the type of bilingualism determines the ability to acquire all of the subtle nuances of meaning in the L2 that are available in the highly skilled L1 (see Segalowitz & Hulstijn, 2005, for a discussion of issues of automaticity in the L2).

2.3.2. Sentence context

The findings from research on both auditory speech and text comprehension are compatible with a fairly open lexical system in which information activation flows across both languages. However, as mentioned previously, most of these findings stem from studies in which stimuli were presented in fairly de-contextualized conditions, such as single-word naming and lexical decision. Is the cross-language flow of activation at all constrained when there is a context such as a sentence that provides additional semantic information? To date there have been very few studies that have addressed this question directly (Elston-Güttler, 2000; Greenberg & Saint-Aubin, 2004; Schwartz & Kroll, 2006; Van Hell, 1998). Findings from these studies suggest that cross-language activation can be indeed constrained by the presence of a sentence context, allowing the system to operate in a more language-selective manner. For example, both Schwartz and Kroll (2006) and Van Hell (1998) found that although lexical access for cognates was facilitated in low-constraint sentences, this facilitation was eliminated when the same cognates were in high-constraint sentences.

What is the mechanism that allows processing to proceed in a more language-selective manner in high-constraint sentences? According to Schwanenflugel and colleagues (e.g., Schwanenflugel, 1991; Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985; but cf. Traxler & Foss, 2000), sentence constraint influences lexical activation through a set of feature restrictions that readers generate as they comprehend sentences. With increasing constraint, an increasing number of restrictions are generated. Lexical entries that conflict with these constrictions are inhibited. Thus, in the bilingual case increasing sentence constraint would lead to the generation of language-specific lexical feature restrictions which would inhibit lexical entries from the non-target language. Support for this account was observed in an eye-movement monitoring
study of Spanish–English bilinguals’ reading of code-switched sentences (Altarriba, Kroll, Sholl, & Rayner, 1996). The eye-movements of Spanish–English bilinguals were monitored as they read high- and low-constraint sentences in English. On half of the trials, one word in each sentence was a code-switched word from the non-target language, Spanish [e.g., He wanted to deposit all of his dinero (money) in the credit union]. Critically, this code-switched word was either a high-frequency word or a low-frequency word in Spanish. Analyses of the first fixation durations on the code-switched words revealed an interaction between sentence constraint and word frequency, such that fixation durations were elevated in high-constraint sentences when the code-switched word was of a high lexical frequency. This suggests that the participants generated both semantic and lexical-level feature restrictions when reading high constraint sentences. That is, when presented with the high frequency, code-switched word (e.g., *dinero*), processing was inhibited because the word met all of the semantic but not the lexical feature restrictions. However, it should be noted that effects of sentence constraint appear to operate at a later point, once processes of initial lexical access have been completed (Greenberg & Saint-Aubin, 2004). In other words, the studies to date cannot rule out an initial non-selective activation of lexical candidates from the non-target language followed by a subsequent inhibition.

An interesting feature of the results on the effects of sentence context, is that a similar pattern has been observed in studies that have examined the effects of non-linguistic factors on cross-language activation, such as instructions, expectations, and working memory resources (e.g., De Bruijn et al., 2001; Dijkstra et al., 2000; Michael, Dijkstra, & Kroll, 2006; Van Hell & Dijkstra, 2002). Just as the results on sentence context suggest that knowing the language of the sentence has little effect on constraining lexical alternatives in the non-target language, the effects of single word context or expectations appear not to restrict the influence of the other language, at least for the recognition of visually presented words.

## 3. CROSS-LANGUAGE INTERACTIONS AT THE SENTENCE LEVEL

Far fewer studies have examined sentence processing in bilinguals relative to the substantial literature on word recognition and lexical processing. Much of the research on sentence processing in non-native speakers has focused on issues of acquisition, asking either to what extent access to the grammar of the L2 is mediated by transfer from the more highly skilled L1 (e.g., MacWhinney, 1997, 2005) or by the age of acquiring the L2 (e.g., Hahne & Friederici, 2001; Johnson & Newport, 1989; Weber-Fox & Neville, 1996). Only a small number of studies have addressed the issue of how language-specific constraints and biases in one of the bilingual’s languages affect processing in the other language (for recent reviews, see Frenck-Mestre, 2005; Kroll & Dussias, 2004). Although there is a rich history in psycholinguistics of using cross-linguistic evidence to assess the universality of language processing mechanisms (e.g., Bates, Devescovi, & Wulfeck, 2001; Cuetos & Mitchell, 1988; Vigliocco, Hartsuiker, Jarema, & Kolk, 1996)
only a small number of studies have asked directly how sentence processing is accommodated to the presence of two languages. Two types of experiments have been conducted to address these issues. In one, grammatical constructions that differ across language are the focus. The question is then how the bilingual resolves the potential conflict between the two alternatives in one and the same mind (e.g., Dussias, 2001, 2003; Fernández, 1998, 2003; French-Mestre, 2005). In the other, a priming paradigm has been used to determine whether structural repetitions that typically facilitate performance within language, also facilitate performance across languages (e.g., Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003).

3.1. Cross-Language Parsing

If a bilingual’s two languages follow different syntactic arrangements and if those differences give rise to distinct parsing preferences, then bilingualism potentially poses a problem for language processing if the two languages are not represented and accessed independently. A number of solutions to this problem are available in theory. One possibility is that the two languages are treated independently and the parsing preferences associated with each language are engaged appropriately as a function of the language context. As the literature on bilingual word recognition makes clear, there is very little evidence at the lexical level that the two languages function independently. Although it might seem that at the level of the grammar it might be easier to separate the two languages, our brief review of sentence context effects on lexical processing suggests that bilinguals do not use the language of a sentence itself as a strong cue to differentiate the two languages. As we will discuss later, bilinguals are also prepared to code switch with other similarly bilingual individuals and to understand code-switched utterances, suggesting the grammars of both languages are available and engaged.

A second solution to the problem is to bias parsing toward the native or more dominant language. This is a solution that has been discussed in detail in the literature on second language acquisition where there is a great deal of empirical support for transfer from the L1 to the L2 (e.g., see MacWhinney, 1997, 2005 for an illustration of this approach embodied within the Competition Model). However, even a strong transfer account such as the Competition Model includes a developmental component whereby successful L2 learners eventually acquire the cues for the weaker L2. The degree of transfer will depend on the relative proficiency of the bilinguals, with greater L1 influence for less than for more proficient bilinguals.

A third solution to the problem is to assume that the bilingual develops a set of parsing strategies that are uniquely bilingual, in that they represent a mix of the preferences used within each language by native speakers. This is an instance of Grosjean’s (1989) well-known admonition that the bilingual is not two monolinguals in one. Recent studies of linguistic convergence (e.g., Bullock & Toribio, 2004; Malt & Sloman, 2003) provide support for the claim that language contact produces a pattern of language use that is distinctly bilingual at all levels of linguistic description.
Although research on bilingual parsing is at an early stage of investigation, recent findings suggest that although it is sometimes possible for advanced L2 learners and proficient bilinguals to acquire native-like parsing preferences for the L2 (e.g., Frenck-Mestre & Pynte, 1997), they may also be slower to process L2 than L1 sentences (e.g., Hoover & Dwivedi, 1998), and more likely to recruit additional working memory resources (e.g., Hasegawa, Carpenter, & Just, 2002; Miyake & Friedman, 1998). What is not yet clear within this emerging body of research is whether the demands on working memory are related to the presence of structural differences between the bilingual’s two languages. That is, it may be more difficult to acquire a new form in L2 when it does not exist in the L1 or when the L1 requires distinctions that are not functional in the L2 (e.g., see Juffs, 1998 for an illustration of how otherwise proficient L2 speakers may be limited by these cross-language distinctions in the case of causative-inchoative constructions, and Jiang (2004), who shows that highly proficient Chinese–English bilinguals are restricted in their ability to comprehend subject–verb agreement in their L2 in an online task although they are able to recognize the correct English forms in an offline measure).

The experiments on sentence parsing that are perhaps most critical to a model of bilingual comprehension are those in which structural preferences in the two languages are in conflict. For example, Dussias (2001, 2003) has examined the resolution of such a cross-language conflict in the case of attachment preferences for temporally ambiguous sentences containing a relative clause. Dussias (2003) uses the sentence below to illustrate how native speakers of English and Spanish make distinct structural commitments.

Peter fell in love with the daughter of the psychologist who studied in California.

Native speakers of English prefer to attach the relative clause, who studied in California, to the immediately preceding noun, psychologist. Thus, in response to the question, “Who studied in California?” a native English speaker would respond, the “psychologist.” This preference has been called low attachment (e.g., Frazier & Rayner, 1982). In contrast, native Spanish speakers prefer high attachment, so their answer to the same question would be, “the daughter” (but see also Gilboy, Sopena, Clifton, & Frazier, 1995). Carreiras, Salillas, and Barber (2004) recently provided evidence for the high-attachment preferences of native monolingual Spanish readers in an ERP study. They found a P600 effect, typically observed in ERP studies of sentence processing in response to a syntactic violation, when Spanish readers processed a sentence that was consistent with low attachment, contrary to their ordinary bias. What happens when both languages are available to highly proficient bilinguals? Dussias reports that Spanish–English bilinguals prefer the low-attachment strategy even when they are reading in Spanish, their native language. That is, the preference for high attachment appears to change once a high level of proficiency is achieved in the L2 (i.e., English). Like the results reviewed earlier on bilingual word recognition, these findings suggest that not only does the L1 affect the L2, but the L2 can come to influence the L1, even at the level of the grammar. Dussias considers the possibility that the dramatic shift to low attachment for the native Spanish speaking bilinguals may be due to nature of their exposure to English in a predominantly English-speaking environment in the US. A critical question, and one currently under
investigation, is whether the frequency of exposure to the L2 or proficiency in the L2 per se is the key factor determining this pattern of performance.

A summary of the results on bilingual sentence parsing at this early stage in the research program suggests that almost all of the possible outcomes can be obtained. Bilinguals can sometimes adopt L2 structures as if they were native speakers of the L2, but in other contexts they are influenced by transfer from the L1 or the L1 itself is modified by the use of the L2. A clear goal in the next stage of research will be to identify the range of factors that modulates these different patterns and particularly the linguistic constraints associated with each outcome. There are a number of existing results that are apparently conflicting and the resolution of these differences will be important to the development of a comprehensive model of bilingual sentence processing. For example, the studies that have examined the effects of age of acquisition on sensitivity to syntax in the L2 (e.g., Hahne & Friederici, 2002; Johnson & Newport, 1989; Weber-Fox & Neville, 1996), suggest that there is a limit to the extent that late second language learners can fully acquire the syntax of the L2, even those who are highly proficient in the L2 and have lived in the L2 environment for many years. However, the bilinguals in the studies we have reported typically acquired their L2 late and yet were able to adopt native-like processing preferences in the L2 or, even more dramatically, to have the L1 take on the properties of the L2. It remains to be seen whether the degree of constraint in L2 sentence processing is determined by the type of structures that have been examined, by the degree of proficiency of the bilinguals, a factor that is difficult to assess perfectly, or by other contextual factors that influence the degree to which the native language is maintained actively when bilinguals live in an L2 environment. However these issues are resolved, however, the fact that it is possible to demonstrate that the L2 can influence the L1 is again consistent with the view that the bilingual’s two languages are open to interactions and that some of those interactions may have long-term consequences.

3.2. Cross-Language Syntactic Priming

Although most syntactic priming studies examine production rather than comprehension, we describe the main results of these studies briefly because those results converge closely with the evidence reviewed above on parsing. Syntactic priming is the phenomenon whereby the production of a target sentence is influenced by the syntactic form of a previously produced prime sentence (Bock, 1986). For example, for a monolingual speaker of English, the probability of producing a sentence describing a picture in active vs. passive voice is a function of whether a spoken prime sentence is active or passive. In the bilingual research, the question has been whether a switch of language, and therefore syntax, from prime to target sentence, will disrupt priming relative to the monolingual case. The few studies that have examined cross-language syntactic priming have reported effects of priming that are very similar to those found within language, suggesting that there is a common basis for this effect across languages (e.g., Hartsuiker et al., 2004), although there are some suggestions as well that the range of priming may be more limited across than within languages (e.g., Loebell & Bock, 2003). Recent work has shown that structural priming can be observed in cases in which lexical priming alone cannot
account for the results and that lower frequency or less-dominant structures are more sus-
ceptible to priming (e.g., Scheepers, 2003). In the cross-language case, this means that
syntactic priming will be more likely from L1 to L2, the more dominant to the less-dom-
inant language (Flett, Branigan, Pickering, & Sorace, 2005). Hartsuiker et al. argue that
the syntactic representations computed for each of the bilingual’s two languages access
the same abstract information, rendering the syntactic level integrated across languages
and open to code switching, the topic to which we turn next.

3.3. Other Approaches to Bilingual Sentence Processing

3.3.1. Understanding code-switched sentences

A surprising gap in the literature on bilingual sentence processing is that very few stud-
ies have investigated the comprehension of spoken code switches between languages.
Code switching is a phenomenon common within bilingual communities. However, most
of the research on this topic has been conducted from a linguistic rather than psycholin-
guistic perspective, with the goal to elucidate the syntactic constraints that govern
allowable switches (e.g., Muysken, 2000; Myers-Scotton, 2002). The few studies that
have examined the consequences of code switching for language processing have focused
on processes occurring primarily at the lexical level rather than the syntactic level. A
number of studies have examined lexical-level code switching while bilinguals read sen-
tences (e.g., Altarriba et al., 1996; Moreno, Federmeier, & Kutas, 2002). Although lan-
guage switching in comprehension tasks is an interesting phenomenon in and of itself for
what it tells us about how effectively bilinguals can use expectations about the language
they are processing to control subsequent language selection (e.g., Thomas & Allport,
2000), it can be argued that mixed language presentations during reading are rare,
whereas code switches in spoken discourse are common. It would therefore seem critical
to investigate this issue in spoken language contexts. Again, the few studies that have
examined code switches in speech, have also focused at the lexical level. To illustrate, a
number of experiments have asked how bilinguals comprehend a guest word spoken in
the non-target language (i.e., not in the language of a preceding sentence context). The
results of these studies support the conclusions of the word recognition research reviewed
earlier in showing that information about the non-target language is available even during
sentence processing, but demonstrate that the scope of activation of non-target alterna-
tives is constrained to some degree by cues available to the listener (e.g., Li, 1996;
Grosjean, 1988; Soares & Grosjean, 1984). An interesting observation is that code
switches in written sentence contexts are in some respects less disruptive to reading than
within-language lexical switches. Moreno et al. (2002) showed that the typical N400
effect observed in the ERP record when a lexical violation is encountered is greater when
within-language synonyms are presented than when a cross-language translation appears.
One interpretation of this result is that code switching is a relatively natural phenomenon,
a conclusion that is compatible with the high degree of parallel activity observed across
languages. It will be critical in the next phase of research to begin to examine syntactic-
level constraints in the online comprehension of code switches, particularly under cir-
cumstances in which the bilingual’s two languages differ syntactically.
3.3.2. **Reading for translation**

A final area of research in which sentence processing has been examined in bilinguals concerns the processes that are engaged when a translator is comprehending a sentence in one language for the purpose of producing it in the other language. The processes that support proficient translators in achieving real time simultaneous translation and interpretation are fascinating in and of themselves because they represent an extraordinary feat of cognition (see Christoffels & De Groot, 2005 for a recent review of the psycholinguistics of translation and interpretation). For present purposes, the question of interest with respect to translation, is how incoming material is comprehended as a function of how it will be used. Earlier research on sentence comprehension within the native language (e.g., Aaronson & Ferres, 1986) has shown that comprehension strategies are adapted to the goals of the task. Macizo and Bajo (in press) showed that when translators read for the purpose of translation, there is activation of information in the target language to be spoken during the comprehension process. When they read only for the subsequent task of repeating what they have read, there is little evidence of cross-language activation. These results suggest that task goals influence the degree to which both languages are active. Of particular interest is that translation makes greater demands on working memory resources than simple reading and those demands are reflected during the comprehension process, prior to actual production. One implication of the findings with translators is that the degree to which the other language is required in the larger discourse context appears to influence sensitivity to both languages during initial comprehension (see Grosjean, 2001 for a related argument about language mode). For translators, that context may be the likelihood of having to produce the currently processed sentence in the other language. For ordinary bilinguals, it may be related to the likelihood that they will interact with other bilinguals who are likely to switch into the other language completely or in part in code switched exchanges.

4. **FACTORS THAT INFLUENCE BILINGUAL COMPREHENSION**

In each area of research reviewed above, we focused on the performance of highly proficient bilinguals. However, comprehension is also open to the influence of a variety of factors that are likely to modulate performance. One factor includes the individual differences that affect the cognitive resources that are available to be recruited during language processing (see Perfetti, this volume for a review of this work within the native language). As we noted in the review of research on bilingual sentence processing, a great deal of evidence suggests that the L2 makes greater demands on memory and attention than the L1, even for relatively skilled bilinguals (e.g., Miyake & Friedman, 1998).

Another factor is the similarity between the two languages. Languages differ syntactically, morphologically, and phonologically and those differences are likely to affect the ease of cross-language comprehension in both listening and reading. In reading, languages also differ with respect to whether they are alphabetic and use the same or a
different script. A Dutch–English bilingual may easily mistake a written word in English
for Dutch, but a Chinese–English bilingual will never make that mistake. Although
differences across languages at each level modulate the pattern of processing (e.g.,
MacWhinney, 1997; Thomas & Allport, 2000; Vaid & Frenck-Mestre, 2002), there is
very little evidence that suggests that once an individual becomes relatively proficient in
the L2 they restrict the degree of cross-language interaction at either the lexical, sub-lex-
ical, or sentence levels or serve as a cue to maintain greater separation across the bilin-
gual’s two languages.

If language status does not readily enable language selection, then how do profi-
cient bilinguals effectively use the intended language without frequent intrusions from
the unintended language? In this chapter, we have focused our review on comprehen-
sion but there is also a recent literature on bilingual production that demonstrates that
language non-selectivity is not restricted to comprehension alone. Although spoken
production in bilinguals is initiated by an idea to be expressed in words and sentences,
a picture to be named or described, and words or sentences to be translated, the con-
ceptually driven nature of production does not itself appear to restrict activation to the
target language alone (e.g., Costa, 2005; Kroll et al., in press). The presence of mutual
activity across the bilingual’s two languages in both comprehension and production
suggests that another mechanism must be in place to allow attention to be appropri-
ately directed so that the correct language choices are made and, at the same time, that
systematic code-switching can be accommodated without incurring a significant pro-
cessing cost.

One solution to this apparent problem is to hypothesize that cognitive mechanisms
outside the linguistic representations themselves function to resolve the observed
cross-language competition, either by modulating the relative activation of the unin-
tended language or by actively inhibiting candidates from the non-target language
(e.g., see Green, 1998 for an illustration of how such an inhibitory mechanism might
operate in production). A model that incorporates an extra-linguistic mechanism for
bilingual word recognition has been described by Dijkstra and Van Heuven (2002).
The BIA+ model, shown in Figure 5, includes both a lexical identification system and
a task schema system. The lexical identification system, an embellishment of the BIA
model seen in Figure 3, represents lexical and sub-lexical information in each lan-
guage and their interactions. The lexical system is hypothesized to be encapsulated in
the sense that language-specific selection within the lexical system itself is possible
in response to linguistic context but not affected directly by more cognitive, non-lin-
guistic factors, such as expectations and instructions. On this view, bilingual word
recognition is fundamentally a data-driven process that is uninfluenced by top-down
factors until quite late in the process. The task schema system controls not only the
output of the lexical identification system with respect to the mapping of language
output to response processes, but also the manner in which language output is
weighted with respect to decision criteria. As such, the model can account for many
of the word-level phenomena we have reported. Bilingual word recognition is funda-
mentally language non-selective and even sentence context per se does not appear to
There remains a great deal of work to be done to determine to what extent the principles embodied within a model like BIA+/H11001 will serve as a general foundation on which a more comprehensive account of bilingual comprehension can be developed. However, it provides a useful basis on which to begin to identify those factors that constrain the basic architecture of the bilingual’s language system and those that reflect the manner in which bilingualism affects cognitive control. The recent work by Bialystok and colleagues (e.g., Bialystok, 2005; Bialystok et al., 2004) provides compelling evidence for the positive cognitive consequences that bilingualism appears to confer to young bilingual children and to elderly bilinguals in the realm of executive function. Bilinguals are not cognitively superior to monolinguals in general, but quite specifically in tasks that require inhibitory control in which irrelevant information or responses must be ignored. It is tempting to speculate that the control mechanisms of the sort included in Dijkstra and Van Heuven’s (2002) BIA+ model, required to
modulate proficient performance, may contribute to the development of the enhanced executive function in bilinguals.

5. CONCLUSIONS

In this chapter, we reviewed some of the recent research on the processing of words and sentences by bilingual speakers. The picture presented by our review suggests a language system that is highly permeable across the bilingual’s two languages, with information about words and grammatical structures activated concurrently even while a bilingual is reading text or listening to speech in one of his or her two languages alone. A very counterintuitive aspect of this body of research is that the activity of the unintended languages is not simply a matter of proficiency. Both languages appear to be active in even highly proficient bilinguals. Although much of this research is at a very early stage of development, it holds important implications for characterizing bilingual performance and for the way in which bilingualism provides a model more generally for investigating constraints and plasticity in language processing. At the heart of this review is the observation that bilinguals themselves are not special. To the contrary, more of the world’s population is bilingual than not and most of the cross-language interactions we have reviewed are related to phenomena observed within language in the presence of ambiguity. However, the presence of two active and competing languages makes the bilingual an especially informative source for psycholinguists interested in how cognitive systems compete and in how the resulting competition is resolved. We are confident that the contribution of this approach will be increasingly valuable in mapping the relations between language and cognition and their neural underpinnings.

ACKNOWLEDGMENT

The writing of this chapter was supported in part by NSF Grant BCS-0418071 and NIH Grant R01MH62479 to Judith F. Kroll.

REFERENCES


