A cognitive view of the bilingual lexicon: Reading and speaking words in two languages

Judith F Kroll
Bianca M Sumutka
Ana I Schwartz, University of Texas at El Paso

Available at: https://works.bepress.com/ana_schwartz/7/
A cognitive view of the bilingual lexicon: Reading and speaking words in two languages

Judith F. Kroll, Bianca M. Sumutka, and Ana I. Schwartz

The Pennsylvania State University

Acknowledgments
This research was supported in part by NSF Grant BCS-0111734 and NIH Grant MH62479 to Judith F. Kroll and NSF Dissertation Grant BCS-0212571 to Ana I. Schwartz and Judith F. Kroll.

Abstract
In this paper we review recent research on experimental psycholinguistic approaches to the bilingual lexicon. The focus in this work is to understand how it is that lexical access in both comprehension and production is fundamentally nonselective with respective to language, yet bilinguals are able to control the use of their two languages with relatively high accuracy. We first illustrate the nature of the data that support the claims of nonselectivity and then consider some of the factors that may modulate the resulting cross-language competition. These include differences in lexical parsing strategies across languages, in lexical cues that signal one language rather than another, in the ability to allocate cognitive resources, and in the nature of the tasks that initiate spoken production. We argue that the competitive nature of processing across the two languages of the bilingual provides an exquisite model to examine cognitive activity and its control.

Key words
- bilingual comprehension
- bilingual production
- lexical access

Introduction
Although proficient bilinguals rarely make the error of speaking words in the wrong language or thinking that they are reading text in a language other than the one intended, recent cognitive research on lexical access in word recognition and in spoken production suggests that information about both languages is active, at least briefly, in even highly skilled tasks such as reading and speaking (e.g., Brysbaert, van Dyck, & van de Poel, 1999; Colomé, 2001; Hermans, Bongaerts, de Bot, & Schreuder, 1998; van Heuven, Dijkstra, & Grainger, 1998). The absence of a simple mechanism to switch off one of the two languages when using the other makes the problem more complicated from the perspective of bilingual performance, but also more interesting from the perspective of elucidating cognitive mechanisms. For this reason, bilingualism has become an important tool for psychologists who wish to model developing systems, the competition

Address for correspondence
Judith F. Kroll, Department of Psychology, 641 Moore Building, Pennsylvania State University, University Park, PA 16802 USA, Phone: 814-863-0126, Fax: 814-863-7002, e-mail: jfk7@psu.edu

'International Journal of Bilingualism' is © Kingstone Press Ltd. 2005
between them, and the consequences for executive control (see Bialystok, this issue). Identifying the codes across the two languages that are activated by input from each provides important information about the degree to which language representations are constrained by the course of acquisition or open to later modification. Likewise, understanding how bilinguals negotiate the inherent nonselectivity of lexical access provides critical evidence for models of selective attention and control.

In this paper we first review the evidence on lexical access in bilingual word recognition and production. Because a number of recent papers and chapters provide extensive reviews of this material, we summarize these findings only briefly (see Brysbaert, 1998; Dijkstra & van Heuven, 2002; Kroll & Dijkstra, 2002; Kroll & Dussias, 2004). We then consider briefly the implications of the findings that we review for models of comprehension and production and examine some of the factors that we believe may be informative with respect to identifying the locus of language selection. These include linguistic information that provides cues to language membership, such as language-specific codes, differences in lexical parsing preferences, the influence of language context and nonlinguistic information that provides cues in the perceived environment of the bilingual. The latter includes the form of the tasks that initiate reading and speaking and the ability of the language user to allocate attention and memory resources to the intended language.

1.1

Lexical access in bilingual word recognition

Psycholinguistic studies of reading and word recognition within the native language have shown that during this process information becomes active not only for the target word itself but also for other words that share aspects of lexical form with the target word. In this sense, word recognition appears to be characterized as a parallel process in which information at different levels interacts until a single lexical candidate emerges (e.g., McClelland & Rumelhart, 1981). What happens then when the reader is a bilingual? Recent studies of bilingual word recognition have extended the demonstration of parallel activation of lexical form representations to information associated with each of the bilingual’s two languages. Thus, a Dutch-English bilingual reading words in English, the second language (L2), will briefly activate orthographic and phonological codes associated with lexical candidates in Dutch, although English is the target language (e.g., Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, van Harsveld, & ten Brinke, 1998; van Heuven et al., 1998). Even more surprising is that similar effects are observed when the more dominant first language (L1) is the target language (e.g., Jared & Kroll, 2001; Jared & Szucs, 2002; van Hell & Dijkstra, 2002) and when the bilingual’s two languages do not share the same alphabetic or orthographic form (e.g., Gollan, Forster, & Frost, 1997).

Word recognition can be characterized by a set of signature empirical results. In general, words are recognized more rapidly when they are frequent (Forster & Chambers, 1973) and when the mapping of spelling to sound is unambiguous (e.g., Jared, McAra, & Seidenberg, 1990). The presence of alternative pronunciations for the same orthographic input has been hypothesized to lead to increased processing time because lexical and/or sublexical codes compete and the process of resolving that
competition is assumed to take time. Likewise, word recognition is sensitive to the number of orthographic neighbors that resemble the target word. In languages like English that have deep orthography and many exception words, larger orthographic neighborhoods typically facilitate response times. In languages like Spanish or Dutch that have shallow orthography, larger orthographic neighborhoods typically inhibit response times (see Andrews, 1997, for a review).

The logic of research on bilingual word recognition has been to exploit the cross-language similarity of words as a means to determine whether these signature effects in word recognition are observed across languages as well. To illustrate, in languages that share the same alphabet and aspects of the orthographic representation, there are often words that resemble one another. Sometimes these words also share the same meaning, in which case they are considered cognates (e.g., bed in English and Dutch). In other cases they correspond to different meanings and are considered homographs or false friends (e.g., room in English means cream in Dutch). To determine whether word recognition for the bilingual is selective with respect to language, past studies have manipulated these cross-language properties to determine whether the unintended version of the word is also activated even when the task requires recognition in one language only. For example, Dijkstra et al. (1998) reported a series of experiments in which Dutch-English bilinguals performed variants of the lexical decision task (i.e., is the input letter string a real word?). When the task was purely in English, the L2 for these bilinguals, there was little evidence of activation of the alternative sense of homographs but there was significant facilitation for cognates relative to controls. When the lexical decision task was modified by including real word distractors from the non-target native language, there was a clear inhibitory effect for the homographs relative to unambiguous controls. Moreover, the word frequency effects that are typically observed in these tasks appeared to modulate the magnitude of homograph interference so that it was particularly difficult for a Dutch-English bilingual to accept an English word that was low frequency in English but high frequency in Dutch. When the task was then modified into generalized lexical decision, with the instruction to accept any letter string that is a real word in either language, the homographs produced facilitation relative to controls. In all, this pattern of results suggests that the bilingual lexicon is fundamentally nonselective with respect to language.

The results of other studies using different aspects of cross-language similarity converge on the same conclusion. When bilinguals read words in one of their two languages, information about the orthography, phonology, and meaning of words in the other language becomes active. The extent of this language nonselectivity appears to depend on a range of factors that influence the relative activation of the two languages and the manner in which the output of the lexical system itself interacts with the goals that are instantiated for a particular task (see Dijkstra & van Heuven, 2002; von Studnitz & Green, 2002).

1.2
Lexical access in bilingual word production
Finding that word recognition is language nonselective is not surprising given the bottom-up nature of early perceptual processes. What is more counter-intuitive is that
similar evidence for nonselectivity has been demonstrated in language production, where spoken words are the result of a process that is initiated by a conceptually-driven event. Logically, it would seem that the intention to speak an idea or to name a pictured object in a specified language should be under the control of the speaker. However, even under these circumstances, there is evidence that the intention to speak a word in one language only does not eliminate the activation of related words in the unintended language. Like word recognition, the evidence for nonselectivity in production is based on demonstrations of cross-language activity. However, unlike word recognition, the top-down conceptually driven nature of production makes it more likely that the competing alternatives that are active in the nontarget language consist of words related to the meaning of the intended utterances, rather than words that share lexical form.

Much of the research on spoken word production in bilinguals has used a cross-language version of the Stroop (1935) task in which a picture has to be named in one language and a distracting word is presented at some point before, during, or after the picture's presentation. Many past studies on lexical access in language production within a single language have used the time course of these distractor effects to map out the stages of production and their associated time course (e.g., Levelt et al., 1991; Schriefers, Meyer, & Levelt, 1990). The typical finding is that semantically related distractors produce interference in picture naming, whereas phonologically related distractors produce facilitation. The issue for bilingual production then is whether the same effects observed within language are also observed across languages. If lexical access is selective, then distractors in the nonselected language should have little effect. However, if lexical access is nonselective, then nontarget language distractors should not only influence performance, but the locus of their effects should tell us something about the level at which cross-language competition is resolved.

A series of bilingual Stroop studies have reported evidence for cross-language interactions. When a picture is accompanied by a spoken or written distractor word, there is interference when the word is semantically related to the picture's name and facilitation when the word is phonologically related to the picture's name, regardless of the language in which the word appears (e.g., Costa & Caramazza, 2000; Costa, Miozzo, & Caramazza, 1999; Hermans, 2000; Hermans et al., 1998). Moreover, the time course of these effects is, for the most part, similar to those observed within language. To isolate the locus of language selection in production, these studies have included distractors that are phonological relatives of the translation of the picture's name. This condition revealed limits in the degree of nonselectivity in that, if anything, phonological relatives of the translation produce interference, like semantic distractors (Costa et al., 1999; Hermans et al., 1998).

Hermans et al. (1998) interpreted their bilingual Stroop results as evidence for a language nonselective model of production in which selection eventually occurs following competition among lexical candidates but prior to specifying the phonology for any of those alternatives. Costa et al. (1999), on the basis of similar data, argued for a language selective model, but one in which there is activation among nontarget alternatives that does not lead to selection.

A problem in reaching firm conclusions on the basis of the bilingual Stroop data is that the Stroop task necessarily involves the presentation of a distractor word in
addition to the primary production task. Because the distractor word itself will initiate bottom-up processing, the resulting pattern of data can be understood as a meeting of the bottom-up process associated with the distractor and the top-down processing associated with the primary task. The concerns associated with the Stroop task make it critical to seek converging evidence using other tasks.

A number of recent studies have examined production outside of the Stroop paradigm. Kroll, Dijkstra, Janssen, and Schriefers (in preparation) developed a cued picture naming paradigm in which bilinguals named pictures in one of their two languages following the presentation of a tone cue. In the mixed language condition of this study, the language of naming was uncertain from trial to trial, so that a high tone signaled one language and a low tone signaled the other. In addition, the tone could be presented at one of three SOAs with respect to the onset of the picture. In the blocked language condition, one of the tones signaled naming consistently in one language, and the other tone signaled a “no” response. The logic of cued picture naming is to force both languages to be active in the mixed condition and then to determine whether this requirement has consequences relative to the blocked condition. If L1 is active when naming in L2, then deliberately forcing it to be active should have little consequence for performance. To determine the locus of language selection, picture naming was compared for pictures whose names are cognate translations with those whose names are noncognate translations (e.g., bed and bed versus bike and fiets in English and Dutch). If both language alternatives are active to the level of the phonology, then pictures whose names are cognates should be named faster than pictures whose names do not share phonology (see Costa, Caramazza, & Sebastian-Galles, 2000, on cognate facilitation in simple picture naming).

In experiments with Dutch-English bilinguals and English-French bilinguals, Kroll et al. (in preparation) found that there were indeed cognate facilitation effects in picture naming under the mixed conditions for both L1 and L2. When picture naming was blocked, however, cognate facilitation was restricted only to L2 and only at short SOAs relative to the picture’s onset. The results suggest that L1 is normally active during L2 production and that activation extends beyond competition among lexical candidates to the specification of phonology. Requiring L1 to be active extends the time course of these effects, but not the presence of cognate facilitation. The fact that cognate facilitation was also observed for L1 under the mixed conditions suggests that even the more skilled L1 is open to the influence of L2. The absence of those influences in blocked naming suggests that language selection can occur for L1 at an early point in time. Overall, these data support a nonselective model of lexical access in production in which the relative dominance of the language determines the time course, and therefore locus, of selection.

Other recent studies also provide support for activation of lexical alternatives in both of the bilingual’s two languages. Colomé (2001) used a phoneme monitoring paradigm in which Catalan-Spanish bilinguals indicated whether a presented letter appeared in the Catalan name of a picture. The critical result was that bilinguals were slower to reject a letter not in the picture’s name when it was present in the translation equivalent. This finding again suggests that the phonological representation of the translation is available even when speaking one language only. Additional evidence for
activity of lexical candidates in both languages during production comes from studies of tip-of-the-tongue states in bilinguals (e.g., Gollan & Silverberg, 2001) and deliberate language switching (e.g., Meuter & Allport, 1999).

**Modulating cross-language competition**

The results we have reviewed suggest that the lexicon allows open exchange between the codes activated by words in each of the bilingual’s languages. At some point, however, that activation must be controlled; otherwise bilinguals would make errors of language and interpretation far more often than they do. To understand how the inherent nonselectivity of lexical access in bilingual word recognition and spoken production is controlled to allow fluent performance in reading and speaking, we examine illustrative models of the bilingual lexicon. We then consider how each of the models addresses the language selection problem and identify a set of factors that may potentially provide a solution to the problem.

2.1 **Models of bilingual word recognition and production**

Figure 1 compares two models of the bilingual lexicon. Figure 1a is the Bilingual Interactive Activation model (BIA) proposed by Dijkstra and van Heuven (1998). In the version illustrated, a Dutch-English bilingual is reading the word *bike* in English. The model, adapted from the McClelland and Rumelhart (1981) model for monolingual word recognition, assumes a similar bottom-up flow of information from visually presented words so that letter features, letters, and words compete for recognition. The BIA model includes an additional feature, the language nodes, to identify the target language. The language nodes collect evidence about the presence of words in each language but also function to inhibit the other language. In this version of the model, the language nodes are sensitive to top-down influences so that information about the context in which the word appears and about the relative dominance of the two languages may affect the activation of the language nodes and influence lexical selection.

The BIA model can handle many of the results that have been reported in the literature on bilingual word recognition. The initial stages of recognition are thought to be perceptually driven. Observed interactions across lexical codes activated in each language cannot be prevented, but only controlled by the later activity of the language nodes. Although the original BIA model was restricted to orthographic interactions across words in the two languages, later research suggests that similar interactions can be observed for the phonology (e.g., Brysbaert et al., 1999; Dijkstra et al., 1999; Jared & Kroll, 2001; Jared & Szucs, 2002). Whether performance appears to be language selective or not in any given task is then thought to be a reflection of the way in which nonlexical processes use the output of the lexical identification system to enable particular responses to be initiated. The lexicon itself is always assumed to be nonselective for language, but its output may be used in a manner that fails to reveal the influence of the unintended language (Dijkstra & van Heuven, 2002).

Figure 1b is a model of language production adapted from those proposed by Poulisse and Bongaerts (1994) and Hermans (2000). Like models of speech production
Figure 1
Two models of the bilingual lexicon. Figure 1a is the Bilingual Interactive Activation (BIA) model (adapted from Dijkstra et al., 1998) of word recognition. Figure 1b is a production model (adapted from Hermans, 2000, and Poulisse & Bongaerts, 1994). It uses the naming of a picture in one language to illustrate the problem of lexical selection.

Models of the bilingual lexicon

Figure 1a
Visual Word Recognition

Figure 1b
Spoken Word Production

within a single language (e.g., Levelt, Roelofs, & Meyer, 1999), the stages of processing that lead up to a spoken utterance are thought to be initiated conceptually, so that planning of the utterance proceeds from meaning to form, over time. In the model that is illustrated in this figure, the production process is initiated by a presented picture. Here, a Dutch-English bilingual is attempting to name a picture of a bicycle as the word ‘bike’ in English. The model assumes that the identification of the picture results in the activation of its respective conceptual representation, the lemmas or abstract lexical representations that correspond to words that may reflect the activated concept, and then the phonology associated with the active lexical forms. In the case of production within
a single language, there is already a debate as to whether selection occurs discretely at each of these levels (e.g., Bloem & La Heij, 2003; Levet et al., 1999; Peterson & Savoy, 1998). For the bilingual, there is the added problem that alternatives may be active in both languages at each of these stages.

Like the language nodes in the BIA model, the bilingual production model includes a language cue, thought to operate at the conceptual level to represent the intention to speak a word in one language or the other. The evidence reviewed above on production demonstrates that the language cue, in and of itself, is not sufficient to restrict activation to the target language only. One possibility is that the language cue simply increases the activation of lexical candidates in the target language but does not actually inhibit lexical candidates in the nontarget language. If the target language is already the more dominant language or L1, then the language cue may suffice to allow production to proceed selectively because the alternatives in L2 will not be active enough to intrude. However, if the target language is the less dominant language, there may still be competition arising from the more dominant alternatives in the L1.

Although models of bilingual word recognition and production such as those shown in Figure 1 do a reasonable job of explaining the evidence for language nonselectivity, there are only a few studies that have systematically considered the manner in which selection occurs, and consequently how a mechanism such as the language nodes of the BIA model or the language cues of the production model actually works. In the section that follows we discuss some of the factors that may be critical in providing clues as to how selection occurs. These include: (1) specific attributes of the bilingual’s languages, (2) processing strategies uniquely enabled by language-specific structures, (3) linguistic and extralinguistic context, (4) characteristics associated with the bilingual individual that may affect the manner in which attention and memory are allocated, and (5) properties of the tasks that are used to assess recognition and production. Although the research on these issues has not been sufficient to draw strong conclusions, our aim is to provide a preliminary framework that may guide future investigations on this problem.

2.2 Factors that modulate cross-language interactions in word recognition

Language attributes

Much of the research on bilingual word recognition has been performed on languages such as Dutch and English that share a common alphabet and aspects of the orthography and phonology. Do these shared properties determine the degree to which lexical access is language nonselective? At one level, it would seem appealing to demonstrate that when a bilingual’s two languages differ in script, or if one is an alphabetic language and the other not, surface features in the perceptual input would serve as a cue to language. Certainly, a Chinese-English bilingual does not experience any overt ambiguity about language status the way a Dutch-English bilingual might when presented with an English word to read. Despite the seemingly obvious differences that may facilitate language selection when a bilingual’s two languages correspond to distinct written forms, the evidence on bilingual word recognition suggests that cross-language interactions persist. For example, Gollan et al. (1997) demonstrated masked priming between Hebrew and
English translation equivalents, with larger effects when the translations were cognates than not. Because Hebrew and English do not share the same alphabet, cognate effects can only be attributable to the activation of shared phonology. Likewise, Jiang (1999) found persistent translation priming from Chinese to English in a masked priming paradigm. If these masked priming effects reflect access to shared semantics across the bilingual's two languages, then perhaps they would not seem terribly surprising. However, the literature on masked priming suggests that the masked priming paradigm itself is not terribly sensitive to semantics, especially when lexical decision is the task performed. These results suggest that information in both languages is active regardless of the presence of surface cues.

Other studies examining languages whose lexical representations differ in more subtle ways have also come to a similar conclusion. An early study on the role of language-specific orthography showed that language switching costs were reduced in a mixed English and French lexical decision task when the words had language specific properties (Grainger & Beaucrillant, 1987). Although that result would seem to suggest that properties of the input can be used to select language, Thomas and Allport (2000) later demonstrated that it was due to the presence of a confounded variable. When this variable was removed, there were independent effects of language switching and language-specific orthography. Although language-specific orthography facilitated response latencies, it did not overcome the cost of switching (see also Rodriguez-Fornells, Rotte, Heinze, Nösselt, & Münte, 2002, and Vaid & Frenck-Mestre, 2002).

Language-specific processing strategies and constraints

Languages differ in the consistency with which orthography can be mapped to phonology. In languages such as English, that have a deep orthography, there is not a consistent mapping between spelling and sound, whereas in languages such as Spanish or Dutch, the presence of a shallow orthography allows a reliable mapping between spelling and sound. To the extent that differences in orthographic depth correspond to distinct lexical parsing strategies, we might expect that bilinguals for whom the two languages vary in orthographic depth would have to adjust their processing strategy to accommodate language differences.

Cross-linguistic studies of native or monolingual speakers reveal differences in the grain size or preferred unit of analysis for languages which differ in orthographic depth. In English, the orthographic rime, the first vowel plus remaining letters in a one-syllable word, appears to be an important unit of processing. A series of statistical analyses have found that English vowel pronunciation is better predicted by the letters that follow the vowel than the letters that precede the vowel and also than the vowel alone (Kessler & Treiman, 1997; Treiman, Mullinnix, Bijelic-Babic, & Richmond-Welty, 1995). Languages that have a more consistent spelling-to-sound mapping do not demonstrate the same orthographic rime effects as English. Martensen, Maris, and Dijkstra (2000) conducted a statistical analysis of Dutch similar to Treiman's. They found that the rime unit did not predict naming latencies of Dutch words above and beyond the smaller sublexical units of the onset, nucleus and coda.

Other cross-linguistic studies provide converging evidence for the idea that differences in orthographic depth may guide different strategies for lexical parsing.
Ziegler, Perry, Jacobs, and Braun (2001) examined cross-language differences for rime neighborhood size and word length between English and German using a reading aloud procedure. Comparing a list of cognates, nearly identical in both languages, they found that English native speakers showed a stronger rime neighborhood effect than German native speakers. However, German speakers were slower to name longer words than English native speakers.

Distinct lexical parsing preferences across languages present an interesting problem for the bilingual and something of a paradox given the evidence on nonselective access during word recognition. On one hand, the bilingual appears to activate information in parallel for both languages. On the other, the two languages may require different analytic strategies that could, in principle, be used as a basis on which the activity of the non-target language might be controlled. Studies of syntactic parsing in bilinguals (e.g., Dussias, 2001) suggest that although there may be some transfer of parsing preferences from the L1 to the L2, there are also influences of L2 proficiency on preferences within the native language itself. A key question is whether lexical parsing preferences are constrained by the structure of the language itself or whether they change in response to the acquisition of a second language.

Only a few studies have examined lexical constraints in bilinguals and most of them have addressed aspects of speech perception. Cutler, Mehler, Norris, and Segui (1989) compared auditory parsing strategies in French-English and English-French bilinguals. The comparison of French and English is of interest because French has clear syllable boundaries whereas English is often ambiguous. Participants were asked to listen for phonemes in French words in which these phonemes fell on a syllable boundary or not. Cutler et al. found French-English bilinguals were faster at identifying the phonemes when the segment was a syllable within the word than when it was not. However English-French bilinguals did not show this benefit. The native English speakers appeared to not have acquired a syllable parsing strategy for French although they were highly proficient French speakers.

In a recent study, Pallier, Colomé, and Sebastián-Gallés (2001) reported similar constraints in an auditory priming task for highly proficient Spanish-Catalan and Catalan-Spanish bilinguals who differed in the order of acquisition of the two languages. The critical stimuli were Catalan primes that differed from a target word by a phonological feature not distinctive in Spanish. Spanish-English bilinguals treated these critical items as if they were homophones. However Catalan-Spanish bilinguals, for whom the contrast was functional, did not show any priming for these critical pairs. The results suggest that these proficient Spanish-Catalan bilinguals were processing Catalan using the phonological constraints of Spanish.

What are the consequences for language selection? One way to reconcile the apparent paradox of language nonselectivity with language-specific processing constraints is to assume that lexical parsing strategies operate over the output of the lexical identification system, following interaction among sublexical units. On that account, parsing would be a late process and therefore open to the decision mechanisms that sort out language membership. However, if the cross-language interactions observed in bilingual word recognition reflect the earliest stages of bottom-up activation and are unaffected by language-specific information, and if the units activated are similar across language,
the reason parsing differences persist would remain unclear (see Dijkstra & van Heuven, 2002, and Green, 2002). In future research it will be important to evaluate the consequences of bilingualism for lexical parsing and to determine whether parsing strategies influence language selection during reading.

**Context**

Most research on bilingual word recognition has used out-of-context tasks in which responses are made to single words. Only a few studies have asked whether the degree of language nonselectivity is constrained by context. The answer to this question is critical for testing assumptions about the architecture of the lexicon and for identifying factors that constrain or modulate cross language interactions.

Dijkstra, de Bruijn, Schriefers, and ten Brinke (2000) examined the question of whether prior knowledge about the task conditions affects the degree of language nonselectivity. In a previous study on interlingual homograph processing (Dijkstra et al., 1998), interference was observed reliably for homographs relative to unambiguous control words in an English lexical decision task when real Dutch words were included among the nonwords. That is, it was difficult for Dutch-English bilinguals to accept a homograph as English when they simultaneously had to reject unambiguous Dutch words as not English. Dijkstra et al. (2000) altered a subtle aspect of the earlier experiment to assess the role of prior knowledge in the results. In the new experiment, bilinguals were told that they would see Dutch words among the distractors. However, without their knowledge, real Dutch words were never presented in the first half of the study. The results showed that in the absence of the Dutch words, there was no homograph interference, but as soon as a single Dutch word was actually presented, the homograph interference effect was obtained. Because the participants expected the Dutch words throughout, Dijkstra et al. (2000) argued that instructions per se are not sufficient to alter processing strategies. In a related study, von Studnitz and Green (2002) showed that informing bilinguals about the presence of interlingual homographs did modulate the magnitude of the effects in a lexical decision experiment similar to Dijkstra et al. (1998). However, a number of differences between these studies make it difficult to determine whether the results are simply conflicting or whether aspects of the participants’ bilingualism influenced the time course of processing and subsequently the degree to which activity of the nontarget language could be controlled.

Two types of studies have examined the effects of linguistic context on cross-language interactions. Semantic priming experiments ask whether presenting a prime word in one language will prime performance in the other language. In this case the context is restricted to a single word in one or the other of the bilingual’s languages. The results of a large number of cross-language semantic priming studies (e.g., Chen & Ng, 1989; Keatley, Spinks, & de Gelder, 1994; Meyer & Ruddy, 1974) demonstrate that semantic priming is possible between languages, even when the prime is masked so that participants are not aware of its presence of language membership. However, these studies do not directly address the question of whether semantic context overrides cross-language interactions at other levels.
Only a few studies have asked whether sentence context overrides the cross-language interactions between lexical codes that have been observed in out of context word recognition (e.g., Altarriba, Kroll, Sholl, & Rayner, 1996; Elston-Güttler, 2000; Schwartz, 2003; van Hell, 1998). Because sentence context contains lexical, syntactic, and semantic information, as well as information about language status, it will be important not only to demonstrate context effects but to isolate the locus of these effects. A preliminary finding in some of these studies suggests that the language of the context per se does not reduce cross-language activation (see also de Brujin, Dijkstra, Chwilla, & Schriefers, 2001) but that sentence context may modulate the influence of the nontarget language (for evidence on comprehending mixed language see Grosjean, 1995; Li, 1996; Moreno, Federmeier, & Kutas, 2002).

Van Hell (1998) and Schwartz (2003) each found evidence that in the context of a highly constrained sentence, cross-language activity was reduced. In the van Hell study, it was the presence of cognate facilitation in lexical decision that was modulated. In the Schwartz study it was the presence of cross-language phonological activation. In each of these studies, the cross-language effects were present out of context and in low constraint contexts. If this pattern proves to be reliable, it suggests that top-down semantic activation may override or dampen the output of these processes. This is contrary to the claims of models such as BIA, according to which interactions among lexical codes are driven only by bottom-up processes (see Dijkstra & van Heuven, 2002, for an illustration of how the BIA + model may be able to accommodate the influence of linguistic context).

**Characteristics of the bilingual**

Most of the research on the nonselectivity in bilingual word recognition has focused on relatively proficient bilinguals. Although there are differences in the nature of the bilingualism of the participants in the studies reported in the literature, in most cases these bilinguals are late bilinguals who acquired L2 after early childhood and are highly proficient in L2 but still dominant in L1. Two characteristics of the bilingual have been linked to the findings on language nonselectivity. One concerns the bilingual’s level of proficiency in the L2 and the other concerns individual differences in memory capacity and the ability to allocate attentional resources during language processing.

How does proficiency in L2 influence the degree of cross-language competition? Only a small number of studies have examined this question directly (e.g., Bijeljac-Babic, Bardeau, & Grainger, 1997; Jared & Kroll, 2001; Sunderman, 2002). Initially, the effects of L1 on L2 are greater than those of L2 on L1. With increasing proficiency, the effects become more similar, although typically still larger when L2 is the primary language of the task. Contrary to the commonsense notion that with increasing skill in a second language there will be greater automaticity and therefore greater independence, the evidence suggests that even highly proficient bilinguals continue to show evidence of the nontarget language’s influence. Sunderman (2002), in a study of the effects of L1 on L2 in a sample of native English speakers learning Spanish, found that all learners, regardless of their level of L2 proficiency, were sensitive to cross-language form relations and also, to some degree, to the meaning of L2 words. However, only learners at the very earliest stages were sensitive to words that were related in form to the translation.
equivalent of the L2 word. Sunderland's results provide support for the data-driven account of lexical form activation described by the BIA model, but further suggest that the mapping of form to meaning changes over the course of development (e.g., Kroll, Michael, Tokowicz, & Dufour, 2002; Kroll & Stewart, 1994).

If the data-driven account of bilingual word recognition is correct, then one might not expect that the degree to which bilinguals have available cognitive resources would modulate the degree of cross-language competition. One recent study has addressed this question directly by asking whether the homograph interference reported in past studies in lexical decision would differ for individuals as a function of their relative memory capacity (Michael, Dijkstra, & Kroll, 2002). Michael et al. replicated Experiment 2 of Dijkstra et al. (1998). In addition, participants were given a memory span task (Daneman & Carpenter, 1980) in Dutch and a translation production task. The results showed that higher span bilinguals were faster than lower span bilinguals to translate words. However, there was no effect of span on the degree of homograph interference in the lexical decision task. These findings suggest that memory span is related to language performance, but does not constrain lexical competition in data-driven tasks, supporting the claims of the BIA and extended BIA+ models.

A further issue to consider is that any factor that potentially affects the speed of processing, as proficiency and cognitive resources surely do, will also affect the time course over which cross-language interactions extend. A challenge for future research will be to develop methods to identify effects that result from an extended time course per se from those that reflect differences in the nature of the information available (see also McElree, Jia, & Litvak, 2000).

Properties of the task
Although bilingual lexical activation appears to be fundamentally language non-selective, the observed effects of this activation vary considerably across different tasks. For example, several studies reviewed earlier included interlingual homographs as critical materials to test for language non-selectivity (Dijkstra et al., 1998; Dijkstra et al., 1999; Gerard & Scarborough, 1989; Jared & Szucs, 2002; von Studnitz & Green, 2002). Those studies showed that the pattern of results for interlingual homographs depends on the list composition within the lexical decision task. By manipulating the conditions of the task itself, it becomes possible to obtain a pattern of results consistent with either language selectivity or nonselectivity. Only by taking the different conditions together is it possible to come to a clear conclusion about the underlying mechanism.

An important aspect of any psycholinguistic task is the nature of the response that is required from the bilingual participant. In lexical decision, participants can make a response without necessarily selecting a single lexical candidate (and see von Studnitz & Green, 2002, for an analysis of how the processing of response mapping may account for some of the observed effects). In naming, however, the specific lexical entry must be selected in order to produce the appropriate phonological code.

Although lexical decision tasks engage phonology (e.g., van Orden, 1987), the requirement to specify the phonology associated with a single lexical entry makes the naming task more likely, a priori, to reflect language-specific properties of the target language. Despite these differences, the evidence on word naming in the L1 and L2
converges, for the most part, with the evidence on lexical decision (e.g., Jared & Kroll, 2001; Jared & Szucs, 2002; Schwartz, Kroll, & Diaz, submitted). For example, Schwartz et al. showed that the time to name cognates in both L1 and L2 was influenced by the match between orthography and phonology across languages. Thus, even in a task which requires that a single lexical entry be specified phonologically, activation of information in the nontarget language influences performance.

2.3 Factors that modulate cross-language interactions in word production

Language attributes

Because production, unlike recognition, is initiated by conceptually driven processes, the manner in which language attributes may influence language selection will necessarily reflect different aspects of language representation. If the conceptual representation of ideas or objects to be named was distinct for the bilingual’s two languages, then that information itself might be used to identify the language to be selected. However, although the cultural and linguistic context in which ideas are generated may differ across languages, there is very little psycholinguistic evidence to suggest that meanings are distinct for the bilingual's languages (e.g., Francis, 1999). Moreover, for the sort of single word production experiments that we reviewed, when naming is initiated by a simple line drawing of an object, there is little information in the object itself that would identify language-specific contexts. It is certainly possible that language selection might be sensitive to such a distinction, but there is little research that has tested that hypothesis. For the pictures whose names are concrete nouns, the empirical results suggest that shared meanings are activated across the bilingual’s two languages (e.g., Costa et al., 1999; Dufour & Kroll, 1995; Hermans et al., 1998, and see de Bot & Schreuder, 1993, for a suggestion about how reduced access to semantics in L2 may impact lexicalization in production).

In cross-linguistic research on language production, there is some evidence, for example, that grammatical gender may sometimes play an important role in constraining selection of the appropriate determiner (e.g., Schiller & Caramazza, 2003). However, the role of a grammatical feature such as gender has not been systematically investigated in the bilingual case. Again, because most of the bilingual research has relied on tasks which require only the naming of bare nouns, there is little evidence to evaluate the manner in which language-specific grammatical properties might influence lexical selection in production. Likewise, there is little evidence regarding the issue of whether bilinguals whose two languages differ in phonological features, such as stress or prosody, can use that information in advance to direct lexical access.

Language-specific processing strategies, constraints, and context

We combine the factors of processing strategies and constraints with context because we believe that this is an area in which there is potential for examining the issue of language selection in lexical production and, at the same time, bringing together research on linguistics and psycholinguistics. Although there has been a great deal of linguistic research on code switching in bilingual production, particularly from a syntactic perspective (e.g., Myers-Scotton, 2002), there has been very little corresponding psycholinguistic research to examine the time course and processing constraints associated
with the production of switches. For example, it will be interesting to understand how the production of deliberate code switches affects the locus of lexical selection. The few psycholinguistic studies of language mixing or switching in context have focused primarily on comprehension rather than production (e.g., Alterrribe et al., 1996; Grosjean, 1995; Li, 1996; Moreno et al., 2002).

Outside of linguistic context, a number of recent studies have examined language switching that is induced by the experimental conditions (e.g., Hernandez, Martinez, & Kohnert, 2000; Meuter & Allport, 1999). Meuter and Allport found a switch cost in number naming that was greater when switching into the more dominant L1 than into the less dominant L2. According to models of inhibitory control (e.g., Green, 1998), the greater switch cost is observed for the stronger language because on the preceding trial the naming in the weaker language required inhibition of the more active L1. On this view, there would be parallel activation of lexical candidates in both L1 and L2, with the degree of activation determined by the dominance of the language and the bilingual's L2 proficiency. Like word recognition studies showing that the language of the context does not affect lexical selection, in language switching, knowledge of an upcoming language switch does not appear to eliminate switch costs, suggesting again that it cannot be used as a basis on which to reduce activation of candidates in the unintended language.

**Characteristics of the bilingual**

The two properties of bilingual individuals we considered relevant to word recognition, proficiency in the L2 and available cognitive resources, are also critical for production. For present purposes, the question is not simply how proficiency and the allocation of cognitive resources affect production, but how they specifically affect the ability to select the language of production. These two aspects of the bilingual are potentially related to one another because as individuals become more proficient in the second language, they also acquire information about the contexts in which each language is used and potentially the skill to attend to distinctive cues.

As individuals become more proficient in L2, the ability to lexicalize concepts into L2 words becomes faster and more accurate (e.g., Kroll et al., 2002). To the extent that stronger lexical alternatives in L1 compete with weaker lexical alternatives in L2, production for L2 will necessarily be slower and more vulnerable to cross-language competition. But does a high level of L2 skill allow earlier selection of language-appropriate candidates? The available studies suggest not, at least for the less dominant L2. Even highly proficient bilinguals appear to activate lexical and phonological information about L1 alternatives.

If cross-language competition persists beyond early the stages of second language learning, then the ability to modulate that competition may be critical in order to successfully achieve proficiency. A recent set of studies has addressed this issue by asking how individual differences in cognitive abilities affect bilingual performance (e.g., Michael & Gollan, 2005). Again, our immediate purpose is not to examine this question broadly, but to consider whether differences in the ability to allocate attentional and memory resources relate to language selection.
Kroll et al. (2002) compared the performance of a group of L2 learners who differed in reading span (Waters & Caplan, 1996). The span task, administered in L1 measured the ability of these learners to process and retain information simultaneously. The learners then performed a simple word naming task and a word translation task. Of interest are the results for translation production. For words whose translation were noncognates, the higher span learners were faster to translate than the lower span learners. However, for words whose translations were cognates, the higher span learners were significantly slower than the lower span learners. In all, the lower span learners showed a significant benefit for cognate translations, whereas for the higher span learners that benefit was substantially reduced. The results suggest that the lower span learners exploited the lexical transparency of the cognates to translate whereas the higher span learners apparently avoided that strategy. On the basis of these results alone, it is difficult to identify the locus of the observed interaction between span and the cognate effect in translation. However, unlike the results reported by Michael et al. (2002) demonstrating no effect on the modulation of interlingual homograph interference in lexical decision, the results of Kroll et al. suggest that span does affect performance in a production task.

*Properties of the task*

Most of the available evidence on production has been generated by only a few tasks. They include simple and Stroop-type picture naming and simple and Stroop-type translation production. The assumption has been that any task that requires that concepts be lexicalized into words in one of the bilingual's two languages will similarly reflect the planning and selection that occurs prior to speaking a single word. Indeed, a number of studies provide evidence that suggests that picture naming and translation share many critical features (e.g., Kroll & Stewart, 1994; Potter, So, von Eckhardt, & Feldman, 1984).

One reason picture naming may so easily reveal the parallel activation of lexical alternatives in both of the bilingual's languages is that the pictures used in these experiments, typically simple dictionary-like line drawings, do not depict culturally-specific features that might serve as a cue to language selection. In the absence of a language cue, the intention to speak the picture's name in one language alone may be insufficient to restrict activation to the alternatives in the target language. In contrast to picture naming, the word that itself initiates the planning of an utterance in the translation task may also serve as a language cue. If the task is to translate the word into the other language, then the bilingual has information about the language not to speak. If the translation task presents a language cue, then we might hypothesize that the selection of a language-appropriate lexical candidate may occur earlier in production for translation than for picture naming.

Miller and Kroll (2002) tested the language cue hypothesis in translation by comparing the two distractor conditions that have typically been compared in the picture-word Stroop task. In the picture task, a distractor word can appear in either of the bilingual's two languages and still produce semantic interference and/or phonological facilitation regardless of the language of naming (e.g., Costa et al., 1999; Hermans et al., 1998). In the translation variant of the Stroop task, a word is presented for translation.
along with a semantically or phonologically related distractor. In past studies (e.g., La Heij et al., 1990), the distractor has always been presented in the same language as the spoken output. The results under these conditions closely replicate those of the picture Stroop task, with semantic interference and phonological facilitation. Miller and Kroll first replicated the pattern reported by La Heij et al. and then asked whether the same pattern would hold when the distractors were presented in the language of the input word rather than the language of the output word. If the language of input can function as a cue to language selection, then distractors in the input language should have little effect on production performance and this is indeed what happened. When the distractor appeared in the language of input, semantic interference and phonological facilitation were eliminated. Miller and Kroll interpreted the results as support for the hypothesis that in the presence of an appropriate language cue, bilinguals can effectively reduce cross-language competition and restrict lexical selection to candidates within the language to be spoken. One implication of these findings is that language cues may be available in a range of linguistic, perceptual, and conceptual contexts.

3 Conclusions

In this paper we reviewed the recent literature on the bilingual lexicon. The evidence suggests that in both word recognition and production there is language nonselectivity such that competition across activated information in the two languages interacts prior to selection. The nature of the activated information differs, however, for recognition and production (see also Kroll & Dijkstra, 2002). In recognition, aspects of lexical form and its representation in the two languages are critical. In production, meaning-based representations in the form of the translation equivalent or semantic relatives are initially active and, depending on the time course in planning the utterance, the phonology of the alternatives may also be active.

This research provides compelling support for the nonselectivity of lexical access in both bilingual reading and speaking. What is less clear is how the competition that results from that nonselectivity is constrained to allow relatively accurate and language-specific performance. We proposed a framework for identifying a range of factors that may be important in determining how these cross-language interactions are eventually modulated. The factors include language attributes, language-specific strategies and constraints, context, characteristics of the bilingual, and properties of the task. At this stage, the available data are not yet comprehensive enough to allow firm conclusions to be drawn, but we believe that an approach that seeks to develop a principled account of how these different sources of information function will not only allow us to test specific predictions about performance, but will also provide a more complete understanding of the relationship between comprehension and production in two languages.

It is beyond the scope of this paper to consider in detail the implications of the lexical system that we have characterized as fundamentally competitive. However, a number of research programs are investigating these consequences as part of a larger effort to understand the role of attentional control and executive functioning in bilinguals (e.g., Bialystok, this volume, 2005; Green, 1998). It is appealing to think that the consequences of negotiating cross-language competition may be to benefit the
acquisition of skills that reach beyond language itself to domain-general functions. In this sense, bilingualism is of interest not only in its own right, but also for the model it provides for cognition in general.

References


The International Journal of Bilingualism


