Journal of Experimental Psychology:
Learning, Memory, and Cognition
2008, Vol. 34, No. 1, 12–31

0278-7393/08/$12.00 DOI: 10.1037/0278-7393.34.1.12

Native Language Influences on Word Recognition in a Second Language: A Megastudy

Kristin Lemhöfer, Ton Dijkstra, Herbert Schriefers, and R. Harald Baayen
Radboud University Nijmegen

Jonathan Grainger
Aix–Marseille University and
Centre National de la Recherche Scientifique

Pienie Zwitserlood
University of Münster

Many studies have reported that word recognition in a second language (L2) is affected by the native language (L1). However, little is known about the role of the specific language combination of the bilinguals. To investigate this issue, the authors administered a word identification task (progressive demasking) on 1,025 monosyllabic English (L2) words to native speakers of French, German, and Dutch. A regression approach was adopted, including a large number of within- and between-language variables as predictors. A substantial overlap of reaction time patterns was found across the groups of bilinguals, showing that word recognition results obtained for one group of bilinguals generalize to bilinguals with different mother tongues. Moreover, among the set of significant predictors, only one between-language variable was present (cognate status); all others reflected characteristics of the target language. Thus, although influences across languages exist, word recognition in L2 by proficient bilinguals is primarily determined by within-language factors, whereas cross-language effects appear to be limited. An additional comparison of the bilingual data with a native control group showed that there are subtle but significant differences between L1 and L2 processing.

Keywords: visual word recognition, bilingualism, progressive demasking task, cross-language influences, multiple regression

Since the introduction of a common European currency, the development of cultural, political, and economic integration within the European Union has shifted into a higher gear. In the course of this development, there is a growing awareness of the importance of speaking a second language next to one's native language. In the European Community as well as in many other parts of the world, English has become the standard language of international communication and one of the most frequently spoken second languages. According to a survey by the European Union in winter 2005, 56% of all Europeans were able to have a conversation in a foreign language (European Commission, 2006). For 38% of the E.U. population, this second language was English. Next in line were German and French (both 14%). In countries such as Luxembourg, Slovakia, Latvia, and the Netherlands, well over 90% of the citizens are bilingual or even multilingual.

This growing awareness of the importance of multilingualism has also stimulated research on bilingualism and multilingualism, including studies on how bilinguals recognize words in their first or second language. A long-standing debate in this area (Kolers, 1966; Macnamara & Kushnir, 1971) was concerned with the issue of whether the bilingual word-recognition process involved the initial activation of word representations from a target language only (*language-selective lexical access*) or whether all words known to an individual, including those from a nontarget language, are considered as potential candidates for recognition (*nonselective access*). A large number of studies have contributed to this discussion, the majority of them demonstrating that the two languages do interact during word recognition. For instance, it has been shown that when bilinguals recognize words in one of their languages, they process orthographically identical *cognates* (words that share form and meaning in two languages, e.g., *art* in French

Kristin Lemhöfer, Ton Dijkstra, Herbert Schriefers, Nijmegen Institute of Cognition and Information, Radboud University Nijmegen, Nijmegen, the Netherlands; R. Harald Baayen, Interfaculty Research Unit for Language and Speech, Radboud University Nijmegen; Jonathan Grainger, Laboratoire de Psychologie Cognitive, Aix–Marseille University, and Centre National de la Recherche Scientifique, Marseille, France; Pienie Zwitserlood, Psychological Institute II, University of Münster, Münster, Germany.

This research was part of Kristin Lemhöfer's PhD project at Radboud University Nijmegen within the framework of a MaGW International Comparative Research project, titled *Lexical Competition in Bilinguals: A Cross-Language Comparison*, granted by the Netherlands Organization for Scientific Research to the Nijmegen Institute for Cognition and Information (400-60-000).

We would like to thank Mona Aicher, Mohamed El Halfaoui, Marije Michel, and Lonneke Bucken for their practical help in preparing the materials and running the experiments. We are also grateful to Marc Brysbaert and Steve Lupker, who commented on earlier versions of this article.

Correspondence concerning this article should be addressed to Kristin Lemhöfer, who is now at Max-Planck-Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, the Netherlands. E-mail: kristin.lemhoefer@mpi.nl

and English) and, under certain circumstances, also *false friends* (words that share their form but not their meaning, e.g., *pain* in French [meaning bread] and English) differently from words that exist in the target language only (Lemhöfer & Dijkstra, 2004; a more detailed overview of this research appears below). Moreover, van Heuven, Dijkstra, and Grainger (1998) demonstrated that bilingual word recognition in a given language is affected by orthographically similar words, so-called orthographic *word neighbors*, from the other language.

These findings have led to the conclusion that bilingual lexical access can most accurately be described as initially language nonselective (for an overview, see Dijkstra & van Heuven, 2002; Kroll & Dijkstra, 2001). It is especially well established that word recognition in the second language (L2) can be affected by the native language (L1), although effects in the other direction have also been demonstrated (e.g., Duyck, 2005; van Hell & Dijkstra, 2002; van Wijnendaele & Brysbaert, 2002). However, hardly any research has investigated the extent to which the size and types of cross-linguistic effects in L2 depend on the characteristics of the respective native language and its relation to L2. The first aim of the present study was to determine the extent to which results obtained from bilinguals with a particular L1–L2 combination generalize to other bilingual populations with the same L2 but different L1s. We addressed this issue by comparing three groups of bilinguals (native speakers of French, German, and Dutch) with respect to their performance on a word recognition task in their common L2, English.

The second aim of the study was to create an experimental setting containing a large and representative set of word stimuli not specifically designed to generate cross-language influences. In contrast, virtually all previous studies on bilingual word recognition have focused on whether cross-language interactions can be observed at all, and have used factorial designs and restricted sets of selected items (such as form-identical cognates) in an effort to make those interactions as visible as possible. However, it is as yet unclear what role cross-language influences play within the spectrum of all possible influences of stimulus characteristics, including the influence of within-language variables. The present study examines how these different variables might simultaneously affect word recognition in a second language (rather than holding within-language variables constant, as in factorial studies).

So far, only little systematic research has addressed the first aim, the impact of different native languages on visual word recognition in L2. The existing studies focus on the difference between L1–L2 combinations in terms of whether the two languages use the same or different writing scripts. The reasoning underlying these studies is that bilinguals might transfer reading strategies they acquired in their native language to their second language. For example, Muljani, Koda, and Moates (1998) observed that native speakers of Indonesian (a language with an alphabetic writing system) performed better in a lexical decision task in their second language, English, than equally proficient participants with Chinese as their first language. Akamatsu (2002) showed that bilingual speakers with Chinese, Japanese, or Persian as L1 and English as L2 displayed differential effects of word frequency but comparable effects of phonological regularity in English word naming. In a study using the semantic categorization task developed by van Orden (1987), Wang, Koda, and Perfetti (2003) found that native speakers of Korean misclassified homophones of English category members (*stare* in the place of *stair*) more often than graphemic controls (e.g., *stars*), whereas this effect was not present for Chinese speakers. The authors concluded that native speakers of Korean (a language with a highly consistent grapheme-to-phoneme mapping) rely more on phonological than on orthographic information, whereas the reverse is true for speakers of a logographic native language such as Chinese.

The lack of studies comparing the influence of different mother languages within the same alphabet on word recognition in L2 hints at a tacit agreement that at least for fairly related languages, the specific L1–L2 combination does not play a crucial role. However, this assumption is not necessarily true and remains to be tested. Even though they share the same script, alphabetic languages differ in many respects, some of which have been claimed to affect how monolingual speakers recognize words. One example is the *orthographic depth hypothesis* (Katz & Frost, 1992), claiming that the regularity of the mapping between spelling and sound in a language (orthographic depth) determines which route to the mental lexicon is used. According to this hypothesis, word recognition in a language where the sound of a word can be directly derived from its spelling (i.e., in "shallow" languages, such as Italian or Finnish) is possible without the involvement of the mental lexicon. In contrast, recognition of words in orthographically deep languages benefits from the use of all sorts of lexical (i.e., frequency or semantic) information. In line with this reasoning, differences in orthographic depth have been found to modulate the effects of word frequency (Frost, Katz, & Bentin, 1987) or of semantic variables (Cuetos & Barbón, 2006) in word recognition tasks, in particular in word naming but also in silent tasks such as lexical decision. Similarly, differences in phonological consistency between Dutch and English appear to cause differences in sublexical clustering during visual word recognition (Martensen, Maris, & Dijkstra, 2000; Ziegler, Perry, Jacobs, & Braun, 2001), which might show up in varying effects of variables such as bigram frequency. If different languages do indeed entail varying processes and strategies during visual word recognition, these processes and strategies might (partially) be transferred from the native to a second language, just as what has been claimed for native speakers of languages with different scripts.

Thus, even within the same alphabet, different language backgrounds might give rise to different reading strategies in a common second language. In the present study, three groups of participants were included for whom their mother language (French, German, or Dutch) and their second language (English) are western European languages sharing the same alphabet (except for accents and special letters) but with different degrees of orthographic depth: German is the most shallow among these orthographies, with almost a one-to-one mapping of orthography and phonology, whereas Dutch is somewhat deeper, followed by French, and English has the deepest orthography of these four languages (Seymour, Aro, & Erskine, 2003). By choosing participants who all had the same L2, we were able to use the same (English) word materials for all three participant groups and thus to single out the differential role of the mother language.

This study was also designed to provide a further test of the nonselective nature of bilingual lexical access. However, rather than orthogonally manipulating the factors in question and carefully selecting a small number of stimuli, as in previous studies, we used a multiple regression design, involving a large stimulus set

(1,025 English words). Regression designs have a number of important advantages over the more commonly used factorial designs. For instance, as pointed out by Balota, Cortese, Sergent-Marshall, Spieler, and Yap (2004), there are a number of problems that arise from the a priori selection of restricted sets of stimuli in factorial designs, such as difficulties with matching the stimuli on all relevant dimensions, the possible occurrence of experimenter biases during stimulus selection, and the sometimes disproportionate use of words that take extreme values on the target dimensions. In contrast, in regression designs, the variables of interest are not used as selection criteria of the stimulus set, thereby avoiding these difficulties. Furthermore, considering that most variables under investigation are continuous (e.g., word frequency, orthographic neighborhood), regression analyses avoid the loss of information that is associated with the categorization of continuous variables. Finally, regression designs allow for establishing not only whether a particular factor has an effect on the dependent variable but also how large the contribution of this factor is in accounting for the measured variance.

Because of these advantages, an increasing number of large-scale psycholinguistic studies have recently used regression designs (e.g., Alario et al., 2004; Baayen, Feldman, & Schreuder, 2006; Balota et al., 2004; Bird, Franklin, & Howard, 2001; Cuetos & Barbón, 2006; Spieler & Balota, 1997). However, to our knowledge, only one study of this kind has so far been conducted within the domain of visual word recognition in bilinguals. De Groot, Borgwaldt, Bos, and van den Eijnden (2002) had Dutch–English bilinguals carry out a lexical decision and a naming task in both Dutch (L1) and English (L2). They examined the differences both between the two tasks and between the languages with respect to the effects of a number of semantic, orthographic, and phonological variables. Generally speaking, the results revealed considerable differences between the tasks (explainable by the varying task demands) and smaller differences between the two languages, especially for lexical decision. In that study, effects of the individual variables were reported in terms of absolute predictor–criterion correlations, which are difficult to interpret owing to the high intercorrelations between the predictors themselves. Leaving this problem aside, the results of that study suggest that among the cross-language variables, cognate status had a facilitatory effect on both lexical decision and naming reaction times (RTs), but only in L2, whereas (for short words) cross-language orthographic neighbors affected naming but not lexical decision RTs in both languages.

In contrast to the study by de Groot et al. (2002), the present study involved a methodological design that allows for the simultaneous assessment of the partial effects of a large number of predictor variables. If the claim of a profoundly nonselective nature of bilingual lexical access holds true, variables that have previously been shown to carry interlingual influences, like cognate and false friend status and between-language orthographic neighborhood, should also show significant effects for the present comprehensive set of stimuli and predictor variables.

The task chosen in the present study was the progressive demasking (PDM) task developed by Grainger and Segui (1990; see Dufau, Stevens, & Grainger, in press, for freely available software for this task). This task has been shown to produce similar patterns of results as the more frequently used lexical decision task (Dijkstra, Grainger, & van Heuven, 1999; van Heuven et al., 1998). It is a variant of the perceptual identification task in which a word that slowly emerges from a pattern mask must be identified as quickly as possible and must be typed in after identification. Presumably, masking slows down the word recognition process, making the task especially sensitive to factors affecting the early stages of word recognition. Of course, the fairly recent emergence of this task as an alternative to other more standard word recognition paradigms may complicate the comparison with the literature. However, as Carreiras, Perea, and Grainger (1997) pointed out, PDM and other perceptual identification tasks represent a purer measure of orthographic word processing than tasks like lexical decision or word naming, because (unlike lexical decision) they require the unambiguous identification of the word and are not influenced by external factors like the nature of nonword foils or articulatory factors. Furthermore, in a bilingual setting, the PDM paradigm seems to be more appropriate than the lexical decision task, because problems associated with selecting a set of nonwords that are "neutral" with respect to the different mother languages involved are avoided. Moreover, note that the naming task is also not ideal for use in a nonnative language setting, because participants' L1 will likely determine how easily they can pronounce L2 phonemes.

This task was carried out on 1,025 monosyllabic English words by three groups of 20 bilinguals each, with French, German, or Dutch as their native languages. Additionally, a control group of 20 native speakers of English performed the same experiment, to enable evaluation of the obtained results with respect to native speakers' performance. We chose to include all variables that current approaches of word recognition regard as important factors, insofar as they could be calculated using the lexical databases British National Corpus (BNC Consortium, 2001) for English; CELEX (Baayen, Piepenbrock, & Gulikers, 1995) for English, Dutch, and German; and LEXIQUE (New, Pallier, Brysbaert, & Ferrand, 2004) for French. In the following, we describe the chosen predictors and how they were expected to influence word recognition performance in our populations of second language speakers.

Within-Language Variables

Word Frequency

The word frequency effect (more frequent words are recognized faster than words with a lower frequency) is one of the most robust findings in the visual word recognition literature (e.g., Howes & Solomon, 1951; Schilling, Rayner, & Chumbley, 1998; Whaley, 1978) and has been incorporated in virtually every monolingual model of word recognition. For instance, interactive activation models of lexical access assume that frequency affects the resting activation levels of word representations (McClelland & Rumelhart, 1981; for a more detailed discussion of frequency effects and their locus in visual word recognition, see Allen, Smith, Lien, Grabbe, & Murphy, 2005; Hino, Lupker, Ogawac, & Sears, 2003). Even though the size of the effect is task dependent (e.g., Balota et al., 2004), it has been reported for all standard tasks of visual word recognition, including the PDM task (Grainger & Segui, 1990; Perea, Carreiras, & Grainger, 2004). For the bilingual domain, some evidence suggests that the frequency effect might even be larger in the second as compared with the first language (van

Wijnendaele & Brysbaert, 2002). However, de Groot et al. (2002) observed different sizes of the frequency effect in L2 relative to L1 only for word naming and not for lexical decision. Furthermore, if languages with varying orthographic depths indeed give rise to differences in the frequency effect, as suggested by Frost et al. (1987), native speakers of Dutch, German, and French (the last having the deepest orthography among the three) might display different frequency effects also in their L2, English, due to a transfer of word recognition processes from L1 to L2. In that case, native speakers of French should show a larger frequency effect than those with more shallow L1 orthographies, such as Dutch and German.

Recently, the idea has emerged that owing to sex differences in verbal memory (Kimura, 1999), females process words differently from males. For example, Ullman et al. (2002) found that the role of word frequency during the production of inflected verb forms is modulated by sex. Here, we included the sex by frequency interaction as a predictor variable to investigate whether sex differences in verbal memory have consequences for the frequency effect, even for uninflected word forms.

Finally, Baayen et al. (2006) demonstrated that the relative frequency in written compared with spoken English (quantified as the ratio between the two) played an important role in both English monolingual lexical decision and word naming: The more frequent a word was in spoken relative to written English, the faster it was recognized. In the present study, we took a slightly different approach and included both written and spoken frequency (as taken from the British National Corpus) as predictors in the regression analyses, to determine their separate roles both for nonnative and for native speakers of English. Our populations of bilinguals, who are not immersed in an English-speaking environment, might be exposed to proportionally more written than spoken English relative to native speakers, possibly leading to a difference in the relative importance of written versus spoken word frequency for first and second language speakers.

Morphological Family Size

Recently, it has repeatedly been shown that the number of derivations and compounds in which a word occurs, the morphological family size, facilitates response latencies in monolingual and bilingual lexical decision (de Jong, Schreuder, & Baayen, 2000; Dijkstra, Moscoso del Prado Martín, Schulpen, Schreuder, & Baayen, 2005; Schreuder & Baayen, 1997). Paradoxically, it seems to be the number of morphological family members that influences recognition latencies, not their frequency. This argues against a purely frequency-based account of the morphological family size effect. Schreuder and Baayen themselves suggest a semantic source of the effect. Whether the effect generalizes across tasks and participant populations is still subject to research, especially because Schreuder and Baayen failed to find the effect in the PDM task.

Word Length

It is hardly surprising that for most word recognition tasks, longer words take longer to recognize (e.g., McGinnies, Comer, & Lacey, 1952, for perceptual word identification; Ziegler et al., 2001, for word naming; Just & Carpenter, 1980, for eye fixation times in reading; see New, Ferrand, Pallier, & Brysbaert, 2006, for a more complex pattern in lexical decision). Of more interest, Ziegler et al. (2001) have shown that word length effects were larger in German as opposed to English, possibly due to German having a more shallow orthography than English. Consequently, owing to the possible transfer of reading strategies from L1 to L2, word length effects may also differ for bilingual readers varying in their L1 when reading words in their common L2.

English Orthographic Neighborhood

Effects of orthographic neighborhood (i.e., of words that differ from the respective word in one letter only; Coltheart, Davelaar, Jonasson, & Besner, 1977) are thought to reflect the activation of multiple word candidates during word recognition (e.g., Andrews, 1997). The relative importance of various neighborhood measures for the different standard word recognition tasks has been discussed extensively (see Andrews, 1997, and Perea & Rosa, 2000, for reviews). In PDM or similar perceptual identification tasks, the number of higher frequency neighbors has repeatedly been found to slow down recognition latencies for the target word, whereas the total number of neighbors had no or only little effect on recognition performance (Carreiras et al., 1997; Grainger & Jacobs, 1996; but see Snodgrass & Mintzer, 1993). Higher frequency neighbors are thought to delay the pass of the recognition threshold for a target word through lateral inhibition (Grainger & Jacobs, 1996).

Bigram Frequency

Average bigram frequency is a variable capturing the orthographic typicality of a word in the context of its own language (Rice & Robinson, 1975). Whether or not it is relevant to the word recognition process has been subject to debates from the first appearance of this variable in monolingual word identification studies (Broadbent & Gregory, 1968; Gernsbacher, 1984; McClelland & Johnston, 1977; Rumelhart & Siple, 1974). All of these studies used perceptual word identification tasks. Those that reported an effect of bigram frequency (which was primarily the case for low-frequency words) found that it was inhibitory, which is probably a consequence of the difficulty to discriminate “typical,” high-bigram-frequency words from similar words. More recently, Westbury and Buchanan (2002) demonstrated that it was minimal bigram frequency (i.e., the frequency of the least likely bigram in a word), representing a “simple sublexical marker of how unusual a word is” (p. 68), that, for high-frequency words, had an (inhibitory) effect on lexical decision latencies. Assuming that effects of bigram frequency indeed reflect sublexical orthographic processing, they may be sensitive to the different language backgrounds of our participant groups, again assuming the transfer of reading strategies from L1 to L2.

Number of Meanings

Research on the relationship between the form and semantic level of word representation has made extensive use of words that have several meanings. However, there is considerable disagreement concerning whether and how this variable influences word recognition (e.g., Borowsky & Masson, 1996; Duffy, Morris, & Rayner, 1988; Gernsbacher, 1984; Hino, Pexman, & Lupker, 2006;

Piercey & Joordens, 2000; Rodd, Gaskell, & Marslen-Wilson, 2004) and whether related word senses have to be discriminated from unrelated word meanings (Klein & Murphy, 2001; Klepousniotou, 2002; Rodd, Gaskell, & Marslen-Wilson, 2002). To our knowledge, the effect of number of meanings has not yet been investigated for the PDM task, in either L1 or L2 processing. Thus, the question is whether native and nonnative speakers are influenced by the number of word meanings during a word recognition task that presumably involves relatively little semantic processing. Considering that representations of L2 words have been regarded as less "richly populated" (i.e., possessing fewer senses) than L1 words (Finkbeiner, Forster, Nicol, & Nakamura, 2004), it is possible that the number of meanings affects word recognition in the first but not in the second language.

Syntactic Ambiguity

In contrast to semantic ambiguity, syntactic ambiguity of words has as yet received relatively little attention in research on isolated words. In a study using event-related potentials, Elston-Güttler and Friederici (2005) demonstrated that initially, both meanings of syntactically ambiguous homonyms presented in sentence context were activated but that the ambiguity was resolved at a later point in time. However, the disambiguation mechanism was shown to be more effective in native than in nonnative speakers. The latter finding would imply that relative to bilinguals, native speakers should show smaller or no effects of syntactic ambiguity, owing to their fast disambiguation process.

Semantic Variables

For a subset of words (n = 659) for which these values were available, we included the semantic variables concreteness, familiarity, and meaningfulness from the MRC Psycholinguistic Database (Wilson, 1988). For all of these variables, the orthographic depth hypothesis would predict smaller effects for orthographically shallow languages (e.g., German) relative to deeper languages (French), and in the case of a transfer of reading strategies from L1 to L2, this might also show up in the present participants' L2, English. Note that English is the deepest among the four languages; therefore, any semantic effects should be largest for the native control group.

Concreteness

Concreteness has been considered one of the major semantic variables that influence the recall and recognition performance of words. Within the monolingual word-recognition literature, some studies have shown that concrete words are recognized more easily than abstract ones in various word recognition tasks (James, 1975; Juhasz & Rayner, 2003; Kroll & Merves, 1986; Richards, 1976), with the effect being more pronounced for words in the lower frequency range. These effects are usually ascribed to a greater richness, associative embedding, or imagistic quality of the semantic representation for concrete compared with abstract words (Samson & Pillon, 2004); this results in faster semantic processing, which can in turn influence the orthographic processing of a word, depending on the task at hand.

Familiarity

Subjective word familiarity has been claimed to play an important role in (monolingual) word recognition (Connine, Mullennix, Shernoff, & Yelen, 1990; Gilhooly & Logie, 1982; Kreuz, 1987; Williams & Morris, 2004). Gernsbacher (1984) even reported that effects of other variables (bigram frequency, concreteness, and number of meanings) on lexical decision latencies disappeared when familiarity was controlled for. Balota, Pilotti, and Cortese (2001) showed that familiarity ratings, as they are usually obtained, are different from subjective ratings of frequency and that they are more related to a semantic variable (meaningfulness). Effects of familiarity (as rated by a bilingual population) have also been reported for bilinguals, in both their L1 and their L2 (de Groot et al., 2002). In the present study, we included familiarity not only to investigate its effects on PDM latencies for the different participant groups but also to test whether effects of seemingly related variables, such as word frequency, bigram frequency, number of meanings, concreteness, or meaningfulness, would persist with familiarity controlled for.

Meaningfulness

Even though meaningfulness, defined as the ease with which a word can be associated with other words, has been included in many lexical databases (Locascio & Ley, 1972; Paivio, Yuille, & Madigan, 1968; Spreen & Schulz, 1966; Toglia & Battig, 1978; Wilson, 1988), few studies have investigated its effect on word recognition. Johnson and Zara (1964) and Johnson, Frincke, and Martin (1961) observed that words with higher meaningfulness ratings had lower visual recognition thresholds than those with low ratings, at least for the low-frequency range. Recently, the variable was found to be an important predictor of subjective frequency ratings (Balota et al., 2001) and to facilitate RTs in lexical decision for young adults (Balota et al., 2004). Presumably, the stronger semantic connections in which meaningful words are embedded in conceptual memory help their recognition, similar to what has been claimed with respect to concrete words.

Between-Language Variables

Orthographic Neighborhood Variables With Respect to L1

Besides the investigation of the effects of word characteristics with respect to English itself, we expanded our approach to cross-language influences and examined the role of the participants' first language in the present experiment. As a first set of cross-language variables, we included orthographic neighborhood variables with respect to L1. Any occurrence of a between-language neighborhood effect would indicate that during the recognition of a word in a second language, word candidates of the (nontarget) native language become active as well and compete for recognition. Thus, such an effect (as well as any other effect of a between-language variable) would provide strong support for a nonselective view of bilingual lexical access. In one of the few studies on between-language orthographic neighborhood effects, van Heuven et al. (1998) found that Dutch–English bilinguals performing PDM and lexical decision in English (L2) reacted more slowly when the number of Dutch (L1) orthographic neighbors increased. In contrast, in a nonfactorial design, as the present one, de Groot et al.

(2002) found null effects for both English and Dutch neighborhood density in an English lexical decision task, when analyzing short words (of three to five letters) only. Thus, the evidence on interlingual neighborhood effects is as yet both scarce and inconclusive.

Interlingual Cognates and Homographs

Investigating how words that overlap in form across two languages—that is, orthographically identical cognates (hereafter simply referred to as *cognates*) and interlingual homographs—are processed by bilinguals in different task and language contexts can provide valuable information on the structure of the bilingual language system. Different RTs and error rates for these words compared with control words that exist in only one language are usually interpreted as a consequence of the coactivation of both readings in the two languages, and thus as support for the nonselective nature of bilingual lexical access (e.g., Dijkstra et al., 1999).

Most studies on the recognition of cognates by bilinguals have demonstrated that cognates are recognized faster and/or with higher accuracy than control words. This has most frequently been shown for lexical decision in L2 (Caramazza & Brones, 1979; Cristoffanini, Kirsner, & Milech, 1986; de Groot et al., 2002; Dijkstra, van Jaarsveld, & ten Brinke, 1998) or L3 (Lemhöfer, Dijkstra, & Michel, 2004). Furthermore, Dijkstra et al. (1999) obtained the facilitatory cognate effect not only for lexical decision but also for PDM.

In contrast to the consistent picture concerning cognate effects, experimental evidence on interlingual homographs, or false friends (without shared meaning in the two languages), is less conclusive. Both Dijkstra et al. (1998) and de Groot, Delmaar, and Lupker (2000) found that whether and in which way such words are processed differently from control words depends on the task, the stimulus list composition, and the target language (L1 or L2). In English (L2) lexical decision tasks with a purely English stimulus list, both studies observed small, nonsignificant inhibitory effects for those homographs that had a low frequency in English and a high frequency in Dutch (L1). Similar but significant inhibitory homograph effects were found by von Studnitz and Green (2002). In contrast to these studies showing null effects or inhibition of interlingual homographs, Dijkstra et al. (1999) observed that homographs that shared their spelling but not their pronunciation in Dutch and English were recognized faster than English control words, both in an English lexical decision task (see also Lemhöfer & Dijkstra, 2004) and in a PDM task.

The discrepancy between the reliable cognate effects and the fragile or variable homograph effects indicates that there are fundamental differences in the processing of these two kinds of words. Thus, whether form-identical or similar interlingual word pairs also share their meaning is crucial. For instance, it is possible that cognates share one orthographic representation in the bilingual lexicon whereas noncognate homographs do not (Gollan, Forster, & Frost, 1997; Sánchez Casas, Davis, & García Albea, 1992).

The three native languages of our participants differ with respect to the number of cognates and false friends they share with English: For instance, Dutch shares more cognates with English than does French (see Table 4 later in article). Possibly, the bilingual language system exploits this sort of lexical similarity more effectively when it occurs more often. Thus, Dutch participants might show a larger cognate effect than native speakers of French, because they might make more use of overlapping language representations.

Method

Participants

Twenty-one native speakers of French, German, or Dutch with English as a second language participated in the experiment, all in their own country. The participants were mostly recruited among university students and staff. Participants were tested for their English proficiency by means of a vocabulary-size test, described in a separate section below.

The data from 1 participant in each country were excluded because of high error rates in the proficiency test, so that the remaining number of participants was 20 per group. The number of men and women among these remaining participants was, respectively, 2 and 18 in France, 6 and 14 in Germany, and 7 and 13 in the Netherlands. The participants' experience with English, as reported in a language questionnaire, is listed in Table 1.

Table 1
Means (and Standard Deviations) in the Language Questionnaire for the Three Participant Groups

Item	French	German	Dutch
Age	22.7 (1.8)	23.7 (4.4)	24.3 (4.3)
Age of first contact with English	12.2 (5.2)	11.5 (3.7)	12.0 (1.8)
Years of experience with English	12.3 (2.2)	12.5 (3.0)	12.7 (4.3)
Self-ratings[a]			
How often do you read English in leisure time?	5.0 (1.9)	4.3 (1.9)	4.1 (1.7)
How often do you read English at work?*	6.4 (1.0)	4.8 (1.5)	6.5 (0.7)
How often do you speak English?	4.3 (1.3)	4.1 (1.8)	4.2 (1.8)
Self-rated reading experience in English	5.7 (0.8)	5.4 (1.3)	6.0 (1.0)
Self-rated writing experience in English	5.6 (0.8)	5.0 (1.2)	5.1 (1.4)
Self-rated speaking experience in English	5.0 (1.7)	5.5 (1.4)	4.9 (1.5)

Note. Asterisk indicates dimension for which there were significant ($p < .05$) differences between the three participant populations, as analyzed by one-way analyses of variance.
[a] Self-ratings were given on a scale from 1 (*low*) to 7 (*high*).

One-way analyses of variance were conducted to examine whether there were significant differences between the three participant populations. Such a difference was found only for the frequency with which they read English at work. Paired *t* tests showed that both French and Dutch participants reported reading more English for their work than did German participants: French > German, $t(57) = 4.46$, $p < .001$; Dutch > German, $t(57) = 4.88$, $p < .001$. None of the other comparisons were significant.

For the control experiment, 20 native speakers of English currently visiting Nijmegen, the Netherlands, who reported having little knowledge of Dutch and not speaking any foreign language regularly, performed the same experiment as the bilingual speakers. They were, on average, 21.9 years old ($SD = 2.6$). Ten of them were male, 10 female. They had grown up in the United Kingdom (6 participants), the United States (11), Canada (2), or Australia (1). Some reported having been in contact with foreign languages other than English (Spanish, French, German, Italian, or Dutch). The mean rating for their degree of experience with those foreign languages was 2.7 on a scale from 1 (*low*) to 7 (*high*); the mean rating for their frequency of usage was 1.8.

Materials

The experimental stimulus set consisted of English words selected from the CELEX database (Baayen et al., 1995). They met the following criteria: They were between three and five letters long; only content words were used (i.e., nouns, verbs, adjectives, and adverbs); they were monosyllabic; each word had only one possible spelling and one pronunciation; the written lemma frequency of the words according to the CELEX database lay between 10 and 10,000 occurrences per million (o.p.m.); and words with more than two different entries in CELEX (i.e., due to the number of syntactic categories the word can belong to) were not included. Exceptions to these restrictions were made for words ($n = 20$) that had been part of the stimulus materials in our earlier studies, to allow us to compare the results with previous ones in the future.

To make sure that words would be known by the participants, we gave a list of the 200 least frequent words to five students (native speakers of Dutch) at the University of Nijmegen, drawn from the intended population of Dutch participants in the main experiment. They were asked to indicate which of the words they did not know. Words that were unknown to one or more persons (80 out of the 200 words) were excluded. The final set of stimuli consisted of 1,025 English words.[1]

For the multiple regression analyses, a number of within- and between-language variables were calculated as predictors. For the native control group, obviously, only the within-language predictors were used.

Description of Predictors of the Regression Analyses

The variables characterizing the English properties of the words were drawn from the English part of the CELEX database, as well as from the British National Corpus (BNC Consortium, 2001).

Within-Language Variables

Word length. Length was coded in number of letters, ranging from three to five. All words were monosyllabic in English.

Word frequency and morphological family size. Written and spoken word form frequency were taken from the British National Corpus (BNC Consortium, 2001) and normalized for corpus size. Word form frequency refers to the frequency of occurrence of a specific surface word form excluding its inflections (e.g., the plural form). Frequency was one of our predictors that was characterized by a highly skewed distribution. To avoid the possibility that a few atypical data points with extreme values would exert undue leverage in the regression, we removed most of this skew by means of a logarithmic transformation. This practice is common in the word recognition literature (e.g., Baayen et al., 2006; Baayen, Tweedie, & Schreuder, 2002; Balota et al., 2001). When a word had several entries in the database, as many words had (e.g., *bite* is both a verb and a noun), the frequencies of the different entries were added before applying the logarithm.

Morphological family size refers to the number of words and compounds that are derived from the word itself. For example, according to CELEX, the word *bride* has five family members (with a frequency larger than 0): *bridal*, *bridecake*, *bridegroom*, *bridesmaid*, and *bride-to-be*. Following the same motivation as for frequency, we logarithmically transformed the family size counts.

English orthographic neighborhood. For English neighborhood characteristics, the number and summed frequency (taken from the CELEX corpus) of orthographic neighbors of a word within English (i.e., words of the same length differing in exactly one letter) were counted. This measure did not include neighbors with a frequency lower than 1 o.p.m. For testing whether higher and lower frequency neighbors would have differential effects on word recognition, the number of orthographic neighbors with a frequency higher or lower than the word itself was calculated for each word. Again, all neighborhood variables were logarithmically transformed.

Bigram frequencies. Several counts of bigram frequency were determined, based on either word types or word tokens in CELEX. For the type count, the number of word forms with the same length, sharing a given bigram in the same position, was taken into account. For the token count, the word form frequencies of these words were summed for each bigram. On the basis of preliminary analyses (hierarchical regression analyses with different variable orders), we selected the two most predictive counts: the token counts of the mean bigram frequency (average frequency of all bigrams in the word) and minimal bigram frequency (frequency of the least frequent bigram in the word). Both variables were logarithmically transformed.

Number of syntactic word categories and meanings. The number of CELEX entries for each word was also entered as a variable. In most cases, this variable reflected the number of syntactic categories a word can adopt. In only two cases (*sake* and *rose*), there were separate entries for two meanings of a word even though they belong to the same syntactic category, because they possess different etymological roots. Because there were few items with more than two entries (see the selection criteria above), the variable was dichotomized (one entry vs. more than one entry).

For the number of meanings of a word, *The Wordsmyth Dictionary* (Wordsmyth, 2002) was used to determine the number of

[1] The list of stimuli as well as the item means for the four participant groups can be obtained from Kristin Lemhöfer.

Table 2
Characteristics of the Word Materials With Respect to English

Variable	*M* (*SD*)	Range
Number of letters	4.17 (0.70)	3–5
Log written word frequency (BNC)	3.38 (1.34)	0.25–7.32
Log spoken word frequency (BNC)	2.94 (1.70)	0–8.93
English morphological family size	12.3 (18.3)	0–247
Number of English orthographic neighbors	6.77 (4.58)	0–22
Number of English lower frequency neighbors	4.34 (3.51)	0–19
Number of English higher frequency neighbors	2.44 (2.61)	0–16
English log summed neighborhood frequency	2.30 (0.92)	0–4.80
Mean bigram frequency (token count)	1,662 (1,405)	9–15,066
Minimal bigram frequency (token count)	497 (539)	3–4,171
Number of entries in CELEX[a]	1.74 (0.47)	1–4
Number of word meanings (*Wordsmyth*)	1.26 (0.56)	1–5
Concreteness[b]	482 (109)	183–670
Familiarity[b]	548 (40)	417–645
Meaningfulness[b]	452 (51)	281–607

Note. Unless stated otherwise, variables are reported in absolute rather than logarithmic values. BNC = British National Corpus.
[a] This variable was dichotomized (equal to 1 vs. greater than 1). [b] Available for only a subset of words (659 items).

semantically unrelated meanings, similar to the method of Rodd et al. (2002). *The Wordsmyth Dictionary* is suitable because it lists the number of unrelated meanings regardless of syntactic category; thus, this variable is not confounded with the previous variable, the number of CELEX entries.[2] For example, the word *host* has three semantically distinct word meanings: (a) a person who entertains guests, (b) a very large number of people or things, and (c) eucharistic bread.

The characteristics of the 1,025 words with respect to the mentioned intralingual variables are shown in Table 2.

Semantic variables. Three semantic variables were taken from the MRC database (Wilson, 1988): concreteness, familiarity, and meaningfulness. These values were available for only a subset (659) of the 1,025 words. The ratings for familiarity and concreteness in the MRC database were merged values from three sets of norms and transformed to values between 100 and 700. The meaningfulness ratings listed in the database were taken from Toglia and Battig (1978) and were also multiplied so that they lay in a range from 100 to 700. The word characteristics with respect to these semantic variables for the subset of 659 words are also shown in Table 2.

Between-Language Variables

For the coding of between-language variables, the following databases were used: the German and Dutch part of the CELEX database for word forms, excluding words with a frequency lower than 1 o.p.m., and the French LEXIQUE, which also contains words with a frequency of 1 o.p.m. or more. Note that for a given word, whereas within-language variables take the same values for all three participant groups, cross-language variables are by definition different for the French, German, and Dutch participants (e.g., a given English word might be a cognate with respect to French but not regarding German or Dutch).

L1 orthographic neighborhood. For each English word, the number and summed frequency of orthographic neighbors in Dutch, German, and French were calculated. In addition, the number of neighbors was split into high- and low-frequency neighbors. Here, an absolute frequency split was used, distinguishing between L1 neighbors above and below a frequency of 50 o.p.m.[3] In analogy to the within-language variables, all neighborhood variables were logarithmically transformed. In Table 3, a summary of the word characteristics is given with respect to the between-language neighborhood variables.

L1 homographs and cognates. We refer to words that share their form but not their meaning in two languages as *(noncognate) interlingual homographs* or *false friends*; words that have both the same spelling and the same meaning are referred to as (orthographically identical) *cognates*. Owing to computational restrictions, only words that are orthographically identical with their reading in the respective other language were counted as homographs or cognates; for example, the English word *bed* was counted as a cognate with respect to Dutch (*bed*) but not with respect to German (*Bett*). Homograph and cognate status between English and the respective L1 were coded as dichotomous (0 or 1) variables. The numbers of interlingual homographs and form-identical cognates for each of the three language pairs are given in Table 4.

Procedure

Because of the large number of trials, the experiment was conducted in three sessions, which were held on different days.

[2] Rodd et al. (2002) distinguished between the number of (unrelated) meanings and the number of (related) senses. We did not include the latter measure because it is correlated with the number of CELEX entries, as the different syntactic categories a word can adopt are counted as senses in *The Wordsmyth Dictionary*.

[3] Because of the impossibility of directly comparing first- and second-language word frequencies for unbalanced bilinguals, an absolute split criterion was used rather than the relative one used in the case of within-language neighbors.

Table 3
Characteristics of the Word Materials With Respect to French, German, and Dutch

Variable	French		German		Dutch	
	M (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range
No. of orthographic neighbors	3.59 (3.98)	0–21	2.38 (2.50)	0–13	4.20 (4.40)	0–21
No. of orthographic neighbors with freq. > 50	0.65 (1.12)	0–7	0.62 (1.07)	0–6	0.84 (1.45)	0–11
No. of orthographic neighbors with freq. < 50	2.93 (3.34)	0–18	1.76 (1.93)	0–11	3.36 (3.42)	0–16
Summed neighborhood freq.	355 (1,276)	0–16,831	776 (3,671)	0–43,895	1,202 (5,229)	0–51,742

Note. All variables are reported in absolute rather than logarithmic values. Freq. = frequency.

Participants were tested individually. The experiment was controlled by software developed in collaboration with the Technical Group of the Nijmegen Institute for Cognition and Information, running on Macintosh computers. The experiments were run in Nijmegen, the Netherlands; Aix-en-Provence, France; and Münster, Germany. The same software was used in all three laboratories. In each lab, the monitor was a 17-in. screen with a refresh rate of 66 MHz and a resolution of 640 × 480 pixels, placed at a distance of about 60 cm from the participants.

Each session began with a practice block (10 trials in the first session, 5 in the second and third ones), consisting of words that were not used in the actual experiment. In the main experiment that followed, 342 (first and second sessions) or 341 words (third session) were presented in seven blocks of 50 words (except for the last block, which was shorter). Every participant within each group received a differently randomized order of words; however, the same 21 randomizations were used in all three labs. An English instruction was given to the participants explaining that they would see a word alternating with a mask, with the word gradually becoming more visible. They were asked to press a button on a button box as soon as they had recognized the word and to type in the word after a prompt had appeared.

Words were presented in black, lowercase Courier letters (size 18 point) in a white window, which was surrounded by a black background. At the beginning of a trial, a fixation cross appeared in the middle of the screen. When the participant pressed the button, one of two checkerboard masks appeared on the screen, consisting of the same number of checkerboard blocks as the word had letters. The mask remained on the screen for 25 refresh cycles (or 378 ms) and was replaced by the word presented on the same location, which was visible for one cycle (or 15 ms). The word was followed by the second mask (the inverse pattern of the first), after which it was presented again, and so on, with the duration of the word presentation increasing by one cycle at every alternation and that of the mask decreasing by the same time (15 ms). The process stopped when the participant pressed the button (otherwise, the time-out deadline would have been after 650 refresh cycles, but this time limit was never reached). After the participant had responded, a window appeared on the screen with the text *please enter the word*, and the participant was to type the word using the computer keyboard. One second after the participant had pressed *Enter* to enter the word, the next trial was started by the participant pressing a button. No feedback was given to the participant concerning whether the answer had been correct.

Every experimental session took about 60 to 75 min. Participants received course credit or money for their participation after the last of the three sessions.

Table 4
Number of Interlingual Homographs and Cognates in the Stimulus Materials

Word type	French	German	Dutch
No. of cross-language homographs	167	170	202
No. of cognates	103	109	137
No. of noncognate homographs	64	61	65

Proficiency Test

At the beginning of the first session, before the experiment, the participants completed an English vocabulary-size test. The results of this test were analyzed immediately to make sure that the person met the proficiency criteria (no more than 20 errors out of the total of 60 items) for the experiment. The test was a nonspeeded lexical decision task on 40 low-frequency English words and 20 highly wordlike nonwords and was derived from an unpublished vocabulary-size test developed for high-proficiency populations by P. Meara (Meara, 1996). This test version is described in more detail in a previous publication (Lemhöfer et al., 2004). The test score was calculated using a percentage correct measure (i.e., the unweighted mean percentage of correctly recognized words and correctly rejected nonwords). The results of the three groups of bilingual participants, as well as those of the native speakers, are summarized in Table 5.

According to a one-way analysis of variance carried out on the data of the four participant groups, there were significant differences with respect to the proficiency scores, $F(3, 76) = 17.38, p <$

Table 5
Means, Standard Deviations, and Ranges of Scores on the Proficiency Test for the Four Participant Groups

Participant group	*M*	*SD*	Range
French	88	6	75–96
German	82	8	71–98
Dutch	84	9	70–99
English	97	4	85–100

Note. Scores were calculated as the unweighted mean of the percentage correct for words ($n = 40$) and for nonwords ($n = 20$), respectively.

.001. Planned comparisons showed that the native speakers of English had higher scores than the three bilingual groups, as evident in a significant difference between the native speakers and the best of the bilingual groups, the French participants, $t(38) = 5.23$, $p < .001$. Furthermore, French participants had higher scores than German participants, $t(38) = 2.59$, $p < .02$, but there was no significant difference between French and Dutch or between German and Dutch participants (both ps $> .06$).

Results

For all statistical analyses, the RTs were logarithmically transformed in order to avoid the "long right tail" of the skewed RT distributions overly influencing the results. The mean RTs on correct trials, error rates, and standard deviations per participant group are shown in Table 6.

Similarity of Item Means Between Participant Groups

The first aim of the present study was to assess the extent to which the results of the three bilingual participant groups were comparable. A first indication of the similarity of their data is given by the between-group correlations. The correlations between the three bilingual participant groups, as well as those with the native control group, calculated across the 1,025 item means, are shown in Table 7.

As can be seen from Table 7, the three nonnative groups correlate to a larger degree with each other than with the native group. Pairwise comparisons of the correlation coefficients using Fisher Z transformations confirmed this observation (all ps $< .02$). For a different view of the similarity between the data of the four groups, we conducted a number of linear multiple regression analyses with the item means of one group as dependent variable and those of the other three groups as predictors, respectively. The resulting proportion of explained variance (R^2) indicates the degree to which the outcomes of one group could be predicted by those of the other three groups, or in other words, how much variance is shared between the groups. Additionally, the standardized regression weights give an indication of the relative contribution of the three predictor groups and the direction of this contribution. The resulting regression weights and R^2 values are reported in Table 8.

The first set of analyses involved all four participant groups; to obtain measures of how much variance was shared between the three bilingual groups, we carried out a second set of analyses involving the three nonnative groups only. The resulting R^2 values (indicating the proportion of shared variance) are also reported in Table 8.

As can be seen from Table 8, a large proportion (around 60%) of the RT variance in each nonnative group could be explained by the item means of the other two bilingual groups. In contrast, the predictability of the native data by the bilingual item means was considerably reduced ($R^2 = .49$). Similarly, the regression weights show that the item means of the English group were not as good a predictor of the nonnative data as those of the other nonnative groups were. This result gives a first indication that whereas the three nonnative participant groups largely processed the English words in the same way, the native group was less similar to the bilinguals. To investigate the similarities and differences between the participant groups in more detail, we included a number of within- and between-language variables in multiple regression analyses.

Regression Analyses Including Within- and Between-Language Predictors (Bilingual Participants)

Because the focus of the present article was on bilingual word recognition, we analyzed the data of the nonnative participants first and compared the results with those of the native group as a second step (described below). For the regression analyses, trials with RTs shorter than 500 ms or greater than 5,000 ms were removed from the data set. A linear mixed-effects analysis of covariance

Table 6
Mean RTs (Absolute and Logarithmic) for Correct Trials (in ms), Error Rates, and Standard Deviations in the Progressive Demasking Task for the Four Participant Groups

	RT		Log RT		Error rate (%)	
Participant group	*M*	*SD*	*M*	*SD*	*M*	*SD*
French	1,698	425	7.41	0.24	1.91	13.7
German	1,723	454	7.42	0.26	2.17	14.6
Dutch	1,619	451	7.36	0.26	2.24	14.8
English	1,607	395	7.35	0.24	1.92	13.7

Note. RT = reaction time.

Table 7
Intercorrelations of Log RTs on Correct Trials (Based on Item Means) Between the Four Participant Groups

Participant group	1	2	3	4
1. French	—	.71	.70	.63
2. German		—	.75	.62
3. Dutch			—	.64
4. English				—

Note. All correlations are significant at the $p < .01$ level ($n = 1{,}025$). RT = reaction time.

Table 8
Results of the Multiple Regression Analyses Regarding the Mutual Predictability of Log RTs (for Correct Trials) Between the Four Participant Groups, Calculated Across Item Means

	Standardized regression weights (predictors)				R^2	
Dependent variable	French	German	Dutch	English	All groups	Bilingual groups only
French	—	.33	.31	.23	.594	.561
German	.32	—	.46	.16	.635	.614
Dutch	.27	.44	—	.20	.644	.620
English	.29	.20	.29	—	.486	

Note. All R^2 values and regression weights are significant at $p < .001$. RT = reaction time.

(Baayen, in press; Bates & Sarkar, 2005; Faraway, 2005; Pinheiro & Bates, 2000; Wood, 2006) with the logarithmically transformed RT as dependent variable and with crossed random effects for subject and item was fit to the data, using a stepwise variable selection procedure. The following variables did not reach significance as predictors (i.e., their regression weights were nonsignificant) and were therefore dropped from the model: mean bigram frequency, number of low-frequency English neighbors, total number and frequency of English neighbors, number of meanings, noncognate homograph status, total number and frequency of L1 neighbors, number of high-frequency L1 neighbors, and number of low-frequency L1 neighbors. Furthermore, the interactions of all remaining variables with participant group were tested but were not significant unless reported otherwise. All *p* values were supported by Markov chain Monte Carlo confidence intervals sampled from the posterior distribution of the predictors.[4]

Inspection of the residuals revealed marked nonnormality due to higher estimation errors for a number of data points with very long RTs. After removing overly influential outliers (defined as having absolute standardized residuals exceeding 2.5) from the data set, we refitted the model. We report only the significant fixed effects for this final model (see footnote 6 for the random effects). Note that because the semantic variables (concreteness, familiarity, and meaningfulness), as taken from the MRC database, were available for only a subset of words, those predictors were tested in a separate analysis. The first analysis was carried out on the complete nonnative data set, but excluding the set of semantic predictors. The resulting regression weights, *t* values, and significance levels are summarized in Table 9.

Table 9
Results of the Mixed-Effects Regression Model Including the Three Bilingual Groups

Effect	β	$t(60368)$	p
Session	−.069	−77.70	<.0001
Trial (within session)	−.0002	−23.19	<.0001
Previous RT	.159	45.33	<.0001
Error	−.126	−4.13	<.01
Sex	.003	0.06	.96
Participant group: French vs. Dutch	.072	1.30	.19
Participant group: German vs. Dutch	.085	1.58	.11
Number of letters	.021	5.06	<.0001
Written word frequency (linear)	−.071	−5.35	<.0001
Written word frequency (quadratic)	.038	4.65	<.0001
Spoken word frequency (linear)	−.009	−5.30	<.0001
Morphological family size	−.012	−2.53	<.02
Minimal bigram frequency	.006	3.47	<.001
Number of high-frequency neighbors	.009	2.66	<.01
Number of CELEX entries	−.011	−2.39	<.02
Cognate status	−.011	−3.62	<.001
Number of CELEX Entries × Participant Group[a]	.008	2.18	<.03
Sex × Written Word Frequency[b]	.006	2.72	<.01
Error × Written Word Frequency[c]	.013	3.21	<.01
Error × Minimal Bigram Frequency[c]	−.011	−2.66	<.01

Note. RT = reaction time.
[a] Additional effect for the French compared with the Dutch group. [b] Additional effect for men compared with women. [c] Additional effect for error trials compared with correct trials.

Control variables in our study were session, trial number (within a session), previous RT, error (vs. correct response),[5] and participant group. Table 9 shows that with each successive session, participants responded faster. Participants also recognized words faster as they progressed through the trials in a session. We included the RT to the preceding trial as a covariate, as previous research has suggested that difficult preceding trials have a spillover effect on the next trial (Taylor & Lupker, 2001; Wurm, Aycock, & Baayen, 2007). In our data, a longer RT to the previous trial indeed implied a longer RT for the target trial. Finally, trials for which participants reported a different word from the word actually shown tended to have shorter RTs, suggesting that on these trials, the button was pressed before the word had been identified correctly. The accuracy of the response also interacted with written word frequency and minimal bigram frequency (further details are provided below).

Written word frequency emerged as a nonlinear predictor, with facilitatory effects on RTs. Nonlinearities for word frequency were also reported by Gordon (1985), Balota et al. (2004), and Baayen et al. (2006). Across these studies, the effect of frequency leveled off for the higher frequency words. Additionally, and in line with the monolingual findings by Baayen et al. (2006), spoken frequency also had an independent, facilitatory effect on RTs, but only as a linear predictor (the quadratic predictor was nonsignificant and therefore excluded from the model). Apparently, the familiarity with a word in speech (besides its use in written language) is important not only for native speakers of English but also for nonnative speakers. Surprisingly, and possibly as a consequence of females having a better verbal memory than males (Kimura, 1999), the frequency effect was stronger for women than for men (whereas there was no main effect of sex). Finally, the facilitation of written frequency was attenuated for error trials. Together with the facilitatory main effect of error, this interaction suggests that words of low frequency sometimes led to faster but premature erroneous responses, probably due to the confusion of the correct word candidate with a similar higher frequency word.

The morphological family size measure revealed a facilitatory effect, as expected. Furthermore, words with more than one entry in the CELEX lexical database were processed faster than words with only one entry. Compared with native speakers of Dutch, this effect was significantly reduced for French but not for German participants ($p > .50$).

[4] Although extreme multicollinearity characterized our original full set of predictors, the condition number (following Belsley, Kuh, & Welsch, 1980) for the item variables in our final model was 32. A condition number this high indicates substantial collinearity. However, exclusion of number of letters from the set of predictors reduced the condition number to 15. Furthermore, the coefficients of a model excluding number of letters were very similar to those in the model including this variable. (The correlation between the two sets of coefficients was .9999.) This allows us to conclude that multicollinearity is not distorting our regression model.

[5] We opted to include error trials in the model (along with the error variable) in order to keep as many data as possible, and also because any possible effects of the lexical variables on error trials might provide valuable insights into the processes at work when an error occurs. An analysis of only the correct trials showed very similar results, with the same variables becoming significant.

As the only cross-language effect that emerged across the complete data set, words that are cognates in the participant's native language elicited shorter response latencies than noncognates.

Word length, measured in number of letters, was inhibitory as expected. Similarly, minimum bigram frequency had an inhibitory effect on RTs, reflecting the difficulty of unambiguously identifying more typical (and thus easily confusable) English words. This effect confirms the findings of Westbury and Buchanan (2002) for high-frequency words in monolingual lexical decision. However, for error responses, the beta weight was reduced. Analogous to the interaction of error with word frequency, this interaction shows that (owing to their high confusability) difficult words (i.e., words with a high minimum bigram frequency) in some cases elicited relatively fast but incorrect responses. Among the variables coding for English orthographic neighborhood, only the number of higher frequency English neighbors was significant: Words with many of these neighbors elicited longer RTs. This replicates what has been observed for the PDM task in monolingual studies before and confirms the notion of higher frequency neighbors being a better predictor of word recognition performance than the total number or summed frequency of neighbors.[6]

Even though (noncognate) homograph status was not significant as a dichotomous variable across the complete data set, we ran an additional analysis to examine whether evidence for a possibly hidden effect was nevertheless present in our data. For a subset of items (see Table 4) including only those words that were noncognate homographs with the respective L1, we investigated whether the frequency of the L1 reading had an influence on RTs. Indeed, in a regression model simplified owing to the small data set, with written and spoken English frequency, number of letters, and number of English higher frequency neighbors as lexical predictors, the frequency of the homographs in L1 had a significant inhibitory effect on RTs ($\beta = .013$), $t(3742) = 2.20$, $p < .03$. This replicates the findings of previous studies reporting that effects of interlingual homographs depend on the relative frequencies of the readings in the two languages (de Groot et al., 2000; Dijkstra et al., 1998; Kerkhofs, Dijkstra, Chwilla, & de Bruijn, 2006).

Analysis of Semantic Variables

The influence of the semantic variables taken from the MRC database was investigated in a second analysis, on 659 out of the 1,025 words. The same variables that had emerged as significant predictors in the previous analysis were entered in this analysis, with the addition of the three semantic variables concreteness, familiarity, and meaningfulness. However, in a subsequent step, the interaction of participant group with number of CELEX entries was dropped from the model, because it was not significant any more for this subset ($p > .30$). Similarly, and in line with what we expected in the light of previous findings, concreteness did not have a significant effect on RTs ($p > .80$) and was excluded from the model as well.

Familiarity and meaningfulness are rating measures that themselves can be predicted from other variables in our data. More than half of the variance in the familiarity ratings (adjusted $R^2 = .60$) is explained by an ordinary least squares regression model with familiarity as dependent variable and spoken frequency, number of higher frequency English neighbors, concreteness, and meaningfulness as (significant) predictors (see Appendix A for the β, *t*, and *p* values). These results are comparable to those reported by Baayen et al. (2006) for the familiarity ratings obtained by Balota et al. (2001). The ratings for meaningfulness can be predicted from spoken and written frequency, concreteness, and the number of CELEX entries. The adjusted R^2 for this model was .12.

To avoid problems with increased multicollinearity, we included the residuals of the models for familiarity and for meaningfulness as predictors in our mixed-effects model for the response latencies in PDM. These residuals are thus corrected for the influence of all variables correlated with the original familiarity and meaningfulness measures. Both residual familiarity and residual meaningfulness had significant facilitatory effects on RTs: familiarity ($\beta = -.0002$), $t(38800) = -1.99$, $p < .05$; meaningfulness ($\beta = -.0001$), $t(38800) = -2.38$, $p < .02$. Meaningfulness also interacted with participant group, with a stronger facilitatory effect of this predictor (by $-.0001$), $t(38800) = -2.86$, $p < .01$, for the French as opposed to the Dutch group. In contrast, the effect of meaningfulness for German participants did not differ significantly from that for Dutch speakers ($p > .20$). All other effects that had evolved in the previous model on the complete data set (apart from the effect of the number of CELEX entries) were present for this subset of the data as well. Note that these bilingual data did not provide evidence for any main effects (e.g., bigram frequency) disappearing once familiarity was controlled for (or included in the regression model), as claimed by Gernsbacher (1984).

Control Group: Native Speakers of English

To investigate whether the monolingual English speakers deviated from the bilinguals in terms of the effects of the lexical variables, we used the same model that had emerged as the one providing the best fit to the bilingual data (but, of course, excluding cross-language variables and the variable participant group) for the group of native English speakers. Most predictors included in the original model were found to be significant again, with the following exceptions: The facilitatory effect of morphological family size failed to reach significance ($p = .30$); there was no significant difference between words with only one and those with more than one entry in CELEX ($p = .66$); and the interactions of error with minimal bigram frequency ($p = .08$) and with written or spoken frequency (both $p > .15$) were no longer significant. Of most interest, however, written frequency was no longer significant for native speakers of English (linear: $p = .32$, quadratic: $p =$

[6] We complete the specification of the mixed-effects model with a report of the random-effects structure. The model contained five random effects. A random effect concerns a random variable that follows a normal distribution and that is characterized by a mean equal to zero and an unknown standard deviation. For each of the five random effects in our model, we therefore have a standard deviation that characterizes its spread around zero. The standard deviation for the random effect of word was 0.052—that is, the by-word adjustments to the intercepts are characterized by this standard deviation. Our model included three random effects involving the participants: by-participant random intercepts ($\sigma = 0.164$) and by-participant random slopes for written frequency ($\sigma = 0.006$) and number of letters ($\sigma = 0.021$). Centering of the data points showed that parameters for the correlations between random effects were not necessary ($p > .10$). All random intercepts and slopes were supported by likelihood ratio tests (all *p*s $< .05$).

.43), and spoken frequency (linear and quadratic) remained as the only significant frequency predictor. Apparently, written frequency is an important factor only for nonnatives who are primarily exposed to written English, but not for native speakers. These insignificant effects were excluded from the model for this group of participants.

All other effects that had been observed for the second language speakers were present in the native data as well. Other variables that had been dropped from the bilingual model because of nonsignificance also turned out to be nonsignificant for the native speakers. Furthermore, the same random-effects structure was found to be present. The table of results of the final model for the monolingual English speakers can be found in Appendix B.

Similarly, the effects of the three semantic variables on the subset of the monolingual data for which these measures were available very much resembled those observed for the bilinguals: Although the facilitatory effect of concreteness just failed to reach significance ($p = .054$), residualized familiarity and meaningfulness both had significant facilitatory effects on RTs: familiarity ($\beta = -.0002$), $t(12938) = -2.71$, $p < .01$; meaningfulness ($\beta = -.0002$), $t(12938) = -3.54$, $p < .001$. This time, two effects turned nonsignificant in this subset analysis, probably as a consequence of the reduced statistical power in the data subset: that of the quadratic term of spoken word frequency ($\beta = .009$), $t(12938) = 1.72$, $p = .08$, and the interaction of sex and spoken word frequency ($\beta = .003$), $t(12938) = 1.66$, $p = .10$.

Discussion

In recent years, many studies have demonstrated that bilinguals are influenced by knowledge of their native language during word recognition in their second language. However, little is known about whether the findings concerning bilingual word recognition are generalizable across bilinguals with different language combinations, especially for languages that share the same alphabet. The first aim of the present study was to investigate this issue by comparing word recognition in English as a second language for three groups of participants that had different western European native languages. Additionally, this comparison included a control group of native speakers of English, to allow us to relate the results for the bilinguals to the "standard" case of language processing by monolinguals and to directly compare L1 and L2 processing. Furthermore, most previous studies have used factorial designs and special categories of stimulus words in order to create the most favorable conditions for cross-language effects to occur. Our second goal was to examine whether these cross-language interactions still play a vital role in the context of a large, representative, and unbiased set of word materials, and with the simultaneous influence of factors within the target language itself. If bilingual word recognition is indeed profoundly nonselective with respect to language, between-language variables that have previously been found to affect recognition performance should account for a considerable amount of variance in the present data.

We now discuss the results of the first set of analyses (correlations and mutual predictability between groups) and those of the multiple regression analysis separately, before evaluating the findings more generally in the light of the two issues at hand.

Correlations and Mutual Predictability Between Groups

All analyses that were carried out to investigate aspects of the first issue, the generalizability of L2 word recognition results across different bilingual groups, showed that the overlap between the results of the three nonnative participant groups was considerable. Descriptively, this is already evident from the high correlations of the item means between these three participant groups (above .70; see Table 7). Furthermore, linear multiple regression analyses (see Table 8) showed that far more than 50% (56%–62%) of the item variance in each group could be explained by the data of the other two groups.

This high degree of similarity between the three bilingual groups is surprising, given that a number of differences between these groups had been expected, as specified in the introduction. One possible implication of this result is that during reading in a second language, word recognition is mainly determined by factors within that language itself, and that it should thus be similar for all users of this language. However, in contrast to the large overlap of the nonnative groups, the group of native English speakers differed considerably from the bilinguals. Apparently, although the specific native language of a bilingual does not make much of a difference, second language processing seems to be different from native language processing. A more detailed investigation of the specific differences between the participant groups with respect to English word recognition was provided by the regression analyses involving a large number of word characteristics as predictors.

Regression Analyses

We now discuss some selected effects of individual variables that were novel or unexpected, or that differed across the participant groups, before we return to the question of the similarity of the data across the groups and the question of the impact of between-language influences during lexical access in L2.

Word Frequency

When we compared the frequency effects for the three bilingual groups, no significant differences emerged. All three groups were influenced by both written and spoken word frequency to comparable extents. We had hypothesized that differences in the frequency effect across the groups might occur as a consequence of the different orthographic depths of the participants' native languages and of the resulting differences in routes of lexical access (Frost, 1994; Frost et al., 1987; Katz & Frost, 1992). The absence of such differences in our data indicates that the nonnative participants used very similar reading strategies in their common second language, English. It is interesting to note, however, that the effect of written frequency completely disappeared for the native group. Only spoken frequency influenced the speed with which these speakers recognized words in their native language. This difference in the relative importance of written and spoken word frequency between native and nonnative speakers reflects the way in which these language users are exposed to English: Compared with native speakers, the nonnatives' experience with English is to a large extent based on reading and writing, with relatively little exposure to spoken English. The emergence of spoken frequency as a predictor that is superior to written frequency for native speakers confirms and extends the findings by Baayen et al. (2006) and questions the present common practice to use only written word frequency measures in psycholinguistic research.

An additional analysis on the complete data set (including native and nonnative speakers) with a new factor, nativeness, revealed that whereas the effect of spoken frequency did not differ significantly between native and nonnative speakers (linear: $p > .90$, quadratic: $p > .45$), the effect of written frequency was significantly larger for nonnatives (linear: $p < .001$, quadratic: $p < .01$). Taken together, these results indicate that the total effect of word frequency is larger for nonnatives than for natives. This difference in the size of the frequency effect for first and second language speakers cannot be caused by the orthographic depth of the native language: With English being the deepest among the four orthographies (Seymour et al., 2003), the frequency effect should, if anything, be largest for the group of English native speakers. Another possibility is that frequency effects are generally larger for second as compared with first language speakers. Indeed, van Wijnendaele and Brysbaert (2002) and de Groot et al. (2002) observed a larger frequency effect in standard word naming in L2 when compared with L1. Such a difference can be accounted for in terms of the nonlinear form of the frequency effect, with an increased frequency sensitivity in the lower frequency range; for unbalanced L2 speakers, the subjective frequency distribution of L2 words is likely to be shifted toward the most sensitive left region, when compared with monolinguals.

Morphological Family Size

With respect to the quite recent finding that word recognition rate is influenced by morphological family size, the three bilingual groups showed a similar pattern of results as in the lexical decision task in Schreuder and Baayen (1997) and in Dijkstra et al. (2005): Morphological family size facilitated RTs above and beyond the effect of word frequency. For the native control group, however, the effect was not significant. Similarly, Schreuder and Baayen also failed to find an effect of morphological family size in the PDM task. Possibly, the effect of family size is magnified in a second language owing to an increased sensitivity to the number of occurrences of a word, similarly to what we observed for effects of word frequency.

Number of Higher Frequency English Neighbors

In line with previous results obtained with the PDM paradigm, we found that the number of higher frequency English neighbors slowed down RTs for both native and nonnative participants, whereas other neighborhood measures (total number of neighbors or their cumulative frequency) did not significantly affect RTs. According to the multiple read-out model (Grainger & Jacobs, 1996), the PDM task is especially sensitive to lateral inhibition by strong competitors, because the target word has to be unambiguously identified (unlike in the lexical decision task, for example). This requires the target word representation to pass an activation threshold, which is delayed by the simultaneous activation of high-frequency competitors. Lower frequency neighbors cause relatively little lateral inhibition and therefore have little effect on the activation rate of the target word (unlike in lexical decision, where global lexical activation is another factor determining RTs).

Semantic and Syntactic Ambiguity Variables

The results for the bilingual participants for the two ambiguity variables included in the present analyses showed a facilitatory effect of the number of word entries—that is, syntactically ambiguous words were recognized more quickly. Thus, when recognizing isolated words, the bilinguals seemed to benefit from multiple syntactic representations, possibly owing to the coactivation of both readings, as demonstrated by Elston-Güttler and Friederici (2005). Whereas the effect was smaller for the French compared with the Dutch and German participant groups, it was completely absent for native speakers. This is in line with Elston-Güttler and Friederici's observation of a more effective disambiguation mechanism in native as opposed to nonnative speakers. Apparently, the more proficient the speakers are, the less they show an effect of multiple syntactic representations. With the French group being the most proficient among the three bilingual groups and the natives obviously being even more proficient than the French, this would account for the observed effect patterns.

The second ambiguity variable, number of unrelated meanings, did not have a significant effect on the RTs for any of the participant groups in this experiment. This is in contrast with a large body of literature that has demonstrated effects of the number of meanings on word recognition (e.g., Borowsky & Masson, 1996; Kellas, Ferraro, & Simpson, 1988). The discrepancy might be due to task differences, as most of the existing literature has made use of the lexical decision paradigm. Note that it has been shown before that semantic ambiguity effects are modulated by task requirements (Hino, Lupker, & Pexman, 2002). Furthermore, the present ambiguity count was different from that in the majority of previous studies; in fact, in many of them, the ambiguity variable is a mixture of the number of unrelated word meanings, the number of related word senses, and the number of syntactic roles a word can take. For instance, more than 40% of the "ambiguous" stimuli of Experiment 2 in the frequently cited study by Borowsky and Masson (1996) have only one unrelated meaning, according to *The Wordsmyth Dictionary*. The roles of the different forms of lexical ambiguity are at present unclear. For the present task and populations, though, we can conclude that the number of strictly unrelated meanings did not influence word recognition in either L2 or L1.

Semantic Variables

The results with respect to the three semantic variables show that whereas concreteness had no significant effect on the RTs, meaningfulness and familiarity both facilitated them. These findings are noteworthy in several respects. First, the results indicate that even a primarily perceptual task like PDM involves some degree of semantic processing; moreover, this is the case not only in L1 but also in L2, where semantic processing might be thought to be weaker and/or slower than in L1 (Ardal, Donald, Meuter, Muldrew, & Luce, 1990; Kotz & Elston-Güttler, 2004). Second, the semantic ratings from the MRC database are taken from first, not second, language speakers of English; nonetheless, meaningfulness and familiarity influenced both native and nonnative speakers. Apparently, the variables captured word aspects that are universal across speakers with different language backgrounds.

Among the three bilingual groups, the results also showed that the French participants were more strongly affected by meaningfulness than the Dutch and the German group. This time, this interaction is in line with the concept of orthographic depth, in particular, with the claim that speakers of orthographically deep languages show enhanced semantic processing (Cuetos & Barbón, 2006; de Groot et al., 2002; Katz & Feldman, 1983): The bilingual group with the deepest L1 orthography, French, showed the largest effect of a purely semantic variable, meaningfulness. Thus, the data suggest that although semantic processing in English was very similar across speakers of different native languages, there was some modulation of the size of semantic effects that is in agreement with the orthographic depth hypothesis.

Between-Language Variables

Previous studies have indicated that the two languages of a bilingual interact during word recognition in a given language. In these studies, two classes of variables have usually been used as indicators for cross-lingual interaction: cross-language orthographic neighborhood variables and interlingual homographs and orthographically identical cognates. Both clusters of variables were included in the present study as well.

The data showed that, similar to the results of de Groot et al. (2002), none of our L1 neighborhood measures (number of high- and low-frequency neighbors, total number and summed frequency of neighbors) turned out to be a significant predictor of response latencies. Thus, in the context of the many within-language influences already present in the model, there was no evidence of cross-language neighbors from the participants' native language becoming active upon the presentation of the English target word.

In contrast, a significant effect of cognate status was observed: Bilinguals recognized cognates faster than noncognates, with similar effect sizes across all three bilingual groups. However, no such effect was found for noncognate homographs or false friends. This pattern of results is consistent with the picture in the literature, as summarized in the introduction: Cognate effects are robust, whereas effects of interlingual homographs are more variable and sometimes even reverse (see also Dijkstra et al., 1999, for effects of cognates and homographs in a PDM task). However, note that the number of noncognate homographs among the stimuli was relatively small (about 60 out of 1,025 words, depending on the respective L1; see Table 4). In contrast to the balanced designs used in the described factorial studies, the imbalance between homographs and nonhomographs might in this case complicate the comparison between these two categories. Therefore, we conducted an additional analysis, including only those words that are noncognate homographs in the participants' L1, and investigated whether the word's frequency in the L1 had an effect on the speed with which it was recognized. Indeed, for this small set of words, there was an inhibitory effect of L1 frequency on RTs. This result confirms the general pattern observed in previous factorial studies: The recognition of noncognate interlingual homographs in L2 depends on their frequency in L1, with, in particular, a high L1 frequency making recognition more difficult. Thus, although a microanalysis was able to confirm the results of previous studies with factorial designs, it also became apparent that in the grand picture of a large representative sample of English word stimuli, interlingual (noncognate) homograph effects do not play an important role during word recognition.

Summary

In summary, the detailed picture of the effects of a host of lexical variables provided a number of important insights on recent and unsettled issues in the mono- and bilingual word-recognition literature—for example, the findings concerning the relative importance of spoken relative to written word frequency in L1 and L2, the difference in the size of the frequency effect both between men and women and between native and nonnative speakers, the lack of an effect of lexical ambiguity, and the greater relevance of minimal as opposed to mean bigram frequency. Moreover, many effects that have previously been reported for (usually English-speaking) monolingual populations and for different experimental tasks have now for the first time been shown to generalize to L2 speakers with different L1s and to the PDM task (e.g., effects of morphological family size, higher frequency orthographic neighbors, and familiarity and meaningfulness).

In general, the multiple regression analysis confirmed the picture obtained in the analyses before: The similarities across the three bilingual groups were substantial. Among the 20 variables that were included in the regression, only 2 (number of CELEX entries and meaningfulness) affected the three groups in significantly different ways; in both cases, this interaction was caused by a difference in the size of the effect, not in the direction of it.

Besides the large overlap between the groups, the multiple regression analyses also showed that the degree of cross-language interaction in bilingual word recognition was limited. Among the between-language variables, only cognate status had an overall measurable effect. Apparently, across a large and unbiased set of words from the second language, bilinguals processed only words that overlap in all respects between the two languages differently from other words. If it is assumed that cognates share the same orthographic and/or semantic representation in the bilingual lexicon (Gollan et al., 1997; Sánchez Casas et al., 1992), the cognate effect might be explained even without the online activation of the native language: The effect would then simply be a consequence of the higher cumulative frequency of the joint representation.

However, whereas all other between-language effects on the whole data set were nonsignificant, a closer look at a subset of words revealed that the native language must have been activated. For noncognate homographs or false friends, the frequency of their reading in the native language had an inhibitory effect on RTs. This is in line with previous studies and supports the view of bilingual lexical access being principally nonselective with respect to language. However, the fact that homograph status had no effect when analyzed across the complete data set shows that globally, whether an English word had the same spelling as a word from the participants' L1 did not play an important role.

The regression analyses also provided some important and surprising insights concerning the comparison of L1 and L2 speakers. On first sight, the pattern of effects found for natives in the regression analysis was fairly similar to that of the nonnatives, with the majority of significant predictors overlapping between the two groups. However, a closer investigation revealed a number of differences as well: Whereas both written and spoken frequency independently affected word recognition times in nonnative speakers, only spoken frequency played a role for native speakers. Moreover, the total effect of word frequency was larger for nonnative than for native speakers. Furthermore, the effects of morphological family size and the number of syntactic representations disappeared for native speakers. Generally, this pattern of differences shows that compared with native speakers, nonnative speakers are more sensitive to the number of occurrences of a word (i.e., how often the word itself occurs or in how many derivations and syntactic classes it occurs). Presumably, this increased sensitivity is a consequence of the lower subjective frequency of words in an L2: Small changes in the number of times or ways a word is encountered have the largest consequences when the degree of experience with the given word is low, which tends to be the case for L2 words more than for words from one's mother tongue.

Thus, the results show that although the overlap between native and nonnative speakers of English is still large, the specific way in which within-language word variables that are related to frequency and ways of occurrence affect performance is different for L1 and L2 word recognition, suggesting that the organization and function of the L1 and L2 language systems is not completely identical.

Conclusions

Three major conclusions can be drawn from the present results. First, bilingual speakers of different native languages process L2 words in largely the same way. Second, and in line with the first conclusion, the extent of cross-language influences in L2 word recognition, when investigated across a large, unbiased set of words, is small. Finally, the data indicate that word recognition in L1 and L2 differ primarily with respect to the sensitivity to frequency-related variables (such as written vs. spoken frequency, morphological family size, and number of syntactic categories). Thus, even though there are not many L1-specific effects on L2 word recognition, L2 speakers differ from monolinguals in terms of frequency-related aspects of the organization of their language processing system.

The high degree of similarity between the three bilingual groups is in contrast with what had been expected in terms of an extension of the orthographic depth hypothesis. Even though the set of languages that we tested admittedly does not include the full variation in orthographic depth (not including extremely deep languages like Arabic or Chinese, for example), it still covers a representative range of orthographies when considering European alphabetic languages, as the present study intended (given that these are still the most intensively studied languages in experimental psycholinguistic research). Such variations are thought to cause differences in lexical processing, which become visible, for instance, in the effects of word frequency, word length, bigram frequency, or semantic variables (de Groot et al., 2002; Frost et al., 1987; Ziegler et al., 2001). Although this has primarily been argued for monolingual speakers, the question arises how bilingual speakers with differently deep native languages deal with a second language. For instance, Wang et al. (2003) demonstrated differences in English word recognition between native speakers of Korean (with a shallow orthography) and Chinese (with a deep writing system). In our data, the only indication for orthographic depth of L1 affecting word recognition in English was the effect of meaningfulness, which differed for the native speakers of French compared with the other bilingual groups in a way that is in line with the concept of orthographic depth. This finding suggests that even though there was an effect of semantic variables for all participants, the degree of their involvement was stronger for participants with an orthographically deep mother tongue. This might be a consequence of speakers of orthographically deep languages being accustomed to basing their responses on the full (including the semantic) word representation rather than on the orthographic code alone. Of course, whether other language combinations with larger contrasts in orthographic depth (e.g., including non-European languages) or a different target language than English might give rise to increased differences in lexical processing, thus lending more support to the orthographic depth hypothesis, remains to be tested. At this point, our data show that for native speakers of different western European languages having English as an L2 (just like, as mentioned in the introduction, 38% of the people within the European Union), L2 processing is very similar and largely unaffected by L1.

One reason for the large degree of overlap of the different bilingual participant groups might be that proficient bilinguals are able to adapt to the requirements of a second language, even though these might differ from those of their native language. In this vein, de Groot et al. (2002) claimed that Dutch–English bilinguals performed Dutch and English word-naming tasks in different ways, owing to the difference in orthographic depth between the two languages. Similarly, Meschyan and Hernandez (2006) reported that in English–Spanish bilinguals, different brain regions were involved while they were reading orthographically transparent Spanish words as compared with less transparent English words. Such an adaptation of bilinguals to the most appropriate lexical access strategies of the target language is likely to develop with increasing L2 proficiency. Further research is needed in order to clarify whether less proficient participants with varying native languages would display larger differences in L2 word recognition than the present groups did.

As a second important main finding, the limited role of between-language predictors evident in the present data qualifies and extends previous findings regarding the selectivity of lexical access in bilinguals. Our results confirm the existence of some cross-language interactions as shown by factorial studies before, and thus indicate that the bilingual word-recognition system is not principally divided by language and that its architecture allows for coactivation of both languages. However, the present analyses also show that for a word recognition task in a single target language involving a proficient bilingual population, a substantial part of the data can already be accounted for by characteristics of the target language itself, with only a small or barely present additional contribution of cross-language influences. This led to only minimal processing differences between participants with different language backgrounds.

The largest group difference we found in the present data was the one between first and second language speakers, which indicates that L2 processing (regardless of specific cross-language influences) is fundamentally different from word processing in L1. This is in accordance with Grosjean's (1989) warning, "Neurolinguists, beware! The bilingual is not two monolinguals in one person."

References

Akamatsu, N. (2002). A similarity in word-recognition procedures among second language readers with different first language backgrounds. *Applied Psycholinguistics, 23,* 117–133.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers, 36,* 140–155.

Allen, P. A., Smith, A. F., Lien, M.-C., Grabbe, J., & Murphy, M. D. (2005). Evidence for an activation locus of the word-frequency effect in lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 713–721.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461.

Ardal, S., Donald, M. W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language, 39,* 187–205.

Baayen, R. H. (in press). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge, England: Cambridge University Press.

Baayen, R. H., Feldman, L., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55,* 290–313.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (Release 2) [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R. H., Tweedie, F. J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon. *Brain and Language, 81,* 55–65.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29,* 639–647.

Bates, D. M., & Sarkar, D. (2005). The lme4 library. Retrieved July 28, 2007, from http://lib.stat.cmu.edu/R/CRAN/.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33,* 73–79.

BNC Consortium. (2001). *The British national corpus* (Version 2) (BNC World). Oxford, England: Oxford University Computing Services [Distributor]. Available from http://www.natcorp.ox.ac.uk

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 63–85.

Broadbent, D. E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior, 7,* 569–571.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society, 13,* 212–214.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 857–871.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1084–1096.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish–English cognates. *Quarterly Journal of Experimental Psychology, 38*(A), 367–393.

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18,* 415.

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47,* 91–124.

de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology, 53A,* 397–428.

de Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes, 15,* 329–365.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41,* 496–518.

Dijkstra, T., Moscoso del Prado Martín, F., Schulpen, B., Schreuder, R., & Baayen, R. H. (2005). A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes, 20,* 7–41.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dijkstra, T., van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51–66.

Dufau, S., Stevens, M., & Grainger, J. (in press). Windows executable software for the progressive demasking task (PDM). *Behavior Research Methods, Instruments, & Computers*.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1340–1359.

Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language, 52,* 256–283.

European Commission. (2006). *Special Eurobarometer 243: Europeans and their languages*. Retrieved July 28, 2007, from http://ec.europa.eu/public_opinion/archives/ebs/ ebs_243_sum_en.pdf

Faraway, J. J. (2005). *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models*. Boca Raton, FL: Chapman & Hall/CRC.

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51,* 1–22.

Frost, R. (1994). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 116–129.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 104–115.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256–281.

Gilhooly, K. J., & Logie, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologica, 50,* 21–34.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122–1139.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language, 24,* 631–645.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics, 47,* 191–198.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36,* 3–15.

Hino, Y., Lupker, S. J., Ogawac, T., & Sears, C. R. (2003). Masked repetition priming and word frequency effects across different types of Japanese scripts: An examination of the lexical activation account. *Journal of Memory and Language, 48,* 33–66.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 686–713.

Hino, Y., Pexman, P. M., & Lupker, S. J. (2006). Ambiguity and relatedness effects in semantic tasks: Are they due to semantic coding? *Journal of Memory and Language, 55,* 247–273.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41,* 401–410.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 130–136.

Johnson, R. C., Frincke, G., & Martin, L. (1961). Meaningfulness, frequency, and affective character of words as related to visual duration threshold. *Canadian Journal of Psychology, 15,* 199–204.

Johnson, R. C., & Zara, R. C. (1964). The influence of word meaningfulness on visual duration threshold at various frequency levels. *Journal of General Psychology, 70,* 235–239.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1312–1318.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 157–166.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Oxford, England: North-Holland.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601–609.

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & de Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research, 1068,* 170–183.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language, 45,* 259–282.

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language, 81,* 205–223.

Kolers, P. A. (1966). Reading and talking bilingually. *American Journal of Psychology, 79,* 357–376.

Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics, 17,* 215–235.

Kreuz, R. J. (1987). The subjective familiarity of English homophones. *Memory & Cognition, 15,* 154–168.

Kroll, J. F., & Dijkstra, T. (2001). The bilingual lexicon. In R. Kaplan (Ed.), *Handbook of applied linguistics* (pp. 301–321). Oxford, England: Oxford University Press.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 92–107.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlexical homographs: Effects of code similarity in language specific and generalized lexical decision. *Memory & Cognition, 32,* 533–550.

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19,* 585–611.

Locascio, D., & Ley, R. (1972). Scale-rated meaningfulness of 319 CVCVC words and paralogs previously assessed for associative reaction time. *Journal of Verbal Learning and Verbal Behavior, 11,* 243–250.

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

Martensen, H., Maris, E., & Dijkstra, T. (2000). When does inconsistency hurt? On the relation between phonological-consistency effects and the reliability of sublexical units. *Memory & Cognition, 28,* 648–656.

McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics, 22,* 249–261.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McGinnies, E., Comer, P. B., & Lacey, O. L. (1952). Visual-recognition thresholds as a function of word length and word frequency. *Journal of Experimental Psychology, 44,* 65–69.

Meara, P. M. (1996). *English Vocabulary Tests: 10k*. Unpublished manuscript, Center for Applied Language Studies, Swansea, Wales.

Meschyan, G., & Hernandez, A. E. (2006). Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation. *NeuroImage, 29,* 1135–1140.

Muljani, D., Koda, K., & Moates, D. R. (1998). The development of word recognition in a second language. *Applied Psycholinguistics, 19,* 99–113.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45–52.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers, 36,* 516–524.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76,* 1–25.

Perea, M., Carreiras, M., & Grainger, J. (2004). Blocking by word frequency and neighborhood density in visual word recognition: A task-specific response criteria account. *Memory & Cognition, 32,* 1090–1102.

Perea, M., & Rosa, E. (2000). The effects of orthographic neighborhood in reading and laboratory word identification tasks: A review. *Psicologica, 21,* 327–340.

Piercey, C. D., & Joordens, S. (2000). Turning an advantage into a disadvantage: Ambiguity effects in lexical decision versus reading tasks. *Memory & Cognition, 28,* 657–666.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus*. New York: Springer.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition, 3,* 513–518.

Richards, L. G. (1976). Concreteness as a variable in word recognition. *American Journal of Psychology, 89,* 707–718.

Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46,* 245–266.

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28,* 89–104.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81,* 99–118.

Samson, D., & Pillon, A. (2004). Orthographic neighborhood and concreteness effects in the lexical decision task. *Brain and Language, 91,* 252–264.

Sánchez Casas, R. M., Davis, C. W., & García Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology, 4,* 293–310.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26,* 1270–1281.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37,* 118–139.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Snodgrass, J. G., & Mintzer, M. (1993). Neighborhood effects in visual word recognition: Facilitatory or inhibitory? *Memory & Cognition, 21,* 247–266.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8,* 411–416.

Spreen, O., & Schulz, R. W. (1966). Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning and Verbal Behavior, 5,* 459–468.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 117–138.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

Ullman, M. T., Estabrooke, I. V., Steinhauer, K., Brovetto, C., Pancheva, R., Ozawa, K., et al. (2002). Sex differences in the neurocognition of language. *Brain and Language, 83,* 141–143.

van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780–789.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition, 15,* 181–198.

van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 616–627.

von Studnitz, R., & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition, 5,* 1–23.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition, 87,* 129–149.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. *Brain and Language, 81,* 66–78.

Whaley, C. P. (1978). Word–nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143–154.

Williams, R. S., & Morris, R. K. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology, 16,* 312–339.

Wilson, M. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavior Research Methods, Instruments, & Computers, 20,* 6–11.

Wood, S. (2006). *Generalized additive models: An introduction with R.* Boca Raton, FL: Chapman & Hall/CRC.

Wordsmyth. (2002). *The Wordsmyth Dictionary.* Retrieved February 29, 2004, from http://www.wordsmyth.net

Wurm, H. L., Aycock, J., & Baayen, R. H. (2007). *Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities*. Manuscript submitted for publication.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12,* 379–384.

Appendix A

Results of the Mixed-Effects Models Predicting Familiarity and Meaningfulness Values, Respectively, for Bilingual Participants

Effect	β	*t*(654)	*p*
Predicted variable: Familiarity			
Spoken word frequency	13.66	21.78	<.001
Number of higher frequency English neighbors	−3.71	−2.60	<.01
Concreteness	0.02	2.57	<.02
Meaningfulness	0.26	13.11	<.001
Predicted variable: Meaningfulness			
Written word frequency	8.29	3.51	<.001
Spoken word frequency	3.94	1.78	<.03
Concreteness	0.11	5.97	<.001
Number of CELEX entries	−11.82	−2.76	<.01

Appendix B

Complete Results of the Mixed-Effects Model for the Control Group of English Native Speakers

Effect	β	*t*(20140)	*p*
Session	−.054	−38.07	<.001
Trial (within session)	−.0001	−9.26	<.001
Previous RT	.210	34.22	<.001
Error	−.061	−7.40	<.001
Sex	−.001	−0.02	.986
Number of letters	.029	5.58	<.001
Spoken word frequency (linear)	−.017	−4.19	<.001
Spoken word frequency (quadratic)	.001	2.14	<.05
Minimal bigram frequency	.005	3.16	<.003
Number of higher frequency English neighbors	.017	5.08	<.001
Sex × Spoken Word Frequency[a]	.004	2.70	<.01

Note. RT = reaction time.
[a] Additional effect for men compared with women.

Received February 7, 2007
Revision received July 13, 2007
Accepted July 26, 2007 ■