Memory & Cognition
2004, 32 (4), 533-550

Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision

KRISTIN LEMHÖFER and TON DIJKSTRA
University of Nijmegen, Nijmegen, The Netherlands

In four experiments, we investigated how cross-linguistic overlap in semantics, orthography, and phonology affects bilingual word recognition in different variants of the lexical decision task. Dutch–English bilinguals performed a language-specific or a generalized lexical decision task including words that are spelled and/or pronounced the same in English and in Dutch and that matched one-language control words from both languages. In Experiments 1 and 3, "false friends" with different meanings in the two languages (e.g., *spot*) were presented, whereas in Experiments 2 and 4 cognates with the same meanings across languages (e.g., *film*) were presented. The language-specific Experiments 1 and 2 replicated and qualified an earlier study (Dijkstra, Grainger, & Van Heuven, 1999). In the generalized Experiment 3, participants reacted equally quickly on Dutch–English homographs and Dutch control words, indicating that their response was based primarily on the fastest available orthographic code (i.e., Dutch). In Experiment 4, cognates were recognized faster than English and Dutch controls, suggesting coactivation of the cognates' semantics. The nonword results indicate that the bilingual rejection procedure can, to some extent, be language specific. All results are discussed within the BIA+ (bilingual interactive activation) model for bilingual word recognition.

In a foreign country where an unfamiliar language is spoken, words such as *hotel*, *taxi*, and *café* can often still be recognized because they possess the same or a similar spelling and meaning across languages. Such words are called *cognates*. However, there may also be misleading words in the foreign language that are identical in spelling but different in meaning to words from one's native language. These items are called *false friends* or (noncognate) *interlingual homographs*. An example is the word *spot*, which means "mockery" in Dutch. Apart from spelling (orthography) and meaning (semantics), a third code thought to play a major role in word processing is sound (phonology), which can also be shared between words of different languages. For example, the English word *cow* is pronounced very much like the Dutch word *kou* ("cold"). Items with similar pronunciations across languages are called *interlingual homophones*.

Interlingual homographs and cognates have been the most important sources of stimulus materials in studies attempting to unravel the process of bilingual word recognition. Through such words, a wealth of studies in the last decade have revealed that during the initial stages of word identification by bilinguals, word candidates from several languages are often coactivated (see Dijkstra & Van Heuven, 2002, for an overview). In accordance with these results, several word recognition models propose that bilingual word recognition involves an initial language-nonselective access process into an integrated lexicon. According to the BIA+ (bilingual interactive activation) model (Dijkstra & Van Heuven, 2002), the visual presentation of a word to a bilingual leads to parallel activation of orthographic input representations in the native language (L1) and the second language (L2). These representations then activate associated semantic and phonological representations, leading to a complex interaction (or *resonance process*) between codes from which the lexical candidate corresponding to the input word emerges and is recognized.

Furthermore, the BIA+ model makes predictions about a number of important issues that are still unresolved and debated in the literature. First, with respect to representational issues, it is still unclear exactly how cognates and interlingual homographs are represented in the bilingual lexicon. The BIA+ model proposes that interlingual homographs have separate representations for each language, whereas it remains possible that cognates have shared representations (Dijkstra & Van Heuven, 2002). This proposal is based on hints in the data from earlier studies (e.g., Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998), but no solid evi-

This research was part of a PhD project performed by the first author within the framework of a MaGW International Comparative Research project entitled "Lexical Competition in Bilinguals: A Cross-Language Comparison," granted by the Netherlands Organization for Scientific Research to the Nijmegen Institute for Cognition and Information (NICI). We thank Herbert Schriefers, Marc Brysbaert, Debra Jared, Dominiek Sandra, and Heike Martensen for valuable discussions during various stages of this research. Correspondence concerning this article should be addressed to K. Lemhöfer, NICI, University of Nijmegen, P. O. Box 9104, 6500 HE Nijmegen, The Netherlands (e-mail: k.lemhofer@nici.kun.nl).

dence supporting this claim is available. The first aim of the present study is to test the BIA+ approach to interlingual homograph and cognate representations.

Second, with respect to word processing issues, it is still unclear to what extent the language (non)selectivity of result patterns is task dependent. To test the generality of the language-nonselective access hypothesis and the dependence of lexical retrieval on task demands, we need cross-task comparisons using identical stimulus materials. According to the BIA+ model, the activation of various lexical representations is constantly monitored by a task/decision system that subserves task execution and decision making (see Green, 1998). The task/decision system systematically uses the activation pattern in the word identification system to optimize responding. The BIA+ model predicts that different tasks will lead to systematically different response patterns, because responding can occur at different moments in time and can be based on different information sources.

This point can be clarified by the following example. In an English language-specific lexical decision task, participants must respond "yes" to English words and "no" to nonwords and words from languages other than English (if any). In contrast, in a Dutch–English generalized lexical decision task, bilinguals respond "yes" to words from either of their two languages and "no" to items that are nonwords in both of them. According to BIA+, the result patterns for interlingual homographs and cognates will differ for the two tasks. For Dutch bilinguals performing the English language-specific lexical decision task, the L2 (English) target reading of an interlingual homograph will become active relatively late. This allows the L1 (Dutch) reading to affect target processing. In the generalized lexical decision task, however, bilinguals can respond to the first reading of the homograph they identify. Because this will often be their L1 reading, the cross-linguistic effect (measured relative to Dutch controls) should be considerably smaller. The second aim of the present study is to test this view on task demands for these two variants of the lexical decision task.

We will first replicate and extend an earlier study involving language-specific lexical decision by Dijkstra et al. (1999). We will use their interlingual homograph and cognate materials, but in separate experiments rather than in only one (as they did). This replication "with a twist" is important for a number of reasons. As we shall see, some doubts have arisen with respect to the stability of the data patterns in that study. Furthermore, performing separate experiments for homographs and cognates will indicate the extent to which the earlier results may have been dependent on stimulus list composition, and it will make the language-specific decision experiments comparable in this respect to the generalized lexical decision task in Experiments 3 and 4. Next, we will include these materials in a generalized lexical decision task (see, e.g., Dijkstra et al., 1998; Van Heuven, Dijkstra, & Grainger, 1998). The comparison of the two lexical decision tasks will provide insight into the effect of variations in task demands on bilingual lexical processing.

Third, with respect to nonword processing, it is unknown to what extent the similarity of nonwords to words from one language or the other affects nonword rejection times and error rates. How nonwords resembling words in one or the other language are rejected can, in turn, inform theories on the structure of the bilingual lexicon. However, no currently available bilingual model makes any predictions about the nonword rejection procedure in lexical decision. The third aim of the present study is to collect evidence on nonword rejection in bilinguals and specify the BIA+ model with respect to this mechanism.

Below, when the experiments of the present study are described, the predictions of the BIA+ model with respect to representational issues, task demands, and nonword processing will be considered in more detail. However, we will first present an overview of studies examining the role of orthographic, phonological, and semantic overlap in the processing of interlingual homographs and homophones in word recognition tasks such as lexical decision and perceptual identification. The predictions of BIA+ are based on these studies. Next, we will focus on the present study by reviewing the earlier study by Dijkstra et al. (1999) and our intended variations in task demands. Finally, we will consider our nonword manipulations and indicate in detail how we manipulated their resemblance to words from the two target languages.

The Orthographic Representation of Cognates and Interlingual Homographs

Studies involving stimulus words with the same spelling in two languages can be distinguished according to whether the items were (homographic) cognates (also having semantic overlap across languages) or interlingual homographs (having different meanings across languages).

Many studies have shown that homographic cognates such as *film* in English and Dutch are processed faster than one-language control words during lexical decision in L2 (Caramazza & Brones, 1979; Cristoffanini, Kirsner, & Milech, 1986; Dijkstra et al., 1998). However, for lexical decision performed in L1, the findings are less clear. Caramazza and Brones failed to find a cognate effect in the dominant-language task, but Van Hell and Dijkstra (2002) reported a response time (RT) advantage for lexical decisions on native language words that were cognates with respect to the second language and, for sufficiently proficient participants, even for decisions on words that were cognates with respect to a third language. Font (2001) also obtained a facilitation effect in the dominant language for Spanish–French cognates. The usually stronger effects from L1 on L2 than from L2 on L1 indicate that L2 representations are generally activated less strongly or less rapidly than L1 representations, implying that they have less chance to affect the response when L1 is the target language.

In contrast, interlingual homographs (or homographic noncognates, such as *spot*) usually led to small or no RT differences relative to one-language control words in lexical decision tasks in which L2 was the target language and the stimulus list was monolingual (De Groot, Delmaar, & Lupker, 2000; Dijkstra et al., 1998; Gerard & Scarborough, 1989). An exception is the study by Von Studnitz and Green (2002), in which inhibition effects were found for homographs relative to controls. However, RT differences between homographs and controls in that study were also found for a control group of monolingual participants, so part of the inhibition in bilinguals may be due to incomplete stimulus matching (note that the relatively slow RTs in bilinguals may lead to inflated RT differences between homographs and controls). Furthermore, in that study only homographs that were of low frequency in the target language were used, which is known to give rise to inhibition rather than facilitation (Dijkstra et al., 1998).

Dijkstra et al. (1998) showed that the recognition of interlingual homographs is sensitive to stimulus list composition and task demands. In a lexical decision study with Dutch–English bilinguals, they manipulated the relative frequencies of homographs in the two languages and investigated how these words were processed in different task contexts. Over all frequency categories, the null effect of homographs in a standard English lexical decision task (Experiment 1) turned into a robust inhibitory effect when Dutch words were included in the stimulus list of items that had to be treated as nonwords (Experiment 2). When the same mixed-language stimulus list was used but Dutch words had to be accepted as words (a generalized lexical decision task—Experiment 3), the homographs were generally recognized faster than the English controls, but there was no effect in comparison with Dutch control words.

In the second experiment, the recognition of interlingual homographs depended on the *relative frequency* of the two readings in the two individual languages. When the homographs had a low frequency in English (the target language) and a high frequency in Dutch, strong inhibition effects occurred for the homographs relative to one-language controls. This finding is incompatible with the notion of a shared orthographic representation in the mental lexicon, because such a word representation would be characterized by a common, cumulative frequency and would not be affected by the relative frequencies in the two languages. However, given other empirical findings and simulation results (see, e.g., Dijkstra et al., 1999) that are most easily explained in terms of a single orthographic node for homographs, further evidence is needed.

In sum, when homographic *cognates* are processed in a second-language context, the first-language reading seems to become active as well and to facilitate recognition. The few available studies on the recognition of cognates in a first-language context indicate that under these circumstances, cognate effects are weaker but still present. For *noncognate interlingual homographs*, the results are more variable. Homograph effects seem to depend on several factors, such as the frequency characteristics of the words, the task requirements, and the mono- or bilingual composition of the stimulus list.

The Phonological Representation of Cognates and Interlingual Homophones

Little research has been conducted to investigate whether the presentation of (heterographic) *interlingual homophones* in one language triggers the activation of their homophonic mates from the other language. Like homographs, words with (almost) the same phonology in two languages can be divided into two categories: (*homophonic*) *cognates*, which share meaning (e.g., *wiel*, which is the Dutch word for *wheel* with about the same pronunciation), and (*noncognate*) *homophones*, which have different meanings (e.g., *cow–kou*, where *kou* means "cold" in Dutch).

Homophonic cognates of the *wheel–wiel* type, however, have almost never been studied within standard bilingual word recognition paradigms such as those involving word naming or lexical decision. An exception is the study by Dijkstra et al. (1999), including all possible types of interlingual homophones and homographs, which will later be described in detail (for another exception concerning cross-script cognates, see Gollan, Forster, & Frost, 1997).

Interlingual homophones of the *cow–kou* type have been investigated somewhat more frequently. In a study with balanced English–Afrikaans bilinguals, Doctor and Klein (1992) included interlingual homographs, interlingual homophones, and English and Afrikaans control words in a generalized lexical decision task. Interlingual homophones were processed more slowly than the interlingual homographs, and RTs for English and Afrikaans control words lay between those for homographs and homophones. Unfortunately, the authors did not report on the comparison between homophones and English or Afrikaans controls, although the 84- and 89-msec inhibition effects seem fairly reliable. The study thus gives some indication of an involvement of the phonological word representations of both languages in bilingual word recognition. However, a large proportion of the nonwords in the study were pseudohomophones, which, according to monolingual studies (e.g., Pexman, Lupker, & Jared, 2001), may have exaggerated the size of the inhibitory homophone effects.

More evidence of cross-language phonological effects (on the sublexical rather than on the lexical level) comes from a study in which masked priming was used in a perceptual identification task with Dutch–French bilinguals (Brysbaert, Van Dyck, & Van De Poel, 1999; see also Van Wijnendaele & Brysbaert, 2002). French target words that were preceded by Dutch homophonic word or nonword primes were recognized more accurately than those that were preceded by nonhomophonic primes.

On the basis of these few studies, we conclude that under specific conditions, homophone effects seem to occur between languages, providing evidence for coactivation in L1 and L2 at the phonological level. Because the bilingual studies conducted on this issue are still too few to justify more definite conclusions, the present study aimed at further examining the interlingual homophone effect.

The Role of Representational Overlap in Cognates, Interlingual Homographs, and Homophones

The effect of representational overlap on the different stimulus types was studied in combination by Dijkstra et al. (1999) for language-specific lexical decision and progressive demasking. Because we will use their stimulus materials in the present study, we will describe them in some detail. Dijkstra et al. (1999) systematically manipulated the degree of semantic, orthographic, and phonological overlap in word recognition in L1 and L2. They created six categories of Dutch–English homographs and homophones. Three categories were classes of cognates: cognates that shared both spelling and sound (e.g., *film* in Dutch and English), those that were spelled but not pronounced the same (e.g., *fruit*, which is pronounced /frœyt/ in Dutch), and those that were homophonic but spelled differently (e.g., *wiel*–wheel). All cognates were translation equivalents by definition. With S referring to shared semantics, O to shared orthography, and P to shared phonology, the three classes of cognates may be labeled SOP, SO, and SP items, respectively. Analogously, three categories of false-friend items were defined that were similar in form but not in meaning: OP items overlapped in both spelling and sound (e.g., *spot*, Dutch for "mockery"), O items agreed only in spelling (e.g., *glad*, meaning "slippery" and pronounced /xlat/ in Dutch), and P items matched only in pronunciation (e.g., *kou*, pronounced like English *cow*). No exclusively Dutch words were included in this study.

Dijkstra et al. (1999) incorporated these false friends and cognates in two English tasks (lexical decision and progressive demasking) performed by Dutch–English bilinguals. In both tasks, orthographic and semantic overlap facilitated word processing relative to English control words, whereas homophony with Dutch words exerted an inhibitory influence on word recognition. Moreover, the overlap of codes seemed to have additive effects on word recognition. For example, in English lexical decision, all homographic cognates (SO and SOP items) were recognized faster than controls, but the effect was more pronounced for cognates without a phonological overlap (SO items) than for those with a phonological overlap (SOP items). For homophonic cognates without shared orthography (SP items), there was no effect.

A similar pattern was observed for false friends. Homographs that shared only their spelling (O items such as *glad*) were processed faster than controls, whereas homographs that overlapped in both spelling and sound (OP items such as *spot*) did not show any RT difference relative to English controls. For pure homophones (P items) such as *cow*, inhibition was observed. This finding suggested that inhibitory effects of phonological overlap, canceling out the facilitatory effects of orthographic overlap, underlay the weak homograph effects found in previous studies. However, more recently, a potential problem with respect to the homophonic P condition has become evident. Jared and Kroll (2001) noted that the inhibition effect in this condition may be untrustworthy, because participants in the monolingual control group produced more errors on the homophones than on the controls. Thus, the test and control words may have differed on properties other than those for which they were matched.

In other respects, the homograph results are in general agreement with findings in the monolingual domain: Words with multiple meanings within the same language have repeatedly been found to be recognized more quickly (Borowsky & Masson, 1996; Pexman & Lupker, 1999; but see also Rodd, Gaskell, & Marslen-Wilson, 2002), which is analogous to a facilitatory effect of orthographic overlap between languages. However, especially in naming, words that have not only several meanings but also several pronunciations (e.g., *wind*) have been found to be processed more slowly rather than faster than controls (Kawamoto & Zemblidge, 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

Dijkstra and Van Heuven (2002) explained the result pattern for interlingual homographs in English lexical decision by assuming that in this task Dutch participants respond on the basis of the English readings of the homographs, but that these generally become available more slowly than the Dutch readings. This proposal was called the *temporal delay hypothesis*, which was assumed to operate at the lexical level. As a consequence, responses to the English readings can be affected by the earlier available Dutch readings. This holds for all three types of lexical codes (orthography, phonology, and semantics).

The Present Study

In the present study, we first wish to assess the stability of the result patterns obtained by Dijkstra et al. (1999) by replicating the language-specific lexical decision experiment in two parts, one for interlingual homographs and one for cognates. First, this will allow us to assess whether or not the potential problems with the P condition are valid, and, second, it will show us to what extent the observed result patterns were a consequence of the combination of cognates and homographs in one experiment (i.e., an effect of stimulus list composition).[1]

Next, we will assess the effect of task demands by conducting a different variant of the lexical decision task. Dijkstra et al. (1999) showed cross-linguistic effects of all three codes in tasks in which the *second* language was the target language. In the present study, we also examine the effects of semantic, orthographic, and phonological cross-linguistic overlap in a generalized lexical decision task in which *both* languages are target languages, by comparing recognition of cognates and false friends with that of both Dutch and English control words. By using the same stimulus materials as did Dijkstra et al. (1999), as well as a comparable group of Dutch–English bilinguals in both an English language-specific and a Dutch–English generalized lexical decision task, the effects of cross-linguistic overlap can be compared for exclusively English and mixed Dutch–English lexical decision variants. The issue under investigation is, therefore, whether the effects found in the English task relative to English (L2) control words can also be demonstrated in comparison with Dutch (L1) control words in a mixed-language task. If so, semantic and orthographic overlap of cognates and orthographic overlap of false friends should cause facilitatory effects on word recognition, whereas phonological overlap should inhibit recognition performance. Alternatively, overlap effects might be task dependent and nonsignificant relative to Dutch control words in a mixed-language task. This would indicate that, whereas the simultaneous activation of an L1 code affects the recognition of words in the second language, the reverse is not the case.

As a special case of the broader issue of orthographic and semantic cross-linguistic interaction, our experiments allow us to test whether interlingual homographs and cognates are characterized by shared or separate orthographic representations across languages. A shared orthographic representation would lead to a maximal degree of interaction of L1 and L2 orthography and, due to cumulative frequency, would produce faster RTs than both L1 and L2 controls. For noncognate homographs, this is not what has been found in most previous studies; however, these studies did not take into account the amount of phonological overlap of the two homograph readings, which may have canceled out any facilitation. In the present study, phonological overlap is controlled. Therefore, the prediction is that if interlingual homographs and cognates have a common orthographic representation across languages, at least items without a large phonological overlap (O and SO items) should be recognized faster than both English and Dutch controls. On the other hand, if no such facilitation is found, the notion of a shared orthographic representation must be rejected. The absence of an effect for O and/or SO items relative to Dutch control words would also indicate that L2 lexical codes are generally activated later (as was argued by Dijkstra & Van Heuven, 2002) or more weakly than L1 lexical codes, and that participants reacted mainly on the basis of L1 representations.

In addition to the variation of ambiguity with respect to language membership in the words, a similar manipulation was realized in the nonword materials: Nonwords either were word-like with respect to both English or Dutch (e.g., *brank*) or were only Dutch-like (e.g., *muig*) or only English-like (e.g., *baint*). On the basis of the classical finding that in a lexical decision task word-like nonwords take longer to be rejected than other nonwords (Coltheart, Davelaar, Jonasson, & Besner, 1977), such a manipulation can be used to investigate whether the rejection procedure for nonwords is dependent on language similarity. In principle, three different outcomes could be expected and accounted for by different notions: (1) no RT differences between the three nonword types, (2) slower rejection for Dutch-like nonwords than for English-like nonwords, and (3) faster rejection for Dutch-like nonwords than for English-like nonwords.

The first prediction (i.e., no effect) can be derived from an interactive activation type of model (such as BIA+) if we assume that a nonword response is given if, after a specified period of time, no word candidate matching the visual input has been recognized. The deadline may be set later for higher levels of word-likeness (Grainger & Jacobs, 1996). Assuming that overall word-likeness were the same for all nonwords, one would expect that nonwords would be rejected with equal speed and accuracy.

The second prediction (i.e., Dutch-like nonwords are rejected more slowly than English-like nonwords) follows from assuming that English word representations on average have lower subjective (i.e., person based and proficiency dependent) frequencies than Dutch words do, given that the bilinguals in the present study were unbalanced. English-like nonwords, activating mainly English word candidates, should therefore cause less global lexical activity than Dutch-like nonwords, activating the "strong" word nodes of the mother tongue. In this way, Dutch-like nonwords would behave as more "word-like" than English-like nonwords. Therefore, Dutch-like nonwords should (in, e.g., the multiple read-out model; Grainger & Jacobs, 1996) elicit slower responses than English-like nonwords.

The third prediction (i.e., Dutch-like nonwords are rejected faster than English-like nonwords) follows if one assumes, in contrast to the second prediction, that different temporal deadlines are set for the acceptance or rejection of English and Dutch (non)words. Due to their lower subjective frequency, English (L2) words will now have later deadlines. This leads to slower rejection times for English-like nonwords than for Dutch-like nonwords. Note that this prediction assumes that somehow language membership information can be involved in bilinguals' nonword rejection procedure.

These issues concerning word and nonword processing by bilinguals were investigated in four experiments in which the word materials used in the earlier study by Dijkstra et al. (1999) were incorporated. In Experiments 1 and 2, we used an English lexical decision task and carefully mimicked the earlier study, except that the homograph and cognate materials were presented separately and new nonwords were used that were English-like or neutral (i.e., similar in structure to the word materials, which comprised English words and false friends). In Experiments 3 and 4, we used the same materials (and more) in a generalized lexical decision task. In Experiments 1 and 3, recognition of three classes of false friends (OP, O, and P items; e.g., *spot*, *glad*, and *cow*, respectively) was compared with that of control words. Similarly, Experiments 2 and 4 involved cognates of the types SOP, SO, and SP (e.g., *film*, *fruit*, and *wheel*, respectively). The nonwords in Experiments 3 and 4 were the same as those in Experiments 1 and 2, except that Dutch-like nonwords were added.

EXPERIMENT 1
English Lexical Decision Involving False Friends

Method

Participants

Twenty right-handed participants (8 men, 12 women), drawn from the same population (mostly undergraduates at the University of Nijmegen with Dutch as their native language) as those involved in Dijkstra et al. (1999), took part in the experiment. (Each participant took part in only one of the experiments of the present study.) They were between 18 and 57 (mean 24.5) years old and had, on average, 12.4 years of experience with the English language. On a scale from 1 (*very little experience*) to 7 (*very much experience*), the mean self-rating of their reading experience in English was 5.0 (*SD* = 1.28). All the participants had normal or corrected-to-normal vision. They were paid for their participation or received course credit.

Stimulus Materials

Words. The false-friend items and their matched English control words were taken from the false-friend conditions (OP, O, and P) in Dijkstra et al. (1999). These words were three to five letters long and possessed a frequency of at least two occurrences per million in English, as listed in the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). All materials are presented in the Appendix.

Each of the three false-friend conditions consisted of 15 words. Both OP and O items were what are traditionally called *interlingual homographs*. For OP words, there was almost complete phonological overlap between the two readings of the item, as for the word *spot* (Dutch meaning: "mockery"), which is pronounced /spɒt/ in English and /spɔt/ in Dutch. For the O items (e.g., *glad*), phonological overlap of the two pronunciations was as small as possible, whereas orthographic identity was maintained. In the P condition, the critical items were exclusively English words that are homophonic but not orthographically identical to a Dutch word, such as *cow*, which is pronounced very similarly to the Dutch word *kou* ("cold"). For all items, the semantic overlap of the English–Dutch competitor pairs was as small as possible. Twelve participants in the study by Dijkstra et al. (1999) had rated the items for the subjective degree of their phonological, orthographic, and semantic overlap with their respective competitors. The rating data supported the a priori classification of the stimuli into OP, O, and P items. The characteristics of the items are given in Table 1 (note that the characteristics of the Dutch control words reported in the table apply to Experiment 3 only).

Dijkstra et al. (1999) had matched each false friend to an English control word as closely as possible in length, English frequency, and consonant–vowel structure (see Table 1 of the present article). The total number of words in the experiment was 90 (15 test words and 15 control words in each of the three conditions). In addition to the item-by-item matching between false friend items and control words, there was a correspondence in frequency and length between the group means of the OP, O, and P conditions (all $ps > .40$).

Nonwords. For use in all experiments, a set of 150 nonwords consisting of 50 "English-like," 50 "Dutch-like," and 50 "neutral" nonwords, was constructed. First, a list of English and Dutch words with lengths of three to five letters was selected and turned into nonwords by replacing one letter with another. All items were orthographically legal in both English and Dutch, and the letter strings were neither homographic nor homophonic with an existing word in either language. Next, 100 "English" nonword items were selected that had more orthographic neighbors and a higher neighborhood frequency (i.e., summed frequency of all neighbors) in English than in Dutch. One hundred "Dutch" nonwords were selected in an analogous way. Finally, 100 "neutral" nonwords were chosen, which consisted of letter strings derived from unused English–Dutch homographs that had about the same number and summed frequency of neighbors in both languages. Ten Dutch–English bilinguals rated the 300 nonwords according to how English- or Dutch-like they were on a 7-point scale. Finally, the 150 nonwords that had been rated as being in best accordance with the a priori classification (as shown by their scores on the rating scale) were selected for the present study. Information about the selected nonwords is given in Table 2 and in the Appendix.

For the present English lexical decision experiment without Dutch words, only the "English-like" and "neutral" nonwords were included. Dutch-like nonwords were not used in order to preserve the "English" character of the experiment. Because the number of nonwords had to match the number of words in the list (90), five nonwords of each category were randomly excluded from the lists.

Procedure

Before the experiment, the participants read a written instruction in English explaining that they would have to decide whether or not a presented letter string formed a correct word in English. The participants were instructed to press a button with the index finger of their preferred hand if they thought the letter string was an English word, and to react with their other hand if it was a nonword. The participants were asked to react as quickly and accurately as possible.

Testing took place individually on a Macintosh Quadra computer controlled by software developed at the Nijmegen Institute for Cognition and Information. The participants were seated at a distance of about 60 cm from the computer screen, where stimuli were presented in black 18-point lowercase letters in Courier font on a white background.

Each trial began with the presentation of an asterisk in the center of the screen for 800 msec. The target stimulus appeared at the same place after another 300 msec. The target remained on screen until the participant responded or until 1,500 msec after stimulus onset. The next trial started 700 msec after the response was given.

Table 1
Frequency and Length Characteristics of the Word Stimuli Used in Experiments 1 and 3 (False Friends)

	Test Words						English Control Words		Dutch Control Words*	
Condition	EF	DF	Length	Sem.†	Orth.†	Phon.†	EF	Length	DF	Length
O	40.2	27.4	4.2	1.6	7.0	2.8	40.4	4.1	27.1	4.3
OP	40.2	27.9	3.9	2.1	7.0	6.0	40.3	3.9	27.7	3.9
P	41.7	29.1	3.9	1.2	2.8	6.0	41.9	3.9	29.5	3.7

Note—EF, English written frequency (per million); DF, Dutch written frequency (per million); Length, word length in letters; Sem., mean score of semantic similarity rating with the competitor word; Orth., mean score of orthographic similarity rating with the competitor word; Phon., mean score of phonological similarity rating with the competitor word. *Dutch control words were present in Experiment 3 only. †The ratings were based on a scale from 1 (*no similarity or overlap*) to 7 (*identity or perfect overlap*).

Table 2
Characteristics of the Nonwords Used in Experiment 1

Nonword Type	Length	EN	DN	ESFN	DFSN	Rating Score*
English-like	4.2	9.3	2.3	610.3	56.2	2.6
Dutch-like	4.2	1.0	8.2	34.8	1497.1	−2.5
Neutral	4.2	4.2	4.1	124.0	126.2	−0.04

Note—Length, word length in letters; EN, number of English neighbors; DN, number of Dutch neighbors; ESFN, summed frequency of English neighbors; DSFN, summed frequency of Dutch neighbors. *On a 7-point rating scale, –3 was assigned to the category *very Dutch-like*, 0 to *equally Dutch- and English-like*, and 3 to *very English-like*.

At the beginning of the experiment, each participant completed a practice block of 30 items with the same proportions of false friends, English words, and nonwords as the experimental stimulus list. Subsequently, the 180 experimental items were presented in two blocks of 90 items each. In addition, the first 2 items of each block were dummy items that were not included in the analyses. Between the blocks, the participants were free to take a break. The order of presentation of trials was determined by a pseudorandomization procedure, with no more than four words or four nonwords in a row.

After the task was performed, the participants filled out a questionnaire about their experience with the English language. In total, each experimental session lasted between 15 and 20 min.

Results

For the analysis of RTs, only correct reactions were considered. The overall error rate was 11.3%. Furthermore, RTs that lay more than two standard deviations away from both the item and the participant mean (for the given condition) were considered outliers and discarded from the analyses (accounting for an additional 2.0%). In total, 13.1% of the data were excluded.

The mean RTs, standard deviations, and error rates are listed in Table 3. (A summary of the effects in this study and in the earlier one is given in Table 10.)

Word Data

Taking into account the exceptional nature of the P condition (i.e., test words were exclusively English words rather than interlingual homographs), we first analyzed the homograph conditions (OP and O) and the P condition separately. For the homographs analysis over participants, a repeated measures analysis of variance (ANOVA) was conducted on RTs and error rates with condition (O vs. OP) and word status (test word vs. English control word) as within-participants factors. For the analysis over items, word status was a repeated measures factor (due to the item-by-item matching of test and control words), whereas condition was treated as a between-items factor.

Next, planned comparisons were carried out on each of the three word types relative to their controls, to enable a comparison with the effects obtained in Dijkstra et al. (1999).

O and OP conditions. O and OP items were analyzed together in a 2 × 2 ANOVA with condition (O vs. OP) and word status (test vs. control word) as factors. For RTs, there was a main effect of word status by participants, but not by items [$F_1(1,19) = 5.30, p < .05; F_2(1,28) = 2.04, p > .15$]. On average, homographs were recognized after 550 msec, relative to 566 msec for control words. In the participants analysis, there was also a main effect of condition, which was also not significant in the items analysis [$F_1(1,19) = 8.53, p < .01; F_2(1,28) = 1.72, p > .20$], with slower responses in the O condition (568 msec) than in the OP condition (548 msec). The word status × condition interaction was also significant in the analysis over participants only [$F_1(1,19) = 4.86, p < .05; F_2(1,28) = 2.21, p > .10$]. Planned comparisons indicated that in the O condition, the (facilitatory) homograph effect was significant over participants [$F_1(1,19) = 7.26, p < .05$] but not over items [$F_2(1,14) = 2.62, p > .10$]. In the OP condition, the RT difference between homographs and controls was not significant [$F_1(1,19) < 1; F_2(1,14) < 1$].

The error analyses revealed that the effect of word status was not significant [$F_1(1,19) = 1.79, p > .15; F_2(1,28) < 1$]. The error rates were 13.8% for the homographs and 11.5% for the controls. However, more errors had been made in the O condition (15.7%) than in the OP condition (9.7%), but this difference was significant in the participants analysis only [$F_1(1,19) = 20.12, p < .001; F_2(1,28) = 1.61, p > .20$]. The interaction between word status and condition (with the homograph effect going in different directions) was significant over participants [$F_1(1,19) = 15.8, p = .001$] and almost significant over items [$F_2(1,28) = 3.75, p < .07$]. Planned comparisons showed that the facilitatory error effect in the O condition was not significant or only marginally so [$F_1(1,19) = 3.57, p < .08; F_2(1,14) < 1$], but the (inhibitory) effect for OP items and their controls was significant [$F_1(1,19) = 14.1, p < .001; F_2(1,14) = 5.73, p < .05$].

P condition. Analyses of the P condition showed that by participants, P items were recognized faster than control words [$F_1(1,19) = 5.61, p < .05$], but this RT difference was not significant by items [$F_2(1,28) < 1$]. There were no differences in terms of error rates [$F_1(1,19) < 1; F_2(1,28) < 1$].

Nonword Data

ANOVAs with nonword type as the only factor (within participants but between items) showed that the participants were slower to reject English-like nonwords (644 msec, $SD = 65$) than neutral nonwords [585 msec, $SD = 56; F_1(1,19) = 77.52, p < .001; F_2(1,88) = 45.22, p < .001$]. Likewise, they made more errors on English-like nonwords (14.7%, $SD = 10.8$) than on neutral nonwords [3.4%, $SD = 5.8; F_1(1,19) = 49.71, p < .001; F_2(1,88) = 29.13, p < .001$].

Table 3
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Words Used in Experiment 1

	Test Words			English Control Words		
Condition	RT	*SD*	%Error	RT	*SD*	%Error
O	552	54	13.3	584	55	18.0
OP	547	41	14.3	548	43	5.0
O and OP	550	48	13.8	566	49	11.5
P	572	54	15.3	586	55	15.3

EXPERIMENT 2
English Lexical Decision Involving Cognates

Method

Participants. Twenty participants (9 men, 11 women) with Dutch as their native language, aged 20–31 (mean 24.4) years, took part in Experiment 2. Three participants were left-handed. They had learned English at school from the age of about 12 years and had, on average, 12.9 years of experience with the English language. Their mean self-rating of reading experience in English (on a scale from 1 to 7) was 5.4 (SD = 1.14). All the participants had normal or corrected-to-normal vision. They received money or course credit for their participation.

Stimulus Materials and Procedure. As in Experiment 1, the cognate items and English control words were identical to the materials in the SOP, SO, and SP conditions in Dijkstra et al. (1999). Again, the materials are presented in the Appendix.

For all cognate stimuli, the two readings were translation equivalents, with SOP and SO pairs being homographic as well. For SOP (e.g., *film*) and SP (e.g., *wheel*) words, the two translation equivalents had almost the same pronunciation. For the SO items, however, the phonological overlap of the two pronunciations was as small as possible. For instance, *fruit* is pronounced /fruːt/ in English, but /frœyt/ in Dutch.

Each of the three cognate conditions comprised 15 cognates. Like the false-friend items, the cognates had been rated for their subjective degree of phonological, orthographic, and semantic overlap by 12 participants in the study by Dijkstra et al. (1999). Item characteristics (e.g., frequency and length) and rating results for the cognates are given in Table 4. (The table also lists the characteristics of Dutch control words, included in Experiment 4 only.) The matching of English controls to cognates, across conditions, was analogous to that of Experiment 1, as was the experimental procedure.

Results

The overall error rate was 9.2%. The percentage of outliers among the correct trials was 2.3%, so that 11.3% of the data were excluded in total.

Word data. The mean RTs, their standard deviations, and the error rates for the word data are listed in Table 5.

SO and SOP conditions. In the RT analysis of the SO and SOP conditions, word status was significant both over participants [$F_1(1,19) = 41.96$, $p < .001$] and over items [$F_2(1,28) = 15.41$, $p < .001$], with faster responses on cognates (546 msec) than on control words (601 msec). Planned comparisons indicated that the cognate effect was significant for both the SO condition [$F_1(1,19) = 34.75$, $p < .001$; $F_2(1,14) = 11.64$, $p < .01$] and the SOP condition [$F_1(1,19) = 14.16$, $p < .001$; $F_2(1,14) = 4.89$, $p < .05$]. The main effect of condition was significant over participants [$F_1(1,19) = 5.95$, $p < .05$] but not over items [$F_2(1,28) = 1.15$, $p > .25$], with shorter RTs in the SO (564 msec) than in the SOP (583 msec) condition. The interaction of the two factors was not significant [$F_1(1,19) < 1$; $F_2(1,28) < 1$].

In the analysis of error rates, significantly fewer errors were made on cognates (3.8%) than on control words [15.0%; $F_1(1,19) = 17.02$, $p < .001$; $F_2(1,28) = 14.66$, $p < .001$]. There was no significant main effect of condition [$F_1(1,19) < 1$; $F_2(1,28) < 1$]. In contrast to the pattern found for RTs, the interaction of condition and word status was significant by participants but not by items [$F_1(1,19) = 10.35$, $p < .01$; $F_2(1,28) = 1.73$, $p > .15$]. Planned comparisons showed that fewer errors were made on SO cognates than on control words [$F_1(1,19) = 30.2$, $p < .001$; $F_2(1,14) = 13.13$, $p < .01$]; however, the difference in the SOP condition was significant in the participants analysis only [$F_1(1,19) = 5.36$, $p < .05$; $F_2(1,14) = 3.19$, $p > .09$].

SP condition. The analysis of the SP condition revealed no significant differences between the recognition latencies for SP cognates and their controls [$F_1(1,19) = 2.95$, $p > .10$; $F_2(1,14) < 1$]. The difference between SP items and their control words with respect to error rates was not significant either [$F_1(1,19) = 2.27$, $p > .10$; $F_2(1,14) = 1.00$, $p > .30$].

Nonword data. The RTs for English-like and neutral nonwords were 667 msec (SD = 58) and 593 msec (SD = 42), respectively. This difference was significant [$F_1(1,19) = 201.29$, $p < .001$; $F_2(1,88) = 48.23$, $p < .001$]. Similarly, more errors were made on English-like nonwords (10.4%, SD = 10.1) than on neutral nonwords [2.85%, SD = 4.1; $F_1(1,19) = 13.75$, $p < .001$; $F_2(1,88) = 22.52$, $p < .001$].

Discussion

Word data. Table 10 provides a summary of the basic effects in Experiments 1 and 2. It also presents the effect sizes for the data by Dijkstra et al. (1999) for the same stimulus materials combined in a single experiment. The data patterns for homographs and cognates in the two

Table 4
Frequency and Length Characteristics of the Word Stimuli (Cognates) in Experiments 2 and 4

	Test Words						English Control Words		Dutch Control Words*	
Condition	EF	DF	Length	Sem.†	Orth.†	Phon.†	EF	Length	DF	Length
SO	43.0	27.9	4.2	6.7	7.0	2.9	43.0	4.2	28.3	4.3
SOP	41.3	32.3	4.0	6.3	7.0	6.1	41.5	4.0	32.9	4.0
SP	42.5	27.2	4.2	6.2	3.3	5.7	42.0	4.2	26.9	4.0

Note—EF, English written frequency (per million); DF, Dutch written frequency (per million); Length, word length in letters; Sem., mean score of semantic similarity rating with the competitor word; Orth., mean score of orthographic similarity rating with the competitor word; Phon., mean score of phonological similarity rating with the competitor word. *Dutch control words were presented in Experiment 4 only. †The ratings were based on a scale from 1 (*no similarity or overlap*) to 7 (*identity or perfect overlap*).

Table 5
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Words Used in Experiment 2

	Test Words			English Control Words		
Condition	RT	*SD*	%Error	RT	*SD*	%Error
SO	533	50	1.7	594	68	16.7
SOP	558	59	6.0	608	71	13.3
SO and SOP	546	55	3.9	601	70	15.0
SP	581	60	15.3	597	57	18.0

studies are very similar, with only one exception (P condition). The similarity in the data arose in spite of differences in stimulus list composition and nonwords in our study relative to the earlier one. The observed pattern of results suggests that Dijkstra et al.'s (1999) interpretation of the data pattern in terms of orthographic and semantic facilitation effects can be maintained. However, with respect to the inhibitory effects of phonological overlap, the present results suggest that the original conclusion may not have been based on completely solid grounds. The earlier study reported a large phonological inhibition effect in the P condition, which completely disappeared in the present study. This confirms Jared & Kroll's (2001) suspicion that the test and control items in that condition may have differed on uncontrolled variables. We will come back to this issue in the General Discussion.

Nonword data. The nonword data pattern in the two experiments was clear and consistent. In both experiments, the participants were slower to reject English-like nonwords than neutral nonwords, and they made more errors on the former type of nonwords as well. This finding suggests that the similarity of nonwords to words in one or two languages plays a role during the rejection procedure. However, a more pertinent conclusion can be drawn after the generalized lexical decision experiments have been conducted, because in those experiments Dutch-like nonwords will be included next to English-like and neutral nonwords.

EXPERIMENT 3
Generalized Lexical Decision Involving False Friends

Method

Participants

Thirty-four students (7 men, 27 women) of the University of Nijmegen with Dutch as their native language participated in the experiment. Three of them were left-handed. The participants were 18–26 (mean, 21) years old. The participants reported having 7–20 years of overall experience with the English language, with a mean of 10.6 years. Asked for the degree of their reading experience in English on a scale from 1 (*very little experience*) to 7 (*very much experience*), their mean response was 5.5. All the participants had normal or corrected-to-normal vision. They were paid for their participation or received course credit.

Stimulus Materials

Words. As in Experiment 1, the P, O, and OP items from Dijkstra et al. (1999) were used. For the generalized lexical decision task conducted here, an additional set of Dutch control words was selected in the same way as the English control words, so that it matched the Dutch readings of the homographs in the cases of the OP and O conditions and the Dutch homophone partner in the case of the P condition. Note that it was not possible to verify the quality of the match by conducting a control experiment with Dutch monolinguals, as was done for the English materials in the Dijkstra et al. (1999) study. This is due to the fact that virtually all Dutch language users have at least some knowledge of English. However, the matching was done on the basis of the CELEX statistics and in exactly the same manner as for the English readings, which turned out to be successful in the earlier control experiment.

Because the P items were exclusively English words, another set of 15 Dutch words were added to the list as filler items in order to maintain a proportion of 50% Dutch and 50% English words in the experiment as a whole. The total set of word stimuli consisted of 150 items, of which 30 were both existing English and existing Dutch words (15 OP and 15 O words), 60 were exclusively English words (15 P items and 45 English controls), and 60 were exclusively Dutch words (15 fillers and 45 Dutch controls). For all stimuli and their characteristics, see Table 1 and the Appendix.

Nonwords. The 150 nonwords (50 English-like, 50 Dutch-like, and 50 neutral) described in the Method section of Experiment 1 were included in this experiment. All nonword materials are listed in the Appendix.

Procedure

Unless stated otherwise, the procedure was the same as that used in Experiments 1 and 2. The written instruction (in English, to keep this factor analogous to that in the previous experiments) informed the participants that they were being asked to decide whether a presented letter string formed a correct word in either English or Dutch, or a nonword in both languages.

The practice block (30 trials) at the beginning of the experiment was similar to the one used in Experiments 1 and 2, but it included Dutch words and Dutch-like nonwords. The 300 experimental items were presented in three blocks of 100 items each, plus 2 dummy items at the beginning of each block that did not enter into the analyses. Four different lists of presentation orders had been created, with no more than four words or nonwords and no more than four English(-like), neutral, or Dutch(-like) items in a row.

The participants filled out the same language questionnaire as did those in Experiments 1 and 2. The experimental session took about 20–30 min.

Results

The data of 4 participants with exceptionally high error rates (above 13%) were excluded from the analyses. For the remaining 30 participants, the overall error rate amounted to 8.8%. For the subsequent analyses of RTs, only correct responses were considered. If a participant made an error on a word (e.g., a homograph), responses to its matched partner items (e.g., the Dutch and English control words) for that participant were also excluded from the RT analysis.[2] By this procedure, another 8% of the data were excluded. In addition, RTs that lay more than two standard deviations away from both the participant mean and the item mean in a condition (1.5% of the remaining RTs) were classified as outliers and (together with their partner items) omitted from analysis. This procedure resulted in an exclusion of in total 17.3% of the data from the RT analyses.

Word Data

The mean RTs, standard deviations, and error percentages of test words and Dutch and English control words are presented in Table 6. A summary of the RT and error effects in the present study and in Dijkstra et al. (1999) is given in Table 10 (see the General Discussion). The data were analyzed using ANOVAs directly comparable to those in Experiment 1, except that word status now had three levels (test word, English control, and Dutch control) rather than two.

OP and O conditions. In the analysis of RTs, there was no main effect of condition [$F_1(1,29) = 3.96, p > .05$; $F_2(1,28) = 1.08, p > .10$]. More importantly, word status exerted a main effect on latencies [$F_1(2,58) = 29.71, p < .001$; $F_2(2,56) = 11.39, p < .001$]. Two planned comparisons showed that homographic test words were recognized significantly faster than English control words [$F_1(1,29) = 64.63, p < .001$; $F_2(1,28) = 21.60, p < .001$] but not faster than Dutch controls [$F_1(1,29) = 2.10, p > .10$; $F_2(1,28) < 1$]. The interaction between condition and word status was not significant [$F_1(2,58) = 1.93, p > .10$; $F_2(2,56) = 1.06, p > .10$].

The error data showed a main effect of condition over participants only [$F_1(1,29) = 12.3, p < .001$; $F_2(1,28) < 1$], with more errors in the O condition than in the OP condition. The main effect of word status was significant in both analyses [$F_1(2,58) = 35.55, p < .001$; $F_2(2,56) = 4.87, p < .01$], with error rates for homographs being lower than for English controls [$F_1(1,29) = 63.51, p < .001$; $F_2(1,28) = 13.0, p < .01$] but not lower than for Dutch controls [$F_1(1,29) = 3.19, p > .05$; $F_2(1,28) < 1$]. This time, in the analysis over participants, the interaction of condition and word status was significant [$F_1(2,58) = 7.04, p < .01$], although it did not approach significance in the items analysis [$F_2(2,56) < 1$]. Further contrasts showed that for both OP and O conditions the difference between homographs and English controls was significant, but that it was larger in the O condition [$F_1(1,29) = 70.14, p < .001$] than in the OP condition [$F_1(1,29) = 9.72, p < .01$].

P condition. The data of the P condition were analyzed separately in a repeated measures ANOVA with word status as the only factor. This factor had a significant effect on recognition latencies [$F_1(2,58) = 18.85, p < .001$; $F_2(2,28) = 3.74, p < .05$]. Planned comparisons indicated that, by participants, RTs on homophonic test words were longer than those on Dutch control words [$F_1(1,29) = 32.26, p < .001$], but the difference was not significant in the items analysis [$F_2(1,14) = 3.07, p > .10$]. In contrast, latencies for the P items did not differ from English control latencies [$F_1(1,29) < 1$; $F_2(1,14) < 1$]. The error rates showed an effect of word status by participants only [$F_1(2,58) = 6.57, p < .01$; $F_2(2,28) < 1$], with error rates for P items being higher than those for both Dutch [$F_1(1,29) = 8.06, p < .01$] and English [$F_1(1,29) = 8.69, p < .01$] controls.

Table 6
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Words Used in Experiment 3

	Test Words			Dutch Control Words			English Control Words		
Condition	RT	*SD*	%Error	RT	*SD*	%Error	RT	*SD*	%Error
O	559	83	4.7	571	61	8.0	633	81	18.7
OP	559	67	5.3	569	70	5.6	604	88	10.9
O and OP	559	75	5.0	570	66	6.8	619	85	14.8
P	622	84	18.0	565	62	11.1	623	91	11.3

Nonword Data

The RT and error data for nonwords are presented in Table 7.

In a one-factorial ANOVA over participants with nonword type as a repeated measures factor in the participants analysis and as a between-groups factor in the items analysis, there was a significant difference between the RTs for the three nonword groups [$F_1(2,58) = 15.27, p < .001$; $F_2(2,147) = 6.07, p < .01$]. Planned comparisons revealed that all pairwise comparisons were significant over participants (Dutch vs. neutral nonwords: $F_1(1,29) = 8.34, p < .01$; Dutch vs. English nonwords: $F_1(1,29) = 9.86, p < .01$; English vs. neutral nonwords: $F_1(1,29) = 22.93, p < .001$], with rejection times for neutral nonwords being the shortest, followed by those for Dutch- and for English-like nonwords. In the items analysis, only the difference between English and neutral nonwords was significant [$F_2(1,98) = 11.32, p < .001$], whereas the other two differences did not reach significance [Dutch vs. neutral nonwords, $F_2(1,98) = 3.74, p > .05$; Dutch vs. English nonwords, $F_2(1,98) = 2.72, p > .10$].

The same pattern held for error rates. There was a main effect of nonword type [$F_1(2,58) = 37.30, p < .001$; $F_2(2,147) = 15.64, p < .001$], and all pairwise comparisons were significant in the participants analysis [Dutch vs. neutral nonwords, $F_1(1,29) = 10.60, p < .01$; Dutch vs. English nonwords, $F_1(1,29) = 28.79, p < .001$; English vs. neutral nonwords, $F_1(1,29) = 54.40, p < .001$]. Analogously to the RT data, the fewest errors were made on neutral nonwords, more errors were made on Dutch nonwords, and the highest error rate was obtained for English nonwords. In the analysis over items, the difference in error rates between Dutch and neutral nonwords did not quite reach significance [$F_2(1,98) = 3.45, p = .07$], but the other two comparisons showed significant differences [Dutch vs. English nonwords, $F_2(1,98) = 11.86, p < .001$; English vs. neutral nonwords, $F_2(1,98) = 25.44, p < .001$].

EXPERIMENT 4
Generalized Lexical Decision Involving Cognates

Method

Participants. Thirty-four students (9 men, 25 women) of the University of Nijmegen with Dutch as their native language participated in the experiment. Five of them were left-handed. The participants were 17–29 (mean 23.2) years old. They had 6–17 years of overall experience with the English language (mean, 11.5 years). Asked for the degree of their reading experience in English on a scale from 1 (*very little experience*) to 7 (*very much experience*), their mean response was 5.4. All the participants had normal or corrected-to-normal vision and were either paid for their participation or received course credit.

Table 7
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Nonwords Used in Experiment 3

Nonword Type	RT	*SD*	%Error
English-like	662	84	12.8
Dutch-like	644	71	5.5
Neutral	630	75	3.3
Total	645	77	7.2

Stimulus Materials and Procedure. In analogy to Experiment 3, the materials consisted of the SP, SO, and SOP items, their English controls, and Dutch controls matched to the Dutch competitors of the items. The same held for the stimulus list composition (15 SO and 15 SOP words with their Dutch and English controls, 15 SP words with their English and Dutch controls, and 15 Dutch filler words). The characteristics of the stimuli are shown in Table 4.

The nonwords and the procedure were also the same as those in Experiment 3.

Results

The data of 4 of the 34 participants were excluded because of error rates higher than 13%. Furthermore, the data for one SP item (*fay*) and its controls (English control, *pox*; Dutch control, *sla*) were removed because of high error rates (above 80%). The frequency matching between conditions was not adversely affected by this (SP items and their English controls did not statistically differ in frequency; neither did the frequencies of the English and Dutch controls for the SP, SO, and SOP conditions differ significantly). For the remaining items and the remaining 30 participants, the overall error rate was 8.1%.

As in Experiment 3, for each participant, matched partner items of incorrectly processed items were also excluded from the RT analysis. By this procedure, an additional 7.5% of the data had to be excluded. The same procedure regarding outliers was used as in Experiment 3, by which 1.9% of the remaining data were excluded. In total, 17.2% of the data did not enter into the RT analyses.

Word data. The mean RTs, standard deviations, and error percentages of test words and Dutch and English control words are presented in Table 8. Table 10 provides a summary of the RT and error effects in this study and the previous study by Dijkstra et al. (1999). The data analysis of Experiment 4 was equivalent to that of Experiment 3: The homographic cognate conditions (SO and SOP) were analyzed separately from the nonhomographic cognate condition (SP).

SO and SOP conditions. In the analysis of RTs of the SO and SOP conditions, there was no main effect of condition [$F_1(1,29) < 1$; $F_2(1,28) < 1$]. More importantly, word status (cognates vs. Dutch controls vs. English controls) had a significant effect on RTs [$F_1(2,58) = 124.10$, $p < .001$; $F_2(2,56) = 29.28$, $p < .001$]. Planned comparisons showed that homographic cognates were recognized faster than both English [$F_1(1,29) = 287.93$, $p < .001$; $F_2(1,28) = 58.57$, $p < .001$] and Dutch [$F_1(1,29) = 30.79$, $p < .001$; $F_2(1,28) = 12.75$, $p < .001$] control words. The interaction between condition and word status was not significant [$F_1(2,58) < 1$; $F_2(2,56) < 1$].

A similar pattern of results was obtained for the error data. The main effect of condition was not significant [$F_1(1,29) < 1$; $F_2(1,28) < 1$]. The main effect of word status, however, was significant [$F_1(2,58) = 117.19$, $p < .001$; $F_2(2,56) = 15.57$, $p < .001$]. Again, planned comparisons showed that fewer errors were made on SO and SOP cognates than on both English controls [$F_1(1,29) = 179.15$, $p < .001$; $F_2(1,28) = 26.72$, $p < .001$] and Dutch controls [$F_1(1,29) = 112.65$, $p < .001$; $F_2(1,28) = 9.51$, $p < .01$]. The condition × word status interaction was not significant [$F_1(2,58) = 2.61$, $p > .05$; $F_2(2,56) < 1$].

SP condition. In the analysis of SP cognates such as *wheel*, word status exerted a significant effect on RTs [$F_1(2,58) = 29.22$, $p < .001$; $F_2(2,26) = 6.73$, $p < .01$]. Planned comparisons indicated that SP items were recognized more slowly than Dutch controls [$F_1(1,29) = 41.11$; $p < .001$; $F_2(1,13) = 9.61$, $p < .01$], but not differently from the English controls [$F_1(1,29) = 2.79$, $p > .10$; $F_2(1,13) < 1$].

In the error rates, there was a trend toward a significant main effect of word status only in the analysis over participants [$F_1(2,58) = 2.94$, $p = .06$; $F_2(2,26) < 1$]. Planned comparisons between SP words and English and Dutch control words were both nonsignificant, indicating that the trend observed for the word status effect was due to a difference between the two classes of controls.

Some SP items also possessed a relatively large orthographic overlap with their Dutch translation equivalents. An additional analysis by items was therefore carried out in which the items were split into two groups, characterized by large or small orthographic overlap with their translations, as measured by subjective ratings in Dijkstra et al. (1999). Items with little orthographic resemblance to their translations (e.g., *fat–vet*) were processed 24 msec more slowly than English controls, a difference that was not statistically significant in paired, two-tailed *t* tests [$t(6) = 1.06$, $p = .33$]. In contrast, items with larger O overlap (e.g., *boat–boot*) were recognized 59 msec faster than English controls, which was significant [$t(6) = 3.16$, $p = .02$]. In the Appendix, it is indicated which items belonged to which group in this analysis. In addition, a significant positive correlation was obtained between the mean ratings for orthographic overlap and the RT difference between the SP items and their English controls ($r = .60$, $p = .02$). The same data pattern held for the English

Table 8
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Words Used in Experiment 4

	Test Words			Dutch Control Words			English Control Words		
Condition	RT	*SD*	%Error	RT	*SD*	%Error	RT	*SD*	%Error
SO	519	61	0.9	553	63	9.8	635	72	23.6
SOP	522	62	3.1	559	76	12.2	637	73	21.1
SO and SOP	521	62	2.0	556	70	11.0	636	73	22.4
SP	606	70	12.4	540	52	8.6	626	78	15.0

lexical decision task used by Dijkstra et al. (1999), whose data we reanalyzed. There was facilitation for large-overlap items [mean effect −54 msec; $t(6) = 3.14, p = .02$] but no effect for small-overlap items [mean effect + 36 msec; $t(6) = .10$].

Nonword data. The RT and error data for nonwords are presented in Table 9.

In the analysis of response latencies on nonwords, there was a significant effect of nonword type [$F_1(2,58) = 66.96, p < .001$; $F_2(2,147) = 29.47, p < .001$]. All pairwise planned comparisons between the three nonword types were significant (Dutch–neutral, $F_1(1,29) = 16.19, p < .001$; $F_2(1,98) = 4.07, p < .05$; English–neutral, $F_1(1,29) = 114.98, p < .001$; $F_2(1,98) = 54.84, p < .001$; English–Dutch, $F_1(1,29) = 46.47, p < .001$; $F_2(1,98) = 25.93, p < .001$]. Neutral nonwords were rejected the fastest, followed by Dutch- and English-like nonwords.

The analysis of error rates revealed the same pattern. There was an effect of nonword type [$F_1(2,58) = 17.04, p < .001$; $F_2(2,147) = 9.96, p < .001$]. Again, all pairwise comparisons were significant except that between Dutch and neutral nonwords in the items analysis [Dutch–neutral, $F_1(1,29) = 10.88, p < .01$; $F_2(1,98) = 2.17, p > .10$; English–neutral, $F_1(1,29) = 26.30, p < .001$; $F_2(1,98) = 22.69, p < .001$; English–Dutch, $F_1(1,29) = 10.57, p < .01$; $F_2(1,98) = 6.66, p < .05$]. The fewest errors were made on neutral nonwords, more errors were made on Dutch-like nonwords, and the most errors were made on English nonwords.

Discussion

Word data. In Experiment 3, interlingual homographs (O and OP conditions) elicited faster and more accurate responses than did English control words, but there were no differences between these items and Dutch controls. The presence or absence of phonological overlap of the homographs (O vs. OP) with their competitors did not yield a difference in RTs. There were more errors on O items than on Dutch control words, but this effect was significant by participants only. In sum, the comparison of O to OP conditions does not provide evidence that phonological overlap affected word recognition in generalized lexical decision.

For English–Dutch homophones (P items), no RT difference was found relative to English controls (because the homophones were presented in English spelling, the comparison to Dutch control words is difficult to interpret). More errors were made on homophones than on English controls, but this effect was not significant over items. This suggests that there was no cross-linguistic phonological effect in this condition. Note that our findings for generalized lexical decision correspond better with our own findings of language-specific decision (Experiment 1) than with those obtained by Dijkstra et al. (1999). Again, this suggests that the result for the P condition in the earlier study was not reliable.

Table 9
Mean Reaction Times (RTs, in Milliseconds, With Standard Deviations) and Error Rates (%Error) for Nonwords Used in Experiment 4

Nonword Type	RT	*SD*	%Error
English-like	673	78	9.3
Dutch-like	618	56	4.6
Neutral	603	58	1.9
Total	631	64	5.3

In Experiment 4, orthographically identical cognates (SO and SOP items) were processed faster and more accurately than both English and Dutch control words. Because Dijkstra et al. (1999) were not able to compare the RTs for their cognates to Dutch controls (given that they used an English lexical decision task), this result generalizes the finding of facilitation effects for cognates to both L1 and L2. There were no clear differences in the size of the cognate effect for SO and SOP cognates in the generalized lexical decision task. This contrasts with Dijkstra et al. (1999), who found reduced degrees of facilitation for cognates with phonological overlap in English lexical decision (SO–SOP effect difference: 18 msec), which was not statistically tested, however. Note that in our English lexical decision experiment (Experiment 2), we observed a nonsignificant difference of 11 msec between the SO and SOP effects.

In contrast to homographic cognates, cognates that share their sound but not their spelling (SP items) were processed 20 msec faster than the matched English controls. However, this effect was statistically nonsignificant. Recall that the additional analysis revealed a correlative relationship for these items between the degree of *orthographic* overlap with the translation and the degree of observed facilitation. Possibly, the effects in the SP items relative to English control words would have been smaller than 20 msec if there had been no orthographic overlap with the competitors at all. In all, this suggests that the SP items should not be considered as cognates in the same sense as the SO and SOP items are.

Nonword data. Although the significance of the differences between nonword types differed a bit between Experiments 3 and 4, the general pattern was clear and consistent with the results of Experiments 1 and 2: Neutral nonwords elicited faster and more accurate responses than did Dutch-like nonwords, which were in turn rejected more quickly and accurately than English-like nonwords. This result pattern indicates that not all nonwords are treated equally. An overall analysis on the nonwords of Experiments 3 and 4 showed that all pairwise comparisons on RTs and error rates were significant ($p < .05$), except for the comparison between error rates on Dutch-like and neutral nonwords in the items analysis.

We will consider a theoretical interpretation of the findings in the General Discussion.

GENERAL DISCUSSION

The present study was performed with three main goals in mind. First, with respect to words, we aimed at clarifying the nature of the representation of cognates and false friends in the bilingual mental lexicon by including them in two versions of the lexical decision task. Second, we tested the BIA+ model's predictions about task demands by varying the task at hand (language-specific vs. generalized lexical decision) while keeping the stimulus materials constant. Third, we examined how nonwords derived from English and Dutch words are rejected in generalized lexical decision, in order to see how language information is used in this task. All three aspects are part of the more general question of how the bilingual mental lexicon is organized.

The Representation of Cognates and False Friends

Let us address the representational issue by examining the overall result patterns for cognates and false friends in language-specific lexical decision and generalized lexical decision. Table 10 shows the basic RT and error effects for such items in the earlier study by Dijkstra et al. (1999) and the present one.

For language-specific English lexical decision, we can compare the result patterns for cognates (SO, SOP, and SP items) and control items in the two studies. The overall patterns were very similar, indicating that changes in design and stimulus list composition did not strongly affect cognate processing. The cognate effects were larger in the present study. According to the BIA+ model, this may be a consequence of an optimization of the decision criteria by the participants in our study, in which cognates and other items were divided across experiments. Our results for false friends (O, OP, and P items) also replicated those of Dijkstra et al. (1999), except in the P condition. However, as we pointed out, in this condition effects arose in monolingual participants, and there may be a problem with respect to the matching of test and control items in this condition. The associated error rates were more equal in test and control conditions in the present study, suggesting that our results may be the more trustworthy ones.

In all, the present results indicate, as suggested by Dijkstra et al. (1999), that cross-linguistic orthographic and semantic overlap in the test items led to facilitation. The data are compatible with the conclusion that cross-linguistic phonological overlap also has an effect: When orthographic and, in the case of cognates, semantic overlap was accompanied by phonological overlap as well, the amount of facilitation was considerably reduced. However, given the relative variability of the result patterns observed in the present study (especially in the P condition), we recommend that future studies test again the role of phonological overlap in individual word recognition using different items and language pairs.

An analysis of the result pattern for cognates and false friends in generalized Dutch–English lexical decision provides a wealth of new data. In the generalized lexical decision task in Experiment 4, responses to homographic cognates (SO and SOP conditions) were faster and more accurate than those to both the English and the Dutch control conditions, whereas null effects were obtained for O and OP interlingual homographs. In other words, shared orthography alone did not lead to facilitation relative to the fastest (Dutch) controls, but adding semantic overlap did. This finding fits well with the proposal by Pexman and Lupker (1999) that both of these types of information are used in the lexical decision task.

There are two possibilities to account for the facilitatory cognate effects. First, orthographically identical cognates may share a *single orthographic representation* across the two languages, as some researchers suggest (De Groot & Nas, 1991; Gollan et al., 1997; Sánchez Casas, Davis, & García Albea, 1992). In this case, the facilitation found for cognates would simply be an effect of cumulative frequency.[3] However, this account does not easily accommodate the fact that facilitatory (though usually weaker) effects have also been found for non-identical cognates such as *tomato–tomaat* (Cristoffanini et al., 1986; Font, 2001; Van Hell & Dijkstra, 2002), for which a completely shared orthographic representation is impossible. Note that our post hoc analysis of the SP items points in the same direction: Items possessing considerable orthographic overlap with their Dutch translations were recognized faster than the English control

Table 10
Reaction Time Effects (RT, in Milliseconds) and Error Effects in the Present Study and in Dijkstra, Grainger, and Van Heuven (1999)

	Dijkstra et al. (1999) Effect (E)		EVLD (Exps. 1 and 2) Effect (E)		GVLD (Exps. 3 and 4) Effect (E)		GVLD (Exps. 3 and 4) Effect (D)	
Condition	RT	%Error	RT	%Error	RT	%Error	RT	%Error
O	−21*	−4.8*	−32*	−4.7	−74**	−14.0**	−12	−3.3*
OP	8	8.7**	−1	9.3**	−45**	−5.6**	−10	−0.3
P	34**	6.3**	−14*	0	−1	6.7**	57**	6.7**
SO	−43**	−15.7**	−61**	−15.0**	−116**	−22.7**	−34**	−8.9**
SOP	−25**	−4.9*	−50**	−7.3*	−115**	−18.0**	−37**	−9.1**
SP	−10	0.4	−16	−2.7	−20	−2.6	66**	3.8

Note—EVLD, English visual lexical decision; GVLD, generalized visual lexical decision; Effect (E), difference between test words and English control words; Effect (D), difference between test words and Dutch control words; %Error, percentage of errors. Planned comparisons in the analysis over participants: *$p < .05$; **$p < .01$. Negative signs indicate facilitation.

words. The size of this facilitatory effect was a linear function of the degree of orthographic overlap, indicating that nonidentical cognates are not processed in a fundamentally different way from identical cognates.

A second possibility that may account for both identical and nonidentical cognate effects involves semantic feedback. If we assume that a cognate is characterized by *two orthographic representations*, the following account of our cognate results can be given: Suppose an orthographically identical cognate (i.e., an SO or SOP item) is presented to the participant. Both orthographic representations are activated, because both perfectly match the orthographic input pattern. Next, both word units activate their (partially) shared semantic representation, which feeds back to the two orthographic representations and thus amplifies their activation (Pexman & Lupker, 1999). One of the two units (presumably the L1 one, in most cases) will thus reach its threshold sooner than would the representation of a noncognate. As a consequence, cognates can be recognized even sooner than the fastest control words—the Dutch ones in our experimental situation.

Now consider the case of orthographically nonidentical cognates, such as words of the *tomato–tomaat* type. When *tomato* is presented, its English orthographic representation becomes activated. However, depending on the degree of orthographic overlap, its cognate partner *tomaat* (Dutch) is also activated to a certain extent. Recognition then proceeds in the same fashion as for identical cognates: Both active units feed the same semantic node, receive feedback from this node, and reach the recognition threshold relatively fast. This leads to facilitation effects for this item type that are smaller than for identical cognates, and of which the magnitude correlates with the degree of orthographic overlap.

Note that this account of the cognate effect is in line with recent studies on monolingual word recognition that also argue in favor of semantic-to-orthographic feedback (e.g., Pecher, 2001; Reimer, Brown, & Lorsbach, 2001).

Yet another possibility would be to claim that lexical decisions take place directly on the semantic level (Plaut, 1999); cognates would then be recognized faster because their semantic representations (receiving input from two rather than one orthographic representation) are activated to a larger degree. However, a "read-out" at the level of semantic representations cannot explain why O items (having two separate conceptual representations) were responded to faster than controls in the English lexical decision task.

For interlingual homographs (O and OP conditions), the generalized lexical decision task (Experiment 3) yielded faster responses than for English but not Dutch controls. The error rates followed the same pattern. These findings are in agreement with the generalized lexical decision study by Dijkstra et al. (1998), in which the same effect pattern was found when we averaged across the relative L1/L2 frequency categories for homographs. Note that the homographs in our experiments had frequencies lying between the high- and low-frequency categories in that study, which makes a comparison of our results with the average result pattern in Dijkstra et al. (1998) interesting for generalization purposes. Furthermore, faster RTs for (O-type) interlingual homographs than for English (L2) controls were also found in the English lexical decision experiment by Dijkstra et al. (1999) and in our Experiment 1, but in these experiments no Dutch controls were available for a comparison to L1. The finding that interlingual homographs were not responded to with greater speed or accuracy than matched Dutch control words stands in contrast to the assumption of shared orthographic representations of interlingual homographs, suggested as one possibility by Dijkstra et al. (1999). If interlingual homographs were indeed characterized by a single orthographic representation, they should be recognized more quickly than both English and Dutch controls, because their frequency, being the sum of the individual readings in the two languages, would be considerably higher than that of both sets of controls. As is shown by the results for the frequency-matched cognates in Experiment 4 and the homograph and cognate results in Dijkstra et al. (1999), the differences in the mean English frequencies of the items were large enough to result in cross-linguistic effects under particular experimental circumstances. In other words, had those effects been due to cumulative frequency effects on a common L1/L2 representation, we should have replicated those effects under the present experimental circumstances. Because no such effects arose, we must conclude that (as is assumed by the BIA+ model) interlingual homographs are represented by *two different orthographic representations* across languages, each of which is connected to its own semantic representation.

The Role of Task Demands in the Processing of Cognates and Interlingual Homographs

The result pattern discussed here not only indicates that there are two (at least partially) separate representations for interlingual homographs, but also confirms the prediction of the BIA+ model that responses to these homographs are based on the *fastest available of the two codes* (L1 or L2) that is appropriate for the task at hand. In the generalized lexical decision task, in which, for a homograph, both the English and the Dutch codes are appropriate, the Dutch (L1) orthographic code is probably the fastest code to become available and, therefore, the one on which the response is based. In the language-specific English (L2) lexical decision task, however, the response cannot be made before the English code has been accessed and its language membership has been verified. One possible explanation of the data pattern is to assume that because the Dutch (L1) codes are available earlier than the English (L2) codes, there is time for them to affect the response to the English readings of the interlingual homographs in an English task. In contrast, the absence of an orthographic facilitation effect in *generalized* lexical decision indicates that the faster Dutch orthographic codes formed the basis of response and that they were generally available before their English counterparts were active enough to affect responding.

We would like to propose a similar rationale for the activation and use of *phonological* codes in the two experimental situations. In Experiment 3, the homograph effects in the O and OP conditions were similar in size. Likewise, the facilitation effects in the SO and SOP conditions in Experiment 4 did not differ. Thus, phonological overlap, in addition to orthographic and (in Experiment 4) semantic overlap, did not alter the response patterns, whereas it did in the English tasks in Experiments 1 and 2. These findings do not imply that L1 phonological representations were not activated in the generalized lexical decision task. Given the large body of evidence in favor of phonological influences in reading (e.g., Ferrand & Grainger, 1994; Jacobs, Rey, Ziegler, & Grainger, 1998), it is unlikely that phonological information did not become active in those experiments. Instead, they fit the time course account proposed by the BIA+ model. This model would argue that the phonological lexical representation in L2 became available too late to affect the response based on already activated L1 codes.

Neither for the homophones in the P condition nor for homophonic SP cognates did we obtain an inhibitory effect of cross-linguistic phonological overlap (independent of orthographic overlap). When English words (e.g., *cow* or *wheel*) sounded like Dutch words, RTs did not differ from those for exclusively English control words. The observed result pattern suggests that participants responded primarily on the basis of the *orthography* of the English items, not on the basis of their phonology.

The lack of a cognate effect in the SP condition stands in contrast with the results of Gollan et al. (1997) and Kim and Davis (2003), who found enhanced masked priming effects for (cross-script) cognates that were similar in phonology. However, the comparability of these studies with the present one is limited not only because of potential differences in the experimental paradigms, but also because Hebrew, Korean, and English make use of different scripts, which might induce differences in word processing (e.g., more reliance on phonology than on orthography). Furthermore, note that script could serve as a bottom-up cue to the language identity of each presented stimulus, which potentially could affect lexical selection.

In sum, in line with the BIA+ model, we suggest that the null effects for phonological overlap in our generalized lexical decision experiments can be accounted for in terms of a delayed time course of the phonological relative to the orthographic code.

Language Membership Information and Nonword Rejection

The generalized lexical decision Experiments 3 and 4 of this study included the same set of nonwords, inducing the following RT and error pattern. In both experiments, English-like nonwords (e.g., *baint*) were rejected more slowly and with a higher proportion of errors than Dutch-like nonwords (e.g., *muig*). Responses to neutral nonwords (e.g., *brank*) were even faster and tended to be more accurate than responses to Dutch-like nonwords (the error effect failed to reach significance when calculated over items).

The finding that English-like nonwords were rejected more slowly and with more errors than Dutch-like nonwords contradicts a completely language-nonselective nonword rejection procedure for lexical decision. One broadly accepted theory of nonword rejection assumes that if at a critical point in time the search for a matching word candidate in the lexicon has remained unsuccessful, a "no" response is given. The deadline is set later if the stimulus is more word-like (Grainger & Jacobs, 1996). However, an assumption of such a language-independent rejection procedure for all types of nonwords cannot explain our results. Because L1-like nonwords should generally activate stronger word candidates than L2-like nonwords, they should result in a later deadline for "no" responses than L2-like nonwords. Therefore, in the present study English-like nonwords should have been rejected more quickly and more accurately than Dutch-like nonwords.

However, the opposite effect was observed. The data suggest that the bilingual word recognition system makes a distinction between "Dutch" and "English" stimuli, and subsequently applies different rejection criteria to Dutch-like and English-like nonwords. This requires that English- and Dutch-like stimuli (nonwords as well as words) can be distinguished *before* their actual recognition or rejection, a suggestion that is also put forward by Lemhöfer and Radach (2003). Such an early language discrimination mechanism could be based either on sublexical information such as the language-specific probability of the orthographic pattern (orthotactics) or on lexical information such as the number and frequency of word candidates that overlap with the presented letter string (orthographic neighbors). Examples of models that make use of lexical information are the BIA and BIA+ models, in which language membership is coded through *language nodes* that are activated by words of the respective language.

After the potential language membership of a stimulus has been identified, different time-outs seem to be applied to English- and Dutch-like items. More specifically, the deadline for English items must be set later than that for Dutch ones, resulting in longer rejection times for English-like nonwords than for Dutch-like ones. The notion of two different time-outs for Dutch and English items is in accordance with the assumption, explained above, of the differential *time courses* of L1 and L2 codes. Because L2 word nodes generally become active more slowly, the best way to distinguish L2-like nonwords from these "slow" word candidates is to wait for a longer period of time until the word/nonword distinction process is completed. Thus, not only word recognition in

L2, but nonword rejection as well, are slowed down relative to L1.

To summarize the basic findings of this study, our results suggest that the time courses of word acceptance and nonword rejection are different for L1 and L2. Upon presentation of a stimulus, word candidates from both languages are activated, with a slower activation of L2 than of L1 lexical codes. As a consequence, cross-linguistic facilitation and inhibition effects are observed either when there is enough time for lexical effects to occur (e.g., when L2 is the target language) or when semantic feedback can amplify the activation pattern arising at the orthographic level (e.g., when the item is a cognate). In accordance with the BIA+ model, the absence of cross-linguistic effects for interlingual homographs in generalized lexical decision indicates that responses are based on the fastest available code (here, Dutch orthography) and that such items are characterized by two representations rather than one. Finally, for nonwords, we have found that the rejection response is based on language-dependent time criteria. These findings have important consequences for our understanding of how participants make bilingual lexical decisions and, more generally, how they recognize words from different languages.

REFERENCES

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. University of Pennsylvania Linguistic Data Consortium.

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

Brysbaert, M., Van Dyck, G., & Van De Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 137-148.

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

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

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

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 Groot, A. M. B., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory & Language*, **30**, 90-123.

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

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

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language & Cognition*, **1**, 51-66.

Doctor, E. A., & Klein, D. (1992). Phonological processing in bilingual word recognition. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 237-252). Amsterdam: North-Holland.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology*, **47A**, 365-382.

Font, N. (2001). *Rôle de la langue dans l'accès au lexique chez les bilingues: Influence de la proximité orthographique et sémantique interlangue sur la reconnaissance visuelle de mots*. Unpublished doctoral dissertation, Université Paul Valery, Montpellier.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 305-315.

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, & Cognition*, **23**, 1122-1139.

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

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language & Cognition*, **1**, 67-81.

Jacobs, A. M., Rey, A., Ziegler, J. C., & Grainger, J. (1998). MROM-*p*: An interactive activation, multiple readout model of orthographic and phonological processes in visual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 147-188). Mahwah, NJ: Erlbaum.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory & Language*, **44**, 2-31.

Kawamoto, A. H., & Zemblidge, J. H. (1992). Pronunciation of homographs. *Journal of Memory & Language*, **31**, 349-374.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory & Language*, **49**, 484-499.

Lemhöfer, K., & Radach, R. (2003). *The role of language context in bilingual word recognition*. Manuscript submitted for publication.

Pecher, D. (2001). Perception is a two-way junction: Feedback semantics in word recognition. *Psychonomic Bulletin & Review*, **8**, 545-551.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology*, **53**, 323-334.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 139-156.

Plaut, D. C. (1999). Computational modeling of word reading, acquired dyslexia, and remediation. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 339-372). Cambridge, MA: MIT Press.

Reimer, J. F., Brown, J. S., & Lorsbach, T. C. (2001). Orthographically mediated inhibition effects: Evidence of activational feedback during visual word recognition. *Psychonomic Bulletin & Review*, **8**, 102-110.

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

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

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, **23**, 383-404.

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 & Language*, **39**, 458-483.

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 & 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 & Cognition*, **5**, 1-23.

NOTES

1. Because words of both languages are target items, the generalized lexical decision task includes relatively many nonword stimuli to keep the word/nonword ratio at 50%. As a consequence, the number of stimuli in a session can be kept within bounds only by dividing the cognate and homograph materials over two experiments. In addition to the motivation to study effects of stimulus list composition, this forms a practical consideration for splitting up the experiment in two parts.

2. Given the item-by-item matching and the relatively large difference between the matched item sets, this set-wise exclusion seemed the appropriate procedure. However, in Experiments 1 and 2, we followed the original analysis method as it was used in Dijkstra et al. (1999), which excluded individual erroneous responses individually. However, additional analyses showed that both procedures led to the same pattern of results.

3. Note that this issue is distinct from the question of whether cognates share the same *conceptual* representation (De Groot & Nas, 1991), which is not explicitly addressed by the present study.

APPENDIX
Test Words, Control Words, and Fillers

For each of the 15 test items in a condition, the following information is given: spelling, English phonology, Dutch phonology, spelling of English control word, and spelling of Dutch control word.

O Condition

stage, /steɪdʒ/, /staʒɜ/, mouth, hulde; glad, /glæd/, /xlɑt/, coat, stom; roof, /ruːf/, /roːf/, sale, pels; boon, /buːn/, /boːn/, hero, zeil; steel, /stiːl/, /steːl/, rough, flits; boot, /buːt/, /boːt/, acre, wang; lover, /lʌvə/, /loːvər/, entry, dadel; fee, /fiː/, /feː/, mud, hes; tube, /tjuːb/, /tyːbə/, lion, lama; angel, /eɪnʒəl/, /ɑŋəl/, elbow, anjer; lap, /læp/, /lɑp/, jar, hak; brave, /breɪv/, /braːvə/, crude, wreed; rug, /rʌg/, /rʉx/, shy, dier; brand, /brænd/, /brɑnt/, gown, fiets; sage, /seɪdʒ/, /saːɣə/, flea, pion

OP Condition

step, /stɛp/, /stɛp/, skin, mout; star, /stɑː/, /stɑr/, king, nors; box, /bɒks/, /bɔks/, gun, dop; spot, /spɒt/, /spɔt/, wing, pand; pink, /pɪŋk/, /pɪŋk/, song, spek; brief, /briːf/, /briːf/, funny, beeld; arts, /ɑːts/, /ɑrts/, twin, plek; bond, /bɒnd/, /bɔnd/, lawn, woud; pet, /pɛt/, /pɛt/, pie, bak; pit, /pɪt/, /pɪt/, fox, mep; stout, /staʊt/, /stɑut/, eagle, stoom; dot, /dɒt/, /dɔt/, cue, mot; rover, /rəʊvə/, /roːvər/, peach, merel; brink, /brɪŋk/, /brɪŋk/, crook, klont; kin, /kɪn/, /kɪn/, ale, kip

SO Condition

type, /taɪp/, /tiːpə/, nice, ziel; wild, /waɪld/, /ʋɪlt/, desk, bron; model, /mɒdl/, /modɛl/, skill, tante; fruit, /fruːt/, /frœyt/, youth, stier; pure, /pjʊə/, /pyːrə/, soil, roze; jury, /dʒʊərɪ/, /ʒyːriː/, wale[A1], vete; code, /kəʊd/, /koːdə/, tale, adel; mild, /maɪld/, /mɪlt/, chin, fors; humor, /hjuːmə/, /hyːmɔr/, fever, kapel; rat, /ræt/, /rɒt/, jaw, gids; oven, /ʌvn/, /oːvə/, chap, ezel; chaos, /keɪɒs/, /xaːɔs/, spine, gevel; ego, /ɛgəʊ/, /eːɣoː/, pea, opa; globe, /gləʊb/, /xloːbə/, torch, sonde; menu, /mɛnjuː/, /meːnyː/, bike, lade

SOP Condition

hotel, /həʊtɛl/, /hoːtɛl/, event, vogel; film, /fɪlm/, /fɪlm/, bird, fles; lip, /lɪp/, /lɪp/, sky, zak; tent, /tɛnt/, /tɛnt/, luck, slok; sport, /spɔːt/, /spɔrt/, guilt, vuist; trend, /trɛnd/, /trɛnt/, pride, stang; storm, /stɔːm/, /stɔrm/, thigh, gunst; fort, /fɔːt/, /fɔrt/, silk, brok; pen, /pɛn/, /pɛn/, fur, som; sofa, /səʊfə/, /soːfaː/, wage, zede; net, /nɛt/, /nɛt/, lad, mus; mist, /mɪst/, /mɪst/, bold, plak; rib, /rɪb/, /rɪp/, cab, bes; torso, /tɔːsəʊ/, /tɔrzoː/, trash, slome; ark, /ɒːk/, /ɒrk/, flu, ets

For each of the 15 test items in the P and SP conditions, the following information is given: English spelling, English phonology, Dutch spelling, Dutch phonology, spelling of English control word, and spelling of Dutch control word. (Items marked by an asterisk were in the high-orthographic overlap group in the additional analysis on SP items.)

P Condition

note, /nəʊt/, noot, /noːt/, army, doek; leaf, /liːf/, lief, /liːf/, fair, dorp; lack, /læk/, lek, /lɛk/, duty, mos; aid, /eɪd/, eed, /eːt/, odd, eik; lake, /leɪk/, leek, /leːk/, holy, grap; lane, /leɪn/, leen, /leːn/, wire, deun; cow, /kaʊ/, kou, /kɑu/, gap, tal; pace, /peɪs/, pees, /peːs/, fate, lont; mail, /meɪl/, meel, /meːl/, pity, goot; core, /koː/, koor, /koːr/, cage, lood; ray, /reɪ/, ree, /reː/, bee, keu; scent, /sɛɑnt/, cent, /sɛnt/, mercy, hemd; dose, /dəʊs/, doos, /doːs/, fame, poot; stale, /steɪl/, steel, /steːl/, alley, preek; oar, /oː/, oor, /oːr/, oat, oom

SP Condition

news, /njuːz/, nieuws, /niːws/, lady, struik; fat, /fæt/, vet, /vɛt/, tea, tas; boat*, /bəʊt/, boot, /boːt/, tall, wens; cool, /kuːl/, koel, /kuːl/, iron, geel; tone*, /təʊn/, toon, /toːn /, suit, tand; wheel, /wiːl/, wiel, /ʋiːl/, chain, mand; clock*, /klɒk/, klok, /klɔk/, giant, bril; cliff*, /klɪf/, klif, /klɪf/, straw, trog; ankle, /æŋkl/, enkel, /ɛnkəl/, unity, dader; soup*, /suːp/, soep, /suːp/, riot, heup; sock*, /sɒk/, sok, /sɔk/, dusk, rit; rack, /ræk/, rek, /rɛk/, brow, gil; cord*, /kɔːd/, koord, /kŏrt/, scar, dwerg; nymph, /nɪmf/, nimf, /nɪmf/, batch, krib; fay, /feɪ/, fee, /feː/, pox, sla

APPENDIX (Continued)

Dutch Filler Words in Experiment 3
berm, dolk, drom, duim, gok, knop, koud, leer, mees, prent, puk, rel, riem, wol, zee

Dutch Filler Words in Experiment 4
borg, dal, hals, jurk, kreet, lepel, lat, luid, nis, pook, rand, rem, stal, strot, zeep

Dutch-like Nonwords in Experiments 3– 4
aleur, bedig, bezel, bloef, brek, dem, doef, doest, eit, geuk, hoel, iem, jocht, kem, klart, klet, kluik, kokel, krein, krus, kuim, meeg, muig, nal, nief, noet, ommer, oos, peeuw, rieg, rodde, roen, sauw, soel, stoeg, troeg, tuik, varig, veef, vik, vliep, voen, voest, wezig, woer, zagel, zel, zeuk, zoeg, zuik

English-like Nonwords in Experiments 1–4
arown, baint, beal, beath, bire, borth, cairy, cleak, corch, coss, darry, dase, dass, deak, deaty, dow, fandy, faw, feak, foat, foint, fure, gine, gless, gue, heam, hise, jeal, lim, loak, loy, ory, pite, pive, potch, poth, pough, pum, rall, rea, rouch, sharn, shill, shont, slace, tead, tir, twipe, vix, yound

Neutral Nonwords in Experiments 1–4
balf, blan, blord, blug, brank, brift, brip, dramp, drist, drus, ela, elim, elind, erap, erm, flust, froom, gleet, gret, gron, ide, lasis, lin, lunt, malf, malm, marst, namp, opin, ori, palon, plap, plert, polt, prug, ragel, rin, sarm, seper, slin, solon, spus, talt, ters, trene, trum, ura, urg, veise, vesse

NOTE

A1. According to the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993), the word *wale* is as frequent as the word *jury* (34 occurrences per million). This frequency count is probably wrong, and might be ascribed to confusion with the word *whale*. However, excluding the item from the analysis does not substantially change the group means of English SO controls (RT, 631 msec instead of 635 msec; error rate, 22.6% instead of 23.6%), nor does it affect the significance of statistical effects.

(Manuscript received June 27, 2002;
revision accepted for publication December 9, 2003.)