Bilingualism, Language and Cognition 1 (1998) 51–66 © 1998 Cambridge University Press

Interlingual homograph recognition: Effects of task demands and language intermixing*

TON DIJKSTRA
NICI, University of Nijmegen
HENK VAN JAARSVELD
ATD, University of Nijmegen
SJOERD TEN BRINKE
NICI, University of Nijmegen

A series of three lexical decision experiments showed that interlingual homographs may be recognized faster than, slower than, or as fast as monolingual control words depending on task requirements and language intermixing. In Experiment 1, Dutch bilingual participants performed an English lexical decision task including English/Dutch homographs, cognates, and purely English control words. Reaction times to interlingual homographs were unaffected by the frequency of the Dutch reading and did not differ from monolingual controls. In contrast, cognates were recognized faster than controls. In Experiment 2, Dutch participants again performed an English lexical decision task on homographs, but, apart from nonwords, Dutch words were included which required a "no" reaction. Strong inhibition effects were obtained which depended on the relative frequency difference of the two readings of the homograph. These turned into frequency-dependent facilitation effects in Experiment 3, where participants performed a general lexical decision task, responding "yes" if a word of either language was presented. It is argued that bilingual word recognition models can only account for the series of experiments if they explain how lexical processing is affected by task demands and stimulus list composition.

How bilinguals recognize words of different languages has long been a major issue in research on bilingualism (Keatley, 1992; Grainger 1993; Dijkstra & Van Heuven, in press). The language nonselective access hypothesis holds that words from both languages of the bilingual are activated when a presented word is recognized, while the language selective access hypothesis assumes that only words of the language targeted in the communicative setting are considered.

In the past decades, a large number of studies have collected evidence that word recognition in one language may be affected by knowledge of words from the other language (Caramazza & Brones, 1979; Altenberg & Cairns, 1983; Nas, 1983; Cristoffanini, Kirsner, & Milech, 1986; Beauvillain & Grainger, 1987; De Groot & Nas, 1991; Grainger & O'Regan, 1992; Van Heuven, Dijkstra, & Grainger, 1998). However, some studies have obtained results that support the language selective access hypothesis (Kolers, 1966; MacNamara & Kushnir, 1971; Scarborough, Gerard, & Cortese, 1984; Gerard & Scarborough, 1989). The prevalence of studies supporting language nonselective access suggests that a selective access hypothesis is no longer tenable. This raises the question of how the seemingly conflicting results can be reconciled. One option is to assume that bilingual word recognition basically involves a language nonselective processing system which may function more or less selectively depending on experimental circumstances.

We can discern at least four factors (adapted from Grosjean, 1997a, 1997b) that may systematically affect the degree of language selectivity observed in experiments. The first factor concerns the relative proficiency of bilingual participants in their second language. If there is a language nonselective processing system, bilinguals should show stronger effects of their second language on word recognition in the mother tongue when their second-language proficiency increases, and increasingly smaller effects in the opposite direction. The underlying reasoning here is that the (competitive) strength of lexical representations of words in the second language, like those in the mother tongue, depends on the frequency with which they are encountered, and thus on the participants' second-language proficiency.

A second factor is the cross-linguistic similarity of the stimulus material. Weaker effects of the non-

* The authors are indebted to François Grosjean, Annette De Groot, Elaine Fox, Judy Kroll, James McQueen, Marcus Taft, and one anonymous reviewer for their extensive comments on earlier versions of this article. They further thank Jonathan Grainger, Walter Van Heuven, and Mark Timmermans for valuable discussion in various stages of this research and Nils Jostmann for his help in running Experiment 3.

Address for correspondence
Ton Dijkstra, NICI, University of Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
E-mail: dijkstra@nici.kun.nl

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

target language on target word recognition may be expected for items that have language-specific cues or are structurally illegal in the other language (cf. Grainger & Beauvillain, 1987). Effects will become stronger, however, when target items become more similar to items in the other language with respect to orthography, phonology, or semantics. To study the effects of this factor many bilingual word recognition studies have incorporated items that are related in spelling and/or meaning across different languages: interlingual homographs and cognates (see, e.g., Beauvillain & Grainger, 1987). Interlingual homographs are defined here as words that have the same written form but different meaning across languages. An example is the English word *room*, which is form-identical to a Dutch word meaning "cream." Cognates are items that share not only their orthographic form but also (most of) their meaning(s) across languages. Examples of English/Dutch cognates are *film* or *lip*. Since interlingual homographs and cognates share at least their orthographic form across languages, the recognition process of these types of items should be more sensitive to effects exerted by the non-target language on target processing than items belonging exclusively to the target language.[1]

A third factor that may affect the degree of language selectivity relates to the experimental task that the bilingual participants perform. For example, Durgunoglu and Roediger (1987) have argued that task variations may induce language nonselective or selective patterns of results. They proposed a distinction between data-driven tasks, which involve a focus on processing the physical properties of the stimuli, and concept-driven tasks, that require processing of the semantic properties of stimuli as well. Durgunoglu and Roediger observed language selective effects for a data-driven recall task such as word fragment completion, but nonselective effects for a concept-driven task such as free recall. Thus, the performance of different bilingual tasks may require different processing and decision operations, which may result in more or less language selective patterns of results.

As a fourth factor we may consider whether experiments incorporate items from one language only or from both languages, which may affect the relative activity of the bilinguals' languages. Grosjean (1997a, 1997b) considers this aspect (which we will call "language intermixing") in his treatment of the "language mode" that is operative and sustained under particular circumstances. For everyday situations, the term "language mode" relates to the relative availability of the bilinguals' languages depending on how often and for how long they find themselves in monolingual or bilingual settings. In experimental situations, bilingual participants may be situated somewhere along the monolingual-bilingual mode continuum depending on whether they are presented with stimuli from two languages or only one.

The experiments to be reported in this paper relate especially to the last two factors just discussed. We examined the degree of language selectivity in the latencies of bilinguals to interlingual homographs in three variants of the lexical decision task when items from another language are present or absent in the stimulus list (language intermixing).

As a first step, we conducted an experiment that had many points in common with Gerard and Scarborough's (1989) often cited study that seemed to support the selective access hypothesis (we will only consider their participants' performance on English target items in part 1 of their study). Following up on an earlier study (Scarborough, Gerard, & Cortese, 1984), the authors had English monolinguals and Spanish/English bilinguals perform a lexical decision task. Stimuli included words of three different types: cognates, homographic non-cognates, and non-homographic control items (called "noncognates" and "noncognate controls"). Cognates and controls were either high-frequency or low-frequency in both English and Spanish. Homographic non-cognates were high-frequency in English and low-frequency in Spanish or vice versa.

Three aspects of the results provided evidence for language selective access according to the authors. First, for the same stimulus material, no significant latency differences were found between Spanish/English bilinguals and English monolinguals, suggesting that they were all effectively operating in a language selective manner. Second, though a significant main effect of word type was found, this was mainly due to slow responses to homographic non-cognates that were of low frequency in the target language. No significant reaction time (RT) differences were found between high-frequency members of the three groups of words. Third, word latencies varied primarily with the frequency of usage in the target language, while the frequency of the word forms in the non-target language did not affect the latencies.

[1] It is currently a matter of debate whether interlingual homographs and cognates are characterized by shared or different orthographic representations across languages (note that a language nonselective access hypothesis does not logically imply only one representation for such items). We will use expressions such as "frequency of the Dutch reading of a homograph/cognate" and "English lexical representation" in an attempt to avoid this difficult issue in this paper (see Dijkstra, Timmermans, & Schriefers, in preparation).

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

In our version of the Gerard and Scarborough study we hoped to find evidence for language selective access, either with cognates or with homographs or with both, so as to have a base to manipulate the two factors under study (task and language intermixing). We did not expect to find exactly the same pattern of results as in their study, as too many variables differ between the two studies. Among these we note languages (linguistic relationship between the bilinguals' languages), stimuli (exact nature of cognates and homographs tested, frequency and syntactic category of the words used in each language, order of presentation, etc.), experimental procedure (language set and language mode), and, of course, the type of bilinguals used. Concerning this last variable, most of Gerard and Scarborough's participants listed Spanish as their native language but at the same time they rated their English vocabulary as being equal to or greater than their Spanish vocabulary (p. 308). We may conclude from this that they were either balanced bilinguals or that they had perhaps started to become functionally dominant in English, their second language. The kind of participants we regularly use in our studies are Dutch/English bilinguals who are clearly dominant in their first language, Dutch, but who have very good proficiency in their second language, English. In addition, our participants appear to come from a much more homogenous group.

As in Gerard and Scarborough's study, Experiment 1 involved an English lexical decision task on interlingual homographs, cognates, control words, and nonwords derived from (non-test) English items. We also used word frequency variation in the target and non-target languages as a diagnostic to make sure that we were tapping into lexical effects, and to identify their language origin. Based solely on the results obtained by Gerard and Scarborough we might expect the following results for our Dutch/English bilinguals: (1) no RT-differences between homographs, cognates, and control words; (2) no effect of the frequency of the non-target language reading of the homographs and cognates on target response time. However, given the many differences we have listed above, we were ready to obtain different results. What was critical was to find some evidence for language selectivity so as to be able to manipulate the two factors of interest, task and language intermixing, in the following experiments.

In Experiments 2 and 3 we investigated in more detail which aspects of an experiment affect bilingual word processing. In Experiment 2 we added Dutch monolingual filler words to the test stimuli and instructed participants to respond "no" to these items, but "yes" to English words. In Experiment 3 we used the same types of test stimuli as in Experiment 2 but we had participants respond "yes" to a word in either language.

Finally, in the "General discussion" we relate the results of all three experiments to the factors mentioned above and integrate them into a theoretical framework that accounts for both processing aspects and task-specific requirements.

Experiment 1

Method

Participants. Forty-one participants participated in the experiment. They had been studying English at Nijmegen university full-time for at least three years. Most participants had ten or more years of experience with the English language, both in terms of written and oral communication. Many students had spent a year or more in England. All participants were native speakers of Dutch and were paid for their participation.

Materials. By means of the English and Dutch CELEX data-base (Baayen, Piepenbrock, & van Rijn, 1993), words were selected that had the same orthographic form in the two languages. All words were three to six letters long and had a lemma frequency of at least one per million. Noun or adjective was always their most frequent syntactic category in English. When syntactically ambiguous, frequency measures were based on summed frequencies across all syntactic categories involved. Words that satisfied these selection criteria were rated for similarity of meaning (high or low) in the two languages by three independent judges. Word forms with high similarity were classified as cognates and word forms with little or no semantic similarity as homographs. A sample of the stimulus material is presented in the appendix.

For homographs, four different frequency categories were formed by an orthogonal combination of word form frequencies (high or low) in English and Dutch. Each of the four categories (HFE-HFD, HFE-LFD, LFE-HFD, LFE-LFD) consisted of fourteen items. For the cognates this orthogonal combination of language frequencies proved not to be feasible. A sufficient number of items could be found only for two frequency categories: high-frequency in both English and Dutch (HFE-HFD) and low-frequency in both English and Dutch (LFE-LFD). Presumably, items in the other two categories (HFE-LFD and LFE-HFD) are relatively rare due to the similarity of the larger cultural contexts in which English and Dutch are used. Each of the two filled categories contained twenty-eight items.

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Table 1. *Mean word length in letters (WL) and log word frequency in English (ELF) and Dutch (DLF) for cognates, homographs, and matched English control items in Experiment 1. Frequency categories are combinations of high and low frequent words in English (HFE;LFE) and in Dutch (HFD;LFD)*

	Word type				
	Cognates			English controls	
Frequency categories	WL	ELF	DLF	WL	ELF
HFE-HFD	3.9	1.94	1.79	3.9	1.93
LFE-LFD	3.9	.83	.86	3.9	.83

	Word type				
	Homographs			English controls	
Frequency categories	WL	ELF	DLF	WL	ELF
HFE-HFD	3.9	1.92	1.78	3.9	1.92
HFE-LFD	3.9	2.05	.64	3.9	2.04
LFE-HFD	3.9	1.10	1.87	3.9	1.10
LFE-LFD	3.8	.91	.75	3.8	.91

For each test item, an English control item was selected that matched the test item in number of letters, syntactic category (noun or adjective), and frequency (as can be seen in the appendix). Frequency data for the different categories of test and control items are presented in Table 1. Nonwords that were orthographically legal in English were constructed by changing one letter in existing English words that had not been selected as test or control items. These nonwords had the same item length distribution as the word stimuli. Care was taken that the nonwords and the control items were not homophonic to Dutch words. The total stimulus set consisted of 448 items, of which 224 items were words (56 cognates, 56 homographs, and 112 control items) and 224 nonwords.

Procedure. Participants were tested individually in a sound-proof room. The experiment was run on an Apple Macintosh LCII with a 70-Hz Triniton monitor. The monitor was placed at a distance of approximately 65 cm from the participants. Stimuli appeared in lowercase Courier (18 points) at the centre of the screen on a white background. Each trial started with the presentation of two small horizontal lines, 6.6 mm (15 pixels) above and below the center of the screen for 800 ms. Then the screen was cleared, and 300 ms later a letter string appeared. This string disappeared after the participant responded or after a time-out of 1500 ms. Inter-trial intervals were 700 ms. Participants received the standard instruction for an English lexical decision task. For left-handed participants, the pairing of "yes" and "no" responses to buttons was reversed. Instructions were written in English and oral communication between participant and experimenter was also conducted in English.

For presentation purposes, the stimulus set was divided into four blocks of 112 items each. These four blocks were preceded by a block of 30 practice trials. Short breaks were introduced between the blocks. At the beginning of each block, the word "ATTENTION" was presented at the center of the screen for 1000 ms. Order of presentation was randomized for each participant. After the experiment, participants filled in a questionnaire to assess their familiarity with the English language. Experimental sessions lasted about 35 minutes.

Results

Data cleaning procedures were mainly based on error rates for items and participants. First, eleven items (four homograph test and seven control items) with error rates larger than 35 percent were discarded. Leaving out these items did not affect the matching between test and control items. Subsequently, data of four participants which had an overall error rate larger than 10 percent or a larger than 10 percent difference between error rates for words and nonwords were also left out, leaving data of thirty-seven participants for the analysis. Response times that were outside the range of two standard deviations from the participant and item mean were considered outliers (2.3 percent of all valid data). The wrong responses and outliers were replaced by the participant mean for each condition of the remaining data. Of the total data set, 4.8 percent were errors (wrong responses and missing data).

Since the pattern of error rates followed the pattern obtained for the latencies, only analyses of the latencies for the correct responses will be reported. Words were reacted to 65 ms faster than nonwords (581 ms and 646 ms respectively) while having comparable error percentages (4.83 percent and 4.81 percent respectively). Means for test and control words in different conditions are presented in Table 2. Analyses of variance were carried out on the mean response times for experimental conditions. All factors are within-participant variables in the participant analyses and between-item variables in the item analyses. Separate analyses were conducted for cognates and homographs in relation to their matched

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Table 2. *Mean latencies and percentages of errors (between brackets) for cognates, homographs, and matched English control words in Experiment 1*

	Word type			
Frequency categories	Cognates	English controls	Homographs	English controls
HFE-HFD	546 (1.8)	567 (2.6)	548 (2.3)	554 (2.3)
HFE-LFD			548 (1.5)	556 (1.2)
LFE-HFD			609 (7.7)	588 (5.4)
LFE-LFD	593 (7.4)	630 (9.1)	620 (7.4)	627 (9.8)

controls. Separate analyses are appropriate because the frequency categories were not the same for the two types of items (see "Materials").

An analysis involving cognate status (cognate/control) and English frequency (high/low; totally confounded with Dutch frequency) compared cognates to their matched controls. Cognate status had a significant effect on latencies [$F_1(1,36)=45.48$, $p<.001$; $F_2(1,103)=15.14$, $p<.001$]. Mean latencies for cognates and their controls were 570 ms and 595 ms, respectively. A significant main effect was found for Frequency as well [$F_1(1,36)=217.27$, $p<.001$; $F_2(1,103)=54.11$, $p<.001$]. Mean latencies for high and low-frequency items were 557 and 612 ms, respectively. The frequency effect appeared to differ somewhat for the cognates and controls (47 ms and 63 ms respectively) as indicated by the trend towards an interaction between cognate status and frequency [$F_1(1,36)=3.27$, $p>.08$; $F_2(1,103)=1.20$, $p>.25$].

In the analogous analysis for the homographs, no significant difference was obtained between homographs and control items (both $F < 1$). Mean latencies for homographs and their controls were both 580 ms. The main effect of English frequency was significant [$F_1(1,36)=136.53$, $p<.001$; $F_2(1,102)=63.30$, $p<.001$]. High-frequency English words were responded to faster (552 ms) than low-frequency English words (611 ms). No interaction was obtained between word type and frequency [$F_1(1, 36)=2.78$, $p>.10$; $F_2(1,102)<1$].[2]

[2] In order to try to understand the 21 ms difference between LFE-HFD homographs (English low frequency-Dutch high frequency) and English controls (means of 609 and 588 ms respectively), we ran a three-way analysis of variance (Word Type, English Frequency, Dutch Frequency) but did not find a three-way interaction as one could have expected based on the means. In addition, close inspection of the RTs indicates that it is not the LFE-HFD homograph condition that deviates from the general pattern, but the matched control condition. Thus, a null-effect interpretation is still the most appropriate here.

On analyzing the homographs separately from their controls, the main effect of Dutch Frequency and the interaction between English and Dutch Frequency were found not to be significant (both $F<1$), but the main effect of English Frequency was [$F_1(1,36)=74.14$, $p<.001$; $F_2(1,103)=36.34$, $p<.001$]. RTs to English high-frequency homographs were 548 ms vs. 614 ms for English low-frequency homographs.

Discussion

The main results are easily summarized. While cognates were responded to significantly faster than their matched purely monolingual control items, no significant difference in RTs was obtained between homographs and their controls. Let us consider the results for these two types of items one by one.

In our discussion of the cognate effect, we must keep in mind that our study differs from other studies that have examined this effect (e.g. Cristoffanini, Kirsner, & Milech, 1986; Gerard and Scarborough, 1989; Nas, 1983) on such variables as languages, stimuli, experimental procedure, and type of bilinguals, as indicated in the Introduction. In what follows, we will propose an explanation that we believe is appropriate for our language pair, experimental set up, stimuli, and type of participants (bilinguals tested in their highly fluent L2).

To start, we may note that the two readings of the cognates overlap semantically, and that this overlap is absent in case of control items and homographs. Thus, the facilitation of the lexical decision response for cognates relative to matched controls and homographs can be taken as evidence that the semantic representation associated with the reading in the target language (English) is also activated on-line by the non-target language reading (Dutch) of the cognate. In other words, the cognate effect can be ascribed to a larger activation of (partially) shared semantic representations in the case of cognates. In addition, the convergence of activation from both the target and non-target language at the semantic level for cognates may subsequently lead to a larger activation feedback from the semantic level to the orthographic level than for monolingual control words. The amount of extra feedback in the case of cognates would be relatively large especially because the non-target language in our experiment (Dutch) was the native language of the participants.

It will be noted that our result is in line with other evidence on cognates in bilingual research (e.g., Cristoffanini, Kirsner, & Milech, 1986). In fact, it has gradually become clear that a proper evaluation of the language selective and nonselective access hy-

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

potheses requires a distinction between a form (orthographic) and content (semantic) level of representation. It is generally agreed upon that representations at a semantic level are (at least partially) shared by different languages (Potter, So, Von Eckhardt, & Feldman, 1984; Snodgrass, 1984; Schwanenflugel & Rey, 1986; Smith, 1991; Kroll & Sholl, 1992; Kroll, 1993; Fox, 1996). The observed facilitation effect for cognates is consistent with this view (which is, in fact, also proposed by Gerard and Scarborough for their Spanish-target result on cognates, p. 313).

In contrast, no significant latency differences were observed between interlingual homographs and control items. The absence of an effect of the overlap in letters for the different readings of the homographs suggests that at the word form level language selective access took place. However, this conclusion is at variance with recent evidence to the contrary (Grainger & Dijkstra, 1992; Bijeljac-Babic, Biardeau, & Grainger, 1997; Van Heuven, Dijkstra, & Grainger, 1998). For instance, Van Heuven et al. (1998) manipulated orthogonally the number of words similar to a target word (so-called "neighbors") within and outside the target language and observed cross-linguistic effects in several experiments. Given these recent findings from the experimental literature, the results for interlingual homographs are puzzling: The two readings of a homograph may be considered to be the strongest neighbors possible, since they consist of the same letter string. If lexical access is nonselective at both semantic and form representation levels, why then did we not obtain any cross-lingual effect on homograph recognition?

We may point out that the bilingual participants in Experiment 1 were explicitly instructed to react to English items only, and that no exclusively Dutch items were included in the experiment. Both factors may have induced a language set in which the English lexicon of the bilingual is much more active than the Dutch lexicon. Indeed, participants may even have tried to suppress the Dutch reading of homographs when they suspected that the experimenter would "sneak in" some purely Dutch items in the stimulus list to fool them. In other words, Experiment 1 may have provided an experimental context in which effects of the activated Dutch readings of homographs did not become apparent because they were relatively small due to the experimental focus on English in terms of instruction and items.

However, alternative interpretations are possible. For instance, the null results for homographs could be a consequence of facilitatory and inhibitory effects canceling each other, due to inter-individual differences in task performance or due to intra-individual shifts in performance strategy in the course of the experiment (suggestion by De Groot, personal communication). For instance, some participants may have been reluctant to use the Dutch readings of the test words as a source of information to speed up their reactions to the test items, because such a strategy would deviate from the experimenter's instructions and might therefore induce errors. This task approach would induce inhibition effects because of the competition between the English and Dutch readings of the test words. Other participants, however, may have felt more confident in using the Dutch readings of homographs and cognates, leading to a facilitation of their reactions to test items relative to controls. Alternatively, it is possible that a shift in performance strategy takes place within individual participants in the course of the experimental session. At the beginning of the experiment, the participant's confidence that Dutch lexical information can be used without detrimental consequences is low, inducing a performance strategy that leads to an inhibition of test items relative to controls. When the participant's confidence gradually increases during the experiment, more and more use is made of Dutch information in target identification, resulting in an increasing facilitatory effect for homographs.

We examined these hypotheses of strategy shifts within and across participants by analyzing the latency differences between homographs and controls for each participant in the first and second halves of the experiment. However, the resulting data patterns did not provide strong support for either account. There was neither a clear development from inhibition to facilitation within participants, nor a bimodal distribution of some participants showing consistent facilitation and others showing consistent inhibition effects. Since our design was not suited to analyze more local shifts in task performance, future studies should test these hypotheses by exerting tight control over the temporal aspects of the experimental design.[3]

Having found evidence apparently supporting selective access with homographs, we were ready to manipulate the task required of participants and the amount of language intermixing in the stimuli. In Experiment 2 below, participants were instructed to

[3] Another possibility would be that orthographic and phonological similarity effects have canceled each other (see Dijkstra, Grainger, and Van Heuven, in preparation). To examine the contribution of phonological similarity effects to the recognition of the homographs, we divided all experimental items (including the cognates) into two groups, with either a relatively high or a low phonological similarity of the English and Dutch item readings. However, no significant RT-differences were observed between the two groups of items for either homographs or cognates.

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

say "yes" exclusively to English target words (as in Experiment 1), but "no" to non-English items (Dutch fillers and nonwords). The inclusion of Dutch filler items in the stimulus list should maintain the Dutch lexicon at a much higher activation level than in Experiment 1. This larger activity in the Dutch lexicon may potentially induce inhibition effects and more errors for the interlingual homographs (relative to purely English control words). If the Dutch reading of an interlingual homograph becomes available to participants earlier than the English reading which is activated in parallel, recognition of the English reading may be delayed and/or participants may be tempted to give a "no"-reaction. This would induce an inhibition effect, the size of which would depend on the relative frequency of the Dutch and English readings of the homograph. Larger inhibition effects are expected for homographs that are characterized by a smaller English/Dutch word frequency ratio. The largest effects are therefore expected for low-frequency English homographs with high-frequency Dutch readings.

Experiment 2

Method

Participants. Twenty-four different participants were paid for their participation in this experiment. They were students of English originating from the same population as the participants in Experiment 1. All were native speakers of Dutch with considerable proficiency in English.

Materials and procedure. First, a set of fifty-six homographs and fifty-six English control items was selected. Most of these items had also been used in Experiment 1 (examples are given in the appendix). Four homographs that had resulted in high error percentages (over 35 percent) in Experiment 1 were replaced, as were matched control items which had resulted in high error rates. Data with respect to matching between test and control items are presented in Table 3.

To these 112 experimental items, 112 monolingual English filler words of about the same length and frequency were added. Next, 56 monolingual Dutch words of 3–5 letters long (mean log frequency: 1.67) were selected to serve as "nonwords in English". Note that these Dutch items required a "no" reaction, so their RTs cannot be directly compared with those of the homographs in this experiment. To this set of 56 "English nonword" items, 168 filler nonwords were added. All but two of these had also been used in Experiment 1. The added nonwords were constructed following the same criteria as before. In all, the total stimulus set consisted of 448 items, of which 224 required "yes" responses and 224 "no" responses. "Yes" responses were to be given to 56 homographs, 56 control items, and 112 monolingual English fillers. "No" responses were correct for the nonword stimuli consisting of 56 Dutch words and 168 filler nonwords.

Table 3. *Mean word length in letters (WL) and log word frequency in English (ELF) and in Dutch (DLF) of homographs and matched English controls in Experiment 2*

	Word type				
	Homographs			English controls	
Frequency categories	WL	ELF	DLF	WL	ELF
HFE-HFD	3.9	1.95	1.83	3.9	1.94
HFE-LFD	3.9	2.05	.64	3.9	2.04
LFE-HFD	3.9	.98	1.81	3.9	.98
LFE-LFD	3.8	.99	.69	3.8	1.00

To control for the effects of language shift between consecutive items (cf. Von Studnitz & Green, 1997), items that preceded test and control items were carefully selected. English fillers (requiring "yes" responses), Dutch fillers and English-like nonwords (requiring a "no" response) were inserted in the stimulus list preceding interlingual homographs or control words (requiring "yes" responses) in such a way that the obtained sequence of "yes" and "no" responses was unpredictable to the participants. Four presentation lists were created to obtain a completely counterbalanced experimental design in which the interlingual homographs and the control words were placed in the same position across the experimental lists.

Participants performed an English lexical decision task similar to the one used in Experiment 1. The language of instruction was English. Participants were notified that exclusively Dutch words could appear, requiring a "no" response. For presentation purposes, the stimulus set was divided into two blocks of experimental trials of 224 items each. A short break was introduced between the blocks. The first four trials of each experimental block were dummy trials.

Results

Data cleaning procedures were similar to those used in Experiment 1. Four items (one homograph, one control, and two fillers) with error percentages larger

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Table 4. *Mean latencies and percentages of errors (between brackets) for homographs and matched English controls in Experiment 2*

Frequency categories	Word type	
	Homographs	English controls
HFE-HFD	603 (12.2)	564 (5.4)
HFE-LFD	578 (2.0)	553 (2.4)
LFE-HFD	661 (19.4)	607 (6.2)
LFE-LFD	632 (15.6)	610 (9.2)

than 50 percent were discarded. Leaving out these items did not affect the matching between test and control items. Data of three participants which had an overall error percentage larger than 15 percent or a larger than 15 percent difference between error rates for words and nonwords were also left out.

Errors and outliers were replaced as before (6 percent of the total data set). The pattern of error rates did not deviate from that obtained for the latencies. Mean RTs for test and control items in different frequency categories are presented in Table 4. Not included in this Table are the RTs to English filler items, Dutch items, and nonwords, which were respectively 568 ms, 648 ms, and 626 ms. An analysis involving homograph status (homograph/control) and English frequency (high/low) compared homographs with their matched controls. The analysis of the latencies for the correct responses showed significant differences between the two types of items [$F_1(1,20)=25.31$, $p<.001$; $F_2(1,106)=18.72$, $p<.001$]. Mean latencies for homographs and controls were 618 ms and 583 ms, respectively. The analysis further showed a significant main effect of English word frequency [$F_1(1,20)=69.45$; $F_2(1,106)=44.29$, both $p<.001$], but the interaction between word type and English frequency was not significant [both $F<1$]. Mean RTs to high-frequency English words (homographs and controls) were 575 ms and to low-frequency English words, 627 ms.

To assess the effects of the Dutch frequency manipulations, an additional analysis was performed on the homograph latencies only. This analysis showed a significant effect of English Word Frequency [$F_1(1,20)=38.30$; $F_2(1,103)=9.89$, both $p<.001$] and of Dutch Word Frequency [$F_1(1,20)=9.73$, $p<.001$; $F_2(1,103)=4.75$, $p=.03$]. High-frequency English items (591 ms) were responded to faster than low-frequency items (647 ms). For the high- and low-frequency Dutch readings of the homographs, mean latencies were 632 and 605 ms, respectively. The interaction between the two frequency factors was not significant [both $F<1$].

Discussion

The inclusion of Dutch filler items requiring a "no" response induced a strong change in the participants" performance. In contrast to Experiment 1, the latency patterns for interlingual homographs in Experiment 2 provide strong support for a language nonselective access hypothesis. Reactions to the English reading of interlingual homographs were strongly inhibited by the presence of Dutch readings. Since for Experiment 2 suppressing activity in the Dutch lexicon would have been beneficial to performance, we must conclude that participants were unable to do so completely. They could not simply identify a homograph as an English item and ignore that it was also a Dutch item. Apparently, when the stimulus list contains both Dutch and English items, the bottom-up activation of the Dutch reading of the homographs is strong enough to override any potential top-down effects exerted by the requirements of the task, and the activated or identified Dutch reading somehow slows down the reaction to the English reading of the homograph.

Furthermore, the size of the inhibition effect depended directly on the frequency of the Dutch reading (high or low). This finding can be explained by assuming that, for both target languages, high-frequency readings become available more quickly than low-frequency readings, and that available Dutch readings slow down the lexical decision to English readings. Therefore, interference effects were found, especially for homographs with high-frequency Dutch and low-frequency English readings. Possible mechanisms underlying the inhibition effects will be discussed in some detail in the General discussion.

In Experiment 3, we will test a consequence of the view that participants cannot ignore the non-target language reading of the interlingual homograph. If this is true, participants might benefit from the presence of a second reading under different experimental circumstances. We tested this hypothesis in a so-called "general lexical decision task," where the bilingual participants had to respond "yes" to both English and Dutch words, and "no" to nonwords. As before, the presence of English and Dutch items in the experiment should ensure that both the English and Dutch lexicons are active. Under these circumstances, participants would be able to react as soon as either of the two readings of a homograph becomes available. As the result of such a "race" between the two readings of the homograph we

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Table 5. *Mean word length in letters (WL), log word frequency in English (ELF) and in Dutch (DLF) of homographs and matched English and Dutch controls in Experiment 3*

	Word type						
	Homographs			English controls		Dutch controls	
Frequency categories	WL	ELF	DLF	WL	ELF	WL	DLF
HFE-HFD	3.9	1.95	1.83	3.9	1.94	3.9	1.83
HFE-LFD	3.9	2.05	.64	3.9	2.04	3.9	.64
LFE-HFD	3.9	.98	1.81	3.9	.98	3.9	1.80
LFE-LFD	3.8	.99	.69	3.8	1.00	3.9	.68

should observe a facilitation effect of homographs relative to their matched monolingual controls. In addition, the degree of facilitation observed should be a function of the frequency of both the English and the Dutch readings of the homograph. The largest benefit to the RT relative to matched English controls should now be observed for the homographs with low-frequency English and high-frequency Dutch readings.

Experiment 3

Method

Participants. Twenty-eight different participants from the same student population as before were paid for their participation in the experiment. They were all native speakers of Dutch. Their proficiency in English was comparable to that of the participants in Experiments 1 and 2.

Materials and procedure. The stimulus set consisted of 336 items. The 168 items that required a "yes" response, consisted of the 56 homographs, 56 matched English control words, and 56 monolingual Dutch control words. While the items in the first two categories were the same as in Experiment 2, the Dutch control items were selected according to the same criteria as the English control words and they were matched with the interlingual homographs on a one-by-one basis. Data with respect to the matching of different word types are presented in Table 5. The 168 items requiring a "no" response were nonwords in both English and Dutch. Of these, 84 items were created by changing one letter in an existing English word and 84 items were created by changing one letter in a Dutch word that was not used in the experiment. All nonwords were orthographically legal pseudowords in English and Dutch. Eighty of the English-like items had also been used in Experiment 2.

Table 6. *Mean latencies and percentages of errors (between brackets) for homographs and matched English and Dutch controls in Experiment 3*

	Word type		
Frequency categories	Homographs	English controls	Dutch controls
HFE-HFD	518 (.3)	569 (3.3)	530 (2.4)
HFE-LFD	546 (2.1)	555 (2.1)	583 (6.7)
LFE-HFD	546 (2.7)	627 (9.7)	524 (2.4)
LFE-LFD	610 (8.7)	626 (11.2)	582 (6.0)

Experimental procedures were generally the same as in Experiments 1 and 2. However, this time participants performed a general lexical decision task, in which they were told to give a "yes" response when the letter string presented was an English or Dutch word. The language of instruction was English. They were to give a "no" response only when the letter string was not an existing word in either language. For presentation purposes, the stimulus set was divided into two blocks of 168 experimental trials, separated by a short break.

Results

Data cleaning procedures were the same as in Experiments 1 and 2. Five items (one homograph and four control items) with error percentages larger than 10 percent were discarded. Leaving out these items did not affect the matching between homograph and control items. Data of five participants which had an overall error percentage larger than 10 percent or a larger than 10 percent difference between error rates for words and nonwords were also left out, leaving data of twenty-four participants for the analysis. The pattern of error rates did not deviate from that obtained for the latencies. The resulting overall RTs to words and nonwords were respectively 566 and 648 ms, with corresponding error percentages of 4.7 percent and 4.9 percent. Means for the four frequency categories of homographs with the corresponding English and Dutch control items in all conditions are presented in Table 6. To assess the effect of word type (homograph/control) and English frequency (high/low), homographs were compared with their English and Dutch controls in two separate analyses. A significant effect of word type was found when homographs were compared with English

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

control items [$F_1(1,23)=33.66$; $F_2(1,106)=20.31$, both $p < .001$]. Latencies for homographs (554 ms) were shorter than for their matched English control words (592 ms). A significant main effect for English frequency was also obtained [$F_1(1,23)=163.11$; $F_2(1,106)=38.19$, both $p < .001$]. The mean RTs obtained for high-frequency and low-frequency English items (homographs and English controls) were 547 ms and 601 ms, respectively. The interaction between word type and English frequency was significant in the participant analysis [$F_1(1,23)=5.28$, $p<.05$], but not in the item analysis [$F_2=1.14$]. The frequency effect for the control items (64 ms) was larger than for the homographs (45 ms).

In the analogous analysis of the latencies for homographs and Dutch controls, the main effect of word type was not significant [both $F<1$]. Mean RTs to homographs and controls were both 554 ms. Furthermore, the main effect of Dutch frequency was significant [$F_1(1,23)=114.07$; $F_2(1,106)=39.81$, both $p<.001$], but the interaction between word type and Dutch frequency was not [$F_1(1,23)=1.69$; $F_2<1$]. High-frequency Dutch items (homographs and Dutch controls) had a mean RT of 529 ms vs. 579 ms to low-frequency Dutch items.

A further analysis of frequency effects for the homographs alone showed a significant main effect for English word frequency [$F_1(1,23)=55.50$; $F_2(1,106)=16.92$, both $p<.001$]. Latencies were longer for items with low English word frequency (577 ms) than with high English frequency (532 ms). The main effect of Dutch word frequency was also significant [$F_1(1,23)=29.55$; $F_2(1,106)=16.34$, both $p<.001$]. Homographs with a high frequency in Dutch had a mean RT of 532 ms vs. 576 ms for those with a low Dutch frequency. The interaction between English and Dutch frequency was significant in the participant analysis [$F_1(1,23)=9.47$, $p<.01$], but not in the item analysis [$F_2(1,106)=2.50$, $p=.12$]. The English frequency effect was smaller for the homographs with a high frequency in Dutch (29 ms) relative to a low frequency (64 ms).

Discussion

In accordance with our predictions, strong facilitatory effects were observed in the general lexical decision task of Experiment 3 for Dutch/English homographs relative to monolingual English control words. This finding indicates that, given the instruction "say yes to English and/or Dutch words," participants were able to use either reading of an interlingual homograph to speed up their decision process. Furthermore, the presence of both English and Dutch word frequency effects suggests that participants reacted on the basis of the first available reading of the homograph. Therefore, the homographs that benefited most from the change in instruction in Experiment 3 were those that suffered most in Experiment 2: homographs with a low-frequency English reading and a high-frequency Dutch reading. Given that the Dutch reading of this type of homograph becomes available quickly relative to the English reading, strong interference effects were found in Experiment 2 and strong facilitation effects in Experiment 3.

Because in Experiment 3 "yes" responses were given to both English and Dutch items, the results for the homographs can be compared not only to their matched English controls but to their matched Dutch controls as well (see Table 6). Relative to these controls, facilitatory effects were observed for homographs with high-frequency readings in English. This suggests that for these types of items the English reading is sometimes accessed before the Dutch reading. If the RT is determined by the first available reading of the homograph, faster RTs would arise than would be expected on the basis of the Dutch reading only. In other words, this finding could be ascribed to "statistical facilitation" (Raab, 1962).

However, close observation of Table 6 indicates that the idea of a simple race with two independent channels (one for the Dutch and one for the English reading of the homograph) cannot explain the complete performance pattern of the bilingual participants. Such a view predicts that the RTs to the homographs should always be as fast as, or even faster than, the minimum of the RTs to the English and Dutch controls (see Raab, 1962; Kawamoto & Zemblidge, 1992; Dijkstra, Frauenfelder, & Schreuder, 1993). As Table 6 indicates, this is true for homographs with high-frequency English readings, but for items with low-frequency English readings RTs are situated between those of Dutch and English controls. The present data therefore are incompatible with a race-model with independent channels.[4]

One admittedly speculative explanation of the current pattern of results is that participants tend to react more often to the English reading of the homograph than would be expected on the basis of a race model. This could result from an attempt on the part of the participants to optimize their task performance by compensating for their weaker language, English. They might, for instance, give some kind of priority to the English lexicon during response selection. This account should, of course, be submitted to further research.

[4] The BIA model presented later could be considered as a race model with complex interactions between the two channels.

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Nevertheless, irrespective of the type of interaction which is assumed between the processing of the English and Dutch readings of the interlingual homographs, we may conclude that the general lexical decision task of Experiment 3 induced the maintenance of activation in both English and Dutch lexicons and resulted in cross-language effects.

General discussion

In this paper we have demonstrated that the degree of language selectivity observed in bilingual word recognition experiments depends critically on the characteristics of the experimental situation at hand. In our experiments, the RTs to one and the same set of interlingual homographs displayed a more language selective or nonselective pattern depending on task demands (e.g., general or language specific lexical decision) and language membership of stimuli in the experimental list (English and/or Dutch). Below we will discuss these results and then integrate them into a processing model of bilingual word recognition that takes into account effects of task demands and language intermixing.

In Experiment 1, all items were English words or nonwords derived from English words. In this experimental context, RTs to homographs did not differ significantly from those to control words, and word frequency in the target language was the main determinant of the RTs. These results, which to some extent replicate those of Gerard and Scarborough (1989), therefore seemed in agreement with a language selective access hypothesis. In contrast, RTs to cognates were faster than to controls, suggesting language nonselective access.

A number of interpretations of the null effect for the homographs have already been presented in the discussion of Experiment 1. The interpretation we prefer is based on the assumption that the English lexicon in Experiment 1 was relatively active, because of the instruction to react to English words and the absence of exclusively Dutch items in the experiment. This active English lexicon quickly suppresses Dutch words, including the Dutch readings of homographs and cognates. Due to this suppression, a null-effect for homographs is obtained. However, the presence of an effect for cognates can still be explained if it is assumed that the Dutch reading of the test items was activated long enough to contribute to semantic activation. Because there is semantic overlap in the case of cognates (but not homographs), activation from both readings of the cognate converges at the semantic level. Furthermore, to the extent that the representation of meaning is language-independent, such semantic information would not be suppressed by the linguistic system. The more active semantic representations could therefore be used by the participants as a basis for response, possibly in combination with the more active orthographic representations that may arise due to subsequent activation feedback from the semantic to the lexical level.

In Experiment 2, two important changes were made with respect to stimulus material and task instruction in addition to removing cognates. First, Dutch filler items were added to the experiment, which resulted in a list of items which were either exclusively English or exclusively Dutch items, or were Dutch/English homographs. Second, participants were instructed to treat the added Dutch words as nonwords. In combination, the two changes made a dramatic difference to the response patterns to interlingual homographs. Now that the required response to the Dutch control items ("no") contrasted with that for homographs ("yes"), strong inhibitory effects were observed. Under these circumstances, task performance would have benefited from a complete suppression of the Dutch lexicon, which would have yielded similar RTs to homographs and exclusively English control words, as in Experiment 1. The inability of the participants to establish such suppression may be ascribed to the presence of Dutch items in the experiment. Such items may ensure and maintain a relatively large activity of the Dutch lexicon. The influence of other-language items in the stimulus list on the recognition of interlingual homographs is most easily understood in terms of a language nonselective view of lexical access.

Several mechanisms underlying the observed inhibition effects in Experiment 2 can be proposed, each locating the origin of the effects at a different stage of the recognition process. An "early" source of competition effects is assumed by the standard Interactive Activation mechanism of (inherent) lateral inhibition between words. Perhaps the inclusion of Dutch words in Experiment 2 induced stronger competition from non-target language lexical candidates. In Experiment 3 this hypothesis was tested by changing the task situation while keeping the type of stimulus material the same. The interference effects disappeared, suggesting that lateral inhibition alone is insufficient to account for the results of Experiment 2.

Alternatively, the task situation of Experiment 2 could be compared with that in other discrimination tasks and an explanation in terms of the "late" mechanism of response competition might be preferred. Response competition may have played a role because the non-target language reading of a homograph was directly linked to the "no" reaction.

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

However, recent findings in our laboratory suggest that the emergence of the inhibition effect is at least not dependent on an *explicit* coupling of non-target language items to a "no" response. Findings that reflect inhibition were also obtained by Dijkstra, Timmermans, and Schriefers (in preparation) for a "go/no-go" paradigm in which the relative frequency of the English and Dutch readings of interlingual homographs was manipulated. Here, participants reacted only when a presented word belonged to a prespecified target language (English or Dutch). In both English and Dutch "go/no-go" tasks, participants were found to react quickly and with few errors to monolingual control words, but they often failed to detect homographs as if they assumed they were words from the non-target language.

These findings suggest a third mechanism which might account for the slower RTs and larger number of errors for homographs relative to controls in Experiment 2. It locates the inhibition effects at an "intermediate" stage of the recognition process, assuming they are due to the occasional identification of the wrong reading of the homograph or to lexical competition effects during late stages of the recognition process. Further research is necessary to test which (combination) of the proposed mechanisms is best suited to account for the data.

Experiment 3 incorporated the same type of stimulus material as Experiment 2, but now the required responses to homographs and monolingual control words were consistent ("yes"). Under these circumstances participants were able to use effectively the activity in both lexicons to speed up their reactions to homographs relative to controls. As in Experiment 2, the size of the effects depended on the ratio of the frequencies of the item readings in the two languages. Apparently, participants were able to take advantage of the change in instruction from Experiment 2 to Experiment 3 (as reflected by the facilitation effects for homographs), but not of that from Experiment 1 to Experiment 2 (as indicated by the observed inhibition).

Furthermore, Experiment 3 suggests that in accordance with task demands participants are to some extent able to base their response on different information sources available from the lexical network (e.g., choosing a "read-out" from the word and/or language level; see Grainger & Jacobs, 1996; Dijkstra & Van Heuven, in press). For instance, the general lexical decision instruction in Experiment 3 resulted in facilitation effects for interlingual homographs relative to purely monolingual controls. It seems participants were able to react as soon as either reading of the homograph was available, or even on the basis of the summed global activity in the lexical system as a whole (cf. Grainger & Jacobs, 1996). Indeed, relative to Experiment 1, the facilitation effects in Experiment 3 were larger especially for homographs with a high-frequency Dutch reading (HFE-HFD and LFE-HFD items; see Tables 2 and 6). This is in line with any processing model that assumes some sort of race between the two readings of the homograph to be recognized. In such a model, the higher frequency reading of a homograph would on average reach an activation threshold for recognition faster than its lower frequency counterpart. This would be the case even though, of course, for our Dutch/English bilinguals the subjective frequency of English (L2) words would overall be lower than for Dutch (L1), implying relatively less activation or a slower activation rate in English than in Dutch.

On the basis of the discussion of the three experiments we can draw a number of conclusions: (1) our homograph results were affected by the relative degree of non-target language activation due to other-language items in the experiment, which can be considered as indirectly operating bottom-up sources of lexical activation (Experiment 1 vs. 2); (2) task demands stimulating a change in the relative activity of the two lexicons (top-down sources) cannot easily override the effect of such bottom-up sources (Experiment 2); but (3) task demands can exploit a state of activation in the two lexicons they are consistent with (Experiment 3).

The BIA model of word recognition

It is hard to see how any bilingual word recognition model could account for the three very different RT patterns to interlingual homographs relative to controls in our series of experiments without making additional assumptions with respect to the effects of task demands and language intermixing. Clearly, in order to account for such systematic variation across the experiments, models must be able to produce different data patterns in reaction to different experimental conditions. The relevance of such a comprehensive account will now be illustrated by discussing one particular model for bilingual word recognition, the BIA model, and the extensions it will need to account for the current series of experiments.

In the Bilingual Interactive Activation (BIA) model, depicted in Figure 1, the bilingual word recognition process is subserved by a network architecture extending that of the well-known Interactive Activation model for monolingual word recognition (McClelland & Rumelhart, 1981; see Dijkstra & Van Heuven, in press, for an overview of the model). The BIA model assumes an integrated lexicon for the two languages of a bilingual. Thus, when a letter string is

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

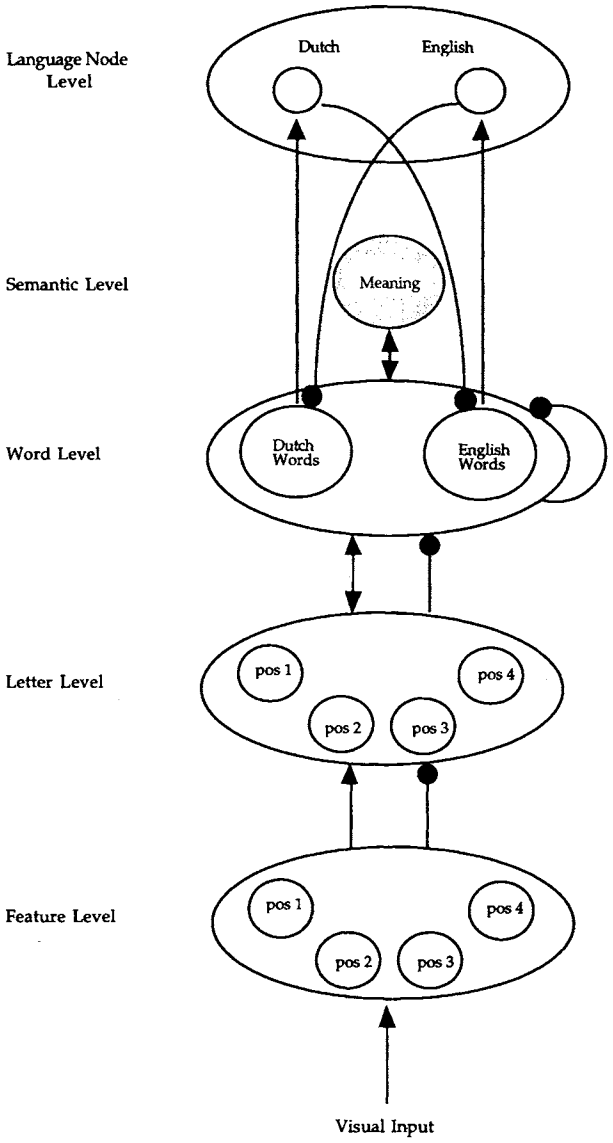


Figure 1. The Bilingual Interactive Activation (BIA) model. The unimplemented extension with semantic representations is shaded.

presented, word candidates in both lexicons are first activated to the extent that they match the input. Next, Interactive Activation mechanisms such as lateral inhibition (between words) and top-down feedback (from words to letters) suppress non-target lexical candidates. This ensures that the proper word candidate is quickly activated the most often and recognized. Thus, the recognition process is initially language nonselective, since item activation is affected by competing word candidates from the same and the other language (cf. Van Heuven, Dijkstra, & Grainger, submitted). However, the incorporation of so-called language nodes makes it possible for language selective data patterns to arise in particular circumstances as well (see Figure 1). A language node collects activation of all words from one lexicon (e.g., English) and suppresses all words in the other lexicon (e.g., Dutch). A relatively monolingual mode of operation of the model may arise if one language node is very active relative to the other.

An extension of the BIA model incorporating semantics (see Figure 1) may explain the finding that in Experiment 1 Dutch/English cognates but not homographs led to faster RTs than English control words as follows. The current model (without semantics) shows that an input letter string initially (e.g., for the first five cycles or so) activates both English and Dutch orthographic representations of cognates and homographs. This happens irrespective of the inhibitory strength of the target's language node to word feedback parameter (Dijkstra & Van Heuven, in press). For cognates, this implies that in the initial phases of recognition, the non-target language reading of the cognate can activate its semantic representation. Assuming that the initial activation spread from the second-language reading of the cognate to its semantic representation is large enough, the observed facilitation effect for cognates follows because the two readings of the cognate share (at least part of) their semantic representation. Thus, for cognates, the suppression by the relatively active English language node arrives too late to cancel the facilitation effect at the semantic level and the subsequent feedback from the semantic to the orthographic level (see Figure 1). However, for interlingual homographs no essential processing difference relative to controls would be observed if the non-target language reading of the homograph is suppressed soon enough. For these items, there is no semantic overlap between the two readings of the item that can be exploited to facilitate the response. As soon as semantic representations are implemented in the BIA model it will be possible to subject these proposals to a precise simulation test.

To account for the different results of Experiments 1 and 2 in relation to task demands and language intermixing effects, the BIA model could make the assumption that the language nodes cannot only be activated on-line on the basis of the input ("bottom-up"), but also by activation from sources external to the architecture presented in Figure 1 ("top-down"). Currently, the BIA model only deals with words presented in isolation. The effect of language intermixing can be accounted for by incorporating a mechanism which carries lexical activation from one trial over to later trials. An obvious way to do this is to assume that the presentation of an item in a particular language activates the corresponding language node and that at presentation of the next item,

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

this activation of the language node has not completely decayed. Indeed, this may explain recent reports of language switching effects in similar experimental situations (Von Studnitz & Green, 1997). Thus, when items from only one language are presented in the experiment, just one language node is re-activated time and again. However, when items from two languages occur, both language nodes may become and stay active in the course of the experiment, each suppressing words in the other lexicon.

The different results of Experiments 1 and 2 can then be understood as follows. The second experiment differed from the first in terms of the Dutch filler items and the task instruction. Due to the presence of Dutch items, the Dutch (non-target) language node is much more active than in Experiment 1. Furthermore, responding in Experiment 2 requires making a distinction between the two readings of the homograph. Inhibition effects for homographs relative to controls may arise because the target reading of the homograph becomes available more slowly and/or because its other reading co-activates the other language node (supporting a "no" response) (cf. discussion of Experiment 2).[5] Thus, it must be specified in the BIA model what happens in the processing system when the non-target reading of a homograph is recognized before the target reading.

A further addition to the BIA model is required to capture the results of Experiment 3. Here it should be possible to react on the basis of either reading of the homograph/cognate rather than just one, or perhaps even on the basis of the global activity generated in the two lexical systems by the two readings of the homographs or cognates. In other words, the processing model must be extended with a decision mechanism that can be made context-sensitive.

This analysis of the BIA model illustrates how a processing model for bilingual word recognition could be extended with an account on task demands and language intermixing effects to provide a coherent and consistent theoretical framework for the interpretation of differently sized and even opposite cross-language effects. Recently, several researchers in the bilingual domain have pointed out how sensitive a bilingual's responses are to various experimental conditions (Grosjean, 1997a, 1997b; De Groot, 1995; Kroll & De Groot, 1997). A detailed task account for bilingual lexical decision that may be compatible with our own is being developed by Green and colleagues (Green, 1986, 1993; Von Studnitz & Green, 1997). Developing, testing, and comparing context-sensitive processing accounts will require further empirical exploration of the effects of task demands and language intermixing on bilingual word recognition. Additional empirical evidence is also needed to constrain the implementation of task-specific factors in a computer model for bilingual word recognition such as the BIA model. But the present series of bilingual experiments has already shown that it may be dangerous and sometimes simply incorrect to interpret empirical evidence of facilitation or inhibition as a straight-forward reflection of the architecture of the underlying processing system.

[5] Effects of language switching between consecutive items in an experimental list would provide extra evidence for this line of reasoning. Such effects have been observed by Von Studnitz and Green (1997) for control items in two experiments that were to some extent comparable to Experiments 1 and 3 in this paper. In our Experiment 2, items preceding target homographs due to design considerations (i.e. the number of consecutive "yes" responses in lexical decision) were either Dutch items or nonwords. Previous items were never English, which in comparison to the Dutch items would have provided the strongest test for language switching. Post-hoc analysis of Experiment 2 indicated a statistical trend for different priming effects on homographs preceded by Dutch items and nonwords in the participant analysis [$F_1(3,20)=1.83$, $p=.15$], but not in the item analysis [$F<1$]. Note that no significant switching effects would be expected when the Dutch lexicon was already maximally activated (ceiling effect).

References

Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, *22*, 174–88.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, 26, 658–72.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition*, 25 (4), 447–57.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, 13 (4), 212–14.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *Quarterly Journal of Experimental Psychology*, 38A, 367–93.

De Groot, A. M. B. (1995). Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning*, 8, 151–80.

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

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Dijkstra, A., & De Smedt, K. (1996) (eds.), *Computational Psycholinguistics: AI and connectionist models of human language processing*. London: Taylor & Francis.

Dijkstra, A., Frauenfelder, U. H., & Schreuder, R. (1993). Grapheme-phoneme interaction in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance*, 19 (5), 1–20.

Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (in preparation). On the role of phonology in bilingual word recognition. Unpublished manuscript, University of Nijmegen.

Dijkstra, A., Timmermans, M., & Schriefers, H. (in preparation). Cross-linguistic effects on bilingual homograph recognition. Unpublished manuscript, University of Nijmegen.

Dijkstra, A., & Van Heuven, W. J. B. (in press). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (eds.), *Localist connectionist approaches to human cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Durgunoglu, A. Y., & Roediger III, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language*, 26, 377–91.

Fox, E. (1996). Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory and Language*, 35, 353–70.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15 (2), 305–13.

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (eds.), *The Bilingual Lexicon* (pp. 11–26). Amsterdam/Philadelphia: John Benjamins Publishing Co.

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology*, 39A, 295–319.

Grainger, J., & Dijkstra, A. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (ed.), *Cognitive Processing in Bilinguals* (pp. 207–20). Amsterdam: Elsevier Science Publishers B.V.

Grainger, J., & O'Regan, J. K. (1992). A psychophysical investigation of language priming effects in two English-French bilinguals. *European Journal of Cognitive Psychology*, 4 (4), 323–39.

Green, D. W. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, 27, 210–33.

Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon. Studies in bilingualism: Vol. 6* (pp. 249–77). Amsterdam: John Benjamins Publishing Co.

Grosjean, F. (1997a). Processing mixed language: Issues, findings, and models. In A. M. B. De Groot & J. F. Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 225–54). Hillsdale, NJ: Lawrence Erlbaum Associates.

Grosjean, F. (1997b). Studying bilinguals: Methodological and conceptual issues. Unpublished manuscript, University of Neuchatel, Switzerland.

Kawamoto, A. H., & Zemblidge, J. H. (1992). Pronunciation of homographs. *Journal of Memory and Language*, 31, 349–74.

Keatley, C. W. (1992). History of bilingualism research in cognitive psychology. In R. J. Harris (ed.), *Cognitive processing in bilinguals* (pp. 15–49). Amsterdam: Elsevier Science Publishers B.V.

Kolers, P. A. (1966). Reading and talking bilingually. *American Journal of Psychology*, 79, 357–76.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon. Studies in bilingualism: Vol. 6* (pp. 53–81). Amsterdam: John Benjamins Publishing Company.

Kroll, J. F. & Sholl, A. (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. J. Harris (ed.), *Cognitive processing in bilinguals* (pp. 191–204). Amsterdam: Elsevier Science Publishers B.V.

Kroll, J. F., & De Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual. In A. M. B. De Groot & J. F. Kroll (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 169–99). Hillsdale, NJ: Lawrence Erlbaum Associates.

Macnamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, 10, 480–87.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, Part 1: An account of basic findings. *Psychological Review*, 88, 375–405.

Nas, G. (1983). Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior*, 22, 526–34.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23, 23–38.

Raab, D. H. (1962). Statistical facilitation of simple reaction time. *Transactions of the New York Academy of Sciences*, 24, 574–90.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 23, 84–99.

Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representation system in the bilingual lexicon. *Journal of Memory and Language*, 25, 605–18.

Smith, M. C. (1991). On the recruitment of semantic information for word fragment completion: Evidence from bilingual priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17 (2), 234–44.

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press

Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior*, 23, 3–22.

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*.

Von Studnitz, R. E., & Green, D. W. (1997). Lexical decision and language switching. *International Journal of Bilingualism*, 1(1), 3–24.

Appendix

Sample of Stimulus Material

The following tables contain examples of the experimental items used in Experiments 1 to 3. HFE: high-frequency English; HFD: high-frequency Dutch; LFE: low-frequency English; LFD: low-frequency Dutch; ELF: English log frequency (per million); DLF: Dutch log frequency (per million).

Examples of Cognates

Note: Cognates were presented in Experiment 1 only.

Frequency category	Test item			English control	
		ELF	DLF		ELF
HFE-HFD	BUS	1.91	1.76	BAG	1.93
	LAND	2.48	2.65	LINE	2.51
	STORM	1.61	1.65	PRIDE	1.61
LFE-LFD	RIB	1.11	1.23	APE	1.11
	ARENA	.85	.70	CROOK	.85
	JEEP	1.00	.70	BIKE	1.00

Examples of Interlingual Homographs

Note: Matched Dutch control items were presented in Experiment 3 only.

Frequency category	Test item			English control		Dutch control	
		ELF	DLF		ELF		DLF
HFE-HFD	BAD	2.52	1.38	LOW	2.33	BOM	1.36
	STEM	1.49	2.55	NAIL	1.49	FEIT	2.55
	SPOT	1.90	1.61	MEAT	1.90	KLEM	1.60
HFE-LFD	LIST	2.06	.78	MILK	2.06	MEST	.78
	STAGE	2.24	.30	DRESS	2.25	SNOEK	.30
	RANK	1.57	.70	FATE	1.56	KRAP	.70
LFE-HFD	BOOM	1.41	2.14	RIOT	1.41	DORP	2.14
	RAMP	.78	1.40	STAG	.78	HEMD	1.40
	VET	.95	1.58	FOG	1.00	RUW	1.58
LFE-LFD	DEN	1.04	.85	NUN	1.04	DAS	.85
	SMART	1.43	.95	VAGUE	1.43	SNOER	.95
	SAGE	.70	.70	DUNE	.70	HULS	.70

https://doi.org/10.1017/S1366728998000121 Published online by Cambridge University Press