More on interlingual homograph recognition: language intermixing versus explicitness of instruction*

TON DIJKSTRA, ELLEN DE BRUIJN, HERBERT SCHRIEFERS and SJOERD TEN BRINKE
University of Nijmegen

We contrasted the effect of instruction-induced expectancies and language intermixing in an English lexical decision task performed by Dutch–English bilinguals. At the start of the experiment, participants were instructed to respond to interlingual homographs and exclusively English words by giving a "yes" response, and to English non-words and exclusively Dutch words by giving a "no" response. In the first part of the experiment the stimulus list did not contain any Dutch words. In the second part of the experiment, Dutch items were introduced. No significant differences were found between interlingual homographs and controls in the first part of the experiment, while strong inhibition effects were obtained for interlingual homographs in the second part. These results indicate that language intermixing rather than instruction-based expectancies drives the bilingual partipants' performance. Consequences for current views on bilingual word recognition are discussed.

One issue that has dominated bilingual research in the last 25 years is whether access to the bilingual lexicon is basically language selective or non-selective in nature. On the basis of an increasing amount of evidence, more and more researchers are convinced that the bilingual word recognition system is basically non-selective in nature, but that, dependent on a number of experimental and non-experimental factors, more or less selective result patterns may be obtained (Dijkstra & Van Heuven, 1998; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Grainger, 1993; Grosjean, 1998, in press). Which factors are most effectively influencing the bilinguals' performance and how they interact has only recently become a focus of systematic study (see Grosjean, in press, for a review of the available evidence). For bilingual word recognition, this paper considers the effect of two of these factors in more detail: task instruction (does the instruction indicate that one or two languages will be important to perform the task?) and stimulus list composition, in particular language intermixing (are words from only one or from two languages present in the experiment?).

Both factors play a role in the notion of a "language mode", developed by Grosjean (1985, 1997a, 1998, in press). In their everyday lives, bilinguals are assumed to find themselves in various general "language modes" that correspond to points on a monolingual–bilingual mode continuum. A mode is a state of activation of the bilingual's languages and language processing mechanisms. At one end of the continuum, if the bilinguals are in a totally monolingual mode (for instance, when they are interacting with monolinguals of one of the languages they know), only the appropriate language is kept active and the other is deactivated. At the other end of the continuum, the bilinguals are in a bilingual language mode if they are communicating with bilinguals who share their two languages. In this case, both languages are active but the one that is used as the main language of communication (called the base language) is more active than the other.

Because a particular mode corresponds to a state of activation of the bilingual's languages and language processing mechanisms, it is assumed to influence both language production (maintenance or change of the base language, amount and type of language intermixing that takes place, etc.) and language perception (speed of processing of a language, access to one or both lexicons, role of the less activated language, etc.) (Grosjean, 1998, 170). Grosjean, 1998, 137) has recently applied the notion of the language mode to language perception and reading as follows:

> During perception, if bilingual listeners who start off in a monolingual mode determine (consciously or not) as they go along, that what they are listening to can contain

* The authors are indebted to Walter van Heuven, Judy Kroll, Marc Brysbaert, and Debra Jared for their valuable comments on an earlier version of this paper.

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/S1366728900000146 Published online by Cambridge University Press

elements from the other language, they will put themselves in a bilingual mode (at least partly), that is, activate both their languages (with the base language being more strongly activated). This is also true of readers, whether they are reading a continuous text or looking at individual lexical items interspersed with items from the other language. Simply knowing that there is a possibility that elements from the other language will be presented (in an experiment, for example) will move the bilingual away from the monolingual endpoint of the continuum. Just one guest word in a stream of base language words can increase this displacement.

In sum, according to Grosjean (a) the bilingual's language mode affects lexical access to the bilingual lexicon during word reading, and (b) the language mode is affected both by the readers' (linguistic) expectations (e.g., due to the instruction for the experiment) and by language intermixing (stimulus list composition).

Recent studies have started to address these issues empirically. Dijkstra et al. (1998) examined the effects of task instruction and language intermixing in two lexical decision experiments that provided the starting point for the present study. In the first experiment, Dutch–English bilinguals performed an English lexical decision task on a list of words including English–Dutch homographs and cognates, as well as exclusively English control words. Interlingual homographs were defined as words having the same orthography but different meanings across different languages. An example is the word LIST, which means "trick" or "guile" in Dutch. Cognates were defined as words with both identical orthography and largely overlapping semantics. An example is HOTEL. The experiment did not result in any significant RT (response time) differences for interlingual homographs relative to exclusively English control words, analogous to earlier findings by Gerard & Scarborough (1989) for Spanish–English bilinguals. For cognates, Dijkstra et al. observed a facilitation effect relative to exclusively English control words. Dijkstra et al. interpreted this finding as evidence in favor of non-selective access. This implies that the null-effect for interlingual homographs was not a consequence of selective access, but of some other characteristic(s) of the experimental situation.

This hypothesis was tested in Dijkstra et al.'s second experiment. Dutch–English bilinguals again performed an English lexical decision on interlingual homographs, exclusively English control items, and non-words, but now the experimental stimuli also included Dutch words requiring a "no" response (i.e., the Dutch words had to be treated as "English non-words"). Here, a dramatic change was observed in the RTs to homographs: relative to English control items, strong inhibitory effects were obtained. The participants apparently found it difficult to respond "yes" to these words. Furthermore, the RT difference between homographs and matched control words was found to depend on the relative frequency difference of the two readings of the homograph: the inhibitory effect was large when the Dutch reading of the homographs had a high frequency relative to the English reading.

Dijkstra et al. (1998) explained the different results in their Experiments 1 and 2 in terms of a difference in the relative degree of non-target (Dutch) language activation. The Dutch items in Experiment 2 activate the non-target language Dutch to a higher degree than in Experiment 1 where no Dutch items occurred in the list. This leads to relatively more competition between the two readings of interlingual homographs in Experiment 2. Thus, according to this account the different results in the two experiments are a consequence of different bottom-up activation processes due to the composition of the stimulus list.

Grosjean (in press) interpreted these results somewhat differently, as (indirect) evidence for an effect of language mode during perception. In Experiment 1, the participants only read English words and non-words (although some words were homographs and cognates) and they were instructed to decide whether the items were English words or not. This would have positioned them towards the monolingual end of the mode continuum, but they did not reach this position totally as they knew they were being tested as bilinguals. Thus, although their Dutch was partly active (which would explain the cognate effect) it was not sufficiently active to create a homograph effect. In sum, Grosjean (in press) assumed that both task instruction (English lexical decision) and degree of language intermixing (encountering mostly English words) affected the bilinguals' performance in the experiment.

Recently, De Groot, Delmaar, & Lupker (in press) replicated the basic results observed by Dijkstra et al. (1998). With respect to the observed differences between Dijkstra et al.'s Experiments 1 and 2 they proposed a theoretical interpretation that is different from both Dijkstra et al.'s and Grosjean's accounts. As they pointed out, in Experiment 1 the task may not always have been carried out according to the instructions. The participants were instructed to perform a "language specific" English lexical decision task, but on some trials they may instead have treated the task as a "language neutral" lexical decision task, i.e. they may have ignored the language a particular item belonged to. The adoption of a "language specific" processing mode by some partici-

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

pants or on some trials would be expected to induce *slower* responses to homographs than to matched controls because, in a non-selective access system, not only the target reading but also the non-target language reading of a homograph would be activated, leading to lexical competition and slower responses (just like in Experiment 2 by Dijkstra et al., 1998). In contrast, in a "language neutral" processing mode the response to a homograph would be based on the availability of any reading, irrespective of language, and homographs could then be responded to *faster* than controls (as in the generalized lexical decision task of Experiment 3 by Dijkstra et al., 1998). In sum, it could be that a mixture of the two processing modes adopted by the participants leads to a mixture of facilitation and inhibition effects for homographs, yielding an overall null-result.[1]

In this context, it is revealing to examine the instruction given to the participants in Experiment 1 by Dijkstra et al. (1998) and the comparable experiments by De Groot et al. (in press), and Gerard & Scarborough (1989). Neither in the first experiment by Dijkstra et al., nor in De Groot et al.'s Experiment 2 (Condition English), were the participants told that some of the to be presented letter strings would be words in both Dutch and English. As far as can be determined on the basis of the article by Gerard & Scarborough (1989), the same was true for Part 1 of their experiment.

It is therefore possible that participants in these experiments sometimes adopted a "language neutral" processing mode simply because they were in an uncertain situation. Moreover, the specifics of the experimental situation allowed the participants to respond with some flexibility at no apparent cost, because a Dutch word was always an English word as well, namely an interlingual homograph. Therefore, the participants in the experiments never made a response error if they responded on the basis of the first available reading of the homograph ("it is a word"; language neutral processing mode) rather than on basis of the English reading only after excluding the Dutch reading ("it is an English word"; language specific processing mode). This uncertainty did not exist for the participants in Experiment 2 by Dijkstra et al., who were explicitly instructed to say "no" to Dutch words, which they encountered from the very beginning of the experiment. Applying the same flexible strategy as in Experiment 1 would immediately have been punished by an increased error rate. Thus, it cannot be excluded that the instruction as formulated in Experiment 1, in combination with the specifics of the stimulus list composition (language intermixing), underlies the null-results in that experiment.

To disentangle the effects of instruction and language intermixing, we designed an experiment that combined features of Experiments 1 and 2 by Dijkstra et al. (1998). Before the start of the experiment, we instructed the participants explicitly that they would encounter interlingual homographs requiring a "yes" response, as well as exclusively Dutch words requiring a "no" response. We also provided them with examples of these stimuli in the practice set. We manipulated language intermixing by including exclusively Dutch words only in the second part of the experiment but not in the first part. Furthermore, in order to examine local effects of the inclusion of Dutch items in the list, we carefully controlled the presence of specific items at the end of the first part and the beginning of the second part.

These design features allowed us to test the following predictions. If the instruction is a determinant of relative language activation in word recognition (as argued by Grosjean, 1998), then bilinguals should activate both languages in part 1 of the present experiment, because both instruction and practice set stress the relevance of both languages in the experiment. This should induce a different pattern of results than in the "implicit instruction" situation of Experiment 1 by Dijkstra et al. (1998), i.e. we should obtain inhibitory effects on the responses to the homographs (similar to those obtained in Experiment 2 by Dijkstra et al.). If, however, language intermixing rather than degree of explicitness in instruction underlies the different results in Experiments 1 versus 2 by Dijkstra et al. (as argued by these authors), part 1 of the present experiment should not yield inhibitory effects (mimicking the results of Experiment 1 by Dijkstra et al., 1998), and part 2 should yield such inhibitory effects (mimicking Experiment 2 by Dijkstra et al., 1998). Such a result would show that it is not the uncertainty of the participants with respect to task requirements or stimulus list composition that determined their performance in the earlier studies. It would also exclude the possibility that participants had resorted to a mix of processing strategies in Experiment 1 because they did not know exactly what kind of stimuli to expect (thus testing one possible motivation for De Groot et al.'s mixed strategy view). Finally, if local stimulus list composition affects the bilingual's performance (e.g., by changing relative language activation or decision criteria), then an immediate effect on homo-

[1] Note that none of these accounts hinges on a null-effect. In fact, some recent studies have found slower RTs for interlingual homographs than for matched controls (Brysbaert, personal communication; Von Studnitz & Green, submitted).

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

graph recognition would be expected to occur as soon as the transition from part 1 to 2 is made (cf. De Groot et al., in press, n.5). Following the first presentation of a Dutch word, response times to homographs should slow down relative to matched exclusively English control words.

Method

Participants

Fifty-six (22 male and 34 female; mean age 24.0 years) native speakers of Dutch participated in the experiment. They had been studying English or Psychology at Nijmegen university for at least two years. All participants had eight or more years of experience with the English language (11.4 years on average), and were paid for their participation.

Materials

Test and control words were identical to those used in Dijkstra et al.'s (1998) Experiment 2. The homographic test words, selected from the English and Dutch CELEX data-base (Baayen, Piepenbrock, & van Rijn, 1993), had the same orthographic form in the two languages but no semantic similarity across languages (form-identical interlingual homographs). All words were three to five 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. The homographs were categorized according to their orthogonal combination of word form frequencies (high frequency, HF, or low frequency, LF) in English (E) and Dutch (D). This resulted in 4 categories of homographs (HFE-HFD, HFE-LFD, LFE-HFD, LFE-LFD) with 14 items in each category. For each of the critical items, an English control item was selected that matched the homograph in number of letters, syntactic category (noun or adjective), and frequency. A sample of the stimulus material is presented in the appendix. Frequency data and mean word length for the different categories of test and control items are presented in Table 1 (for some phonological characteristics of this stimulus set see Dijkstra, Grainger, & Van Heuven, 1999).

Non-words 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 non-words had the same length as the word stimuli. The non-words and the control items were not homophonic to Dutch words.

Table 1. *Mean word length in letters (WL) and log word frequency (per million) in English (ELF) and in Dutch (DLF) of homographs and matched English controls*

	Word type				
	Homographs			English controls	
Frequency categories	WL	ELF	DLF	WL	ELF
HFE-HFD	3.9	2.25	2.20	3.9	2.29
HFE-LFD	3.9	2.22	.75	3.9	2.20
LFE-HFD	3.9	1.08	1.98	3.9	1.08
LFE-LFD	3.8	1.07	.79	3.8	1.07

Furthermore, two sets of exclusively English and exclusively Dutch filler words were selected that were comparable to the test items in mean word length (3.9 letters) and word frequency (mean log frequency 1.63). The total stimulus set consisted of 448 items (56 homographs, 56 controls, 112 English fillers, 28 Dutch fillers, and 196 non-words).

Design

Unknown to the participants, the experiment was divided into two parts. The stimuli were assigned to the parts as follows. Both parts consisted of 28 blocks of eight stimuli including one homograph and one control item. The order of items of the different frequency categories was random across blocks, except for the last homograph item in part 1 and the first one in part 2, that were both members of the LFE-HFD category. In part 1 the remaining six item slots of each block were randomly filled with English fillers or non-words. Twenty-eight lists were constructed, of which 14 contained a homograph of a particular frequency category in a given block, and the other 14 contained their matched control items. For example, if a homograph appeared at a certain position in list 1, the matched control item was put at that position in list 2. Furthermore, 14 lists had 7 homographs and 7 controls of each frequency category in part 1, while the other 14 lists had these controls and homographs in part 2; for part 2, the list's composition was just the opposite (e.g., controls matched to homographs from part 1 were presented in part 2, and homographs matched to controls from part 1 were presented in part 2). To control for the effects of language shift between consecutive items (cf. Von Studnitz & Green, submitted), items that preceded test and control items were carefully selected. English fillers (requiring "yes" responses) and

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

English-like non-words (requiring a "no" response) were inserted in the stimulus list preceding 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.

The Dutch words did appear only in part 2. After the first 224 experimental items, a Dutch word was presented, followed by an English filler or non-word and then by a homograph or a control of the LFE-HFD category. From that list position onwards, Dutch words followed by a homograph or a control item appeared once in every block of eight stimuli. To keep the number of "yes" and "no" responses equal, 28 extra non-words were included in part 1 to match the 28 "no" responses to the Dutch words in part 2.

Procedure

Participants were tested individually in a sound-proof room. The experiment was run on an Apple Macintosh Power PC 7200/95. The 15-inch monitor was placed at a distance of approximately 65 cm from the participants. Stimuli were presented in black lowercase Courier (24 points) at the center of the screen on a white background.

Participants received the standard instruction for an English lexical decision task, except that they were explicitly notified (using examples) that (1) word forms could occur that existed in both English and Dutch (homographs) and that these words required a "yes" response; and (2) words only belonging to Dutch could also appear, requiring a "no" response. The "yes" responses were always given with the preferred hand. 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 three blocks. Short breaks were introduced between the blocks. There was no break between parts 1 and 2. The blocks were preceded by one block of 30 practice trials.[2] At the beginning of each block, the word "ATTENTION" was presented at the center of the screen for 1000 ms. Each trial started with the presentation of a fixation point at the centre 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. After the experiment, participants filled in a questionnaire to assess their familiarity with the English language. Experimental sessions lasted about 35 minutes.

[2] To ensure that the task requirements were absolutely clear to the participants, the practice set did not only contain examples of interlingual homographs, but also two Dutch items, which occurred early in the set. The majority of these items did indeed induce a "no" response.

Results

Data cleaning procedures were mainly based on error rates for participants and items. We first removed the data from nine participants whose test words had an error rate larger than 30 per cent. Next, the data from five items (three homograph test items and two control items) with error rates larger than 50 per cent were also excluded. Leaving out these items did not affect the matching between test and control items. For the remaining data points, response times outside the range of two standard deviations from a participant's and item's mean were considered outliers. Outliers and incorrect responses (7.6 per cent of the data) were excluded from the analyses.

Means for test and control words in different conditions are given in Table 2. Not included in this table are the RTs to English filler items, non-words, and Dutch items. For part 1 the RTs to English filler items and non-words were 573 ms and 630 ms, respectively. For part 2 the RTs to English filler items, non-words, and Dutch items were 592 ms, 629 ms, and 680 ms, respectively.

An analysis of variance with the factors Part (part 1/2), Word Status (homograph/control), and Frequency Category (HFE-HFD/HFE-LFD/LFE-HFD/LFE-LFD) yielded main effects for Part [$F_1(1,46)=17.98$, $p<.001$; $F_2(1,99)=44.94$, $p<.001$] and Frequency Category [$F_1(3,44)=66.67$, $p<.001$; $F_2(3,99)=15.42$, $p<.001$]. Mean RTs in part 1 (578 ms) were faster than in part 2 (602 ms). Word Status did not result in a main effect [$F_1(1,46)=3.08$, $p=.09$; $F_2<1$]. Mean latencies to test and control words were 591 ms and 586 ms, respectively. Critically, Part interacted significantly with Word Status [$F_1(1,46)=8.24$, $p<.01$; $F_2(1,99)=14.83$, $p<.001$]. Mean RTs to homographs were 575 ms in part 1 and 613 ms in part 2, and those for their matched controls were 581 ms and 592 ms, respectively. The interaction is graphically represented in Figure 1, together with the corresponding data of Dijkstra et al. (1998). Frequency Category only marginally interacted with Part [$F_1(3,44)=2.10$, $p=.11$; $F_2(3,99)=2.90$, $p<.05$] and did not interact with Word Status [$F_1(3,44)=1.20$, $p=.32$; $F_2<1$]. The interaction between Part, Word Status, and Frequency Category did not reach significance [$F_1(3,44)<1$; $F_2(3,99)=1.26$, $p=.29$].

To obtain more information about the nature of the interactions of Word Status and Frequency Category with Part, we performed separate analyses of

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

Table 2. *Mean latencies, Standard Errors and percentages of errors (in parentheses) for homographs and matched English controls in Parts 1 and 2*

	Word type			
	Homographs		Controls	
Frequency categories	Part 1	Part 2	Part 1	Part 2
HFE-HFD	552 (11.9, 3.9)	588 (11.0, 9.2)	562 (9.6, 6.5)	578 (8.2, 6.3)
HFE-LFD	536 (8.3, 2.8)	571 (9.4, 2.8)	543 (9.0, 2.8)	547 (8.1, 2.0)
LFE-HFD	602 (15.7, 17.4)	669 (17.8, 26.8)	605 (12.6, 24.7)	624 (11.3, 25.0)
LFE-LFD	610 (16.7, 24.0)	623 (14.8, 29.1)	613 (11.1, 15.1)	619 (9.6, 13.5)
Over-all	575 (10.9, 11.7)	613 (10.5, 17.0)	581 (9.0, 12.6)	592 (7.9, 11.2)

Note: Control items were allocated to the Dutch frequency category of the interlingual homograph they were matched to.

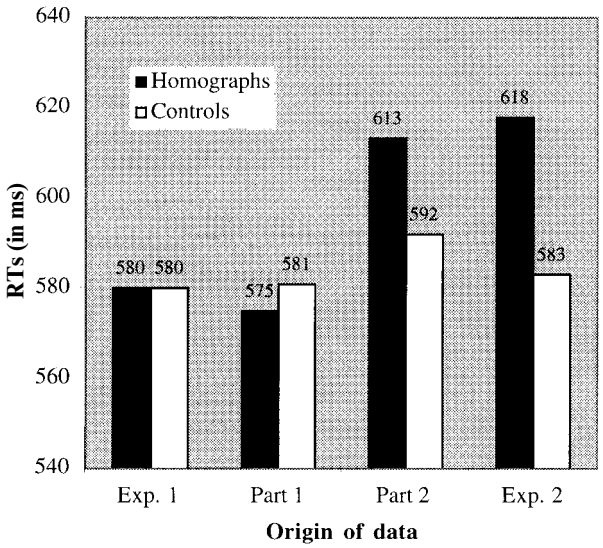


Figure 1. Response times to interlingual homographs and matched English control words in Part 1 and Part 2 of the present study (middle left and right), and in Experiment 1 (left) and 2 (right) by Dijkstra et al. (1998). Part 1 and Experiment 1 of Dijkstra et al. did not contain exclusively Dutch words, Part 2 and Experiment 2 did.

variance on latencies in each of the parts. An ANOVA on part 1 with the factors Word Status and Frequency Category yielded a main effect for Frequency Category only [$F_1(3,44)=47.34$, $p<.001$; $F_2(3,99)=12.74$, $p<.001$; for Word Status both $Fs<1$]. The interaction between the two factors was not significant [both $Fs<1$]. In contrast, a corresponding ANOVA on part 2 yielded significant main effects for both factors [Frequency Category: $F_1(3,44)=36.85$, $p<.001$; $F_2(3,99)=13.78$, $p<.001$; Word Status: $F_1(1,46)=12.15$, $p=.001$; $F_2(1,99)=5.23$, $p<.05$], but no significant interaction [$F_1(3,44)=2.20$, $p=.10$; $F_2(3,99)<1$].

Next, we examined the change in response times for homographs and control items across parts in separate ANOVAs. An analysis with the factors Part and Frequency Category on the latencies of the homographs revealed a main effect of Part, reflecting the slowing down of responses from 575 ms in part 1 to 613 ms in part 2 [$F_1(1,46)=19.15$, $p<.001$; $F_2(1,49)=39.46$, $p<.001$]. Frequency Category also led to a significant main effect [$F_1(3,44)=27.36$, $p<.001$; $F_2(3,49)=7.45$, $p<.001$] and showed no interaction with Part [$F_1(3,44)=1.43$, $p=.25$; $F_2(3,49)=2.58$, $p=.06$].

In contrast, the increase in response latencies to control items from part 1 to part 2 was only marginally significant (581 ms vs. 592 ms, respectively), as indicated by a significant main effect of Part by item only [$F_1(1,46)=3.61$, $p=.06$; $F_2(1,50)=6.75$, $p<.05$]. Frequency Category remained significant across parts [$F_1(3,44)=71.70$, $p<.001$; $F_2(3,50)=8.30$, $p<.001$] and there was no interaction with Part [both $Fs<1$].

English and Dutch frequency for interlingual homographs

For the homographs, the effects of the crossed factors English frequency (high/low) and Dutch frequency (high/low) across parts were analyzed in additional ANOVAs. In the overall analysis, Part yielded a significant main effect [$F_1(1,46)=19.15$, $p<.001$; $F_2(1,102)=47.62$, $p<.001$], as did English Frequency [$F_1(1,46)=71.86$, $p<.001$; $F_2(1,102)=28.62$, $p<.001$]. Dutch Frequency resulted in a significant effect by participants only [$F_1(1,46)=7.59$, $p<.01$; $F_2(1,98)=1.30$, $p=.26$]. While Part did not interact significantly with English Frequency [$F_1(1,46)<1$; $F_2(1,102)=2.40$, $p=.13$], it did interact with Dutch Frequency [$F_1(1,46)=4.04$, $p=.05$; $F_2(1,102)=4.79$,

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

Table 3. *Mean latencies and error rates for homographs (Hom) and controls (Con) at the end of Part 1 and the beginning of Part 2, and for the Dutch item at the beginning of Part 2*

Item position	Item category	Mean RT	Error rate
Last test item in Part 1	Hom LFE-HFD	581	17.4
	Con LFE-HFD	591	20.0
First Dutch item in Part 2	Dutch	822	34.8
First test item in Part 2	Hom LFE-HFD	663	30.4
	Con LFE-HFD	615	23.8

Note: Intermediary positions were filled with non-words and English fillers.

p<.05]. There was no interaction between English Frequency and Dutch Frequency [both Fs<1]. The triple interaction Part by English Frequency by Dutch Frequency only yielded a marginal effect in the participant analysis [$F_1(1,46)=3.48$, $p=.07$; $F_2(1,102)=1.73$, $p=.19$].

A separate analysis for Part 1 showed a significant main effect for English Frequency [$F_1(1,46)=36.84$, $p<.001$; $F_2(1,102)=23.05$, $p<.001$] but not for Dutch Frequency [both Fs<1], and no significant interaction between these factors [$F_1(1,46)=1.63$, $p=.21$; $F_2(1,102)=1.93$, $p=.17$]. Interestingly, the corresponding analysis for Part 2 yielded a significant main effect for English Frequency [$F_1(1,46)=42.25$, $p<.001$; $F_2(1,102)=22.77$, $p<.001$] and marginally for Dutch Frequency [$F_1(1,46)=8.91$, $p<.01$; $F_2(1,102)=3.16$, $p=.08$], while there was still no interaction [$F_1(1,46)=2.36$, $p=.13$; $F_2(1,102)<1$].

Transition from Part 1 to Part 2

The experimental design also allowed us to examine the transition from Part 1 to Part 2 at a more local level (see Design section). Table 3 contains the mean RTs and error rates for the last test items of Part 1 and the first of Part 2. RTs were slowed down significantly for the LFE-HFD homographs in part 2 relative to part 1 (82 ms, one-tailed $t(33)=1.79$, $p<.05$), but not for the matched English control items (24 ms, one-tailed $t(30)=.60$, $p>.10$). The error rates for these items showed a similar pattern.

Discussion

In the present lexical decision experiment, participants knew they would have to respond "no" to words having only a Dutch reading, but such items were presented only in the second part of the experiment. No significant RT differences were found between the interlingual homographs and matched English control items in the first part of the experiment.[3] However, strong inhibitory effects for interlingual homographs relative to control words were observed in the second part. Examination of the transition from part 1 to part 2 showed that, as soon as Dutch items started to come in, the RTs to interlingual homographs were considerably slowed down compared to control words.

The absence of an inhibition effect for interlingual homographs in part 1 of the experiment parallels the results obtained by Dijkstra et al. (1998, Experiment 1) and by De Groot et al. (in press, Experiment 2a). However, in contrast to those earlier experiments, the instruction of the present experiment clearly indicated that Dutch words requiring a "no" response might appear. Nevertheless, no inhibition for interlingual homographs was obtained in part 1, showing that the participants' performance was not affected by whether the instruction did or did not mention the possibility that Dutch words would be presented.

These results have clear consequences for the theoretical views described in the introduction. First of all, they clarify the relative importance of a number of factors that have been proposed by Grosjean (1998, in press) as potentially relevant for the participants' language mode and hence for their performance in this experiment. In our experiment, explicitness of instruction did not affect the observed pattern of results. At first sight, that would seem to be incompatible with Grosjean's (1998, in press) view as exposed in the Introduction. However, in line with Grosjean (1998, in press) one could assume that the participants in our experiment were already in an intermediary bilingual mode when they entered the experimentation room, because, as in Experiment 1 by Dijkstra et al. (1998), they knew that the experiment was going to be bilingual in nature. It then follows from the similar results in the two studies that making explicit in our instruction *how* both languages would be relevant for responding, did not move the participants' language mode any further towards the bilingual end of the continuum. In contrast, the appearance of the first item belonging exclusively to the other language (Dutch) in part 2 did immediately affect the participants' behavior, inducing inhibition effects like in Experiment 2 by Dijkstra et al. (1998). Thus, within this type of

[3] These null-effects in Part 1 arose despite the presence of two Dutch items early in the preceding practice set (see n.2). Apparently, the effects of these Dutch items had decayed by the time the "real" experiment began or had become too small to be detected.

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

account, our experimental results can be interpreted as evidence that the participants' implicit general expectations (i.e. that two languages will be relevant) and language intermixing are more important determinants of the language mode than specific expectancies introduced by the explicit instruction.

Alternatively, it could be assumed that the participant's expectancies, whether induced by the specific instruction or not, did not affect the bilingual word recognition process at all, at least not after the participant had generated a task schema (Green, 1998) for the task at hand. This view would be much more in line with Dijkstra et al.'s (1998) account. According to this view, the presence of inhibitory effects in Part 2 of the experiment only was a consequence of a more strongly activated Dutch lexicon due to the presence of exclusively Dutch items.[4]

However, recent evidence by Dijkstra et al. (1999) suggests it may not even be necessary to assume a difference in the activation of the Dutch lexicon in Part 1 and Part 2 of our experiment to account for the overall data pattern. In an experimental situation with a stimulus list composition directly comparable to that of Part 1 of our experiment, these authors observed facilitatory effects of orthographic identity between Dutch and English readings of interlingual homographs, but inhibitory effects of cross-linguistic phonological similarity. Interlingual homographs that were similar across languages only in terms of their orthography (e.g., STAGE, pronounced differently in Dutch and English) were recognized faster than matched control items, but no RT differences were found for homographs that were similar both in terms of orthography and phonology (e.g., STEP). In line with these findings, Dijkstra et al. proposed that the null-result for homographs in Dijkstra et al.'s (1998) Experiment 1 originated from opposite effects of cross-linguistic orthographic and phonological overlap between the two readings of the interlingual homographs.

Applying this recent explanation to the null-effects in Part 1 of our study, the presence of homograph effects in Part 2 of our study and their absence in Part 1 can then be considered as the consequence of a context-sensitive adaptation of the participants' decision behavior to the local stimulus list composition, rather than as a difference in the relative activation of the Dutch lexicon in the two parts. Lexical candidates from both lexicons would always be completely activated on the basis of word input. Adaptation of decision criteria would be a consequence of the introduction of Dutch items and it would affect especially the latencies of the homographs because they have a Dutch reading as well.

However, a reconciliation of the two positions is also possible, by assuming that the amount of activation required to identify a particular item (its recognition threshold) is strategy and context-sensitive. In Part 2, more activation of an English item would be required for identification in order to avoid mistakenly identifying a Dutch word as a target item. This approach could easily be implemented as a variant of the BIA model (Dijkstra & Van Heuven, 1998).

Strategic aspects of processing also played a role in the third position described in the Introduction. De Groot et al. (in press) would explain the null-results in the first part of the experiment as a cancellation of facilitation and inhibition effects associated with "language neutral" and "language specific" processing modes. The present results indicate that, irrespective of the processing mode that the participants adopted either consciously or unconsciously, this mode did not arise because they were uncertain about the task demands. Furthermore, the results go against a different hypothesis proposed by De Groot et al., namely that the stimulus set (list composition) did not play the important role that Dijkstra et al. (1998) assigned to it. Our participants clearly changed to a different processing mode as soon as the first Dutch test item had been presented. Thus, one has to conclude that it is the change in stimulus list composition (degree of language intermixing) between Part 1 and Part 2 that led to the change in performance by the participants.

This conclusion finds further support in the finding that Dutch frequency played a statistically significant role only in the second part of the experiment, which also shows that the introduction of Dutch items in the second part indeed underlies the change in performance by the participants. The theoretical views by De Groot et al. (in press), Dijkstra et al. (1998), and Grosjean (in press) all have their own mechanism to account for this result. For instance, in line with De Groot et al. (in press) the Dutch items may somehow have forced the participants into a more language specific processing mode. Participants may have become more strict in paying attention to the language label of each encountered word. Alternatively, in line with Dijkstra et al. (1998) and Grosjean (in press), the presence of Dutch items in Part 2 of the experiment may have activated the Dutch lexicon to a larger extent than in Part 1.

Whichever interpretation is favored, it should also

[4] The relatively large Standard Errors and error rates for the interlingual homographs relative to the controls (see Table 2) can be seen as additional evidence of increased competition between the two readings of the homographs in Part 2. Convergent distributional analyses of homograph and control latencies are presented in Dijkstra, Timmermans, & Schriefers (in press).

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

take into account our finding that the transition from Part 1 to Part 2 was rather abrupt in its effect on RTs to interlingual homographs. Immediately after the first Dutch item came in, RTs to interlingual homographs but not to the controls became considerably slower (see Table 3). Indeed, response times to controls were affected only marginally in our experiment by the presence or absence of Dutch items (also compare the control conditions in the three experiments by Dijkstra et al., 1998). Furthermore, a comparison of Table 2 (results for Part 1 and Part 2 as a whole) and Table 3 (results for the last trials of Part 1 and the first trials of Part 2) suggests that the homograph inhibition effect reached its maximal value immediately after presentation of the first Dutch item (LFE-HFD homograph items became 67 ms slower over-all, and 82 ms at the transition from part 1 to part 2). Future research should test this hypothesis that the presentation of just one Dutch word already had a maximal inhibitory effect on the participants' responses to interlingual homographs and that there was no further increase of inhibition after multiple presentations of Dutch items.

Related issues to be explored are how long-lasting (i.e. for how many subsequent trials) the effects of a presented non-target language item on homograph processing are and whether they can be strategically modulated by the participants. A recent study by Von Studnitz & Green (submitted) suggests that the observed increased interference for homograph after the introduction of other-language items can indeed be strategically reduced over time.

To conclude, for language-specific lexical decision we have shown that explicitly informing participants about the presence of non-target language items in the stimulus list did not change their performance relative to a situation with implicit instruction (Dijkstra et al., 1998, Experiment 1). In contrast, encountering non-target language items in the stimulus list had immediate and severe effects on the participants' performance. These results are in line with language non-selective access models of bilingual word recognition that attribute a strong role to bottom-up factors, like the BIA model (Dijkstra & Van Heuven, 1998; Dijkstra et al. 1998; Van Heuven, Dijkstra, & Grainger, 1998). In addition, the results indicate that participants try to optimize their performance in a particular experimental situation on the basis of their recent experience with list items, adapting their bilingual word recognition process depending on the characteristics (e.g., language membership) of the encountered items. Interpreted in this way, the present findings are directly compatible with the common view on monolingual word recognition, namely that word recognition is an automatic process in which multiple lexical candidates are activated in parallel, a process that can be affected by shifts in parameter settings on the basis of recent experience (Lima & Huntsman, 1997; Grainger & Jacobs, 1996).

References

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

De Groot, A.M.B., Delmaar, P., & Lupker, S.J. (in press). The processing of interlexical homographs in a bilingual and a monolingual task: support for nonselective access to bilingual memory. *Quarterly Journal of Experimental Psychology*.

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

Dijkstra, A., Timmermans, M., & Schriefers, H. (in press). On being blinded by your other language: effects of task demands on interlingual homograph recognition, *Journal of Memory and Language*.

Dijkstra, A., & Van Heuven, W.J.B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (eds.), *Localist connectionist approaches to human cognition*, pp. 189–225. Hillsdale, NJ: Lawrence Erlbaum.

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

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition* 15, 305–313.

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon*, pp. 11–26. Amsterdam: John Benjamins.

Grainger, J., & Jacobs, A. (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 and Cognition* 1, 67–81.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development* 6, 467–477.

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. 225–254. Mahwah, NJ: Lawrence Erlbaum.

Grosjean, F. (1997b). The bilingual individual. *Interpreting: International Journal of Research and Practice in Interpreting* 2, 163–187.

Grosjean, F. (1998). Studying bilinguals: methodological

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press

and conceptual issues. *Bilingualism: Language and Cognition* 1, 131–149.

Grosjean, F. (in press). The bilingual's language modes. In J. L. Nicol (ed.), *One mind, two languages: bilingual language processing*. Oxford: Blackwell.

Lima, S. D., & Huntsman, L. A. (1997). Sequential dependencies in the lexical decision task. *Psychological Research* 60, 264–269.

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

Von Studnitz, R. & Green, D. (submitted). Interlingual homograph interference in German–English bilinguals: its modulation and locus of control.

Appendix

Sample of Stimulus Material

The following table contains examples of the experimental items. 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).

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

Received July 2, 1999 Revision accepted September 9, 1999

https://doi.org/10.1017/S1366728900000146 Published online by Cambridge University Press