Bilingualism: Language and Cognition 9 (3), 2006, 281–297 © 2006 Cambridge University Press doi:10.1017/S136672890600263X

Naming interlingual homographs: Variable competition and the role of the decision system

ERICA SMITS
Center for Psycholinguistics, Antwerp University
HEIKE MARTENSEN
Center for Psycholinguistics, Antwerp University
TON DIJKSTRA
Nijmegen Institute for Cognition and Information, Radboud University, Nijmegen
DOMINIEK SANDRA
Center for Psycholinguistics, Antwerp University

To investigate decision level processes involved in bilingual word recognition tasks, Dutch–English participants had to name Dutch–English homographs in English. In a stimulus list containing items from both languages, interlingual homographs yielded longer naming latencies, more Dutch responses, and more other errors in both response languages if they had a high-frequency Dutch reading. Dutch naming latencies were slower than or equally slow as English naming latencies. In a stimulus list containing only English words and homographs, there was no homograph effect in naming latencies, although homographs did elicit more errors than control words. The results are interpreted as the consequence of list-induced variability in the competition between lexical items of the two languages involved. In addition, two additional decision processes have to be assumed: a language check, and a response deadline for non-target-language responses.

A few years ago, the metal band Megadeth issued a CD with the title *Rust in Peace*. For the English reader, the switch from REST to RUST involves a significant change in meaning. For Dutch–English bilinguals, the word play may be even more intriguing, because RUST is the Dutch translation of REST. Bilinguals are often confronted with INTERLINGUAL HOMOGRAPHS like RUST, which have an identical orthographic form but a different meaning in each of their languages. Such homographs have turned out to be very informative on the question whether lexical access is language-specific or language-independent: if access is language-independent, the recognition of an interlingual homograph in one language should either suffer interference from or be facilitated by its other language reading, because both will be accessed simultaneously. Previous research has indeed found such interference and/or facilitation effects, contributing to the broad consensus that lexical access is language-independent (e.g. Altenberg and Cairns, 1983; Nas, 1983; Bijeljac-Babic, Biardeau and Grainger, 1997; Dijkstra and Van Heuven, 1998; Dijkstra, Van Jaarsveld and Ten Brinke, 1998; Van Heuven, Dijkstra and Grainger, 1998; Dijkstra, Grainger and Van Heuven, 1999; De Groot, Delmaar and Lupker, 2000; Dijkstra, De Bruijn, Schriefers and Ten Brinke, 2000; Dijkstra, Timmermans and Schriefers, 2000; Dijkstra and Van Heuven, 2002; Von Studnitz and Green, 2002; Lemhöfer and Dijkstra, 2004).

Although the homograph effect has been replicated in several studies, it has also been found to vary according to stimulus list composition and task demands. Different studies report a sizeable effect of homograph interference when the stimulus list contains non-homographic words from the participants' two languages, but a considerable reduction of the effect when the list contains only L2 words (e.g. Dijkstra and Van Heuven, 1998; De Groot et al., 2000; Dijkstra, De Bruijn et al., 2000; Von Studnitz and Green, 2002). Dijkstra et al. (1998) observed such a reduction in two language-specific English lexical decision experiments with Dutch–English participants. The stimulus lists of their experiments contained Dutch–English homographs, English control words, and English filler words requiring a YES-response, and nonwords, which required a NO-response. The second experiment also contained Dutch filler words, which required a NO-response. Whereas there was a clear inhibitory effect in the mixed experiment with Dutch filler words, there was no obvious homograph interference in the all-English experiment. Apparently, removing Dutch filler words from the stimulus list caused a reduction in the homograph interference effect. In two similar experiments, De Groot et al. (2000) observed the same pattern.

There are two possible explanations for the stimulus list composition effect in lexical decision: in terms of (i) processing strategies and (ii) variable competition. The goal of the current paper is to test these possible mechanisms, and to establish which mechanisms allow for responding to the target-language reading of an interlingual homograph. In the remainder of this introduction,

Address for correspondence
Erica Smits, CPL Antwerp University, Prinsstraat 13, 2000 Antwerpen, Belgium
E-mail: *erica.smits@ua.ac.be*

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

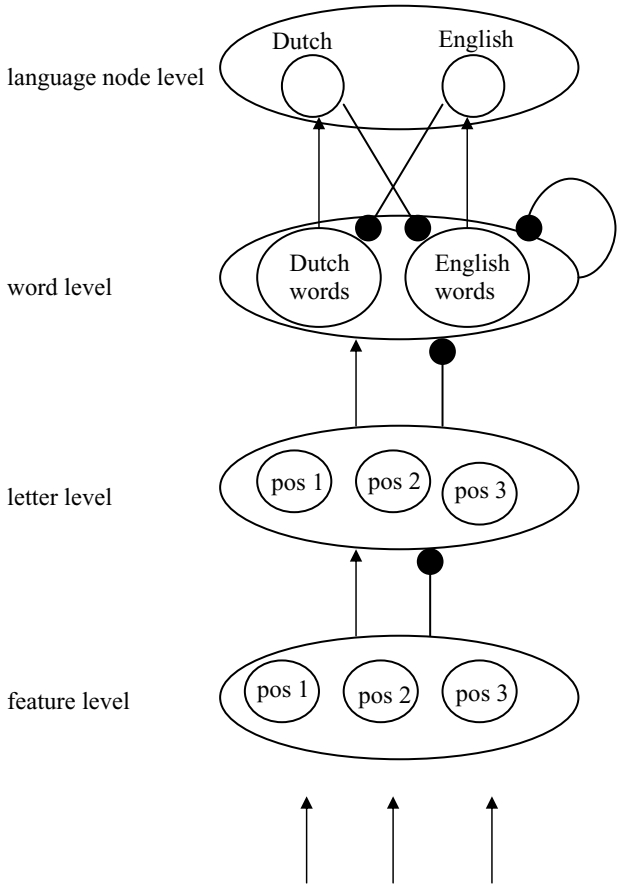


Figure 1. The bilingual interactive activation model (BIA).

we will give a short description of the two hypotheses and outline the current experiments.

De Groot et al. (2000) suggested that the reduction of the homograph effect may have been caused by DIFFERENT PROCESSING STRATEGIES amongst the participants. In bilingual lexical decision, participants are asked to judge the word or nonword status of a string of letters, either in one of their languages (i.e. language-specific lexical decision), or irrespective of language (i.e. language-independent lexical decision). In the studies by Dijkstra et al. (1998) and De Groot et al. (2002), participants were asked to judge whether the target items were English words or not. This entailed responding NO to Dutch words in the mixed stimulus list. In the all-English list, however, participants never had to give NO-responses to purely Dutch words, as there were no such items in the list. This made it possible for the participants to ignore instructions, and to treat the task as language-independent lexical decision, because performance would not be affected negatively by responding YES to the Dutch readings of the interlingual homographs. Such a strategy would cause facilitation for some participants, which – together with an inhibitory effect for other participants – might lead to an overall null effect.

A second account of the stimulus list composition effect is the assumption that the degree of competition between the two readings of an interlingual homograph may vary according to task demands. Such variation may be the result of changes in the lexical activation levels of bilinguals' languages (i.e. relative language activation), or changes in the decision criteria that are applied to the lexical output (i.e. dynamic decision criteria). According to one theoretical view on relative language activation, the activation that is caused by bottom-up processes in either of the bilinguals' languages may be reduced by top-down factors such as the participants' expectancies towards the task demands, their expectancies towards the language of the targets based on the items encountered so far, or feedback activation from previous responses (i.e. pronouncing an item in a particular language might increase the activation level of the lexical representations for this language). In the following, we will give a short description of the Bilingual Interactive Activation (BIA) model as a representative of models of bilingual visual word recognition that include relative language activation. In the original Bilingual Interactive Activation model (Dijkstra and Van Heuven, 1998, see Figure 1), relative language activation

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

was implemented by means of language nodes. All word representations of one language were linked to a language node that inhibited the word representations connected to the other language node. Language nodes were activated bottom-up through the activation of their associated word representations, but language node activation could also be suppressed top-down on the basis of task demands. This model and other models that assume relative language activation predict a decrease in the activation of the Dutch language node when a Dutch–English bilingual performs a lexical decision task in English and the stimulus list contains only English words and nonwords. This results in a reduction of the homograph effect.

However, changes in the degree of competition between two homograph readings may also be situated at the decision level. For example, Dijkstra and Van Heuven (2002) suggested that the activation levels of bilinguals' languages do not vary with stimulus list composition, but that the degree of competition is determined by DYNAMIC DECISION CRITERIA. 'Dynamic' means that the decision criteria may change during the task itself, on the basis of the participants' experiences and their expectancies towards the remainder of the experiment. This account was incorporated in a later version of the BIA-model, the BIA+ model (see Figure 2). With respect to the stimulus list composition effect in Dijkstra and Van Heuven (2002), Dijkstra et al. (2002) suggested more specifically that the link (response binding) between the Dutch lexical representations and their associated response (NO) is weaker in an all-English stimulus list than in a mixed list, because it is activated less often by the input, and participants do not expect to respond to non-target-language words. Consequently, the NO-response that is normally associated with the non-target reading of a homograph does not gather enough activation to interfere with the target-language response (YES), and the homograph effect is reduced.

In the current study, we set out to distinguish the accounts in terms of processing strategies and variable competition. In lexical decision, it is difficult to do so, because the reported data patterns support both hypotheses. To exclude differences in processing strategies as an explanation for the stimulus list effect, we can attempt to replicate the effect in a task that does not allow for a language-independent processing strategy in an all-English stimulus list. We predict that the stimulus list composition effect should disappear in such a task, if it were merely the consequence of some participants treating lexical decision as a language-independent task in a one-language list. If the effect is replicated, however, this would indicate that the degree of between-language competition varies according to the language context and/or the participants' expectancies.

In the present study, Dutch–English participants were asked to name Dutch–English homographs in English, either in an all-English or a mixed English–Dutch stimulus list. In the mixed list, the participants were instructed to respond in Dutch to Dutch words and in English to interlingual homographs. Word naming is a language-specific task by definition – the correct articulatory response has to be selected, simply because one cannot articulate two different responses simultaneously. Therefore, a strong reduction of the homograph interference effect in the all-English list as opposed to the mixed list cannot be interpreted as the result of interference and facilitation effects canceling each other out, as was the case for lexical decision.

The word naming task provides us with the additional advantage that the nature of the error-responses can be investigated. Above, it is argued that an effect of stimulus list composition in an English word naming task is probably caused by variable competition, either at the lexical level or the decision level. By analyzing the probability and speed of non-target-language responses, we can discern whether additional mechanisms have to be assumed. The relative language activation account and the decision criteria account both predict that the degree of competition between two homograph readings mainly depends on their occurrence frequencies. Errors should occur more often when the non-target-language reading is more frequent than the target-language reading, and these error responses should be faster when the non-target-language reading has a high frequency than when it has a low frequency. However, the frequency properties of the two homograph readings might not be the only determinants of the response outcome. When task demands assign priority to one language (i.e. when participants have to give L2 responses to interlingual homographs, then L2 is given priority), wrong language responses to interlingual homographs might be delayed to allow for responding in L2, despite an overall lower activation level in L2. Our data will indeed indicate that, at the decision stage, one has to assume the existence of an additional waiting mechanism for responses in the inappropriate language. In order to test to what extent the occurrence frequencies of the homograph readings determine the reaction times of correct and incorrect responses, the test items in this study were divided into four frequency categories, on the basis of their relative frequencies: high frequency in both English and Dutch (HFE–HFD), high frequency in English and low frequency in Dutch (HFE–LFD), low frequency in English and high frequency in Dutch (LFE–HFD), and low frequency in both languages (LFE–LFD). The homographs were heterophonic, i.e. their pronunciation differed across languages, e.g. ROOM that is pronounced /ro.m/ in Dutch, and they were presented in a stimulus list containing only English words (pure list) or in a mixed list, in which one third of the items were Dutch filler words.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

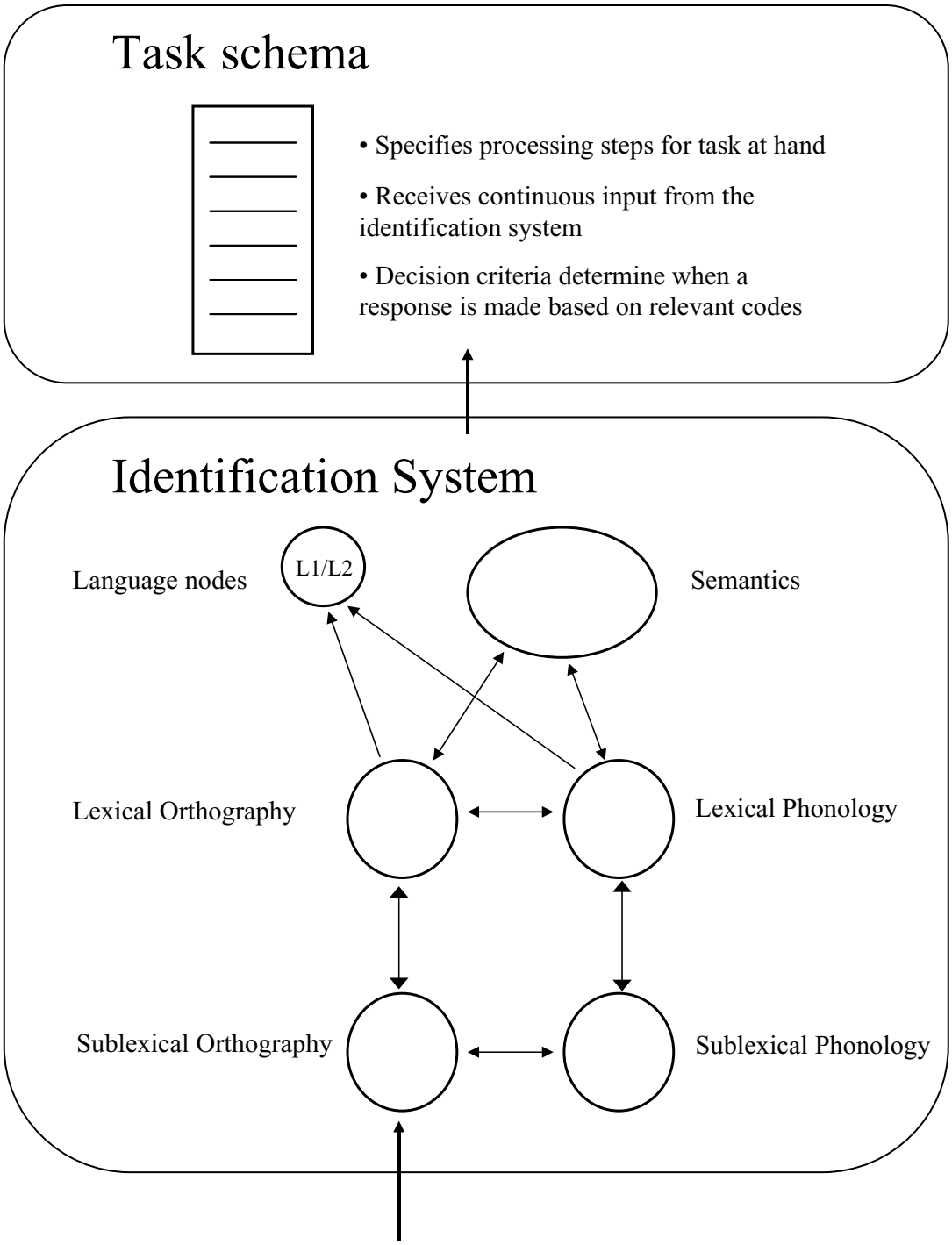


Figure 2. The extended bilingual interactive activation model (BIA+).

Method

Participants

The mixed stimulus list condition involved 34 participants, the pure list condition 24 participants. All participants were studying English and Dutch at the time of the experiment at the University of Antwerp in the Dutch speaking part of Belgium. They were on average 18 years old and all of them had been receiving English instruction in secondary school as of the age of 13. Participation was voluntary.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Materials

A list of 3–6-letter monosyllabic Dutch and English words was extracted from the CELEX-database (Baayen, Piepenbrock and Van Rijn, 1993). Only nouns, adjectives, and verbs with a printed-lemma frequency of at least one occurrence per million were included. From these two lists, all Dutch–English non-cognate interlingual homographs were extracted. We selected 32 items with little phonological overlap (heterophonic homographs; e.g. ROOM, which is pronounced /ro:m/ and means 'cream' in Dutch) between English and Dutch. We constructed four categories of eight interlingual homographs each by orthogonally varying the relative log frequencies of the Dutch and English item readings: high frequency in both languages (HFE–HFD), high frequency in English and low frequency in Dutch (HFE–LFD), high frequency in Dutch and low frequency in English (LFE–HFD), and low frequency in both languages (LFE–LFD). T-tests showed no significant differences in mean log frequency in one language between the two high frequency categories in English or Dutch (**HFE**–HFD vs. **HFE**–LFD; HFE–**HFD** vs. LFE–**HFD,** both p's > .44) or the two low frequency categories in English or Dutch (**LFE**–HFD vs. **LFE**–LFD; HFE–**LFD** vs. LFE–**LFD,** both p's > .18), and significant differences between each English high frequency category and its corresponding low frequency category (**HFE**–LFD vs. **LFE**–LFD; **HFE**–HFD vs. **LFE**–HFD, both p's < .005) and between each Dutch high frequency category and its corresponding low frequency category (LFE–**LFD** vs. LFE–**HFD**; HFE–**LFD** vs. HFE–**HFD,** both p's < .001). We were forced to include one bisyllabic item (ANGEL) in the list to obtain satisfactory frequency matching across categories. Each interlingual homograph was matched to a purely English control item extracted from the same CELEX database. The control items were matched pairwise to the homographs with respect to number of letters, onset phoneme(s), and English log frequency. T-tests were performed to make sure that log word frequency and number of letters were properly matched between interlingual homographs and English controls. No significant differences were observed (all p's > .21). An ANOVA was conducted to check for differences in number of letters between the four relative frequency categories. The number of letters differed significantly as a function of Dutch frequency (LFD or HFD) [$F(1,64) = 8.46$, $p < .01$]. Items in the LFD-categories (LFE–LFD and HFE–LFD) had fewer letters than items in the HFD-categories (LFE–HFD and HFE–HFD): 3.56 versus 4.06, respectively. Note that this difference in number of letters could affect the main Dutch frequency effect for the homograph items, but that it cannot affect the homograph effect, because the number of letters was matched for controls and homographs within each frequency category. The item selection was severely constrained and matching in terms of word length between the frequency categories could not be improved.

Another set of 32 homographs with considerable phonological overlap between Dutch and English was selected, e.g. PET, which is pronounced /pEt/ in Dutch and means "cap", and matched to monolingual English control items, but these could not be divided into four frequency categories. Table 1 shows the mean log frequencies and number of letters for all items.

We constructed eight stimulus lists in both the pure and mixed list condition with fixed positions for the test items across versions. All versions had the same pseudo-randomized order of words (fillers) in between test items, with the restriction that there were never more than three items of the same type in a row. Two lists contained a homograph of a particular frequency category at a certain position, whereas another two lists had the matched control item at that position. All four resulting versions were divided into four blocks, which were presented in two different orders: either 1–2–3–4 or 3–4–1–2. All lists consisted of 400 items each. The mixed stimulus lists contained all heterophonic homographs (n = 32), their English controls (n = 32), 139 purely English fillers, 133 purely Dutch fillers, and the 32 interlingual homophonic homographs and their matched monolingual English words. The homographs and all English words made up two thirds of the items in the list and Dutch words made up one third of the items. The pure list contained all homographs, their English controls, and 272 English filler items, such that all words in the list were English.

Procedure

Participants received all instructions on screen in English, asking them to read aloud each presented word as quickly as possible and with as few errors as possible. The instruction language was English to stress the primacy of the English language in the task at hand. For both list conditions, the participants were informed that next to English words, Dutch words and interlingual homographs might be presented. They were asked to respond in English to every item that had an English reading (i.e. to purely English words and homographs) and in Dutch otherwise (i.e. to purely Dutch words). After instructions, the participants could practice on a list of 20 items, which contained three homographs. Six English words were replaced by a Dutch word for the practice list of the mixed condition.

At the beginning of each trial a fixation sign (+) appeared at the center of the screen for 300 ms. After another 300 ms, the test item appeared and remained on the screen until the participant's response triggered the voice key or for 1500 ms if the participant did

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Table 1. *Mean Dutch and English log word frequency, number of letters, and corresponding standard deviations (sd) for all test items and their controls*

Heterophonic homographs										
	Homographs						Controls			
Frequency category	Number of letters	sd	English log frequency	sd	Dutch log frequency	sd	Number of letters	sd	English log frequency	sd
HFE–HFD	3.88	0.64	2.01	0.54	1.97	0.40	3.88	0.64	2.03	0.57
HFE–LFD	3.75	0.71	2.13	0.55	0.77	0.28	3.63	0.74	2.11	0.51
LFE–HFD	4.25	1.04	1.31	0.22	2.19	0.60	4.50	0.93	1.33	0.23
LFE–LFD	3.38	0.74	1.15	0.22	0.92	0.29	3.63	0.52	1.14	0.25
Homophonic homographs										
	Homographs						Controls			
	Number of letters	sd	English log frequency	sd	Dutch log frequency	sd	Number of letters	sd	English log frequency	sd
	3.78	0.60	1.65	0.61	1.31	0.80	3.94	0.61	1.67	0.65

not respond. In the latter case the participant's attention was drawn to this fact by a beep. At 1000 ms after the participant's response or after the beep a new trial was initiated. After each block of 100 items, there was a break. By pushing a button, the participant could go on to the next block. Each block began with two dummy items. Participants saw each item in the list only once.

Participants were tested individually in a soundproof room. Presentation of the visual stimuli and recording of RTs was programmed in PsyScope (Cohen, MacWhinney, Flatt and Provost, 1993) and conducted on a Power Macintosh 8200/120 computer. The monitor was placed at a distance of approximately 40 cm from the participants. Stimuli appeared in black lower case Courier (24 points) at the center of the screen on a white background. A voice-key registered the latencies. The participants' responses were recorded with a Shure BG 3.1 microphone. They were recorded on audiocassette and scored online. In cases of doubt the tapes were consulted. Experimental sessions for the mixed list took about 45 minutes. For the pure list they lasted 30 minutes.

Results

Responses to homographs were categorized as English responses, Dutch responses and other errors. The 'other errors' category included responses in which the English and the Dutch pronunciations were mixed, stuttering, and non-existent pronunciations.

Data cleaning

All items except the heterophonic homographs and their English controls were removed before determining outliers. The responses to the homograph ARTS and its control word ASH were removed, because ARTS was considered to be a homophonic homograph in the study by Dijkstra et al. (1998). Then, the reaction times (RTs) to erroneous responses were removed. Errors were Dutch responses to homographs, mixed responses, non-existent pronunciations, hesistations, repetitions and so on. Participants never responded in other languages than Dutch or English. The mean RTs per item and per participant were calculated. All responses that were more than 2.56 standard deviations (SDs) longer or shorter than their item and participant mean were removed from the analyses.[1] All data points that were shorter than 200 ms were removed. If a homograph RT was removed, the RT to its matched control word in the same participant was removed as well, and vice versa. After removal of the outliers, the overall mean RT and error rate were calculated. In both the pure and the mixed list, the mean RT to the item PAL was more than 3 SDs removed from the overall mean RT. Therefore, all responses to PAL and its control PEAR were removed from the analyses.

We first conducted analyses on the reaction times (RTs) to heterophonic homographs,[2] in which a distinction was

[1] Analyses with other standard deviation cut-off points (2 or 3 standard deviations) revealed the same results pattern.

[2] We do not report any analyses on homophonic homographs, because we could not always distinguish between English and Dutch responses

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Table 2. *Mean English and Dutch reaction times (rt) and their standard deviations (sd) to heterophonic homographs in four relative frequency categories in both list conditions*

		English RTs					Dutch RTs				
		Homographs		Controls			Homographs		Dutch control words		
	Frequency	RT	sd	RT	sd	Effect	RT	sd	Frequency	RT	sd
Mixed	HFE–HFD	639	175	567	97	72	606	123	HFD	549	104
list	HFE–LFD	564	110	565	104	–1	576	150	HFD	563	126
	LFE–HFD	645	170	571	105	74	623	147	LFD	570	157
	LFE–LFD	554	109	550	101	4	613	172	LFD	576	137
Pure	HFE–HFD	545	96	545	79	0					
list	HFE–LFD	516	91	525	107	–9					
	LFE–HFD	544	114	531	89	13					
	LFE–LFD	496	76	503	85	–7					

made between the RTs of English responses and the RTs of Dutch responses. Next, we conducted analyses of the error rates for all errors combined and also for the Dutch responses separately.

Reaction times

First, an overall analysis was performed on the RTs of English responses to heterophonic homographs and their English control words in both list conditions. Second, separate analyses were done for the pure and mixed lists. For the mixed list, we analyzed the RTs of Dutch responses to homographs and Dutch control words. For an overview of RTs, see Table 2.

English RTs

An analysis of variance including the within-participant and within-item factor Word Status (homograph or control), the within-participant and between-item factors English Frequency (high or low) and Dutch Frequency (high or low), and the between-participant and within-item factor List Condition (pure and mixed) was conducted on the RTs of English responses to heterophonic homographs and their controls. Word Status had a significant effect: homographs elicited slower responses than their controls [$F1(1,55) = 11.49$, $p < .005$; $F2(1,26) = 12.74$, $p < .005$]. Responses were slower in the mixed list than in the pure list [$F1(1,55) = 9.61$, $p < .005$; $F2\,(1,26) = 101.24$, $p < .001$]. Responses were slower in the categories with high English frequencies than in categories with low English frequencies, but this was only significant in the participant analysis [$F1(1,55) = 5.93$, $p < .05$; $F2 < 1$]. Responses to items with a high Dutch frequency were slower than to items with a low Dutch frequency [$F1(1,55) = 56.44$, $p < .001$; $F2(1,26) = 9.41$, $p < .01$]. There was a significant interaction between Word Status and List Condition, such that the homograph effect was larger in the mixed list than in the pure list [$F1(1,55) = 11.68$, $p < .005$; $F2(1,26) = 26.07$, $p < .001$]. There was also a significant interaction between Word Status and Dutch frequency, such that the homograph effect was larger for homographs with a high Dutch frequency [$F1(1,55) = 19.00$, $p < .001$; $F2(1,26) = 15.75$, $p < .005$]. The interaction between Word Status and English frequency was not significant [$F1 < 1$; $F2 < 1$]. There was a significant three way interaction between Word Status, Dutch Frequency, and List Condition, such that the Word Status effect depended on Dutch Frequency in the mixed list, but not in the pure list [$F1(1,55) = 5.72$, $p < .05$; $F2(1,26) = 13.85$, $p < .005$]. No other interactions were significant.

A by-items analysis with the covariate number of letters was conducted on the same data, because the number of letters could not be matched across the four frequency categories. In this analysis, Word Status was treated as a between-items factor, to allow for the inclusion of the numbers of letters of both homographs and control words. The analysis showed that the main effect of English Frequency was not significant with the number of letters as a covariate [$F2 < 1$]. All other effects remained significant: Word Status [$F2(1,51) = 5.88$, $p < .05$], List Condition [$F2(1,51) = 4.41$, $p < .05$], Dutch Frequency [$F2(1,51) = 8.37$, $p < .01$], the interaction between Word Status and List Condition [$F2(1,51) = 18.69$, $p < .005$], the interaction between Word Status and Dutch Frequency

to these items. An analysis of the obvious English RTs to heterophonic and homophonic homographs showed that there was no significant difference in the size of the homograph effect between the two types of homographs [$F1 < 1$; $F2 < 1$].

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

[F2(1,51) = 7.05, p < .05], and the threefold interaction between Word Status, List Condition, and Dutch Frequency [F2(1,51) = 9.93, p < .005].

A separate analysis of the PURE LIST showed that homographs did not lead to significantly slower responses than controls [F1 < 1; F2 < 1]. English Frequency showed a significant effect in the participant analysis, but not in the item analysis: High frequency homographs and controls elicited somewhat slower responses than low frequency items [F1(1,23) = 8.34, p < .01; F2 < 1]. Homographs and controls in the high Dutch frequency categories had longer RTs than items belonging to the low Dutch frequency categories [F1(1,23) = 29.89, p < .001; F2(1,26) = 3.73, p = .06]. There was a significant interaction between Word Status and Dutch Frequency in the participant analysis: Homographs elicited faster responses than control words when the Dutch frequency was low, but slower responses when it was high [F1(1,23) = 5.39, p < .05; F2(1,26) = 1.62, p = .21]. The interaction between Word Status and English Frequency was not significant [F1 < 1; F2(1,26) = 1.06, p = .31], nor was the interaction between English and Dutch Frequency [F1 < 1; F2 < 1]. In a by-item analysis including the covariate number of letters, the main effect of Dutch Frequency was no longer significant [F2(1,59) = 2.57, p = .12]. The interaction between Dutch Frequency and Word Status was not significant either [F2 < 1].

A separate analysis of variance of the MIXED LIST showed that homographs had significantly slower responses than controls [F1(1,32) = 21.08, p < .001; F2(1,26) = 28.32, p < .001]. English Frequency did not show a significant effect [F1 < 1; F2 < 1]. The interaction between Word Status and English Frequency was not significant [F1 < 1; F2 < 1]. Homographs in a high Dutch frequency category and their controls had longer RTs than items belonging to the low Dutch frequency categories. This effect was significant [F1(1,32) = 33.41, p < .001; F2(1,26) = 15.52, p < .005]. The difference between homographs and controls was significantly larger for items in Dutch high frequency categories than for items in Dutch low frequency categories, as indicated by the significant interaction between Word Status and Dutch Frequency [F1(1,32) = 18.09, p < .001; F2(1,26) = 24.02, p < .001]. The interaction between English and Dutch Frequency was not significant [F1(1,32) = 2.62, p = .12; F2 < 1]. The triple interaction was not significant [F1 < 1; F2 < 1]. In a by-item analysis including the covariate number of letters, the effect of Word Status remained significant [F2(1,59) = 18.89, p < .001], as did the effect of Dutch frequency [F2(1,59) = 13.58, p < .005], and the interaction between Word Status and Dutch frequency [F2(1,51) = 16.06, p < .001]. No other effects were significant.

The results from the analysis with the number of letters as a covariate and the significant interaction suggest that in the mixed list – as opposed to the pure list – there was a true effect of Dutch frequency on the response times to homographs. Within each frequency category, the number of letters and the onsets were matched between homographs and their English controls. Therefore, the difference between homographs and controls as a function of Dutch frequency is unlikely to be an artifact caused by these factors.

We conclude that there is a significant homograph effect in the RTs of English responses to heterophonic homographs in the mixed list that depends on the Dutch frequency of the homographs. There was no homograph effect in the pure list.

Dutch RTs

For an overview of RTs, see Table 2. In the mixed list, 33% of all responses were Dutch responses. An analysis of variance including the within-participant and between-item factors English Frequency (High or Low) and Dutch Frequency (High or Low) was conducted on the RTs of the Dutch responses to homographs in the mixed list. Only 17 participants had responses in all four design cells and were included in the analysis. One of the homographs did not elicit any Dutch responses and was excluded from the analysis. English Frequency did not have a significant effect [F1(1,17) = 2.24, p = .15; F2(1,28) = 2.97, p = .10]. The effect of Dutch Frequency was not significant [F1(1,17) = 3.01, p = .10; F2(1,26) = 3.37, p = .08]. The interaction of English and Dutch Frequency was not significant [F1 < 1; F2 < 1]. A by-items analysis with the covariate number of letters showed the same results.

In the low Dutch frequency categories, the number of Dutch responses was rather low (22%). Therefore, we also present an analysis of variance of the RTs of the Dutch responses in the high Dutch frequency categories only, including the within-participant and between-items factor English Frequency. These categories had 45% Dutch responses. Thirty-one participants had Dutch responses in these categories. English Frequency did not have a significant effect [F1 < 1; F2(1,14) = 1.425, p = .26].

An analysis of variance including the within-participant factor and within-item factor Response Language (English or Dutch), and the within-participant and between-item factors English Frequency (High or Low) and Dutch Frequency (High or Low) was conducted on the RTs to heterophonic homographs. Only 17 participants had Dutch responses in all frequency categories and were included in the analysis. One of the homographs did not elicit any Dutch responses and was excluded from the analysis. Dutch responses to homographs were slower than English responses [F1(1,16) = 3.48, p = .08; F2(1,25) = 4.41, p < .05]. The effect of English frequency was not significant [F1(1,16) = 3.95, p = .07; F2(1,25) = 1.06, p = .31]. The interaction of English Frequency with

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Table 3. *Mean percentage Dutch responses and error rates, and the corresponding standard deviations (sd) for the heterophonic homographs in four relative frequency categories in both list conditions*

		Dutch responses		Total errors				
	Frequency	Homographs	sd	Homographs	sd	Controls	sd	Homograph effect
Mixed	HFE–HFD	37	48	43	50	1	11	42
list	HFE–LFD	15	36	19	39	5	21	14
	LFE–HFD	53	50	58	49	3	18	55
	LFE–LFD	28	45	31	46	4	19	27
Pure	HFE–HFD	13	33	16	37	0	0	16
list	HFE–LFD	3	16	5	22	4	20	1
	LFE–HFD	14	35	18	39	2	13	16
	LFE–LFD	2	16	9	28	1	8	8

Response Language was significant in the item analysis, but not in the participant analysis: Dutch responses were slower than English responses in the English low frequency categories, but faster than English responses in the Dutch high frequency categories [$F1 < 1$; $F2(1,25) = 4.32$, $p < .05$]. Responses to homographs with a low-frequency Dutch reading were faster than to homographs with a high-frequency Dutch reading [$F1(1,16) = 11.48$, $p < .005$; $F2(1,25) = 13.43$, $p < .005$]. This difference was much smaller in Dutch than in English as shown by the significant interaction between Response Language and Dutch Frequency [$F1(1,16) = 3.75$, $p = .07$; $F2(1,25) = 4.54$, $p < .05$]. The threefold interaction between Response language, Dutch Frequency, and English frequency was not significant. A by-items analysis with the covariate number of letters showed the same results.

Again, the small number of Dutch responses for the homographs in the low Dutch frequency categories may have skewed the results. Therefore, we performed an analysis on the English and Dutch responses in the high Dutch frequency categories only. Only 31 participants had both English and Dutch responses in both high Dutch frequency categories. Response Language had a significant effect in the item analysis [$F1 < 1$; $F2(1,13) = 12.46$, $p < .005$]: Dutch responses were faster than English responses. English Frequency did not have a significant effect [$F1 < 1$; $F2 < 2$], nor did the interaction between English Frequency and Response Language [$F1 < 1$; $F2(1,13) = 1.52$, $p = .24$].

To allow for a comparison between the RTs to purely English and purely Dutch words, we selected 30 Dutch words from the Dutch filler words that were matched to the English control words for number of letters and log frequency.[3] Note that the Dutch frequency of these control items was matched to the English frequency of homographs and English controls and consequently not to the Dutch frequency of the homographs. So, a low frequency Dutch control word was matched to a low frequency English control word, and a high frequency Dutch word was matched to a high frequency English word. We then conducted an analysis of variance with the within-participant and between-items factor Word Status (Dutch control or English control) on the RTs. The reaction times to Dutch words did not differ significantly from the reaction time to English words [$F1 < 1$; $F2 < 1$]. A by-items analysis with the covariate number of letters and the between-items factor Dutch frequency was performed on the RTs to Dutch control words. Dutch frequency (high or low) did not have an effect [$F2(1,30) = 1.90$, $p = .18$].

To summarize, there is a trend in the data that Dutch responses to homographs were somewhat slower than English responses. Dutch control words were responded to as slowly as English control words. The reaction times of English responses to homographs depended strongly on the Dutch frequency of the homographs. Dutch responses to homographs and Dutch control words, however, were unaffected by their Dutch frequency.

Error rates

All errors

For an overview of error rates, see Table 3. An analysis of variance with the between-participant and within-item factor List Condition, the within-item and within-participant factor Word Status, and the within-participant and between-item factors English Frequency and Dutch Frequency, was conducted on the percentages incorrect responses to heterophonic homographs and their control words. There were more errors in the mixed than in the

[3] Two control words (PEAR and ASH) had been removed from the analysis earlier, because their respective matched homographs PAL and ARTS had been removed in the outlier procedure. Therefore, only 30 Dutch words had to be selected.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

pure list [F1(1,56) = 64.90, p < .001; F2(1,26) = 45.46, p < .001]. Homographs elicited more errors than their control words [F1(1,56) = 203.18, p < .001; F2(1,26) = 66.20, p < .001]. There was a significant interaction between Word Status and List Condition, such that the homograph effect was larger in the mixed than in the pure list [F1(1,56) = 57.04, p < .001; F2(1,26) = 41.97, p < .001]. English Frequency had a significant effect in the participant analysis: low frequency words elicited more errors than high frequency words [F1(1,56) = 28.33, p < .001; F2(1,26) = 2.12, p = .16]. The difference between English low frequency and high frequency words was larger in the mixed than in the pure list, but this was only significant in the participant analysis [F1(1,56) = 17.61, p < .001; F2(1,26) = 2.47, p = .13]. Low English frequency homographs showed a larger Word Status effect than high frequency homographs, but again this was only significant in the participant analysis [F1(1,56) = 21.28, p < .001; F2(1,26) = 2.13, p = .16]. There was a significant threefold interaction between List Condition, Word Status, and English Frequency in the participant analysis, such that the difference in the size of the homograph effect between low and high frequency homographs was larger in the mixed than in the pure list [F1(1,56) = 7.01, p < .05; F2(1,26) = 1.68, p = .21]. Dutch Frequency had a significant effect: words in the high Dutch frequency categories elicited more errors than words in the low Dutch frequency categories [F1(1,56) = 67.60, p < .001; F2(1,26) = 9.47, p < .01]. There was a significant interaction between Dutch Frequency and List Condition, such that the difference between the low and high Dutch frequency categories was larger in the mixed list than in the pure list [F1(1,56) = 13.66, p < .005; F2(1,26) = 3.16, p = .09]. The Word Status effect was larger for homographs in the high Dutch frequency categories than for homographs in the low Dutch frequency categories [F1(1,56) = 100.93, p < .001; F2(1,26) = 12.93, p < .005]. The threefold interaction between List Condition, Word Status, and Dutch Frequency was significant: The difference in the size of the homograph effect between low Dutch frequency categories and high Dutch frequency categories was larger in the mixed than in the pure list [F1(1,56) = 14.83, p < .001; F2(1,26) = 4.28, p < .05]. None of the other interactions were significant. A by-items analysis with the covariate number of letters showed the same results. The threefold interaction between List Condition, Word Status, and Dutch Frequency was marginally significant in this analysis [F2(1,51) = 3.66, p = .06].

A separate analysis of the PURE LIST showed that Word Status had a significant effect: Homographs elicited more errors than control words [F1(1,23) = 23.53, p < .001]. Dutch Frequency also had a significant effect, such that there were more errors in the high Dutch frequency categories than in the low Dutch frequency categories [F1(1,23) = 9.42, p < .01; F2(1,26) = 3.29, p = .08]. There was also a significant interaction of Word Status and Dutch Frequency: The homograph effect was larger when the homograph had a high frequency Dutch reading than when it had a low frequency Dutch reading [F1(1,23) = 15.92, p < .005; F2(1,26) = 5.35, p < .05]. English frequency did not have a significant effect [F1 < 1; F2 < 1]. There were no other significant interactions. In a by-items analysis with the covariate number of letters, the main effect of Dutch Frequency was not significant [F2(1,51) = 16.06, p = .09], but the interaction between Dutch Frequency and Word Status was [F2(1,51) = 5.40, p < .05].

A separate analysis of the MIXED LIST showed that homographs elicited more errors than control words [F1(1,33) = 254.32, p < .001; F2(1,26) = 78.41, p < .001]. English low frequency words elicited more errors than high frequency words [F1(1,33) = 50.84, p < .001; F2(1,26) = 3.09, p = .09]. The interaction between Word Status and English Frequency was significant in the participant analysis: the homograph effect was larger for low frequency homographs than for high frequency homographs [F1(1,33) = 24.80, p < .001; F2(1,26) = 2.73, p = .12]. Words in Dutch high frequency categories elicited more errors than homographs in Dutch low frequency categories [F1(1,33) = 81.68, p < .001; F2(1,26) = 8.58, p < .01]. The homograph effect was larger for homographs with a high Dutch frequency reading than for those with a low Dutch frequency reading [F1(1,33) = 119.04, p < .001; F2(1,26) = 12.50, p <.005]. There were no other significant interactions. A by-items analysis with the covariate number of letters showed the same results.

To summarize, homographs elicited more errors than control words. This effect was larger in the mixed than in the pure list. The size of the homograph effect depended on the Dutch frequency of the homographs in both lists, and on the English frequency of the homographs in the mixed list.

Dutch responses

For an overview of the percentages Dutch responses, see Table 3. An analysis of variance with the between-participant and within-item factor List Condition and the within-participant and between-item factors English and Dutch Frequency was conducted on the number of Dutch responses to heterophonic homographs. List condition had a significant effect: there were more Dutch responses in the mixed list than in the pure list [F1(1,56) = 65.37, p < .001; F2(1,26) = 60.98, p < .001]. English Frequency had a significant effect in the participant analysis [F1(1,56) = 54.75, p < .001; F2(1,26) = 2.92, p = .10]: items with a low English frequency elicited more Dutch responses than homographs with a high English frequency. The interaction between List Condition and English

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

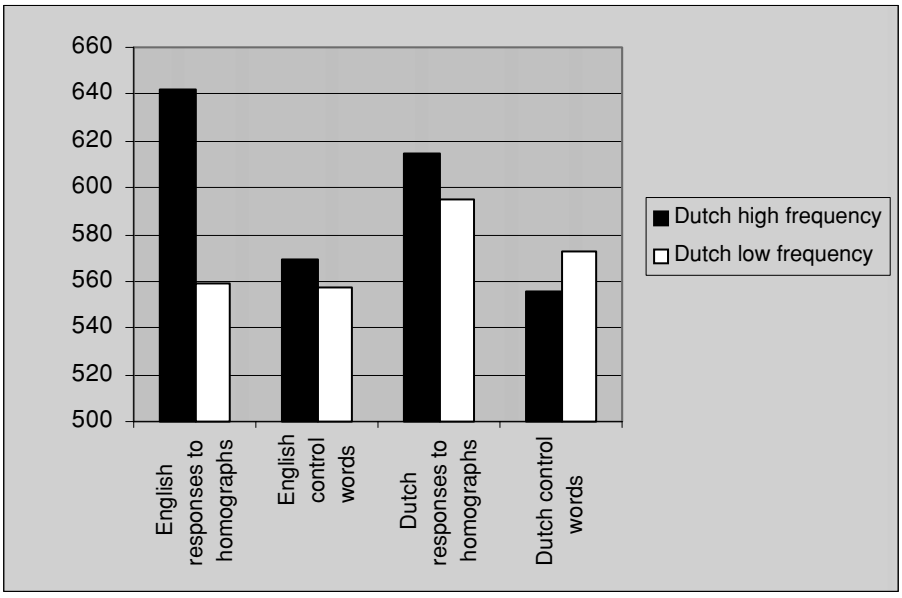


Figure 3. The reaction times in milliseconds in the mixed list for English responses to homographs and English control words, and Dutch responses to homographs and monolingual Dutch words, per Dutch frequency category.

Frequency was significant: The difference between the frequency categories was larger in the mixed list than in the pure list [$F1(1,56) = 19.81$, $p < .001$; $F2(1,26) = 4.61$, $p < .05$]. Homographs in Dutch high frequency categories elicited more Dutch responses than homographs in Dutch low frequency categories [$F1(1,56) = 76.36$, $p < .001$; $F2(1,26) = 15.95$, $p < .001$]. The interaction between Dutch Frequency and List Condition was significant in the participant analysis [$F1(1,56) = 10.21$, $p < .005$; $F2(1,26) = 3.68$, $p = .07$]: the difference between the frequency categories was larger in the mixed list. All other interaction effects were not significant. In a by-items analysis with the covariate number of letters, the main effect of English Frequency was no longer significant [$F2(1,25) = 2.97$, $p = .10$].

English and Dutch frequency were tested as within-participant and between-item factors in analyses on the percentage Dutch responses in the pure and the mixed list separately. In the PURE LIST, English Frequency did not have a significant effect [$F1 < 1$; $F2 < 1$]. There were more Dutch responses when the Dutch Frequency was high than when it was low. This was significant [$F1(1,23) = 17.03$, $p < .001$; $F2(1,26) = 11.00$, $p < .005$]. The interaction between the two frequency factors was not significant [$F1 < 1$; $F2 < 1$].

In the MIXED LIST, English Frequency had a significant effect such that there were more Dutch responses when the English Frequency was low than when it was high [$F1(1,33) = 42.92$, $p < .001$; $F2(1,26) = 4.35$, $p < .05$]. There were significantly more Dutch responses when the Dutch Frequency was high than when it was low [$F1(1,33) = 74.13$, $p < .001$; $F2(1,26) = 11.60$, $p < .005$]. The interaction was not significant [$F1 < 1$; $F2 < 1$].

In short, there were more Dutch responses to homographs in the mixed list than in the pure list. In the pure list and the mixed list there were also more Dutch responses when the Dutch frequency was high than when it was low.

General discussion

In a word naming experiment with Dutch–English bilinguals, we investigated whether the stimulus list composition effect observed in several other studies involving bilinguals (Dijkstra et al., 1998; De Groot et al., 2000; Von Studnitz and Green, 2002) generalizes to word naming, and what causes this effect. More specifically, the word naming task was used to exclude any possible effects of a language-independent response strategy in a pure (all-English) stimulus list. As such, a replication of the stimulus list effect would indicate that variation in the degree of competition between two homograph readings is an inherent characteristic of the bilingual word recognition system and not merely a consequence of differences in participant strategies.

Dutch–English homographs were presented either in a pure (all-English) list or in a mixed list that also contained purely Dutch words. In both conditions, the homographs had to be named in English. As can be seen in Figure 3, in the MIXED LIST, homographs were named more slowly and less accurately in English than their purely English control words. This homograph effect was larger when

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

the homograph had a high-frequency Dutch reading than when it had a low-frequency Dutch reading. Incorrect Dutch responses to homographs tended to be slower than correct English responses, although this result was significant only in the item-analysis. Dutch control words were named as slowly as English control words. Neither Dutch responses to homographs nor correct responses to purely Dutch filler words were affected by their Dutch frequency. In the PURE LIST, no homograph interference effect was found in RTs for English responses. There was, however, a significant homograph effect in the error rates. Homographs with a high-frequency Dutch reading elicited more Dutch responses than homographs with a low-frequency Dutch reading. These incorrect responses were much slower than the English responses and also slower than the correct Dutch responses in the mixed list.

Before addressing the results pattern in more detail, we would like to note that word naming can lead to correct responses via two routes: the correct pronunciation of a word is either accessed in the lexicon or generated by means of spelling-to-sound rules. In order to investigate the contribution of lexical and sublexical factors, we conducted several regression analyses involving frequency as a lexical predictor and different measures of phonological consistency (e.g. rime consistency) as sublexical predictors. Word frequency is assumed to determine how fast a lexical representation gathers enough activation to reach its recognition threshold and how strongly it suppresses other lexical representations. If participants' responses are based on lexical representations, word frequency should correlate highly with response speed. A factor like rime consistency, however, should determine the speed with which the orthographic input is converted into a phonological representation at the sublexical level. Hence, it should correlate with response speed if responses are mainly based on sublexical codes. Stepwise regressions with word frequency and different measures of rime consistency may assess the relative contribution of these factors to the variance in the data. The analyses were conducted on the response times to all English items in the mixed list (i.e. homographs, control words, and English filler words). The items' log frequency was calculated from the CELEX database (Baayen, Piepenbrockand and Van Rijn, 1993). To calculate the consistency variables, all monosyllabic English and Dutch words with a frequency of at least one occurrence per million were extracted from the CELEX database. Onset and rime consistency were calculated as the ratio of the number of words sharing the spelling and pronunciation of the relevant unit (e.g. INT in MINT and HINT) to the number of all words sharing the spelling of the unit, but not necessarily the pronunciation (e.g. INT in MINT and PINT). Two measures of consistency were obtained for each unit: intralingual (considering only English words) and interlingual (considering both English and Dutch words). English log frequency correlated significantly with the response times [$r = -283$, $p < .001$]. Intralingual onset consistency in English did not correlate with response times [$r = -06$, $p = .36$], nor did intralingual rime consistency [$r = .005$, $p = .93$]. The correlation between interlingual onset consistency and response times was marginally significant [$r = -.11$, $p = .07$]. Interlingual rime consistency did not correlate with response times [$r = -.10$, $p = .12$].

To summarize, English frequency had a significant effect on response times, whereas the effect of consistency was not reliable.[4] The English frequency effect indicates that the bilinguals in our study did indeed access their lexicon. The small and unreliable consistency effects show that sublexical information was applied at best to a limited extent, although we cannot exclude individual differences amongst participants. To conclude, for the majority of our bilinguals, the homograph interference effect seems to be situated mainly at the lexical or postlexical level, whereas sublexical processes play a minor role.

In the following section, we will discuss the stimulus list composition effect in the current study in relation to the two accounts mentioned in the introduction, i.e. participant strategies and variable competition.

Overall, we observed that the homograph effect was much smaller in the pure list than in the mixed list. With respect to RESPONSE LATENCIES, homographs did not show a homograph interference effect in the pure list, but did show a significant effect in the mixed list. With respect to ERROR RATES, both list conditions showed a significant homograph effect, but this effect was much larger in the mixed list. Our results pattern is in line with previous studies that used lexical decision to investigate the stimulus list composition effect (Dijkstra et al., 1998; De Groot et al., 2000; Dijkstra, De Bruijn et al., 2000; Von Studnitz and Green, 2002). This generalization of the stimulus list composition effect across different tasks suggests that the effect is caused by an inherent characteristic of the bilingual word recognition system, rather than by specific task-induced participant strategies alone. As described in the introduction, the latter had been proposed by De Groot et al. (2000). According

[4] We performed stepwise regressions with number of letters as the first variable, followed by first log frequency and then consistency or vice versa. Number of letters was introduced to control for all effects that could be explained by either letter length or word frequency, because it was correlated with English log frequency ($t = -0.15$, $p < .05$). The stepwise regressions showed that taking up log frequency as a variable increased the amount of explained variance significantly (all p's $< .001$), independent of the order of introduction of the variables. Taking up interlingual consistency led to a marginally significant increase in the explained variance ($.05 < p < .10$) also independently of the order of introduction. These results suggest that lexical (frequency) and sublexical (consistency) effects were fairly independent of each other.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

to their suggestion, the null-effect in the pure list is a combination of facilitation for participants who treated the task as language-independent lexical decision (YES to English and Dutch items) and inhibition for participants who treated it as language-specific lexical decision (NO to Dutch items). However viable this account might be for lexical decision, it cannot explain the reduced homograph interference effects in the present naming study, because a naming task is language-specific by definition. Our participants were instructed to respond in English to interlingual homographs and a failure to comply with this instruction by responding on the basis of the Dutch reading resulted in an error, unlike the situation in a lexical decision task with a pure list. Therefore, the presence of a similar stimulus list effect in word naming and lexical decision suggests that some inherent characteristic of the bilingual word recognition system is the cause of this finding. Previously, it has been suggested that the stimulus list effect is caused by reduced between-language competition in a one-language context. Two loci for this reduction of competition have been proposed; (i) relative language activation with variable competition at the lexical level, and (ii) dynamic decision criteria with variable competition at the decision level.

For the present experiments, RELATIVE LANGUAGE ACTIVATION as implemented in the original BIA model (Dijkstra and Van Heuven, 1998) predicts that the presence of only English words in the pure list raises the activation in the English language node to such an extent that substantial suppression of the Dutch word representations is possible. At the same time, the activation in the Dutch language node may be suppressed top-down on the basis of the instruction to read aloud in English. Both mechanisms reduce between-language inhibition of the English homograph reading by its Dutch reading. This results in faster English responses and fewer errors. In the mixed list, the presence of Dutch words raises activation in the Dutch language node, resulting in an increase of competition for the English homograph readings, leading to slower English responses and a higher number of Dutch responses. Moreover, the model predicts that responses to L1 words (and L1 homograph readings) are not necessarily faster than responses to L2 words in a mixed list (and may even be slower, as is observed in the current study), because L1 word candidates may be suppressed by the English language node, as the instructions emphasize naming in English.

Although the relative language activation account is in agreement with our data, previous research casts doubt on some of its assumptions. First, this variant of the relative language activation account assumes that bilinguals can control the activation levels in their lexicon on the basis of task demands, at least to some degree (for an extensive overview, see Dijkstra, 2005). However, this assumption is in contradiction with results from experiments in which non-target-language items had to be treated as nonwords (Nas, 1983; Dijkstra, De Bruijn et al., 2000) or had to be ignored altogether (e.g. in a go/no-go task, Dijkstra, Timmermans and Schriefers, 2000). Those experiments showed homograph interference effects that were comparable to those in the present study, suggesting that the participants were unable to suppress activation in the non-target language, although this would have optimized performance and would have been in accordance with task demands. Moreover, several studies have shown between-language interference effects in one-language contexts where the relative language activation account would predict substantial suppression of the non-target-language lexical representations (e.g. the interlingual neighborhood-interference effects in the study by Van Heuven et al., 1998 and see also Nas, 1983; Grainger and Dijkstra, 1992; De Moor, 1998; Dijkstra and Van Heuven, 1998; Dijkstra et al., 1999; Font, 2001; Von Studnitz and Green, 2002). Thus, although bilinguals' non-target-language word representations may be less active in a single language context, several studies have shown that they are at least not deactivated altogether. Moreover, it is debatable whether a reduced activation in the non-target language would be the consequence of top-down factors such as the task demands or the participants' expectancies.

In a later version of the BIA model, the BIA+ model, Dijkstra and Van Heuven (2002) suggested an account in terms of DYNAMIC DECISION CRITERIA for the stimulus list effect in lexical decision: the response binding with the non-target-language reading of an interlingual homograph is weaker in a pure list than in a mixed list. Such an account may apply here as well. In the pure list, any potential response binding with the Dutch part of the lexicon would be very weak as compared to the same response binding in the mixed list. Therefore, fewer Dutch responses will emerge in the pure list than in the mixed list, and the reaction times for English responses to homographs will be faster in the pure list than in the mixed list. Moreover, this account can also explain why Dutch responses to homographs and purely Dutch words are not faster than English responses: the response binding with the Dutch lexical representations is somewhat weaker than that with the English lexical representations because there are still more English than Dutch responses required in the mixed list and responding in English is given primacy in the instructions.

In short, both relative language activation and dynamic decision criteria predict the current result pattern. Although our data do not allow us to distinguish the two accounts, both models hold that it is competition that ultimately leads to the correct response, and that the reduced homograph effect in a pure list is the result

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

of reduced competition, either at the lexical or at the decision level. The question remains, however, whether competition alone can account for the entire data pattern. In the following, we will further examine two aspects of our data that are not in accordance with a pure competition account, and formulate an explanation in terms of a time criterion for responding in the non-target language.

The first observation that is not in line with a pure competition based account is the absence of a Dutch frequency effect in the latencies of Dutch responses to homographs and Dutch filler words. As high-frequency words reach their recognition threshold sooner and will inhibit other lexical representations more strongly, the absence of a frequency effect in these response data is surprising. At the same time, Dutch frequency affected the size of the homograph effect, both in the naming latencies for English responses (mixed list) and in the number of Dutch responses (pure and mixed lists). This brings us to an important question: If interlingual homograph competition is affected by Dutch frequency, why then is this not reflected in the latencies of Dutch responses? As an explanation, we suggest that the word recognition system checks the language membership of each available response, starting with the response option or lexical representation that is most active. If this response meets the language requirements of the task, it may be released. If the language information does not meet the task requirements, the response is delayed until a certain time criterion. Such a deadline is independent of the activation speed of Dutch lexical representations and if the time criterion is set quite late, all Dutch responses will be released at about the same time. Thus, frequency effects that arise at the lexical level, cannot surface in the Dutch naming latencies.

For this hypothesis to apply to our data, a second assumption has to be made, however, because the Dutch responses to homographs are considerably slower than the responses to Dutch control words, i.e. not all Dutch responses take equally long. This implies that the decision to release a Dutch response depends on whether or not an alternative response option is available as well. When only one response option is available, a Dutch response will be released faster than when another response option is quite active as well. In addition, a Dutch response to a homograph is delayed by the competition process between both readings.

A similar process may explain the occurrence of some Dutch responses in the pure list. If no English reading becomes available in time, the decision system might release the Dutch response in an ad-hoc manner. The temporal deadline for this release is probably of a different type than that in the mixed list, where a Dutch response is potentially appropriate. In contrast, responding in Dutch would be the last resolve in a pure list. It is conceivable that such a "last-resolve" deadline would be situated later than the time criterion in a mixed list. The response times for the few Dutch responses in the pure list seem to support this hypothesis: They are on average more than 200 ms slower than the English responses and also slower than the English or Dutch responses in the mixed list.

The second observation that is difficult to account for in a pure competition based account, is that there is no homograph interference effect on the response SPEED in conditions that do show a clear effect in terms of THE NUMBER OF ERRORS. In the current study, there is no delay in correct responses to homographs with a low-frequency Dutch reading in the mixed list and to all homographs in the pure list, although these conditions show a clear homograph effect in their substantial number of Dutch responses. Again, the assumption of a time criterion for non-target-language responses to interlingual homographs might explain this. Arguably, responses to homographs are seldom later than this time criterion, as homographs offer the possibility to respond in Dutch when the English response option is not active enough at that moment. For control words, however, there is no alternative to the correct English response. Therefore, the system simply has to wait longer to make a response in those instances where the English response is not active enough yet, i.e. wait beyond the deadline for non-target-language responses. Consequently, homographs may lead to many erroneous Dutch responses (because often the English reading will be inhibited by the Dutch reading and not reach the required activation level fast enough), but not to English responses beyond the non-target-language deadline, whereas English control words will elicit fewer errors (as they suffer less competition and there are no alternative responses), but sometimes give rise to very slow English responses. This difference between response options for homographs and controls at the postulated deadline could increase the average reaction time for controls and thus mask the difference in access speed between these two word types. As a result, a homograph effect could appear in the errors while being absent in the reaction times.

To conclude, the stimulus list composition effect appears to be the consequence of reduced competition in a pure list, which could be situated either at the lexical or the decision level. The latencies for incorrect responses in the non-target language suggest, however, that competition cannot be the only mechanism that controls in which language one responds. We suggest that there might be a time criterion for responding in the non-target language that masks the speed with which non-target-language lexical representations reach their recognition threshold. Thus, we argue that any competition-based account needs to be supplemented with a decision level, at which task demands, language membership information, and context factors determine the conditions for making both target and non-target-language responses.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Appendix 1. *The heterophonic homographs and their controls with their mean log frequencies, number of letters and corresponding standard deviations (sd)*

	Homographs				English control words			Dutch control words		
Frequency category	Word	Number of letters	English log frequency	Dutch log frequency	Word	Number of letters	English log frequency	Word	Number of letters	Dutch log frequency
HFE–HFD	pad	3	1.41	1.80	pope	4	1.34	neef	4	1.45
	arts	4	1.59	1.97	ash	3	1.67	—	—	—
	boot	4	1.62	1.83	bore	4	1.62	mes	3	1.62
	trap	4	1.72	2.07	tight	5	1.72	vak	3	1.72
	glad	4	1.81	1.58	guest	4	1.81	nauw	4	1.80
	bad	3	2.52	1.40	boy	3	2.57	uur	3	2.63
	kind	4	2.68	2.98	keep	4	2.83	naam	4	2.62
	stand	5	2.72	2.13	side	4	2.67	jaar	4	3.06
Mean and standard deviation		3.88 sd 0.64	2.01 sd 0.54	1.97 sd 0.48		3.88 sd 0.64	2.03 sd 0.57		3.57 sd 0.53	2.13 sd 0.63
HFE–LFD	fee	3	1.51	0.60	fist	4	1.46	ijs	3	1.46
	peer	4	1.62	1.04	pig	3	1.67	hoed	4	1.62
	map	3	1.66	1.00	mad	3	1.70	jas	3	1.70
	roof	4	1.79	0.48	row	3	1.81	dak	3	1.77
	sleep	5	2.33	0.30	sign	4	2.29	thuis	5	2.35
	meet	4	2.52	0.95	month	5	2.51	jong	4	2.55
	room	4	2.73	0.78	run	3	2.72	deel	4	2.59
	put	3	2.89	1.04	play	4	2.73	oog	3	2.91
Mean and standard deviation		3.75 sd 0.71	2.13 sd 0.55	0.77 sd 0.28		3.63 sd 0.74	1.93 sd 0.51		3.63 sd 0.74	2.12 sd 0.54
LFE–HFD	ramp	4	0.85	1.41	ranch	5	0.90	raaf	4	0.85
	brand	5	1.23	1.66	bump	4	1.26	drang	5	1.23
	strand	6	1.23	1.72	splash	6	1.26	naald	5	1.23
	rug	3	1.26	2.26	robe	4	1.23	aap	3	1.32
	gang	4	1.41	2.28	grade	5	1.43	grap	4	1.40
	breed	5	1.45	2.12	bleed	5	1.41	jacht	5	1.43
	toe	3	1.52	2.92	toy	3	1.52	doek	4	1.51
	heel	4	1.53	3.11	hunt	4	1.65	meid	4	1.56
Mean and standard deviation		4.25 sd 1.04	1.31 sd 0.22	2.19 sd 0.60		4.5 sd 0.93	1.33 sd 0.23		4.25 sd 0.71	1.32 sd 0.22
LFE–LFD	rap	3	0.85	0.78	raft	4	0.70	rel	3	0.70
	pal	3	0.90	0.95	pear	4	0.90	—	—	—
	mug	3	1.04	0.85	nag	3	1.04	wol	3	1.04
	log	3	1.11	0.90	ray	3	1.26	pijl	4	1.23
	wig	3	1.18	0.48	wink	4	1.11	darm	4	1.18
	slap	4	1.34	1.45	shy	3	1.32	kaak	4	1.30
	angel	5	1.40	0.78	yell	4	1.41	verf	4	1.45
	lap	3	1.41	1.15	lamb	4	1.38	riem	4	1.36
Mean and standard deviation		3.38 sd 0.74	1.15 sd 0.22	0.92 sd 0.29		3.63 sd 0.52	1.14 sd 0.25		3.71 sd 0.49	1.18 sd 0.25

Note: The Dutch control words were selected from the Dutch fillers in the experimental list and were matched to the English control words in objective frequency. As two homographs and their controls were removed from the analysis, these were not matched to a Dutch control word.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

Appendix 2. *The homophonic homographs and their controls with their mean log frequencies, number of letters, and corresponding standard deviations (sd)*

Homographs				Controls		
Word	Number of letters	English log frequency	Dutch log frequency	Word	Number of letters	English log frequency
bit	3	2.35	0.48	buy	3	2.40
brief	5	1.74	2.30	blame	5	1.75
door	4	2.61	3.62	day	3	2.96
dot	3	1.18	0.48	drip	4	1.15
drift	5	1.72	1.23	dust	4	1.74
drop	4	2.25	0.48	dry	3	2.14
kin	3	0.60	1.52	cute	4	0.60
last	4	2.83	1.86	life	4	2.93
lid	3	1.32	2.36	leak	4	1.28
list	4	2.06	0.85	risk	4	2.02
long	4	3.02	1.32	like	4	3.31
mop	3	1.00	0.90	mule	4	1.00
nut	3	1.40	1.45	melt	4	1.41
pet	3	1.34	1.30	prey	4	1.30
pink	4	1.72	0.85	cure	4	1.79
rest	4	2.42	2.06	rise	4	2.41
rust	4	0.90	1.94	roach	5	0.90
sip	3	1.38	0.30	sail	4	1.45
slim	4	1.18	1.43	fox	3	1.23
slip	4	1.91	0.78	sky	3	1.95
slot	4	1.00	1.86	frog	4	1.00
snip	4	0.48	0.30	swine	5	0.48
spin	4	1.48	1.00	fry	3	1.43
spit	4	1.30	0.48	spy	3	1.26
spot	4	1.90	1.11	shoe	4	1.93
star	4	2.05	1.11	skin	4	2.05
stem	4	1.51	2.49	slice	5	1.52
step	4	2.23	0.30	sell	4	2.16
strip	5	1.67	0.60	sigh	4	1.68
vest	4	1.00	1.04	weird	5	0.95
wet	3	1.85	2.27	wing	4	1.81
wit	3	1.43	2.51	weed	4	1.43
Mean log frequency	3.78	1.65	1.33		3.94	1.67
Standard deviation	0.60	0.61	0.80		0.61	0.65

References

Altenberg, E. P. & Cairns, H. S. (1983). The effect of phonetic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22, 174–188.

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

Bijeljac-Babic, R., Biardeau, A. & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition*, 25, 447–457.

Cohen, J. D., MacWhinney, B., Flatt, M. & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research: Methods, Instruments and Computers*, 25(2), 257–271.

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press

De Groot, A. M. B., Delmaar, P. & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to the bilingual memory. *The Quarterly Journal of Experimental Psychology*, 2, 397–428.

De Moor, W. (1998). Visuele woordherkenning bij tweetalige personen. Master Thesis, University of Ghent.

Dijkstra, A. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, pp. 178–201. Oxford: Oxford University Press.

Dijkstra, T., De Bruijn, E., Schriefers, H. & Ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition*, 3, 69–78.

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

Dijkstra, T., Timmermans, M. & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, 42, 445–464.

Dijkstra, T. & 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. Mahwah, NJ: Lawrence Erlbaum Associates.

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

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

Font, N. (2001). Rôle de la langue dans l' acces au lexique chez les bilingues: Influence de la proximité orthographique et sémantique sur la reconnaissance visuelle de mots. Ph.D. dissertation, Université Paul Valéry, Montpellier, France.

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. 207–220. Amsterdam: Elsevier Science Publications.

Lemhöfer, K. & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in generalized lexical decision. *Memory and Cognition*, 32, 4, 533–550.

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, 526–534.

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

Von Studnitz, R. E. & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, 5, 1–24.

Received April 5, 2005
Revision received December 2, 2005
Accepted February 7, 2006

https://doi.org/10.1017/S136672890600263X Published online by Cambridge University Press