THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2000, *53A* (2), 397–428

The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory

Annette M.B. de Groot and Philip Delmaar
University of Amsterdam, Amsterdam, The Netherlands
Stephen J. Lupker
The University of Western Ontario, London, Ontario, Canada

In three experiments we looked at the processing of interlexical homographs by Dutch–English bilinguals. In Experiment 1 we employed the translation recognition task, a task that forces the participants to activate both language systems simultaneously. In this task the processing of interlexical homographs was inhibited substantially compared to the processing of matched control words, especially when the homograph reading to be selected was the less frequent of the homograph's two readings. In Experiments 2 and 3 we used the lexical decision task: In one condition we asked the participants to categorize letter strings as words or nonwords in Dutch; in a second condition we asked them to do so in English. The makeup of the stimulus set in Experiment 2 permitted the participants to ignore the instructions and to instantiate the task in a language-neutral form—that is, to categorize the letter strings as words in either Dutch or English. Under these circumstances a small, frequency-dependent inhibitory effect for homographs was obtained, but only in condition Dutch. In Experiment 3 the participants were forced in a language-specific processing mode by the inclusion of "nonwords" that were in fact words in the non-target language. Large frequency-dependent inhibitory effects for homographs were now obtained in both language conditions. The combined results are interpreted as support for the view that bilingual lexical access is non-selective.

Two interrelated but separable disputes regarding the nature of bilingualism have dominated psycholinguistic research into bilingualism over the past 15 years. The first concerns the memory organization of the two languages of the bilingual, the specific question

Requests for reprints should be sent to A.M.B. de Groot, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. Email: pn_groot@macmail.psy.uva.nl

We owe much gratitude to Marije Bierlaagh, Leon Ouwerkerk, Nikos Plakas, Steven Scholte, Marc Swart, and Sylvia Witteveen, who collected the data of Experiment 3. We also thank Judith Kroll and an anonymous reviewer for their invaluable comments on an earlier version of this paper. The work reported in this article was presented at the International Symposium on Bilingualism, April 9–12, 1997, Newcastle upon Tyne, U.K.

http://www.tandf.co.uk/journals/pp/02724987.html

posed being whether bilingual memory contains two language-specific stores or instead one integrated, language-independent memory structure with memory nodes shared between the two languages. The second is about the way bilinguals gain access to their language system (or systems, depending upon the outcome of research into the first question), selectively or non-selectively. The selective-access view holds that a language input is processed only by the contextually appropriate language system whereas the non-selective–access view posits that both language systems respond to the input.

These two questions are interrelated because the nature of the memory organization determines what access procedures can plausibly be differentiated. That is, given a completely integrated bilingual memory structure with all memory nodes shared between the two languages, language-selective access is not a feasible option and, therefore, the question of whether bilingual lexical access is selective or non-selective would become moot. However, if at least one level of representation in the memory system can be discovered where (at least a substantial number of) the stored units are unique to just one of the two languages, rather than being shared between the languages, it does make sense to ask the selective- versus non-selective–access question: Does the input activate elements at this particular representational level irrespective of the language context of the input so that elements of the contextually inappropriate language are also among the activated set, or does it selectively activate only the elements of the contextually appropriate language system?

Recent work on bilingual memory organization has produced substantial support for a model of bilingual memory that contains (at least) two levels of representations; namely, one that stores the forms of words in a language-specific manner (possibly with the exception of cognates, that may be stored in language-independent memory structures), and a second that stores the meanings of words in representations that are largely shared between the bilingual's two languages (see De Groot, 1998; Kroll & De Groot, 1997, for recent reviews of the relevant work). The support for this model, with at least one layer of largely language-specific representations, legitimizes the question of whether or not bilingual lexical access is selective. Indeed, this question has been the main subject of investigation in a number of studies (e.g. Beauvillain & Grainger, 1987; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Gerard & Scarborough, 1989; Grainger & Beauvillain, 1987; Grainger & O'Regan, 1992; Macnamara & Kushnir, 1971; Van Heuven, Dijkstra, & Grainger, 1998; the last article also discusses the relation between memory organization and the access procedure in bilinguals). It is also the major topic of this paper.

In the three experiments reported here, lexical access in bilinguals was investigated by comparing the processing of so-called interlexical homographs and matched non-homographic controls. Interlexical homographs (henceforth also referred to as "homographs") are words with the same written form but different meanings in the two languages of the bilingual (e.g. the word *glad*, meaning *slippery* in Dutch, in a Dutch–English bilingual). A difference in terms of response time, percentage errors, or both, between homographs on the one hand and their controls on the other hand will be regarded as support for the involvement of the non-target language; in other words, it will suggest non-selective access. As pointed out later, the absence of any differences between the response patterns for interlexical homographs and their controls may be more difficult to interpret, but the preferred interpretation of such null effects is that the non-target language was not implicated, thus providing support for a selective model of bilingual lexical access. A number of earlier studies have also exploited interlexical homographs to clarify the nature of lexical access in bilinguals (Beauvillain & Grainger, 1987; Dijkstra et al., 1998; French & Ohnesorge, 1996; Gerard & Scarborough, 1989). Of these, the study by Dijkstra et al. (1998) was most similar to the present investigation and is therefore introduced here. The remaining studies are briefly discussed later.

Dijkstra et al. (1998) reported three experiments in which Dutch–English bilingual participants performed the lexical decision task. Different versions of the task were used, and the stimulus materials varied across the three experiments. In Dijkstra et al.'s Experiments 1 and 2, the participants were asked to perform English lexical decisions—that is, to categorize letter strings as words or nonwords in English. According to the authors, the most critical difference between experiments was that in Experiment 1 all nonword letter strings were nonwords both in English and in Dutch, whereas in Experiment 2 a subset of the letter strings that required a "no" response (because they were not words in English) were in fact Dutch words. Homographs and matched controls were responded to equally rapidly in Experiment 1, whereas in Experiment 2 response times were longer for homographs. In Experiment 3 the task was changed such that participants were now asked to perform "general" lexical decisions; that is, they had to respond "yes" if the presented letter string was a word in either language, English or Dutch, and "no" otherwise. The experimental materials consisted of equal proportions of Dutch and English words and Dutch-like and English-like nonwords. Under these circumstances the homographs were processed more rapidly than the controls.

Dijkstra et al. (1998) suggested that the different data patterns across the three experiments were the result of differences in task demands and language intermixing. They argued that the more elements of the non-target language system that are included in the stimulus set, the larger the level of activation of the non-target language system will be. Thus, because words from the non-target language, Dutch, were included among the stimuli that required a "no" response in Experiment 2 but not in Experiment 1, the non-target language system would have been activated more in Experiment 2 than in Experiment 1. This non-target language activation is assumed to inhibit the responses to homographs in (what we will call) the "language-specific" version of the lexical decision task, where letter strings have to be categorized as belonging to one language in particular (the target language, here English). The reason is that the activated non-target reading of the homograph must be suppressed or rejected, which causes a delay in processing.

Both Experiment 1 and Experiment 2 of Dijkstra et al. (1998) were language-specific lexical decision tasks, and in Experiment 2 homographs were indeed responded to more slowly than controls. The fact that, in contrast, in Experiment 1 homographs and controls were processed equally rapidly, suggests that in that experiment the non-target language was not activated sufficiently strongly to interfere with processing.

Finally, in Dijkstra et al.'s (1998) Experiment 3, due to the large number of words from both languages in the stimulus set, both language systems would have been activated strongly. However, because the task demands in this experiment were different from those in Experiments 1 and 2 in that the recognition of any word, irrespective of language, would provide a sufficient basis for a "yes" response, the activation of an additional word reading in the case of homographs could now benefit performance. As a result, homographs were processed faster than controls.

Dijkstra et al. (1998) adhere to the view that bilingual lexical access is non-selective, and they explain the apparent evidence for selective access in their Experiment 1 in terms of rapid suppression of activation in the non-target–language system. The suppression is in fact so fast that it produces no noticeable inhibitory effect. However, the fact that their Experiments 2 and 3 have shown that homograph processing can either be easier or more difficult than processing matched controls suggests another explanation for this null effect, one that is also compatible with the non-selective–access account: The facilitatory and inhibitory effects produced by the activated non-target reading of the homograph may have cancelled each other out, producing the observed null effect of homography. This proposal is discussed more fully later.

The present experiments

In Experiment 1 we employed the "translation recognition" task, a task that has not been used before in studies on the processing of interlexical homographs. A critical feature of this task is that it necessarily requires the simultaneous activation of both of the bilingual's lexicons. On each trial in this task two words are presented, one in the participants' own language (here Dutch) and the second in their other language (here English; see De Groot & Comijs, 1995, for a detailed analysis of this task). Typically, the two words presented on a trial are either translation equivalents or do not share any relationship with one another. The participants are asked to categorize each word pair as either a translation pair or a non-translation pair. (Sometimes the "no" stimuli are related as well, semantically or otherwise; Altarriba & Mathis, 1997; Talamas, Kroll, & Dufour, 1998.)

This task can, of course, only be performed if the participants activate both of their language systems. If both are indeed activated, an interlexical homograph (e.g. *glad*, meaning "slippery" in Dutch, in the case of Dutch–English bilingualism) is likely to activate its meaning in both languages. If the interlexical homograph is then one of the two terms in a true translation pair (e.g. the Dutch–English pair *glad–slippery*), its inappropriate reading (here the English reading) will cause a mismatch with the meaning of the other word in the pair (*slippery*) and bias the participants towards a "no" response. This tendency to produce a "no" response either will have to be overcome, slowing down the response, or will result in an error. In sum, we expect that, as compared to matched controls, interlexical homographs in translation pairs will cause inhibition in translation recognition, due to the fact that both systems must be activated in order to perform the task.

In contrast, Experiments 2 and 3 employed the lexical decision task. Although designed to run independently from Dijkstra et al.'s (1998) experiments, these two experiments were to a large extent similar to their Experiments 1 and 2. As in their experiments, in both of our experiments language-specific task instructions were given to the participants, and the two experiments differed from one another in that words from the non-target language were included among the "no"-response materials in Experiment 3 but not in Experiment 2. However, the stimulus materials used and their composition differed between their study and ours, and we tested performance not only in our participants' non-native language English, but also in their native language Dutch. The critical difference between the language-specific lexical decision task employed in the present Experiments 2 and 3 and the translation recognition task employed in Experiment 1 is that the former per se does not require the simultaneous activation of the bilingual's two lexicons. The central question that Experiments 2 and 3 allow us to address is whether nonetheless both lexicons are activated during task performance.

If the non-target language system is not activated at all when performing a lexical decision task (i.e. access is language selective), only the target language reading of the homograph will become available, and homographs would be processed the same way as matched non-homographic controls. In both cases just one reading will be activated. Equally long response times and equal error rates for homographs and controls would be the result. If, however, both languages are activated (i.e. access is non-selective), task performance will depend on how the participants actually perform the task. If the participants perform the task as instructed—that is, they only respond affirmatively when the stimulus is a word in the target language (rather than in either language)—the inappropriate reading of the homograph may delay the response. As a consequence, interlexical homographs will be responded to more slowly than matched controls. However, if, contrary to the instructions, the participants accept the availability of *any* meaning, irrespective of its source, as the basis for a positive response (cf. Experiment 3 of Dijkstra et al., 1998), a state of activation of two meanings instead of just one would benefit performance on interlexical homographs (because access to either of the two lexical representations would provide a basis for responding). Participants adopting this processing mode would essentially be performing the general lexical decision task (which we will refer to as "language-neutral" lexical decision). The participants were tacitly permitted to do so in Experiment 2 (and in fact also in Dijkstra et al.'s Experiment 1) but not in Experiment 3 (nor in Dijkstra et al.'s Experiment 2), where they would be penalized by high error rates on the "nonwords" that were in fact words in the non-target language.

An additional point to make is that if access is non-selective, the relative levels of activation of the target and non-target reading of a homograph are likely to affect the size of the non-target reading's effect, inhibitory or facilitatory, on homograph processing. For instance, when the non-target reading is only slightly activated and the target reading is highly activated, the effect of the non-target reading will be small or even non-existent (in the latter case the data will in fact mimic language-selective access). In contrast, when the non-target reading is activated to a higher level than the target reading, the effect of the former will be substantial.

The relative activation levels of the two homograph readings may be expected to depend on (at least) the following two factors: the relative frequency of use of the homograph's two readings, and the relative proficiency of the bilingual in his or her two languages (see also Dijkstra et al., 1998; Van Heuven et al., 1998). Especially when the non-target reading has a much higher frequency of use than the target reading, it should greatly affect processing (see Beauvillain & Grainger, 1987, for support for this contention). The basis for this prediction is that high-frequency words have higher baseline levels of activation and are therefore more available than low-frequency words. Similarly, when the non-target reading of the homograph belongs to the participant's dominant language—the language with the highest baseline activation level and, consequently, the most available words—it is likely to cause a larger effect than when it is part of the bilingual's weaker language. We tested these predictions by looking at the effects of language (native, dominant Dutch vs. non-native English) and relative frequency of the homographs' readings on the size of the homograph effect.

EXPERIMENT 1

Translation Recognition

Method

Participants

The participants were 72 first-year psychology students from the University of Amsterdam, The Netherlands. They were all unbalanced bilinguals with Dutch as their native and strongest language and English as their strongest foreign language. Despite being unbalanced bilinguals, their fluency in English was high. Since the age of 12, the participants had been instructed in English for 3 to 4 hours a week, and the majority of their university textbooks were in English.

Design

Of the participants, 36 always received the homograph in the first position of a translation recognition stimulus (e.g. *glad–blij* for English–Dutch pairs/*glad–slippery* for Dutch–English pairs), whereas the remaining 36 participants got the homograph in second position (e.g. *blij–glad/slippery–glad*). For 18 participants within both of these groups, the target reading of the homograph was its meaning in Dutch (e.g. *glad–slippery/slippery–glad*), whereas for the remaining 18 participants in both groups the target reading was its English meaning (e.g. *glad–blij/blij–glad*). Assignment of participants to each of the four language (of the target reading of the homograph: Dutch vs. English) by position (of the homograph in a translation pair: first vs. second) conditions was random.

Materials

The experimental materials were four lists of 192 word pairs each. There was a different list for each of the four groups of participants. Of these 192 word pairs per list, 96 were translation pairs and thus required a "yes" response. The remaining 96 pairs consisted of words that were not translations of one another (nor did they share any other obvious relationship) and thus required a "no" response. Out of the 96 translation pairs of each list, 48 were critical, whereas the remaining 48 served as fillers. Of the 48 critical pairs per list, 24 contained a homograph (henceforth: "homograph pairs"); the remaining 24 (henceforth: "control pairs") served as control items for these homograph pairs. None of the 48 filler translation pairs contained a homograph. Out of the 24 homographs, 11 were also near homophones, sharing not only the consonants but also the vowels across the two languages (although there were length differences between some of the shared vowels across languages). The remaining 13 homographs were not also homophones; their vowels clearly differed across the two languages.[1] In addition to the experimental materials, every list contained 12 translation pairs and

[1] For all three experiments of this study we performed analyses with this homophone variable (same or different pronunciation across languages) as a separate variable. In none of these analyses did this variable interact with the effect of the critical homograph variable, nor did it interact with the other variables ($F < 1$ in most cases, and $p > .20$ in three cases). Because this variable does not appear to qualify the effects of the other variables, for the sake of brevity these analyses are not reported.

12 non-translation pairs to be presented for practice. None of the practice pairs contained a homograph.

The set of 24 interlexical homographs used in this study was selected from the Dutch and English data bases in the Nijmegen (The Netherlands) CELEX corpus (for details, see Baayen, Piepenbrock, & Van Rijn, 1993). Homographs with a very low frequency of occurrence in English were not selected because these would be unlikely to be known in that language by our participants and would, therefore, effectively be non-homographs. Homographs that were conjugated verb forms were not selected either. For each of the 24 selected homographs two control words were chosen, one Dutch and one English. The selected Dutch control word was matched in word-form frequency with the Dutch reading of the homograph in the Dutch corpus. Similarly, the selected English control word was matched in word-form frequency with the English reading of the homograph in the English corpus. Two *t* tests were performed to see whether this matching procedure had been successful. This turned out to be the case: The mean word-form frequency of both the Dutch readings of the homographs and their (Dutch) controls was 50 per million (with *SD*s of 68 and 70, respectively). The mean word-form frequency of the English readings of the homographs was 85 per million (*SD*: 168), whereas that of their (English) controls was 89 per million (*SD*: 176), also a non-significant difference ($p > .10$). The length of the homographs did not differ significantly from either their Dutch controls ($p > .80$) or their English controls ($p = 1.00$). The mean length of the homographs, the Dutch controls, and the English controls was 4.00, 3.96, and 4.00 letters, respectively.

Of the 24 homographs, 13 had a more frequent reading in Dutch than in English, whereas 11 had a more frequent English reading. The mean Dutch word-form frequency of the 13 homographs with a more frequent reading in Dutch was 82 per million, whereas that of their (Dutch) controls was 83 per million. The mean English word-form frequency of these 13 homographs and of their (English) controls was 13 per million. The mean Dutch word-form frequency of the 11 homographs with a less frequent reading in Dutch was 11 per million; their Dutch controls had exactly the same mean word-form frequency. The mean English word-form frequency of these 11 homographs was 171 per million; for their (English) controls it was 178 per million.[2]

One final test was done to see whether the English reading of the homographs and their controls were indeed matched on all relevant characteristics. We presented the homographs, their English controls, and a set of English-like pseudowords to a group of Canadian monolinguals who, like the participants in the present experiments, were also university students. They performed lexical decisions on these stimuli, categorizing each of them as a word or a nonword. For English-speaking monolinguals these interlexical homographs carry just one meaning, just as the controls do. If our matching procedure was successful, homographs and controls should therefore have been responded to equally fast. This indeed turned out to be the case: Homographs and controls showed response times of 563 msec and 555 msec, respectively, a difference that failed to approach significance in an analysis either by subjects ($p > .30$) or by items ($p > .50$). The analogous experiment could not be run on the Dutch stimuli, for the simple reason that no Dutch university students can be found who are not bilingual in Dutch and English.

The 48 filler translation pairs of each list were of the same type as the control pairs—that is, they did not contain a homograph. Unlike the experimental controls, these fillers were not frequency-matched with the homographs. (The mean word-form frequency of the Dutch fillers was 54 per million; the mean word-form frequency of the English fillers was 65 per million.) The purpose of the fillers was to decrease the overall proportion of homograph pairs among the stimulus materials (12.5%; 24 out of the 192). In so doing, the chances were increased that any effect of interlexical

[2] Homographs were matched with their controls on word-form frequency, not on lemma frequency. Nevertheless, neither of the two groups of homographs did differ significantly from their respective controls in terms of lemma frequency.

homography on processing would be due to natural language processes in bilinguals, not to an ad hoc processing strategy developed by the participants in the course of the experiment.

Apparatus and procedure

The experiment was run on an Apple Macintosh Plus ED computer, with a response box with two push buttons, one for "yes" responses and one for "no" responses, connected to it. Stimulus presentation and response registration were controlled by computer software developed in our laboratory. The word pairs were presented in black lower-case letters against a light-grey background on the computer screen. Prior to the presentation of the practice trials, the participants were told that they would be presented with pairs of words, one word of each pair in English and the other in Dutch, and that they were to decide for each pair whether or not the two words were translations of one another. They were also told about the position of the Dutch word and English word in the pairs. No mention was made of the fact that a number of the presented words were interlexical homographs. The participants were asked to push the right-hand button with their right forefinger if they thought the word pair consisted of translations, and to push the left-hand button with their left forefinger otherwise. They were encouraged to push the button as rapidly as possible, while at the same time keeping the number of errors as low as possible. The second word of a translation pair was always presented 240 msec after the first (an interval presumably too short to retrieve the first word's translation from memory before the second word appeared, but sufficiently long to recognize the first word), and it appeared on the screen immediately below the first word (which remained in view). The reason for presenting the words in a pair successively rather than simultaneously was to obtain better control of the processing order of the two words. This was important in order to have a viable manipulation of the position (of the homograph) factor.

Every first word of a pair was preceded by a fixation stimulus (an asterisk), which lasted 1 sec. The interstimulus interval between the fixation stimulus and the first word of a pair was 20 msec. The participant's response was immediately followed by feedback on the screen, consisting of the word *correct*, *wrong*, or *slow* (in Dutch). *Slow* appeared if the response had taken longer than 1200 msec. The feedback remained on the screen for 2 sec, at which point both the word pair and the feedback disappeared from the screen. The next fixation stimulus appeared 1 sec later. The duration between the onset of the second word of a pair and the button press was registered as the reaction time (RT).

The total set of trials was presented in groups of 24, each followed by a brief pause in which the number of errors and the mean RT for that group were shown on the screen. The word pairs were presented in a random order so that there was a unique presentation order for every participant. There was a short break between the practice session and the test session.

Results

For each participant in each of the four conditions formed by the two levels of the variables position (of the homograph in the translation pair; Position 1 vs. Position 2) and language (of the target reading of the homograph; Dutch vs. English), four mean RTs were calculated: one for the homographs in the high-frequency condition (where the target reading of the homograph was the most frequent of the two homograph readings); a second for the frequency-matched controls of these "high-frequency" homographs; a third for the homographs in the low-frequency condition (where the target reading of the homograph was the least frequent of the homograph's two readings); and a fourth for the frequency-matched controls of these "low-frequency" homographs. We also calculated a mean RT for each critical translation pair collapsed across the 18 participants in each of the four position-by-language conditions. Only correct responses were included in these calculations.

We performed two 2 (stimulus type; homograph pairs vs. control pairs) by 2 (position) by 2 (language) by 2 (frequency; high vs. low) analyses of variance (ANOVAs) on these mean RTs, one by subjects and one by items. This same set of analyses was also performed on the error data. Table 1 shows the mean RT and mean percentage of errors for all 16 conditions in Experiment 1. It also presents all homograph effects. These effects are calculated by subtracting the value for the control condition from the value in the corresponding homograph condition. Not shown are the mean RTs and error rates for the (non-homographic) filler materials. These were 593 msec and 7.29% errors for the Dutch filler trials (that is, the filler trials with Dutch words in second position), and 586 msec and 6.38% errors for the English filler trials (with English words in second position). These scores are in the range of those for the critical control conditions (see Table 1).

Interactions

The second-order interaction between stimulus type, position, and frequency was marginally significant in one analysis and significant in the remaining three: $F_1(1, 68) = 3.25$, $p = .08$, and $F_2(1, 88) = 4.82$, $p < .05$, for RT; $F_1(1, 68) = 16.75$, $p < .0001$, and $F_2(1, 88) = 6.71$, $p < .05$, for errors (F_1 and F_2 concern the analyses by subjects and by items, respectively). As shown in Table 1, an especially large homograph effect was obtained when the low-frequency reading of the homograph had to be selected, and when at the same time the homograph was in Position 1 of the translation pair. The large effects in the error analyses in this particular condition are due to the large error rates for the homographs in this condition (see Table 1). These error rates, near chance level, may suggest that the participants did in fact not know the meanings of these words in the

TABLE 1
Mean response times[a] and error rates[b] for all stimulus type × language × position × frequency conditions (Experiment 1)

		Dutch reading				*English reading*			
		Position 1		*Position 2*		*Position 1*		*Position 2*	
Reading		*RT*	*ER*	*RT*	*ER*	*RT*	*ER*	*RT*	*ER*
High-frequency	Homographs	641	7.27	662	10.27	705	18.71	604	10.11
	Controls	583	4.28	604	5.13	580	2.53	586	2.53
	Effect	58	2.99	58	5.14	125	16.18	18	7.58
Low-frequency	Homographs	782	44.46	730	32.34	823	49.14	710	29.09
	Controls	628	4.55	695	15.17	616	5.99	640	5.99
	Effect	154	39.91	35	17.17	207	43.15	70	23.10

[a] Response times in milliseconds; [b] error rates in percentages.

target language, a conclusion that would invalidate an interpretation of the effect in terms of interference caused by the non-target reading. However, the fact that the error rates for the equally unfamiliar control pairs were considerably lower than those for the corresponding matched homograph pairs argues against this alternative interpretation of the homograph effect in this condition.

This second-order interaction between stimulus type, position, and frequency qualified both the first-order interaction between stimulus type and position ($p < .001$ in all four analyses) and the theoretically important interaction (see the Introduction) between stimulus type and frequency. The latter was significant in three of the analyses ($p < .01$ or lower) and marginally significant in the fourth (the RT analyses by items; $p = .07$); the results are depicted in Figure 1.

As can be seen in Figure 1, the size of the homograph effects differs between the two frequency conditions. The effects were 65 msec and 8.0% errors in the high-frequency condition versus 117 msec and 30.8% errors in the low-frequency condition. The direc-

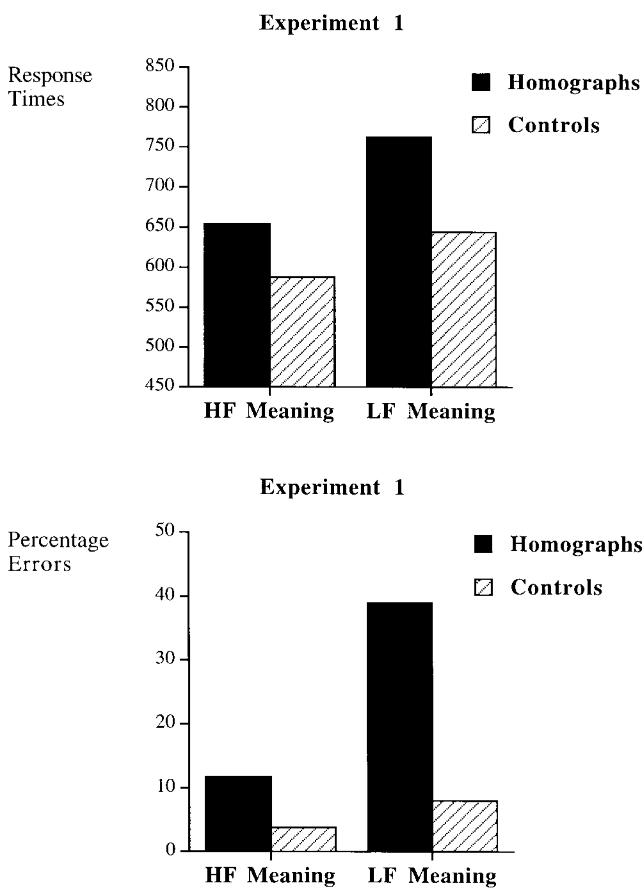

FIG. 1. Mean translation recognition RT (in msec) and errors as a function of stimulus type and frequency.

tion of the differences, with much more inhibition in the low-frequency condition, clearly supports the idea that the inhibitory effect of the non-target reading is particularly large when this reading is the more frequent of the two readings of the homograph (that is, when the least frequent reading has to be selected).

Figure 1 also shows that the mean RTs and error percentages for the high- and low-frequency control conditions differed from one another (an effect that was significant at $p < .05$ in the analyses by subjects, but non-significant in the item analyses). This result is expected, and it merely reflects the fact that the controls for the homographs in the high-frequency condition (matched in frequency with the high-frequency reading of the homograph) had a higher overall language frequency than the controls for the homographs in the low-frequency condition (matched in frequency with the low-frequency reading of the homograph; see Materials section). In other words, the differences in RT and error percentage between the control pairs in the two frequency conditions reflect the common language–frequency effect.

The second-order interaction between stimulus type, language, and position was significant in the RT analysis by items, $F_2(1, 88) = 4.51$, $p < .05$, but not in the remaining analyses ($p > .10$ in all cases): When the homograph was in Position 1 of the stimulus, the homograph effect was much larger when the English reading of the homograph had to be selected (as in the homograph pair *glad–blij*) than when the Dutch reading had to be selected (as in the homograph pair *glad–slippery*); in contrast, when the homograph was in Position 2 (*blij–glad*; *slippery–glad*), the homograph effect was independent of language. The first-order interaction between language and stimulus type was significant in the error analysis by subjects, $F_1(1, 68) = 6.52$, $p < .05$, but not in the remaining three analyses ($p > .10$ in all cases). This interaction pointed to a relatively large homograph effect when the English reading of the homograph had to be selected.

Main effects

The above interactions qualified the three main effects that were significant in at least a subset of the analyses: The main effect of stimulus type was significant in both the RT and the error analyses: $F_1(1, 68) = 68.57$, $p < .0001$, and $F_2(1, 88) = 28.42$, $p < .0001$, for RT; $F_1(1, 68) = 255.29$, $p < .0001$, and $F_2(1, 88) = 54.51$, $p < .0001$, for errors. Overall, RT was 91 msec longer for homograph pairs than for control pairs (707 msec and 616 msec, respectively), and 19.40% more errors were made to homograph pairs (25.17% and 5.77%, respectively). The main effect of position was non-significant in the RT analysis by subjects, $F_1(1, 68) = 0.38$, $p > .10$, but it was significant in the corresponding item analysis, $F_2(1, 88) = 4.54$, $p < .05$. In the error analyses the position effect was significant in the analysis by subjects and marginally significant in the item analysis, $F_1(1, 68) = 4.50$, $p < .05$, and $F_2(1, 88) = 3.53$, $p = .06$. Overall, RT was 26 msec longer when the homograph was in Position 1 than when it was in Position 2 (695 msec and 669 msec, respectively), and 3.29% more errors were made with the homograph in Position 1 than with the homograph in Position 2 (17.12% and 13.83%, respectively). Finally, the main effect of frequency was always significant: $F_1(1, 68) = 70.86$, $p < .0001$, and $F_2(1, 88) = 19.82$, $p < .0001$, for RT; $F_1(1, 68) = 170.73$, $p < .0001$, and $F_2(1, 88) = 35.86$, $p < .0001$, for errors. Overall, RT was 83 msec longer in the low-frequency condition than in the high-frequency condition (703 and 620 msec, respectively) and 15.74% more errors were made in the low-frequency condition (23.34% and 7.60%, respectively). This effect of frequency (as well as the corresponding effects in Experiments 2 and 3) confounds the common language–frequency effect and the effect that is particularly relevant in view of the main questions posed in this article—that is, the relative frequency of the homograph's two readings. Therefore, the interaction between word type and frequency is much more important than this main effect of frequency. Finally, the main effect of language was not significant, $F < 1$ in all four analyses.

Discussion

In line with the predictions, the data showed that the non-target reading of an interlexical homograph is activated in translation recognition and causes inhibition. This inhibition was relatively small when the homograph was the second word within a homograph pair. The cause of this position effect presumably is that a non-homographic first word biases the participant toward the target reading of the subsequent homograph. For example, when *blij* in Position 1 activates its meaning, this increases the probability that the participant will interpret the homograph *glad* as an English word that has the same meaning as *blij* rather than a Dutch word that has a different meaning. The data furthermore showed that the inhibitory effect depends on the relative frequency of the homograph's two readings: The effect was much larger when the non-target reading of the homograph on a particular trial was the more frequent of the homograph's two readings (in other words, when the least frequent reading had to be selected). Because the baseline activation levels for the high-frequency readings are higher than those for the low-frequency readings, a homograph's high-frequency reading will be more available than its low-frequency reading. Thus, there will be a relatively larger delay in processing when the low-frequency reading is the reading to be selected.

There was also a hint in the data that—with the homograph in Position 1—the inhibition was larger when the English reading of the homograph had to be selected than when the Dutch reading was targeted. The cause of this language effect presumably is the fact that in our participants Dutch is the stronger language. The baseline activation levels for Dutch words will therefore generally be higher than those for English words. As a consequence, other things being equal, the homograph will trigger its Dutch reading more easily than its English reading, causing extra inhibition when the English reading has to be selected.

The results very clearly provided a positive answer to the main empirical question addressed in Experiment 1—that is, whether translation recognition (a task that necessarily requires the simultaneous activation of the bilingual's two languages) is hampered when one of the terms in the translation pairs is an interlexical homograph. Thus we can conclude that, in line with the predictions, lexical access is non-selective when participants are performing a translation recognition task. Our next question was whether interlexical homographs are also processed differently from matched controls in language-specific lexical decision—that is, a task that per se does not require the simultaneous activation of the two languages. An affirmative answer would provide additional, and indeed stronger, support for the view that bilingual lexical access is non-selective. The most parsimonious interpretation of a null effect of homography would be that access is language selective in the language-specific lexical decision task. However, as pointed out earlier, there is a way to reconcile an apparent null effect with the non-selective–access view, namely by assuming opposing facilitatory and inhibitory effects that both result from non-selective access.

EXPERIMENT 2

Lexical Decision

Method

Participants

A total of 40 new participants drawn from the same population as that used in Experiment 1 took part; 20 performed the Dutch version of the task and 20 performed the English version.

Materials

The experimental materials were two lists of 192 letter strings each, one for each of the two language conditions, Dutch and English. Of these 192 letter strings per list, 96 were words and 96 were nonwords. The nonwords were all orthographically legal letter strings in the language of that particular condition. However, the nonwords in the Dutch condition were never words in English and, similarly, the nonwords in the English condition were never words in Dutch.

The 96 words within both lists were derived from the experimental materials of Experiment 1. Each of the two lists contained the 24 homographs of that experiment. To these 24 homographs in the Dutch list we added the 24 Dutch words that had served as their controls in Experiment 1. Similarly, to these same 24 homographs in the English list we added the 24 English words that had served as their controls in Experiment 1. Finally, 48 filler words were added to each list, all Dutch words in the Dutch list and all English words in the English list. The 48 filler words of each list were of the same type as the controls (that is, they were not interlexical homographs), but unlike the experimental controls, they were not matched on frequency with the homographs. Again they were included to decrease the overall proportion of interlexical homographs among the stimulus materials. The Dutch and English filler words had served as the Dutch words and the English words, respectively, of the filler translation pairs in Experiment 1.

In addition to the 192 experimental stimuli, every participant was presented with 24 stimuli for practice: 12 words and 12 nonwords, and all Dutch (or Dutch-like, in the case of nonwords) or all English (or English-like), depending upon the condition the participant was in. The practice set did not contain any homographs.

Apparatus and procedure

The apparatus was identical to that used in Experiment 1. The procedure was also as in Experiment 1, except for the following: The participants saw just one letter string per trial instead of a word pair; the language of the feedback corresponded to the language of that condition, Dutch or English; and the instructions were adapted to the specifics of the lexical decision task. The participants in condition Dutch were told that they would be presented with a number of letter strings, and that for each string they would have to decide whether or not it was a Dutch word. Similarly, in condition English the participants were asked to decide for each string whether or not it was an English word.

The participants were not told in advance that a number of the presented letter strings were words in both Dutch and English. Explicit mentioning in the instructions of the occurrence of such stimuli might have had the effect that the non-target language system would have been activated even before presentation of the first stimulus, thus frustrating our goal of finding out whether selective access is at all possible.

Results

For each participant in both language conditions four mean RTs were calculated, one for each of the four cells formed by the two levels of the variable word type (homograph vs. control) and frequency (high vs. low). We also calculated a mean RT for each homograph and control item in both language conditions, collapsing across all participants in a language condition.

On these mean RTs we performed two 2 (language) by 2 (word type) by 2 (frequency) ANOVAs, one by subjects and a second by items. This same set of analyses was performed on the error percentages. Table 2 shows the mean RTs and error rates for all eight cells of these analyses. The data for the (non-homographic) filler materials are not shown. The mean RT and error rate for the fillers in condition Dutch were 517 msec and 2.6% errors, respectively; the corresponding values for the fillers in condition English were 618 msec and 10.9%.

Interactions

The first-order interactions between language on the one hand and word type and frequency on the other were generally non-significant ($p > .10$), with one exception: The interaction between language and word type was significant in the error analysis by subjects, $F_1(1, 38) = 7.20$, $p < .05$. In this same analysis the interaction between language and frequency was marginally significant, $F_1(1, 38) = 3.85$, $p = .06$. The interaction between word type and frequency was non-significant in the RT analyses ($F < 1$ in both cases) and in the error analysis by items ($p > .10$), but it was significant in the error analysis by subjects, $F_1(1, 38) = 10.63$, $p < .01$. These first-order interactions were, however, qualified by the second-order interaction between language, word type, and

TABLE 2
Mean response times[a] and error rates[b] for all word type × language × frequency conditions (Experiment 2)

	Dutch Reading				*English Reading*			
	High-frequency		*Low-frequency*		*High-frequency*		*Low-frequency*	
	RT	*ER*	*RT*	*ER*	*RT*	*ER*	*RT*	*ER*
Homographs	508	1.92	567	22.27	563	5.01	590	12.32
Controls	516	2.69	548	6.36	560	3.64	612	11.94
Effect	−8	−0.77	19	15.91	3	1.37	−22	0.38

[a] Response times in milliseconds; [b] error rates in percentages.

frequency, which was significant for both RT and error percentages in the subject analyses, and marginally significant in the item analyses: $F_1(1, 38) = 6.00$, $p < .05$, and $F_2(1, 88) = 3.06$, $p = .08$, for RT; $F_1(1, 38) = 13.45$, $p < .001$, and $F_2(1, 88) = 2.92$, $p = .09$, for errors. The interaction data (see Table 2) suggest different data patterns for the two language conditions, which are evaluated further later in separate analyses of the two language conditions.

Main effects

The main effect of language was significant in the RT analyses but not in the error analyses: $F_1(1, 38) = 6.39$, $p < .05$, and $F_2(1, 88) = 18.01$, $p < .001$, for RT; both F_1 and $F_2 < 1$ for errors: Overall, the participants responded faster in their native language, Dutch, than in their second language, English (535 msec and 581 msec, respectively). The main effect of frequency was significant in all four analyses: $F_1(1, 38) = 57.72$, $p < .0001$, and $F_2(1, 88) = 21.27$, $p < .0001$, for RT; $F_1(1, 38) = 85.60$, $p < .0001$, and $F_2(1, 88) = 15.61$, $p < .001$, for errors. The participants were faster and made fewer errors in the high-frequency condition (537 msec and 3.31%, respectively) than in the low-frequency condition (579 msec and 13.22%, respectively). Finally, the main effect of word type was significant in the error analysis by subjects, $F_1(1, 38) = 11.46$, $p < .01$, but non-significant in the RT analyses ($F < 1$ in both cases) and in the error analysis by items ($p > .10$). The second-order interaction between language, frequency, and word type (see earlier) qualified these main effects.

Analyses of the separate language conditions

To clarify the role of language, the data of the two language conditions were analysed separately. The results were strikingly different for these two conditions. In condition Dutch (see Table 2, left part) the important interaction between word type and frequency was significant in three out of the four analyses: $F_1(1, 19) = 4.34$, $p = .05$, and $F_2(1, 44) = 2.24$, $p > .10$, for RT; $F_1(1, 19) = 38.19$, $p < .0001$, and $F_2(1, 44) = 4.31$, $p < .05$, for errors: The homograph effect was never significant in the high-frequency condition, but in all three analyses that did show a significant interaction, the homograph effect in the low-frequency condition was statistically significant: Homographs were responded to slower and less accurately than their controls. This interaction qualified the main effect of word type that was statistically significant in the error analysis by subjects, $F_1(1, 19) = 19.00$, $p < .001$, and marginally significant in the error analysis by items, $F_2(1, 44) = 3.51$, $p = .07$, but that was non-significant in the analogous analyses of the RT data, $F_1(1, 19) = 1.16$, $p > .10$, and $F_2 < 1$.

In contrast, in condition English (see Table 2, right part) the interaction between word type and frequency was never significant: $F_1(1, 19) = 2.16$, $p > .10$, and $F_2(1, 44) = 1.01$, $p > .10$, for RT; both F_1 and $F_2 < 1$ for errors, and the same held for the main effect of word type ($F < 1$ in all cases). In other words, no homograph effects occurred in the English condition (simple-effects analyses showed that even the 22-msec facilitation for homographs in the low-frequency condition was not significant). Finally, in both language conditions the main effect of frequency was always significant ($p < .05$ or better in all cases), with faster and more accurate responses in the high-frequency condition than in the low-frequency condition.

Discussion

At first sight, the results of Experiment 2 appear consistent with the idea that bilingual lexical access can be both selective and non-selective, depending upon the target language of the task: The Dutch condition showed statistically significant, frequency-dependent inhibition effects for interlexical homographs, suggesting non-selective lexical access. This finding extends the results of Dijkstra et al. (1998, Experiment 1), who did not include this language condition. In contrast, statistically, the English condition demonstrated no differences between homographs and controls, suggesting language-selective lexical access. This result is essentially the same as the one obtained by Dijkstra et al. for this condition, even though superficially the data of the two studies appear to be conflicting. That is, whereas our study demonstrated a small but non-significant facilitation effect for homographs in the low-frequency condition, Dijkstra et al.'s study showed a small but also non-significant inhibition effect for homographs in the analogous condition.

What is actually a rather puzzling aspect of these results is that support for non-selective access was obtained in condition Dutch but not in condition English. In theory, it should have been easier to block off the weaker language, English, than the stronger language Dutch, because the English words should have had lower baseline levels of activation (see the Introduction). Given that the pattern of results is opposite from the expected pattern, alternative accounts of the data should be considered.

At least one such alternative exists. It holds that the present language-specific lexical decision task was not always carried out according to the instructions, but that some of the time participants treated the task as a language-neutral lexical decision task (presumably without being aware of doing so). As pointed out previously (see Introduction), if a stimulus activates two meanings rather than just one and a response is being based on the availability of any meaning, irrespective of language, homographs will be responded to faster or more accurately than controls. Indeed, the language-neutral lexical decision experiment of Dijkstra et al. (1998, Experiment 3) provided support for this contention. The facilitation obtained for homographs on these trials might counteract any inhibitory effect for homographs on the trials in which the participants actually performed the task as instructed, that is, language specifically. In other words, the absence of a homograph effect does not conclusively support the selective-access view; it is also compatible with the view that bilingual lexical access is in fact always non-selective but that evidence for non-selectivity can fail to emerge due to a mixture of processing modes adopted by the participants.

One additional assumption has to be made in order to provide a complete account of the data obtained in Experiment 2—namely, that in condition Dutch the language-specific task mode was adopted more often than the language-neutral mode (resulting in a small, net inhibition effect for homographs), whereas in the English condition the two processing modes were adopted about equally often (resulting in an apparent null effect). The reason for a more extensive use of the language-neutral mode in condition English than in condition Dutch might lie in the fact that for our participants condition English was the harder of the two language conditions. The participants in condition English may have involuntarily compensated for the extra demands of that condition by applying more often the least constraining, and hence easier, version of the lexical-decision task—that is, the language-neutral version. Although merely hypothetical, some support for this alternative account can be found in the data. The nature of this support will be presented later, when discussing the combined results of Experiments 2 and 3.[3]

The participants in Experiment 2 were implicitly allowed to perform the lexical decision task contrary to the instructions (i.e. in its language-neutral form), because none of the nonwords in either language condition was a word in the non-target language. Thus, there was no penalty, in terms of higher error rates on nonwords, for performing the language-neutral version of the task. However, in Experiment 3 words from the non-target language were included among the stimuli that required a "no" response, thus discouraging the participants from adopting the language-neutral processing mode.

EXPERIMENT 3

Lexical Decision

Method

Participants

A total of 40 new participants drawn from the same population as that used in Experiments 1 and 2 took part: 20 in condition Dutch and the remaining 20 in condition English.

Materials

The experimental materials were identical to those used in Experiment 2, except that 32 of the 96 nonwords in the Dutch list of Experiment 2 were replaced by English words and 32 of the 96 nonwords in the English list were replaced by Dutch words. Experiment 3 thus involved mixed lists of English and Dutch items. In addition to the experimental stimuli, every participant was presented with 40 stimuli for practice. The various types of stimuli occurred in about the same proportion in the practice set as in the experimental set, except that no homographs occurred in the practice set.

Apparatus and procedure

The apparatus and procedure were as in Experiment 2.

[3] Dijkstra et al. (1998, p. 56) tested a slightly different version of this idea, at the time following our suggestion that these counteracting processes might have been operative. The version they tested involved the use of *conscious* strategies on the part of the participants to employ either language-neutral or language-specific lexical decision. They did not find supporting evidence for the use of such conscious strategies, but the possibility that participants involuntarily mix between language-neutral and language-specific lexical decision under circumstances that permit them to do so remains as an option.

Results

In order to see whether the participants complied with the instructions, a first analysis was performed on a subset of the stimuli that required a "no" response. The subset selected for each language condition included all the 32 "no"-response stimuli that were in fact words in the non-target language as well as 32 real nonwords that were drawn from the remaining 64 "no"-response stimuli and that matched the non-target language words in length. If the participants indeed performed a language-specific rather than a language-neutral lexical decision task, the error rate for the non-target language words should not deviate noticeably from the error rate for the matched set of real nonwords. If, however, the participants at times performed a language-neutral lexical decision task, the non-target language words would have invited incorrect "yes" responses and should therefore show a larger error rate than the real nonwords. Analyses of the error percentages by subjects and by items indicated that the participants had indeed complied with the instructions. Collapsed across the two language conditions, real nonwords and non-target language words showed error percentages of 6.9% and 5.7%, respectively, a difference that was not significant ($p > .10$). This held for both language conditions. The only significant result was a main effect of language ($p < .01$), with fewer errors made in the Dutch condition (3.9%) than in the English condition (8.7%).

Having thus verified that our participants had indeed, as instructed, performed the language-specific version of the lexical decision task, the same 2 (language) by 2 (frequency) by 2 (word type) analyses as those performed for Experiment 2 were run on the word data. The results of these analyses differed substantially from those obtained for Experiment 2. Table 3 shows all the relevant cell means. The mean RTs and error scores for the (non-homographic) filler trials are not shown. These were 543 msec and 3.9% errors for the fillers in condition Dutch, and 596 msec and 7.1% errors for those in condition English.

Interactions

As in Experiment 2, the first-order interactions between language on the one hand and word type and frequency on the other were generally non-significant ($p > .10$). The only exception was the interaction between language and frequency, which was significant in

TABLE 3
Mean response times[a] and error rates[b] for all word type × language × frequency conditions (Experiment 3)

	Dutch Reading				*English Reading*			
	High-frequency		*Low-frequency*		*High-frequency*		*Low-frequency*	
	RT	*ER*	*RT*	*ER*	*RT*	*ER*	*RT*	*ER*
Homographs	548	3.85	628	37.73	609	9.09	692	31.92
Controls	528	1.15	551	6.82	577	2.27	584	5.38
Effect	20	2.70	77	30.91	32	6.82	108	26.54

[a] Response times in milliseconds; [b] error rates in percentages.

the error analysis by subjects, $F_1(1, 38) = 5.81$, $p < .05$ (but not in the remaining three analyses).

However, unlike in Experiment 2, the first-order interaction between word type and frequency was always significant: $F_1(1, 38) = 18.95$, $p < .0001$, and $F_2(1, 88) = 10.93$, $p < .01$, for RT; $F_1(1, 38) = 101.18$, $p < .0001$, and $F_2(1, 88) = 21.06$, $p < .0001$, for errors. In the high-frequency condition the homograph effect was significant in the error analyses only by subjects ($p < .05$). In contrast, in all four analyses the homograph effect in the low-frequency condition was statistically significant, with homographs being responded to slower and less accurately (RT: 660 msec; errors 34.82%) than their controls (567 msec and 6.10%, respectively).

Finally, a marginally significant second-order interaction between language, word type, and frequency occurred in the error analysis by subjects, $F_1(1, 38) = 3.18$, $p = .08$. In the remaining three analyses this second-order interaction was not significant ($p > .10$ in all cases). In this respect the data differed critically from those of Experiment 2, where this interaction had been significant for both RT and errors in the analyses by subjects and marginally significant in both analyses by items.

Main effects

As in Experiment 2, the main effect of language was significant in the RT analyses but not in the error analyses: $F_1(1, 38) = 5.66$, $p < .05$, and $F_2(1, 88) = 22.33$, $p < .001$, for RT; both F_1 and $F_2 < 1$ for errors: Overall, the participants responded faster in their native language, Dutch, than in their second language, English (564 msec and 615 msec, respectively). The main effect of frequency was (as in Experiment 2) significant in all four analyses: $F_1(1, 38) = 36.98$, $p < .0001$, and $F_2(1, 88) = 24.36$, $p < .0001$, for RT; $F_1(1, 38) = 134.67$, $p < .0001$, and $F_2(1, 88) = 39.30$, $p < .001$, for errors: The participants responded faster and more accurately in the high-frequency condition (RT: 565 msec; errors: 4.09%, respectively) than in the low-frequency condition (RT: 614 msec; errors: 20.46%, respectively). Finally, whereas in Experiment 2 the main effect of word type had been significant only in the error analyses, it was now significant in all four analyses: $F_1(1, 38) = 55.38$, $p < .001$, and $F_2(1, 88) = 35.04$, $p < .0001$, for RT; $F_1(1, 38) = 90.50$, $p < .0001$, and $F_2(1, 88) = 41.08$, $p < .0001$, for errors: Homographs were processed more slowly and less accurately (RT: 619 msec; errors: 20.65%) than their controls (RT: 560 msec; errors: 3.91%). The main effects of frequency and word type were qualified by the interaction between these two variables (see earlier).

Analyses of the separate language conditions

The above analyses suggest that, unlike in Experiment 2, the two language conditions produced essentially the same pattern of results. This conclusion is substantiated by the results of a set of analyses performed on the data of the two language conditions separately. The main effects of word type and frequency were significant in both language conditions, and in all four analyses ($p < .01$ or better in all cases). Furthermore, the interaction between these two variables was significant in all four analyses performed on the English data set and in three out of the four performed on the Dutch data set ($p < .01$

or better in all cases); it was marginally significant ($p = .06$) in the fourth analysis of the Dutch set (the RT analysis by items). In both language conditions interlexical homographs were more difficult to process than their matched controls, but only when their least frequent reading had to be selected.

Discussion

The occurrence of robust, frequency-dependent homograph effects supports the idea that under the circumstances of this experiment lexical access is language non-selective, both when the participants' native and strongest language, Dutch, is the target language and when their strongest non-native language, English, is the target language.

The low error percentage for "no" response stimuli that are in fact words in the non-target language indicates that the participants in Experiment 3 have performed a language-specific lexical decision task, as they were instructed to do. Earlier it was suggested that in Experiment 2 a mixture of language-specific and language-neutral task performance may have taken place. Given non-selective access, the former would have caused inhibition for homographs and the latter would have caused facilitation with the result being an overall small or null homograph effect. We also suggested that the least demanding, because it is the least constraining, language-neutral mode may have been adopted (presumably not consciously but involuntarily) more often in condition English than in condition Dutch to compensate for the fact that processing non-native English is harder than processing native Dutch. This could explain why a small frequency-dependent inhibitory effect on homographs occurred in condition Dutch of Experiment 2 but not in condition English. The following two analyses provide some support for the validity of these hypotheses.

Comparison of Experiments 2 and 3

Homograph effects. We calculated the size and direction of the homograph effects for all participants in the low-frequency condition of Experiments 2 and 3 (the condition that showed large effects of the word type manipulation). In Experiment 2, 18 out of the 40 participants showed facilitation for homographs, and 22 showed inhibition; of the 18 participants showing facilitation, 12 had had English stimuli, and only 6 had had Dutch stimuli. Experiment 3 showed a different pattern, with 36 out of the 40 participants showing inhibition for homographs (17 in condition Dutch and 19 in condition English). These data point out that in the Dutch condition of Experiment 2, the non-selective language-neutral processing mode was adopted relatively often compared to Experiment 3, but less often than in the English condition of Experiment 2.

These suggestions were supported in a 2 (experiment: Experiment 2 vs. Experiment 3) by 2 (language) ANOVA on the homograph effects of all participants in the low-frequency conditions of Experiments 2 and 3. The main effect of experiment was highly significant, $F_1(1, 76) = 36.29$, $p < .0001$, with an overall facilitatory homograph effect of 1 msec in Experiment 2 and an overall inhibitory homograph effect of 93 msec in Experiment 3. The main effect of language was non-significant, $F < 1$, with overall inhibitory homograph effects of 49 msec and 43 msec in condition Dutch and condition

English, respectively. Most importantly, a significant interaction was obtained between experiment and language, $F_1(1, 76) = 5.39$, $p < .05$. Simple contrasts analysis on the means of the four cells of the interaction showed that the inhibitory effects of 77 msec and 108 msec in the Dutch and English conditions of Experiment 3 did not differ significantly from one another, $F_1(1, 76) = 1.92$, $p > .10$, suggesting that participants performed the task in the same way (language specifically) in the two language conditions of Experiment 3. In contrast, in this analysis the difference between the inhibitory effect of 19 msec in condition Dutch of Experiment 2 and the facilitatory effect of 22 msec in condition English of Experiment 2 nearly approached statistical significance, $F_1(1, 76) = 3.60$, $p = .06$. This finding supports the idea that the participants behaved differently in the two language conditions of this experiment, with, as hypothesized, more participants performing the tasks predominantly language neutrally in condition English. Finally, for both condition Dutch and condition English the homograph effects differed significantly between Experiments 2 and 3 ($p < .01$ in both cases), a finding that provides support for the idea that the participants performed the task differently in the two experiments (i.e. more often language neutrally in Experiment 2 than in Experiment 3).

Processing time. If indeed the participants in Experiment 2 often ignored the instructions and performed language-neutral lexical decisions, whereas those in Experiment 3 generally performed language-specific lexical decisions, homograph processing should have taken longer in Experiment 3 than in Experiment 2. The basis for this prediction is that only in language-specific lexical decision does the language of the accessed meaning of the homograph have to be evaluated, a processing stage that would, presumably, follow lexical access (Grainger & Dijkstra, 1992). The extra processing stage in Experiment 3 should therefore have increased processing time. To examine these ideas, a further set of analyses was run that combined the RT data of Experiments 2 and 3. These analyses involved 2 (experiment) by 2 (word type) by 2 (frequency) by 2 (language) ANOVAs.

The main effect of experiment was significant in both analyses, $F_1(1, 76) = 4.88$, $p < .05$, and $F_2(1, 88) = 32.47$, $p < .001$. The participants were faster in Experiment 2 than in Experiment 3 (558 msec vs. 590 msec). The interaction between experiment and word type was also significant, $F_1(1, 76) = 36.84$, $p < .0001$, and $F_2(1, 88) = 31.11$, $p < .001$. It showed that only the homographs were affected by the extra constraint of Experiment 3: The control stimuli were processed equally rapidly in Experiments 2 and 3 (559 msec and 560 msec, respectively), whereas homographs were responded to faster in Experiment 2 (557 msec) than in Experiment 3 (619 msec). Thus, the important result of the present analysis is that it suggests that homographs are processed differently in our Experiments 2 and 3. Although Dijkstra et al. (1998) did not report overall analyses of their Experiment 1 (which permitted language-neutral task performance) and Experiment 2 (which required language-specific task performance), the RT data that they reported for these experiments separately converge with the present data: The overall RT for the control stimuli was 580 msec in their Experiment 1 and 583 msec in their Experiment 2; the overall RT for the homographs was 580 msec in their Experiment 1 and 618 msec in their Experiment 2.

The other results of these analyses substantiate the conclusions that we drew from the analyses of the separate experiments. They are therefore not repeated here.

GENERAL DISCUSSION

The present experiments support the view that lexical access in bilinguals is non-selective not only when the bilingual participants perform translation recognition, a task that can only be performed if the bilingual's two language systems are simultaneously activated, but also when they perform lexical decision, a task that per se does not require the simultaneous activation of both language systems. In all three experiments, statistically significant inhibitory effects for homographs were obtained, although in Experiment 2 the effect occurred only in condition Dutch and was relatively small. Furthermore, in all cases, the effect was dependent on the relative frequency of the homograph's two readings: It was particularly large when the low-frequency reading of the homograph had to be selected, and, in fact, in Experiment 2 the effect occurred only in the low-frequency Dutch condition. It thus appears that the relative baseline level of the activation of the homograph's two readings determines the pattern of responses. A large inhibitory effect occurs, especially when the targeted reading is the one with the lower baseline activation level. This conclusion is also consistent with the trends in the data of Experiment 1, which indicated that the inhibitory effects were especially large when the weaker of the two languages was targeted.

We suggested that the different patterns of results obtained in Experiments 2 and 3, in both cases involving a lexical decision task, may have arisen from a language-neutral processing mode adopted by the participants in Experiment 2 on at least a subset of the trials. If bilingual lexical access is non-selective, this processing mode should result in facilitation for homographs, an effect that would counteract any inhibitory effects arising on the trials where the language-specific processing mode was adopted. The language-neutral processing mode was not a feasible option in Experiment 3, where the participants would be penalized by high error rates on "no"-response stimuli that were words in the non-target language. A comparison of the homograph effects and of response times in Experiments 2 and 3 was consistent with this explanation of the differences between experiments. The important conclusion is that the non-selective access view can account for the complete pattern of results obtained in this study.[4]

---

[4] According to the present assumption, participants instructed to perform language-specific lexical decisions in fact carry out the language-neutral version of the task at least some of the time when the experimental circumstances permit them to do so (as in Experiment 2). The opposite situation, that participants instructed to perform language-neutral lexical decisions in fact carry out the language-specific version of the task, may also arise. This would explain an unpublished result obtained by French and Ohnesorge (1996). In one of their experimental conditions 45 French–English homographs were embedded in a long list of exclusively French words and French-like nonwords (360 in all). In this condition the RT for homographs of low frequency in French and high frequency in English was considerably longer than the RT for homographs of high frequency in French and low frequency in English (912 msec and 742 msec, respectively). This result suggests that in this particular condition the participants adopted, contrary to the instructions, a language-specific French task mode, and they were thus hindered by the high-frequency English reading of the homographs. Had they adopted the language-neutral mode, the two types of homographs should have produced equally long RTs (similar in size to

Because language-neutral lexical decision was also an option in Experiment 1 of Dijkstra et al.'s (1998) study, the null effect of homography obtained by these authors may also have arisen from the mixed use of the two counteracting processing modes assumed here, both of which reflect non-selective access. Indeed, these authors acknowledged this possibility and tested then rejected one particular version of this idea (see Footnote 3). The interpretation of the null effect that they preferred is that it resulted from fast suppression of initially activated word readings belonging to the non-target language (see Dijkstra et al., 1998, p. 61). The suppression was in fact so fast that the non-target reading of the homograph did not produce a noticeable delay in processing. The authors argued the suppression could be this fast because, due to the instructions and the fact that no stimuli were presented that were words in the non-target language exclusively, the target language system was relatively highly activated.

In Dijkstra et al.'s (1998) account of the different results of their Experiment 1 (a null effect of homography) and Experiment 2 (inhibition for interlexical homographs) the exact composition of the stimulus sets thus seems to play an important role. However, a detailed analysis of these stimulus sets suggests that the difference in stimulus list composition between these two experiments was rather subtle, possibly too subtle to produce the different patterns of results. Specifically, 25% of all the stimuli in both Experiment 1 and Experiment 2 were, exclusively or non-exclusively, words in the non-target language. In Experiment 1 these were the interlexical homographs and a set of cognate words—that is, words that, as interlexical homographs, share form across languages but that also, unlike interlexical homographs, share cross-language meaning. In Experiment 2 these were the interlexical homographs and the stimuli that belonged to the non-target language exclusively. It is thus plausible that the non-target language system was activated to the same extent in these two experiments. It is therefore not immediately obvious how the activation in the non-target language system could have been suppressed without causing a noticeable inhibitory effect in one of these two experiments but not in the other.

One result from the present experiments in particular indicates that indeed, as suggested here, the make-up of the stimulus set may not play a crucial role in obtaining a homograph effect: Although we did not obtain a significant effect in condition English of our Experiment 2, we did obtain a statistically significant inhibitory effect in the Dutch condition of that experiment. This result was achieved despite the fact that the composition of the stimulus materials was exactly the same in the two language conditions (12.5% homographs in a set of 192 stimuli, words and nonwords combined). Also, note that the inhibitory effect for homographs in condition Dutch occurred despite the fact that the proportion of target language words that were at the same time words in the non-target language (here English) was substantially smaller than the analogous proportion in

the shortest of the RTs in the language-specific mode), because, given a horse-race model of word recognition, the most frequent reading, whether English or French, should have determined the response. In fact, the latter response pattern was indeed obtained in a condition where the homographs were embedded in a list with equal proportions of English and French stimuli. This suggests that in this "mixed" condition the task was performed according to the instructions—that is, in a language-neutral mode.

Dijkstra et al.'s Experiment 1 (12.5% vs. 25%). If the make-up of the stimulus set was (partly) responsible for the null effect of homography in Dijkstra et al.'s (1998) Experiment 1, our condition Dutch, with more exclusively Dutch items and, hence, a relatively high level of activation of the Dutch language system, should also have produced a null effect of homography.

In sum, it appears that Dijkstra et al.'s interpretation of the null effect of homography in terms of rapid suppression of the non-target language homograph readings cannot comfortably be in agreement with the fact that inhibition was obtained in condition Dutch of our Experiment 1. The present account of the null effects as resulting from a combination of language-neutral and language-specific processing modes that counteract one another seems a plausible alternative. The most important aspect of this view is that it holds that bilingual lexical access has been non-selective even when the homograph data suggested selectivity of access. Maybe the strongest support of this strong non-selective–access position is the fact that large inhibitory effects occurred in the present Experiment 3 and in Experiment 2 of Dijkstra et al. (1998), where shutting off the non-target language would have benefited performance greatly. The occurrence of these effects indicates that shutting off the non-target language is not a feasible option (see also Dijkstra et al. 1998, p. 58).

Other support for non-selective access

A semantic-priming study by Beauvillain and Grainger (1987) has also provided support for the non-selective–access view. These authors presented French–English interlexical homographs and non-homographic French words as primes preceding English targets. Even though the participants were led to believe that all primes were words in French only, the English readings of the homographic primes exerted a semantic priming effect on lexical decisions to English target words (e.g. prime: *coin*, meaning "corner" in French; target: *money*) that were presented 150 msec after the onsets of the primes. A relevant further finding was that the effect disappeared when the onset asynchrony between primes and targets was lengthened to 750 msec. These data suggest that initially both of the homograph's two meanings were activated, and that later the contextually inappropriate English meaning was somehow suppressed or rejected. The results in fact square with the outcome of studies on the processing of words that are ambiguous *within* a language. Those studies (e.g. Gernsbacher & Faust, 1991; Hino, Lupker, & Sears, 1997; Swinney, 1979) have also shown that early in processing both meanings of a polysemous word are available and that later only the contextually appropriate reading is still available.

Even the results of a language-specific lexical-decision study by Gerard and Scarborough (1989), that led the authors to propose, contrary to the present view, that bilingual lexical access is selective, seem to a large extent compatible with the present results and account. A major finding was that their (English–Spanish) bilingual participants responded either quickly or slowly to the same interlexical homograph, depending on whether the homograph was a high-frequency word or a low-frequency word in the target language. This result is, of course, consistent with the non-selective view defended here: When the high-frequency reading must be selected, the response can be fast because this reading will be highly available due to its high baseline activation level. When, however, the low-frequency reading has to be selected, the inappropriate reading, more highly available because of its higher baseline activation level, will delay responding. A further finding was that low-frequency homographs were recognized slower than low-frequency controls and low-frequency cognates. The authors attributed this effect of word type to poorly matched word frequency between the three word groups, but in fact the non-selective bilingual-access view readily accounts for this result as well.

Further support for the non-selective–access view comes from three studies that have shown, in different ways, that language stimuli give rise to automatic phonological activation of both of the languages of a bilingual, even under circumstances where only one of the languages is targeted; in other words, when language-specific task performance is invited (Altenberg & Cairns, 1983; Nas, 1983; Tzelgov, Henik, Sneg, & Baruch, 1996). Similarly, Lukatela, Savic, Gligorijevic, Ognjenovic, and Turvey (1978) obtained support for automatic phonological activation of both the targeted and the non-targeted alphabetical system in bi-alphabetic language users, and Lam, Perfetti, and Bell (1991) obtained the analogous result for bidialectal language users.

Another source of evidence that bilingual lexical access is non-selective comes from a number of studies that extended monolingual research on the effects of a word's neighbourhood characteristics on word recognition (e.g. Coltheart, Davelaar, Jonasson, & Besner, 1977; see Andrews, 1997, for a review) to the bilingual case (Grainger & Dijkstra, 1992; Grainger & O'Regan, 1992; Van Heuven et al., 1998). In the pertinent monolingual studies a word's neighbourhood is typically defined as all the other words in the target language that share all but one letter with the target word. In bilingual neighbourhood studies a word's neighbours in both languages are taken into account. Both Grainger and Dijkstra (1992) and Van Heuven et al. (1998) have shown that the performance of bilinguals in the target language is affected by the number of neighbours in the non-target language.

Finally, lexical decisions in the non-dominant language of a bilingual (Caramazza & Brones, 1979; Dijkstra et al., 1998, Experiment 1; Van Hell, 1998, chap. 4) and within-language word associations (Van Hell & De Groot, 1998) are produced faster to cognates than to non-cognate controls. These findings are also consistent with the non-selective–access view, because they suggest an involvement of the lexicon of the non-target language in task performance. However, the view advanced by several authors (e.g. Kirsner, Lalor, & Hird, 1993; Sánchez-Casas, Davis, & García-Albea, 1992) that cognates but not non-cognates share a representation between the two languages in bilingual memory, suggests that a word frequency effect may in fact underlie the effect of cognate status: Given a shared representation in bilingual memory for cognates—a representation that is contacted each time each of the two words in a cognate translation pair is encountered—the frequencies of use in the separate languages would need to be summed up when determining the usage frequency for cognates. As a consequence, frequencies of cognates will always be higher than those of non-cognate control words matched with the cognates on the basis of language-specific frequency information. The cognate data are therefore consistent with the non-selective–access view, but do not support it conclusively.

Support for language-selective access

The support for the selective-access view comes primarily from studies that presented paragraphs or sentences in either a unilingual form or in a mixed ("code-switched") form (where the input alternates between the participants' two languages (e.g. Kolers, 1966; Macnamara & Kushnir, 1971; Soares & Grosjean, 1984). These studies showed that processing was more difficult when the stimuli (paragraphs or sentences) were presented in the mixed form. The authors interpreted these results as support for the existence of an automatically operating input switch that directs the incoming information to the appropriate linguistic system. The switching operation takes time, hence the longer comprehension times for mixed sentences and paragraphs.

Grainger and Beauvillain (1987) obtained the analogous result when presenting the participants with words and nonwords in a lexical-decision experiment. Except when the stimulus contained orthography-specific information (see later), lexical decision times were longer in the mixed condition than in the unilingual condition. This finding is consistent with the idea that the language-context information provided by the unilingual lists always directs the incoming information straight to the appropriate language system, which is then searched immediately. In the mixed lists the incoming information may initially be directed to the inappropriate language system, and only after a search through this system has failed will the information be directed to the appropriate language system. Response times should therefore be longer in the mixed condition than in the unilingual condition.

The finding of longer response times in the mixed condition of Grainger and Beauvillain (1987) thus appears, at first glance, to provide additional support for the input switch hypothesis, and, by implication, for language-selective access. Grainger and Beauvillain (1987) and Grainger (1993), however, developed two accounts of their results that are completely compatible with the view of non-selective bilingual lexical access. One account is in terms of the Bilingual Interactive Activation (BIA) model, which plays a prominent role in much of the more recent work on bilingual lexical access (Dijkstra & Van Heuven, 1998; Dijkstra et al., 1998; Van Heuven et al., 1998; see e.g. Grainger, 1993, for the second account). The BIA model extends the monolingual Interactive Activation model of McClelland and Rumelhart (1981) to the bilingual case. One of the extensions concerns an additional representational layer, which contains just two "language nodes", one for each of the bilingual's two languages. Irrespective of language context, a word activates (through the feature and letter levels) word representations from both languages. In the newest, implemented, version of the model (Van Heuven et al., 1998), a language node receives activation from activated word representations from the corresponding language and sends back inhibitory excitation to the word representations of the other language. On a trial where the language of the stimulus word is different from that of the previous trial the language node for the non-target language will be in a state of higher activation at the beginning of the trial than the language node for the target language. This will inhibit processing of the target as compared to processing the target on a no-switch trial. Grainger and Beauvillain (1987) observed that the cost of switching languages disappeared when the stimulus word contained language-specific orthographic information. The reason presumably is that orthography-specific words only activate word representations from their own language (Grainger, 1993).

The important point of Grainger and Beauvillain's (1987) account of the effects of language switching is that it assumes that bilingual lexical access is initially non-selective. In other words, even the language-switch data, which have been regarded by some as the quintessential support for the selective-access position, may in fact be compatible with the non-selective–access view.

Processing in translation recognition and lexical decision: The role of meaning

Given the conclusion that bilingual lexical access is non-selective, the question should be raised as to exactly how multiple access affects processing in the tasks used in the present set of experiments. In this section we present a hypothetical account of our results in which meaning representation and meaning activation play a pivotal role. One reason that we assign such a central role to meaning in interlexical homograph processing is that word meaning has been shown to play a central role in studies that look at lexical ambiguity resolution within a language. It is most parsimonious to regard cross-language ambiguity and within-language ambiguity as two special cases of lexical ambiguity in general and, hence, to argue that these two types of ambiguity are resolved in similar ways (i.e. based on the activation of meaning representations). A second reason for focusing on meaning is the fact that assigning meaning to language input is the primary purpose of receptive language use. If the goal of running experiments is to be informed about natural language processes, the role that meaning plays in the experimental tasks that we use to become so informed will ultimately have to be taken into account. Finally, a third reason is that both the translation recognition task and the lexical decision task have been shown to involve the processing of meaning (see e.g. Altarriba & Mathis, 1997; De Groot, 1992b; De Groot & Comijs, 1995; Talamas et al., 1998, for translation recognition; Balota, 1994; Balota, Ferraro, & Connor, 1991, for lexical decision).

In all then there are enough grounds for assigning to meaning a pivotal role in the processing of interlexical homographs. It should, however, be noted that the present experiments were not designed to test this account, and that it therefore awaits empirical validation.

Translation recognition

In the translation recognition task the word-form representations of the native and foreign words presented on a trial are first activated (presumably following two earlier stages where the visual features and the letter representations for the constituent letters are activated; e.g. McClelland & Rumelhart, 1981). These word-form representations subsequently activate the meaning (or "conceptual") representations onto which they map. For purposes of discussion, let us assume distributed conceptual representations in memory, where the meaning of a word is represented across a set of nodes each of which represent one semantic feature of the word (Masson, 1991). If the two words are translation equivalents, the word-form representations of the two words will map onto a set of semantic feature nodes that is shared between the two words in the pair (De Groot, 1992a; Kroll & De Groot, 1997; Taylor & Taylor, 1990; Van Hell, 1998). (In addition, each of the words in a translation pair is likely to map onto a small set of language-specific, non-shared features, due to the fact that the meaning of translation "equivalent" words is seldom, if ever, completely the same.) As a result, the two words in a translation pair will, together, presumably activate this set of shared feature nodes to a larger extent than when the set of feature nodes is activated by a single source. In contrast, the two words of a non-translation pair will map onto two different sets of feature nodes. Thus, each set of feature nodes will be activated to a smaller degree than the set activated from two sources in the case of translation pairs. A positive decision could then be based on a high level of activation in one and the same set of nodes, whereas a negative decision could be based on a relatively small amount of activation in two different sets.

However, under the specific circumstances of Experiment 1, a relatively small amount of activation in two different semantic feature sets does not always constitute a solid basis for a correct negative decision. Specifically, in the case of a low-frequency homograph pair, the inappropriate high-frequency reading of the homograph (e.g. the English reading of *glad*) will activate its associated semantic feature set first, whereas the other word in this stimulus pair (*slippery*) activates a different feature set. In this case, responding on the basis of activation in two different feature sets would result in an incorrect "no" response.

Therefore, a second requirement for correct performance (under the specific circumstances of our Experiment 1) is that memory nodes are activated that represent the information that two different languages are involved, and that this information is taken into account in the decision process. The relevant nodes may either be the "language nodes" assumed in the BIA model discussed previously (Dijkstra & Van Heuven, 1998; Dijkstra et al., 1998; Grainger, 1993; Van Heuven et al., 1998) or nodes representing so called "language tags" (e.g. Green, 1986, 1998; Poulisse, 1997). If a word pair in our Experiment 1 gives rise to activation in two sets of semantic feature nodes, the participant can only safely conclude that the word pair is not a translation pair if at the same time the nodes indicating that two languages are involved are also activated.

The large number of errors on homographs in Experiment 1, especially in the low-frequency condition, strongly suggests that our participants often failed to take language information into account. More specifically, it suggests that a negative decision was often prematurely based on the presence of activation across two different sets of semantic feature nodes—one corresponding to the inappropriate, high-frequency reading of the homograph (a pattern of activation that will be established before the set of semantic features corresponding to the homograph's low-frequency reading will be activated), and a second corresponding to the non-homographic second word of the translation pair. The relatively long response times for correct responses to homograph pairs (especially low-frequency pairs) suggest that on trials where the language information *is* taken into account, the earlier activation in the inappropriate semantic feature set interferes with processing. Whether this inhibition results from inhibitory connections that are hard-wired within the bilingual language system (cf. the BIA model's account of the effects of language mixing, as discussed earlier), or from a time-consuming attentional suppression mechanism external to the language system (cf. Green, 1998), is an issue that remains to be resolved.

Lexical decision

We have argued that a mixture of language-neutral and language-specific task performance caused the pattern of data obtained in Experiment 2, and that predominant language-specific task performance underlay the pattern of results obtained in Experiment 3. In terms of the present hypothetical theoretical framework—which assigns an important role to the specific patterns of activation at the conceptual level of representation in task performance—a language-neutral lexical decision response could be based on the occurrence or absence of a critical amount of activation in any set of semantic feature nodes. If a critical amount of activation is detected, irrespective of the source of activation, the participant can conclude that the stimulus represents a word, without taking language membership into account. Given non-selective bilingual lexical access, the semantic feature set representing the high-frequency reading of the homograph will be activated before that representing the low-frequency reading. Furthermore, in the low-frequency condition, the semantic feature set of a homograph's non-targeted high-frequency reading will generally be activated in less time than the (only) feature set that stores the meaning of the corresponding non-homographic control word (matched with the targeted low-frequency reading of the homograph). This state of affairs results in faster processing of homographs than controls in language-neutral task performance.

In contrast, by definition, language *is* taken into account in language-specific lexical decision, as is suggested, for instance, by the high percentage of errors to homographs in the low-frequency condition of Experiment 3 (but also by the pattern of responses to the stimuli that required a "no" response, see earlier): The set of semantic feature nodes corresponding to the high-frequency reading of the homograph is first activated, and the corresponding language information (the language node or the language tag) is subsequently also activated and taken into account in the decision process. Because the activated language node or language tag is not the one that was targeted, the participant may prematurely respond "no". On other trials the semantic feature set corresponding to the targeted low-frequency reading and the memory nodes containing the associated language information are both activated in time—that is, before the (incorrect) response based on the earlier activation associated with the inappropriate reading has already been executed. On these trials a correct "yes" response is emitted. The fact that—as in the low-frequency translation-recognition condition—correct responses to homographs in the low-frequency condition are slow (as compared to the responses to their frequency-matched controls) again suggests that the earlier activation in a non-targeted part of the system interferes with their processing. As pointed out above, either hard-wired inhibitory connections in the language system or a time-consuming external suppression mechanism may be responsible for this.

Conclusion

Our main conclusion is that the complete pattern of results that we obtained can be readily interpreted in terms of the non-selective access view. A secondary conclusion is that the composition of the stimulus set does not appear to play the important role that Dijkstra et al. (1998) assigned to it, and that the present, relatively small, differences between stimulus sets do not differentially affect the activation levels of the bilingual's two languages. Instead, the differences between the data patterns of the three experiments can all be attributed to differences in task demands. Although completely compatible with the non-selective access view, our data do not yet warrant the much stronger conclusion that bilingual lexical access is *always* non-selective, either in comprehension tasks or in production tasks. On the basis of the present data, for instance, the possibility could not be ruled out that an even smaller proportion of words from the non-target language might allow the non-target language system to be blocked off completely. On the other hand, if, under such extreme circumstances, there was still evidence that access was non-selective, the conclusion would be warranted that in comprehension in a "monolingual" processing mode (Grosjean, 1994, 1995) the non-target language system cannot be completely deactivated. Further, if a condition with very few non-target language words would produce an equally large influence of the non-target language as a condition with more non-target language words, one could conclude that the activation levels of the non-target language are completely beyond the participants' control. Such a finding would, of course, provide the strongest type of support for the non-selective access view.[5]

REFERENCES

Altarriba, J., & Mathis, K.M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language*, *36*, 550–568.

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

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

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

Balota, D.A. (1994). Visual word recognition: The journey from features to meaning. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 303–358). San Diego: Academic Press.

Balota, D.A., Ferraro, F.R., & Connor, L.T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P.J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, *26*, 658–672.

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

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). New York: Academic Press.

De Groot, A.M.B. (1992a). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 389–412). Amsterdam: North Holland.

De Groot, A.M.B. (1992b). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1001–1018.

[5] In fact, recently collected new data indeed suggest non-selective access under such extreme circumstances. Dutch–English bilinguals each named 440 English words, one of which happened to be an interlexical homograph with a Dutch reading that was more frequent than its English reading (*boot*). Nearly all of the participants generated the Dutch pronunciation of this word, which has a different vowel sound than its pronunciation in English.

De Groot, A.M.B. (1998). La répresentation lexico-sémantique et l'accèss lexical chez le bilingue. *Psychologie Française*, *43*, 297–312.
De Groot, A.M.B., & Comijs, H. (1995). Translation recognition and translation production: Comparing a new and an old tool in the study of bilingualism. *Language Learning*, *45*, 467–509.
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). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
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.
French, R.M., & Ohnesorge, C. (1996). *The interactive-activation model of bilingual memory: Support from noncognate interlexical homographs and nonwords.* Unpublished manuscript.
Gerard, L.D., & Scarborough, D.L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 305–315.
Gernsbacher, M.A., & Faust, M.E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 245–262.
Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 11–25). Amsterdam: John Benjamins.
Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology*, *39A*, 295–319.
Grainger, J., & Dijkstra, A. (1992). On the representation and use of language information in bilinguals. In R.J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: North-Holland.
Grainger, J., & O'Regan, J.K. (1992). A psychophysical investigation of language priming effects in two English–French bilinguals. *European Journal of Cognitive Psychology*, *4*, 323–339.
Green, D.W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210–223.
Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.
Grosjean, F. (1994). Individual bilingualism. *The encyclopedia of language and linguistics* (pp. 1656–1660). Oxford: Pergamon Press.
Grosjean, F. (1995). A psycholinguistic approach to code-switching: The recognition of guest words by bilinguals. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages* (pp. 259–275). Cambridge University Press.
Hino, Y., Lupker, S.J., & Sears, C.R. (1997). The effects of word association and meaning frequency in a cross-modal lexical decision task: Is the priming due to "semantic" activation? *Canadian Journal of Experimental Psychology*, *51*, 181–194.
Kirsner, K., Lalor, E., & Hird, K. (1993). The bilingual lexicon: Exercise, meaning and morphology. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 215–248). Amsterdam: John Benjamins.
Kolers, P.A. (1966). Reading and talking bilingually. *The American Journal of Psychology*, *79*, 357–376.
Kroll, J.F., & De Groot, A.M.B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A.M.B. de Groot & J.F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Lam, A.S.L., Perfetti, C.A., & Bell, L. (1991). Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics*, *12*, 299–311.
Lukatela, G., Savic, M., Gligorijevic, B., Ognjenovic, P., & Turvey, M.T. (1978). Bi-alphabetical lexical decision. *Language and Speech*, *21*, 142–166.
Macnamara, J., & Kushnir, S.L. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, *10*, 480–487.
Masson, M.E.J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1: An account of basic findings. *Psychological Review*, *88*, 375–407.
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.

Poulisse, N. (1997). Language production in bilinguals. In A.M.B. de Groot & J.F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 201–224). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, *12*, 380–386.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.

Talamas, A., Kroll, J.F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second language vocabulary. *Biligualism: Language and Cognition*, *2*, 45–58.

Taylor, I., & Taylor, M.M. (1990). *Psycholinguistics: Learning and using language.* Englewood Cliffs, NJ: Prentice-Hall.

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional word reading via the phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 336–349.

Van Hell, J.G. (1998). *Cross-language processing and bilingual memory organization.* PhD Thesis, University of Amsterdam.

Van Hell, J.G., & De Groot, A.M.B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*, 193–211.

Van Heuven, W.J.B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*, 459–483.

Original manuscript received 17 April 1998
Accepted revision received 8 March 1999