Knowledge of a second language influences auditory word recognition in the native language

Evelyne Lagrou, Robert J. Hartsuiker, and Wouter Duyck

Ghent University

Running Head: Bilingual auditory word recognition

Correspondence Address:

Evelyne Lagrou

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent, Belgium

E-mail: evelyne.lagrou@ugent.be

Tel.: +32 (0)9 264 64 05

Abstract

Many studies in bilingual visual word recognition have demonstrated that lexical access is not language selective. However, research on bilingual word recognition in the auditory modality has been scarce, and it yielded mixed results with regard to the degree of this language nonselectivity. In the present study, we investigated whether listening to a second language (L2) is influenced by knowledge of the native language (L1) and, more importantly, whether listening to the L1 is also influenced by knowledge of a L2. Additionally, we investigated whether the listener's selectivity of lexical access is influenced by the speaker's L1 (and thus his/her accent). With this aim, Dutch-English bilinguals completed an English (Experiment 1) and a Dutch (Experiment 3) auditory lexical decision task. As a control, the English auditory lexical decision task was also completed by English monolinguals (Experiment 2). Targets were pronounced by a native Dutch speaker with English as the L2 (Experiment 1A, 2A, and 3A) or by a native English speaker with Dutch as the L2 (Experiment 1B, 2B, and 3B). In all experiments, Dutch-English bilinguals recognized interlingual homophones (e.g., *lief* (sweet) – *leaf* /li:f/) significantly slower than matched control words, whereas the English monolinguals showed no effect. These results indicate that (a) lexical access in bilingual auditory word recognition is not language selective in L2, nor in L1, and (b) language-specific sub-phonological cues do not annul cross-lingual interactions.

Key words: bilingualism, auditory word recognition, lexical access, accented speech

Research on bilingual word recognition has mainly focused on the visual modality. However, bilinguals of course do not only use their second language (L2) when reading, but also when listening to speech in that L2. Moreover, bilinguals often need to recognize spoken words in L2 not only when these words are spoken by a native language (L1) speaker, but also when these words are spoken by someone who is also using a L2 (e.g., when Dutch-English bilinguals have a conversation with for instance a Hebrew-English bilingual). These issues raise the question of whether bilinguals represent their languages in functionally/structurally independent systems when listening to speech. Perhaps the simplest point of view is that bilinguals have two separate language systems and lexicons, so that a bilingual only accesses words of the currently relevant lexicon (e.g., Gerard & Scarborough, 1989). However, in the bilingual visual word recognition literature, there is now ample evidence supporting a language nonselective account of lexical access, with bilinguals activating both language systems and lexicons in parallel (e.g., Dijkstra & Van Heuven, 1998; Duyck, 2005; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011). According to this account, lexical representations from both languages always get activated to a certain degree, even if only one language is task-relevant.

In visual word recognition, the assumption of language nonselectivity has been studied in several ways, for example by investigating the recognition of homographs (e.g., the Dutch-English homograph *room,* which means *cream* in Dutch). Dijkstra, Timmermans, and Schriefers (2000) for instance, observed longer reaction times (RTs) when reading homographs in a language decision task (i.e., one button was pressed when an English word was presented and another button was pressed for a Dutch word) as well as in a go/no-go task (i.e., participants responded only when they identified either an English word (English go/no-go) or a Dutch word (Dutch go/no-go)). Similarly, in a study by Dijkstra, Moscoso del Prado Martín, Schulpen, Schreuder, and Baayen (2005), Dutch-English bilinguals completed (a) an English visual lexical decision task, (b) a Dutch visual lexical decision task, and (c) a generalized visual lexical decision task (requiring yes-responses for words from both languages). Interlingual homographs were recognized slower than control words when participants completed the English (L2) or the Dutch (L1) visual lexical decision task, but faster than control words in the generalized visual lexical decision task. This implies that lexical access is language nonselective, but that the direction of the homograph effect is task-dependent. When interlingual homographs are presented in a monolingual context (English or Dutch visual lexical decision task), activation of the language-irrelevant phonological representation should yield a no-response for the language-relevant lexical decision, and will therefore compete with the yes-response triggered by the activation of the language-relevant phonological representation of the homophone. This competition causes a delay in responding to homographs. However, when interlingual homographs are presented in a bilingual context (generalized visual lexical decision task), the activation of the Dutch and English phonological representation both activate the yes-response, which causes facilitation.

Second, many recent studies used cognate words (e.g., the Dutch-English cognate *tomato*) to investigate lexical access in bilinguals. A recurring observation is that participants respond faster to cognates than control words in a lexical decision task; this is called the *cognate (facilitation) effect* (e.g., Caramazza & Brones, 1979; Dijkstra, Grainger, & Van Heuven, 1999; Duyck et al., 2007; Lemhöfer & Dijkstra, 2004; Libben & Titone, 2009; Van Assche et al., 2009; Van Hell & Dijkstra, 2002). This effect is considered as a reliable marker for language nonselective lexical activation, and is commonly explained by convergent activation spreading from the cognate's similar semantic, orthographic, and phonological representations across languages. In the study of Dijkstra et al. (1999), cross-lingual overlap with respect to semantics (S), orthography (O), and phonology (P) was systematically manipulated. They observed that orthographic and semantic overlap (SOP and SO items), as in cognates, resulted in response facilitation, whereas the recognition of words that only shared phonology (P) across languages (interlingual homophones) resulted in response inhibition. Initially, the cognate facilitation effect was only found in L2 word processing (e.g., Caramazza & Brones, 1979). However, recent evidence does support language nonselectivity in L1 by demonstrating a bidirectional L1-L2 cognate facilitation effect. Van Hell and Dijkstra (2002) tested Dutch-English-French trilinguals in a Dutch lexical decision task. Both L1-L2 cognates and L1-L3 cognates yielded faster reaction times than control words. However, the L1-L3 cognate facilitation effect was only observed with high proficient L3 speakers. This is in line with an account of language nonselective lexical access, but highlights the importance of language proficiency before cognate effects become noticeable in L1 processing. More recently, Van Assche et al. (2009) demonstrated in an eye-tracking study that even in a L1 sentence context Dutch-English bilinguals read cognates faster than control words.

Third, Duyck (2005) demonstrated that recognition of L2 target words (e.g., *corner*) can be facilitated by L2 primes which are phonologically equivalent to L1 words (e.g., *hook*, a homophone of the L1 word *hoek*, which means *corner*). This implies that the phonological representation of the L2 prime /huk/ activates both its L1 (*corner*) and L2 (*hook*) meaning. But when the language of primes and targets was switched, the phonological representation accessed by a L1 prime did not activate both its L1 and L2 meaning.

Interestingly, it is less clear whether lexical access in the auditory modality is also language nonselective (see below). Indeed, there are good reasons why the degree of language selectivity during bilingual lexical access might differ across (visual vs. auditory) modalities: whereas visual stimuli in the languages typically tested use the same or similar letters (so that the words do not contain information about their language membership), speech contains phonemic and sub-phonemic cues about the language in use: there are many phonemes that occur in English but not Dutch (e.g., /æ/) or vice versa (e.g., /y:/); and while some phonemes overlap between the two languages (e.g., /r/), many of them sound different because of allophonic variation. Indeed, Grosjean (1988) showed that bilinguals are able to judge language membership of so-called guest words pronounced as either code-switches or borrowings, solely on the basis of the words' initial phonemes. Given that bilingual listeners are sensitive to these cues, they might use them to restrict lexical access to the language in use. This would actually be a very efficient strategy to speed up lexical search, because such a selection mechanism would considerably diminish the number of lexical candidates for recognition. Before we turn to the present study, we will first provide an overview of the few studies that have been conducted in the auditory domain.

Auditory Word Recognition by Bilinguals

In contrast to the large body of research in bilingual visual word recognition, evidence in favor of a language nonselective account of lexical access in auditory word recognition has been relatively scarce. One interesting series of studies that reported evidence for a language nonselective view of auditory word recognition was carried out by Marian and colleagues (e.g., Marian, Blumenfeld, & Boukrina, 2008; Marian & Spivey, 2003; Marian, Spivey, & Hirsch, 2003; Spivey & Marian, 1999). These authors used an eye-tracking technique in which participants were instructed in their L2 (English) to pick up a real-life object (e.g., "*Pick up the marker*"). These participants were late Russian-English bilinguals with high L2 proficiency, living in a L2 dominant environment[1]. There were more fixations on competitor objects with a name in the irrelevant L1 that was phonologically similar to the target (e.g., a stamp; *marka* in Russian) than on distracter objects with a name in L1 that was phonologically unrelated to the L2 target.

Additionally, in a study by Schulpen, Dijkstra, Schriefers, and Hasper (2003) Dutch-English bilinguals completed a cross-modal priming task in which primes were presented auditorily and in which targets were presented visually. Visual lexical decision times were longer when the target was preceded by an interlingual homophone than when the target was preceded by a monolingual control (e.g., responses after the pair /li:s/ – LEASE were slower than after /freIm/ – FRAME; /li:s/ is the Dutch translation equivalent for *groin*). The observation of longer reaction times after interlingual homophone pairs suggested that bilinguals activated both the Dutch and the English meaning of the homophone. Furthermore, the authors observed that the auditory presentation of the English pronunciation of the interlingual homophone led to faster decision times on the related English target word than the Dutch version of the interlingual homophone. This indicates that sub-phonemic differences between homophones affect the degree of cross-lingual activation spreading, which is also relevant for the present study (see further).

Further evidence for nonselective lexical access was provided by Weber and Cutler (2004). In that study, Dutch-English bilinguals were instructed to click on one of four pictures presented in a display, and move it to another location on the computer screen (e.g., "*Pick up the desk and put it on the circle*"). The authors observed longer fixation durations on competitor objects with a phonemically similar L1 onset than to distracter objects (e.g., when instructed to pick up the *desk*, participants fixated longer on a picture of a *lid* than on control pictures, because *lid* is the translation equivalent of the Dutch word *deksel*, phonologically overlapping with the L2 target *desk*). Moreover, these authors demonstrated a L1-L2 phonetic similarity effect: Dutch listeners hearing English speech fixated longer on competitor items with names containing vowels that are phonologically confusable with the vowels in the target item (e.g., *pencil*, given target *panda*); Dutch does not have the vowel contrast /ɛ/ - /æ/).

However, when we consider the few studies investigating whether knowledge of a L2 also interferes when bilinguals are listening in L1, the results are less clear. Evidence in favor of language nonselectivity in L1 processing comes from studies by Spivey and Marian (1999) and by Marian et al. (2003). When Russian-English bilinguals received the instruction: *"Podnimi marku"* ("Pick up the stamp"), they looked more often to interlingual competitor objects *(marker)* than to distracter objects. Analogous to the findings with the English instructions, this can be explained because the English translation equivalent of *marka (stamp)* is more phonemically similar to the Russian target word *marku* than to the distracters[2].

However, this effect has not been consistently replicated by other authors. Although Weber and Cutler (2004) observed longer fixation times on competitor pictures with a phonemically similar L1 onset than to dissimilar distracter pictures when listening in L2, they did not find an analogous effect when these Dutch-English bilinguals were instructed in their L1. Distracters that were phonologically related to English targets were fixated longer than phonologically unrelated distracters (*deksel,* given target *desk*), but when stimuli were translated into Dutch (*desk*, given target *deksel*), the cross-lingual interference effect disappeared. To complicate things further, Ju and Luce (2004) found that Spanish-English bilinguals fixated interlingual distracters (nontarget pictures whose English names shared a phonological similarity with the Spanish targets) more frequently than control distracters, but only when the Spanish target words were altered to contain English-appropriate voice onset times. This is very interesting for the present study because such an interaction indeed seems to suggest that bilingual listeners use fine-grained, sub-phonemic, acoustic information to regulate cross-lingual lexical activation. The latter finding is also interesting because Ju and Luce did not observe the L1 interference effect (as observed by Marian and colleagues) when Spanish targets had Spanish voice onset times, even though these bilinguals had been living in a L2 dominant environment since birth or a very young age.

Another factor that seems to constrain parallel language activation is language proficiency. For example, in an eyetracking study using the visual world paradigm, Blumenfeld and Marian (2007) tested German-native and English-native bilingual speakers of German and English when listening in English. On each trial, both groups of participants were instructed to click on a target. Each display contained (a) a target, that could either be a German-English cognate or an English-specific target, (b) a German competitor with phonologically similar word-onsets, and (c) two filler items. The results demonstrated that both bilingual groups fixated more on the German competitor item than on the filler items while processing cognate words, but that only German-native bilinguals, and not the English native bilinguals, co-activated German competitors while processing English-specific words. This demonstrates that parallel language activation is boosted for high-proficient bilinguals. However, in a study by Chambers and Cooke (2009) interlingual competition did not vary according to participant's proficiency, but was instead influenced by the prior sentence context. In this study, using the visual world paradigm, English-French bilinguals with varying proficiency levels listened to L2 sentences, and were instructed to click on the display that represented the last word of the sentence. Each display contained an image of the final noun target (e.g., *chicken*), and an interlingual near-homophone (e.g., *pool*) whose name in English was phonologically similar to the French target (e.g., *poule*). The results demonstrated that interlingual competitors were fixated more than unrelated distracter items when the prior sentence information was compatible (i.e., both the French target and the interlingual near-homophone are plausible in the sentence context) with the competitor (e.g., *Marie va décrire la poule* [Marie will describe the chicken]), but not when this sentence information was incompatible (i.e., only the French target, but not the interlingual near-homophone is plausible in the sentence context) with the competitor (e.g., *Marie va nourrir la poule* [Marie will feed the chicken]). Taken together, even although research on bilingual auditory word recognition is relatively scarce, the evidence in favor of a language nonselective account of lexical access is mixed, and needs further exploration.

The Present Study

The present study was set up to investigate whether lexical access in auditory word recognition by bilinguals is language specific or not. This is especially important for L1 word recognition, given the inconsistent previous findings on this issue, discussed above. As a conservative test of this language nonselectivity hypothesis, we tested this with a sample of proficient, but unbalanced bilinguals, who live in a L1 dominant environment. This contrasts with the work of Marian and colleagues (Marian et al., 2003; Spivey & Marian, 1999), who found interference from L2 to L1 processing with bilinguals living in a L2 dominant setting (but see Ju & Luce, 2004). Our Dutch-English bilinguals completed an English (Experiment 1) and a Dutch (Experiment 3) lexical decision task in which the same Dutch-English homophones (e.g., *lief* (sweet) – *leaf* /li:f/) were presented. The English lexical decision task was also completed by a group of monolingual English participants (Experiment 2). Our first question was whether Dutch-English bilinguals use sub-phonemic cues inherent to speech to decide which lexicon has to be activated and accessed. It seems plausible that bilinguals can exploit such cues, given that Grosjean (1988) showed that bilinguals can accurately judge language membership of words based on just the initial phonemes, and given that Ju and Luce (2004) found that fine-grained acoustic information such as non-target language (English) voice onset times inserted in auditorily presented target language (Spanish) words affect cross-lingual lexical activation.

Most (monolingual) models of native spoken word recognition seem to be compatible with effects of sub-phonemic information on the activation of language-specific lexical items; these models include the Distributed Model of Speech Perception, Shortlist, NAM, and TRACE. According to the Distributed Model of Speech Perception (Gaskell & Marslen-Wilson, 1997), which is the successor of the Cohort model (Marslen-Wilson, 1987; 1990; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978), there are no discrete units for each lexical entry to represent lexical knowledge, as was the case in the original Cohort model, but there are distributed representations that use the same nodes for all lexical entries. This implies that lexical selection is influenced by the pattern and the amount of activation across a lexical representation. The Neighborhood Activation Model (NAM) of Luce and Pisoni (1998) is similar to the Distributed Model of Speech Perception (Gaskell & Marslen-Wilson, 1997), but also accounts for the influence of frequency of occurrence on processing. In contrast with these bidirectional models of spoken word recognition, the strictly bottom-up model Shortlist (Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997) assumes that bottom-up information first determines a set of candidate words before the short-listed candidates compete with each other for recognition. However, there are no bidirectional connections between nodes at adjacent levels. For the case of bilingualism, these models would predict that sub-phonemic (e.g., allophonic variations) information (or even the activation of language-specific phonemes) related to the native accent of the speaker leads to larger activation in the lexical representations of the target language (*bottom-up*), resulting in smaller cross-lingual interactions. For instance, hearing an allophonic variation of /r/ by a native speaker, could provide sufficient bottom-up information to further activate only English lexical candidates during lexical search.

Alternatively, it could be the case that such information activates language nodes that then regulate activation in lexical representations belonging to a given language through *top-down* facilitation or inhibition, which is what would be predicted by the fully interactive TRACE model (Elman & McClelland, 1988; McClelland & Elman, 1986). This is a connectionist model with bidirectional connections between the feature level, the phoneme level, and the word level. This model is similar to the Shortlist model, but does assume influence from higher levels onto lower levels, which implies that activation does not only spread from the phoneme level to the word level, but also vice versa. According to TRACE, the presence of sub-phonemic information would result in lexical search processes that appear to be language-specific to a larger extent. Finally, a verbal model that is much less known than the computational models we mentioned above, but that is specifically designed to account for auditory word recognition in bilinguals, is Léwy and Grosjean’s (1997) Bilingual Interactive Activation Model of Lexical Access (BIMOLA). This model assumes both bottom-up and top-down activation spreading between the different layers of the model. In contrast to the assumptions of the BIA+ model of bilingual visual word recognition, top-down activation is spreading as a function of the bilingual listener’s language mode, and of the higher order linguistic information of a syntactic or semantic nature. Hence, if lexical access in L1 auditory word recognition would be language selective, this could be modeled in BIMOLA by means of top-down inhibition of L2 lexical representations, following activation in language-specific (L1) phoneme representations.

Secondly, while our first question is concerned with the effects of target language cues, inherently present in spoken words, our second question is about sub-phonemic information conveyed by particular speakers. Languages differ of course in their phonological systems and only very few bilinguals manage to speak a second language without a persistent accent that provides many sub-phonemic (and sometimes even phonemic) cues about the speaker's first language. Previous work has demonstrated that listeners have some difficulty processing speech with a non-native accent (Adank, Evans, Stuart-Smith, & Scott, 2009) but it is not clear whether such an accent cues an irrelevant language that subsequently affects the degree to which lexical representations of that irrelevant language become active during lexical search. Our second question was therefore whether the accent of the speaker (i.e., Dutch or English), could influence the degree of language selective lexical access. With this aim, the native accent of the speaker was systematically manipulated with the same set of stimuli and both in L1 and L2.

Again, there are at least two ways in which the sub-phonemic cues in (accented) speech of non-native speakers, now referring to a non-target language, could influence the selectivity of lexical access. The sub-phonemic information, referring to the non-target language, could trigger bottom-up activation in the lexical representations of that language, resulting in larger cross-lingual interaction effects. For instance, if a Dutch native speaker produces a more Dutch-like /r/ than a native speaker of English producing the correct allophonic variation, the system might also consider Dutch lexical candidates.

This would be predicted by bottom-up models of spoken word recognition such as the Cohort model, the Distributed Model of Speech Perception, NAM, and especially Shortlist. The second possibility is that activation of irrelevant language representations lead to larger activation of the corresponding “language node” (see for instance TRACE, BIMOLA, or the early BIA model in visual word recognition). This would imply smaller non-target language inhibition than when listening to a native speaker, implying larger cross-lingual effects.

Experiment 1: English Lexical Decision Task with Dutch-English Bilinguals

In Experiment 1 we investigated whether L1 knowledge affects lexical access when listening to L2. With this aim, Dutch-English bilinguals completed an English lexical decision task in which interlingual homophones were presented auditorily. To answer the question whether this effect is sensitive to sub-phonemic/allophonic differences between native and nonnative speakers, targets were pronounced by a native Dutch speaker (Experiment 1A) or by a native English speaker (Experiment 1B).

Method

Participants

Thirty-four students from Ghent University participated in Experiment 1A for course credits or a monetary fee. All were native Dutch speakers and reported English as their L2[3]. They started to learn English around age 14 at secondary school, and because they were regularly exposed to their L2 through popular media, entertainment, and English university textbooks, they were all quite proficient in their L2. After the experiment, participants were asked to rate their L1 (Dutch) and L2 (English) proficiency with respect to several skills (reading, writing, speaking, understanding, general proficiency) on a 7-point Likert scale ranging from “very bad” to “very good”. We also assessed general L3 (French) proficiency. Means are reported in Table 1. Mean self-reported L1 (M = 5.94), L2 (M = 4.91), and L3 (M = 4.03) general proficiency differed significantly (dependent samples t-tests yielded ps < .001).

INSERT TABLE 1 ABOUT HERE

Participants were not informed that their L1 knowledge would be of any relevance to the experiment. Two participants made more than 20 % errors in the L2 lexical decision task, and were excluded from all analyses. In Experiment 1B there were 35 new participants. They met the same criteria as the participants in Experiment 1A, and were also asked to rate their L1, L2, and L3 proficiency. Means are reported in Table 1. Mean self-reported L1 (M = 6.37), L2 (M = 5.20), and L3 (M = 4.29) general proficiency differed significantly (dependent samples t-tests yielded ps < .001). One participant made more than 20 % errors and was excluded from all analyses.

Stimulus materials

The target stimuli consisted of 440 items: 44 interlingual Dutch-English homophones (e.g., *lief* (sweet) – *leaf* /li:f/), 44 matched English control words, 132 English fillerwords, and 220 nonwords. All targets were three to seven phonemes long. Interlingual homophones were selected from the stimulus lists of Dijkstra et al. (1999, 2005), Schulpen et al. (2003), or were selected from the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). Using the WordGen stimulus generation program (Duyck, Desmet, Verbeke, & Brysbaert, 2004), a control word was generated for each interlingual homophone, matched with respect to number of phonemes and L2 word frequency. The selected homophones and their matched control words are included in the Appendix. WordGen also generated the fillerwords and nonwords. Fillerwords did not differ from homophones and control words with respect to the matching criteria mentioned above. Nonwords were phoneme strings with no Dutch or English meaning, but with a legal English phonology. These were also matched with homophones and control words with respect to word length. In Experiment 1A each target was pronounced by a native Dutch speaker who was also a high-proficient English speaker, and in Experiment 1B each target was pronounced by a native English speaker who was also a high-proficient Dutch speaker. Fourteen targets (seven homophones, one control word, and six fillerwords) were translated incorrectly in a backward translation test following the experiment by more than 30 % of the participants; together with their matched stimuli these were removed from further analyses. Using WaveLab software, stimulus materials were recorded in a sound-attenuated booth by means of a SE Electronics USB1000A microphone on a sampling rate of 44.1 kHz and a 16-bit sample size. Target durations were measured with WaveLab software.

Speakers

The speaker in Experiment 1A was a 25 year old female with Dutch as L1 and English as L2. She had 12 years of experience with her L2. Her English was very fluent but characterized by a clear Dutch accent. The speaker in Experiment 1B was a 45 year old female with English as L1 and Dutch as L2. She had experience with her L2 since she moved to the Dutch-speaking part of Belgium 15 years ago. Her Dutch was very fluent but characterized by a clear English accent. Audio excerpts of both speakers are provided on http://expsy.ugent.be/research/Rdocuments.

Procedure

Participants received written instructions in L2 to perform a L2 lexical decision task. They were instructed to put on a headphone through which targets would be presented auditorily. Before the experiment, a practice session of 24 trials was completed. Each trial started with a 500 ms presentation of a fixation cross in the center of the screen. After another 200 ms the target was presented auditorily. Then participants had to decide whether they heard an English word or a nonword. When a word (nonword) was presented, participants used their right (left) index finger to press the right (left) button of a response box. Visual feedback (i.e., when an error was made the screen turned red, when the response was correct “OK!” appeared on the screen) was always presented on the screen during 200 ms. The next trial started 500 ms later. After the experiment, participants completed a questionnaire assessing self-ratings of L1 and L2 proficiency (reading, speaking, writing, understanding, and general proficiency), and general L3 proficiency on a 7-point Likert scale, and a backward translation test to verify that they knew the L2 words.

Results

Results Experiment 1A: native Dutch speaker

On average participants made 10.64 % errors (*SD* = 2.71). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant’s mean RT after target offset for word targets were excluded from the analyses. As a result, 12.91 % of the data were excluded from the analyses.

Latencies. In all experiments, reported latency analyses are based on reaction times measured from (auditory) target offset[4]. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were recognized significantly slower than control words[5], $F1(1,31) = 24.23$, $p < .001$; $F2(1,36) = 7.64$, $p < .01$ (see Table 2).

INSERT TABLE 2 ABOUT HERE

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants made more errors on homophones than on control words, $F1(1,31) = 71.82$, $p < .001$; $F2(1,36) = 5.40$, $p < .05$ (see Table 2).

Results Experiment 1B: native English speaker

On average, participants made 9.53 % errors ($SD = 3.95$). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant's mean RT after target offset for word targets were excluded from the analyses. As a result, 11.39 % of the data were excluded from the analyses.

Latencies. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were recognized significantly slower than control words, $F1(1,33) = 25.48$, $p < .001$; $F2(1,36) = 5.93$, $p < .05$ (see Table 2).

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants made more errors on homophones than on control words, $F1(1,33) = 9.81$, $p < .01$; $F2(1,36) = 5.19$, $p < .05$ (see Table 2).

Comparison Experiment 1A-1B

We also compared the results of Experiment 1A, in which targets were pronounced by a native Dutch speaker, and the results of Experiment 1B, in which targets were pronounced by a native English speaker (see Figure 1).

INSERT FIGURE 1 ABOUT HERE

An ANOVA with target type (interlingual homophone vs. control) as the independent within-subjects variable, speaker (native Dutch vs. native English) as the independent between-subjects variable, and RTs as the dependent variable demonstrated a main effect of speaker, $F1(1,64) = 14.76$, $p < .001$; $F2(1,72) = 23.89$, $p < .001$, which showed that RTs were slower when targets were pronounced by the native Dutch speaker than when they were pronounced by the native English speaker. The main effect of target type was also significant, $F1(1,64) = 55.65$, $p < .001$; $F2(1,72) = 13.39$, $p < .001$, indicating that participants recognized homophones slower than control words. More importantly, the interaction between target type and the L1 of the speaker was not significant, $F1 < 1$ and $F2 < 1$.

Discussion

The results of the L2 lexical decision task were very straightforward: auditorily presented homophones were recognized more slowly than matched control words. This provides evidence that lexical representations of more than one lexicon are activated during monolingual L2 auditory word recognition. This implies that bilingual listeners do not use sub-phonemic cues inherent to speech as a cue to restrict lexical search to a single language.

In our comparison of the results of Experiments 1A and 1B there was a main effect of speaker: reaction times were slower when targets were pronounced by the native Dutch speaker. There are two possible explanations for this effect. First, because the native English speaker did not have a Dutch accent, her pronunciation may provide a closer match to the listener’s stored lexical representations, so that the threshold for word recognition was exceeded faster, yielding faster word/nonword decisions. This explanation is also compatible with the results of Adank et al. (2009), who observed longer RTs when participants listened to a speaker with a nonnative accent. Second, we noted that the English speaker tended to stretch the pronunciations (particularly the final phonemes) more than the Dutch speaker; indeed, target word durations were significantly longer when spoken by the English speaker (a dependent samples t-test yielded $p < .001$) (see Table 3). Because of the longer word duration, lexical activation has more time to accumulate as speech unfolds, so that participants can respond more quickly at speech offset. However, if we analyzed the RTs from target onset, the main effect of speaker remained significant, which makes this explanation less plausible.

INSERT TABLE 3 ABOUT HERE

Finally, if sub-phonemic cues can influence the degree of language selectivity, one would expect an interaction between target type and speaker. Because the L2 and L1 pronunciations of the targets differ at the sub-phonemic level, we assumed that participants could use this information as a cue to indicate which language is in use. Considering that the nonnative target pronunciation contains sub-phonemic cues of both languages, one may predict that the homophone effect should be larger when targets were pronounced by the nonnative speaker. However, there was no trace of an interaction between target type and speaker, demonstrating that bilinguals do not use these sub-phonemic differences between languages as a strict cue for language selection.

In Experiment 2, the same set of stimuli was tested with a group of English monolinguals. Because these participants do not have any knowledge of Dutch, we expected no difference between reaction times on homophones and on control words.

Experiment 2: English Lexical Decision Task with English Monolinguals

Although the stimulus materials used in Experiment 1 were carefully controlled item by item, it is not impossible that certain stimulus characteristics that were not taken into account could have influenced the results. In order to ensure that the observed homophone effect in Experiment 1 is indeed due to L1 activation during L2 listening, a control experiment with English monolinguals was carried out[6]. In this experiment the same stimulus materials as in Experiment 1 were presented to these monolingual participants. If the homophone effect in Experiment 1 is caused by cross-lingual interactions when listening to L2, this effect should disappear for English monolinguals.

Method

Participants

Thirty students from the University of Southampton participated in Experiment 2A and received course credit or a monetary fee. They were all native speakers of English and had no knowledge of Dutch. Two participants made more than 20 % errors and were excluded from all analyses. In Experiment 2B, 30 further students took part in the experiment. They met the same criteria as the participants in Experiment 2A. One participant made more than 20 % errors and was excluded from all analyses.

Stimulus Materials, Speakers, and Procedure

Stimulus materials, speakers, and procedure were identical to Experiment 1, except that the participants did not complete a questionnaire assessing self-ratings of L1 and L2 proficiency, and a backward translation test after the experiment.

Results

Results Experiment 2A: native Dutch speaker

On average participants made 13.90 % errors (*SD* = 2.62). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant's mean RT after target offset for word targets were excluded from the analyses. As a result, 16.21 % of the data were excluded from the analyses.

Latencies. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were not recognized significantly slower than control words, *F1* < 1 and *F2* < 1 (see Table 4).

INSERT TABLE 4 ABOUT HERE

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants did not make more errors on homophones than on control words, $F1(1,27) = 2.65$, $p > .05$; $F2(1,43) = 1.44$, $p > .05$ (see Table 4).

Results Experiment 2B: native English speaker

On average, participants made 7.41 % errors ($SD = 2.56$). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant's mean RT after target offset for word targets were excluded from the analyses. As a result, 9.30 % of the data were excluded from the analyses.

Latencies. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were not recognized significantly slower than control words, *F1* < 1 and *F2* < 1 (see Table 4).

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants did not make more errors on homophones than on control words, *F1* < 1 and *F2* < 1 (see Table 4).

Comparison Experiment 2A-2B

We also compared the results of Experiment 2A, in which targets were pronounced by a native Dutch speaker, and the results of Experiment 2B, in which targets were pronounced by a native English speaker (see Figure 2).

INSERT FIGURE 2 ABOUT HERE

An ANOVA with target type (interlingual homophone vs. control) as the independent within-subjects variable, speaker (native Dutch vs. native English) as the independent between-subjects variable, and RTs as the dependent variable demonstrated a main effect of speaker, $F1(1,55) = 66.77$, $p < .001$; $F2(1,86) = 293.89$, $p < .001$, which showed that RTs were slower when targets were pronounced by the native Dutch speaker than when they were pronounced by the native English speaker. The main effect of target type was not significant, $F1 < 1$ and $F2 < 1$. More importantly, the interaction between target type and the L1 of the speaker was not significant, $F1 < 1$ and $F2 < 1$.

Comparison Bilinguals – Monolinguals

Next, we compared the results of our Dutch-English bilinguals with the results of our English monolinguals. When targets were pronounced by the native Dutch speaker, an ANOVA with target type (interlingual homophone vs. control) as the independent within-subjects variable, speaker (native Dutch speaker vs. native English speaker) and sample (bilinguals – monolinguals) as the independent between-subjects variables, and RTs as the dependent variable demonstrated a significant interaction between target type and the sample, $F1(1,119) = 22.03$, $p < .001$; $F2(1,158) = 13.27$, $p < .001$. The interaction between speaker and sample was also significant, $F1(1,119) = 9.98$, $p < .001$; $F2(1,158)$

= 30.16, $p < .001$, indicating a larger speaker effect for our monolinguals than for our bilinguals (i.e., our English monolinguals responded faster than our bilinguals when targets were pronounced by the native English speaker).

Discussion

In this control experiment, English monolinguals completed the same English auditory lexical decision task as the Dutch-English bilinguals in Experiment 1. The results showed that these monolingual participants recognized homophones equally fast as control words. In line with the results of Experiment 1 we also observed a main effect of speaker, with faster reaction times when targets were pronounced by the native English speaker. This effect was even larger in our monolingual than in our bilingual group of participants, which can be explained by the fact that monolingual speakers of English are less familiar with the Dutch pronunciation of English than our bilinguals. As in Experiment 1, the main effect of speaker is probably due to the fact that the native speaker provides speech input that matches the listener's stored lexical representation more closely. The absence of the homophone effect in this group of participants, together with the significant interaction between target type and the sample, ensures that the observed homophone effect in Experiment 1 indeed resulted from their bilingual knowledge, and more specifically from the cross-lingual activation of L1 when listening in L2.

In the next experiment, we investigated whether we could also provide evidence for a view of language nonselective lexical access in auditory word recognition when listening in the native language. In line with the results of Experiment 1, we expected an equivalent homophone effect with both L1 and L2 speakers. The next experiment using

Dutch targets also allows dissociating between two possible explanations for the observation that RTs are faster when targets are pronounced by the native English speaker: if the effect depends on the articulation times of the particular speaker, one would predict that RTs are again faster for the English speaker. Instead, if the effect originates from the speaker's accent (non-native speech that mismatches stored lexical representations somewhat), we would predict slower RTs for the English speaker.

Experiment 3: Dutch Lexical Decision Task with Dutch-English bilinguals

Experiment 3 tested whether L2 knowledge affects lexical access when listening to L1. With this aim, Dutch-English bilinguals completed a Dutch auditory lexical decision task in which interlingual homophones were again the crucial targets. To answer the question of whether the homophone effect is sensitive to sub-phonemic differences between native and nonnative speakers, targets were again pronounced by a native Dutch speaker (Experiment 3A) or by a native English speaker (Experiment 3B).

Method

Participants

Thirty-two students from Ghent University participated in Experiment 2A for a monetary fee. All of them met the same criteria as the participants of Experiment 1. Mean proficiency ratings are reported in Table 1. Mean self-reported L1 (M = 6.28), L2 (M = 5.34), and L3 (M = 4.06) general proficiency differed significantly (dependent samples t-tests yielded ps < .001). Participants were not informed that their L2 knowledge would be of any relevance to the experiment. Because virtually all university students in Ghent are sufficiently L2 proficient to participate, knowledge of English was not mentioned as a participation criterion in recruitment. In Experiment 3B there were 29 further participants. They met the same criteria as the participants in Experiment 3A and they were also asked to rate their L1, L2, and L3 proficiency. Means are reported in Table 1. Mean self-reported L1 ($M = 6.45$), L2 ($M = 5.34$), and L3 ($M = 4.48$) general proficiency differed significantly (dependent samples t-tests yielded ps < .001). Two participants made more than 20 % errors and were excluded from all analyses. None of the participants participated in Experiment 1.

Stimulus materials

The target stimuli consisted of 440 items: 44 interlingual Dutch-English homophones which were phonologically equivalent to the homophones in Experiment 1, 44 matched Dutch control words, 132 Dutch fillerwords, and 220 nonwords. Nonwords were phoneme strings with no Dutch or English meaning, but with a legal Dutch phonology. All targets met the same selection criteria as in Experiment 1. The selected homophones and their matched control words are included in the Appendix. The same native Dutch (Experiment 3A) and native English (Experiment 3B) speakers as in Experiments 1 and 2 pronounced all targets for the Dutch lexical decision task. Six homophones were translated incorrectly in the forward translation test following the experiment by more than 30 % of the participants; together with their matched stimulus these were removed from further analyses. The same procedure as in Experiment 1 was followed for material recording and processing.

Procedure

The procedure was identical to Experiments 1 and 2, except that participants now received written instructions in Dutch and that they had to decide whether they heard a Dutch word or a nonword.

Results

Results Experiment 3A: native Dutch speaker

On average participants made 6.48 % errors (SD = 2.69). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant's mean RT after target offset for word targets were excluded from the analyses. As a result, 8.34 % of the data were excluded from the analyses.

Latencies. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were recognized significantly slower than control words[7], $F1(1,31) = 72.21$, $p < .001$; $F2(1,37) = 18.07$, $p < .001$ (see Table 5).

INSERT TABLE 5 ABOUT HERE

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants made more errors on homophones than on control words, $F1(1,31) = 45.99$, $p < .001$; $F2(1,37) = 12.16$, $p < .01$ (see Table 5).

Results Experiment 3B: native English speaker

On average, participants made 11.34 % errors (*SD* = 2.39). Errors, trials with RTs faster than 300 ms after target onset, and trials with RTs more than 2.5 standard deviations above the participant's mean RT after target offset for word targets were excluded from the analyses. As a result, 13.70 % of the data were excluded from the analyses.

Latencies. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and RTs as the dependent variable demonstrated that homophones were recognized significantly slower than control words, $F1(1,26) = 18.24$, $p < .001$; $F2(1,37) = 7.68$, $p < .01$ (see Table 5).

Accuracy. An ANOVA with target type (interlingual homophone vs. control) as the independent variable and error percentage as the dependent variable mirrored RT analyses by revealing that participants made more errors on homophones than on control words, $F1(1,26) = 31.18$, $p < .001$; $F2(1,37) = 8.77$, $p < .01$ (see Table 5).

Comparison Experiment 3A-3B

We also compared the results of Experiment 3A, in which targets were pronounced by a native Dutch speaker, and the results of Experiment 3B, in which targets were pronounced by a native English speaker (see Figure 3).

INSERT FIGURE 3 ABOUT HERE

An ANOVA with target type (interlingual homophone vs. control) as the independent within-subjects variable, speaker (native Dutch vs. native English) as the independent between-subjects variable, and RTs as the dependent variable demonstrated there was a significant main effect of speaker, $F1(1,57) = 4.03$, $p < .05$; $F2(1,74) = 16.64$,

$p < .001$, indicating that RTs were slower when targets were pronounced by the native English speaker than when they were pronounced by the native Dutch speaker. The main effect of target type was also significant, $F1(1,57) = 78.44$, $p < .001$; $F2(1,74) = 22.56$, $p < .001$, indicating that participants recognized homophones slower than control words. As in Experiment 1, there was no interaction between target type and the L1 of the speaker, $F1 < 1$ and $F2 < 1$.

Discussion

The goal of this experiment was to provide further evidence for a language nonselective account of lexical access in auditory word recognition, demonstrating that bilinguals are not only influenced by their L1 when listening to their L2, but also that a weaker L2 can influence the dominant, native L1. The results were straightforward: interlingual homophones were recognized slower than control words, both when targets were pronounced by a native Dutch speaker and when they were pronounced by a nonnative Dutch speaker. This is an important addition to the literature, especially because this cross-lingual lexical interaction effect in auditory word recognition was observed using unbalanced, late bilinguals living in a L1 dominant environment. It therefore extends previous work of Spivey and Marian (1999) and of Marian et al. (2003), who reported phonological similarity effects in the visual world paradigm with bilinguals immersed in a L2 setting. Moreover, this study demonstrated that this cross-lingual effect in L1 is equally large with L2 speakers as with L1 speakers. The present findings (at least those for L1 recognition) are inconsistent with Weber and Cutler (2004), who only reported cross-lingual effects during L2 listening, using the same eyetracking paradigm as Marian and colleagues.

This experiment also allows dissociating between two possible explanations for the main effect of speaker (i.e., faster RTs when targets were pronounced by a native English speaker) that we observed in Experiment 1. If this effect would have originated from the articulation times of the particular speaker, one would again have predicted faster recognition when listening to the native English speaker. Instead, if recognition was slower because a foreign accent yields speech that matched stored lexical representation less closely, one would now predict slower RTs for the English speaker (speaking Dutch). Comparing the results of Experiment 3A and Experiment 3B, we again observed a main effect of speaker, but now recognition was faster overall when targets were pronounced by the native Dutch speaker, in contrast with Experiments 1 and 2. These results support the suggestion that non-accented speech results in a better match with lexical representations than accented speech, and therefore speeds up word recognition. As in Experiment 1, there was no interaction between target type and the speaker, indicating that bilingual listeners do not use cues provided by sub-phonemic differences or allophonic variations between languages to restrict lexical search to a single language when listening in L1.

General Discussion

The goal of this study was to answer two major questions: First, we wanted to investigate whether access to lexical representations in auditory word recognition is nonselective with respect to language. We examined homophone interference effects from L2 to L1 and vice versa, using unbalanced, late Dutch-English bilinguals, living in a

L1 dominant environment. Our second research question concerned whether bilingual listeners are sensitive to sub-phonemic differences (e.g., allophonic variations) between the pronunciation of words by native and nonnative speakers, and whether they use these differences as a restrictive cue to limit lexical search to a single language. With this aim, we constructed a L2 and L1 lexical decision experiment in which targets were pronounced by a native Dutch or by a native English speaker. The results were clear-cut: We found that homophones were recognized slower than control words when participants were listening in L2 (Experiment 1), but also when they were listening in their dominant/native L1 (Experiment 3). Furthermore, this homophone effect was independent of the speaker’s native language. In a control experiment with English monolinguals (Experiment 2) we did not find slower RTs for homophones than for control words. Therefore, we can safely conclude that the observed homophone effects with Dutch-English bilinguals are indeed due to lexical interactions in the bilingual lexicon, and not as a consequence of some confounded variable between homophone and control conditions. Moreover, to avoid the possibility that language-ambiguous stimuli may boost dual-language activation (Blumenfeld & Marian, 2007; Elston-Güttler, Gunter, & Kotz, 2005) in this study, only 10 % of the trials were interlingual homophones. Because this small percentage would probably not be much higher than the proportion of language-ambiguous stimuli in natural language use, we believe the experimental setting does not constitute an artificial bilingual context, and is not perceived as such by the participants. We will discuss the theoretical implications of these findings in the next paragraphs.

First of all, the fact that we do not only observe a homophone effect when listening in L2, but also in L1, provides convincing evidence for language nonselective lexical access. It also shows that such interactions are quite robust, given that activation in lexical representations from a weaker language is strong enough to influence word recognition in the dominant language. Previously, a few studies in bilingual auditory word recognition have shown that L1 knowledge influences recognition of words in L2 (e.g., Schulpen et al., 2003; Weber & Cutler, 2004), but only two studies revealed some effects of L2 knowledge on the auditory recognition of words in L1, both using bilinguals that were immersed in a L2 speaking environment and/or started learning L2 at a young age (Marian et al., 2003; Spivey & Marian, 1999). This study demonstrated the effect of L2 knowledge on L1 auditory word recognition in a population of unbalanced, late bilinguals, living in a L1 dominant environment. Moreover, we investigated L2-L1 and L1-L2 effects in two samples of a very homogeneous population, using the same homophones, which allows a quite direct comparison of the effects. These findings are inconsistent with Weber and Cutler (2004), who only observed cross-lingual lexical interactions during L2 recognition, but not in L1 recognition. The crucial difference between that study and the present one may lie in the strength of the cross-lingual phonological overlap manipulation. Whereas the present study used interlingual homophones with almost complete overlap across languages, the eyetracking paradigm used by Weber and Cutler (and by Marian and colleagues) used items that shared only a few phonemes (the onset) across languages. This may explain why the cross-lingual phonological interference effects of Weber and Cutler are absent in L1 recognition.

The main finding of the present study, the interlingual homophone effect, can be explained within the Distributed Model of Speech Perception (Gaskell & Marslen-

Wilson, 1997), NAM (Luce & Pisoni, 1998) and Shortlist (Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997), if these models are extended with the assumption that L2 representations are part of the same system as, and interact with, L1 representations. These models predict competition, and therefore slower recognition of interlingual homophones, under the above assumption. In general, the cross-lingual interactions observed here are also in line with the core assumption of BIMOLA (Léwy & Grosjean, 1997) that the auditory presentation of a homophone activates phoneme representations, and then lexical representations in both languages. When words from two languages are activated, the bilingual listener needs additional information (and hence more time) to make a final decision about which word is selected. This assumption was confirmed by the results of both experiments. These findings imply that sub-phonemic information referring to the target language does not imply selective bottom-up activation of lexical representations belonging to (only) that language, nor top-down regulation of target language (facilitation) or non-target language (inhibition) lexical representations. This also implies that there is no need for an additional mechanism that uses language nodes (e.g., as in the visual BIA model (Dijkstra & Van Heuven, 1998)) or language schemes (e.g., as in BIA+ (Dijkstra & Van Heuven, 2002)) to inhibit the non-relevant language, because the same mechanisms for within-language competition may be used to account for between-language competition. Thus, at a theoretical level, the results of the present study are compatible with models of auditory word recognition that support language nonselective bottom-up activation with a very limited role for top-down connections.

Second, the present findings offer a clear answer to the question of whether bilinguals are sensitive to sub-phonemic differences or allophonic variations between the pronunciation of words by native or nonnative speakers, and whether they use these differences to modulate activation of lexical representation belonging to a single language. Such modulation of cross-lingual effects may also have at least two different origins: it may be the case that sub-phonemic cues from a non-target language trigger activation in lexical representations belonging to that language (*bottom-up*), so that a foreign accent yields larger homophone effects. On the other hand, such modulation may also occur through a different mechanism, by which hearing sub-phonemic cues of an irrelevant language causes a ‘dual-language mode’ (in BIMOLA’s terms), triggering *top-down* activation of all representations belonging to that language, again yielding larger homophone effects. However, our results demonstrated that the size of the homophone effect was equivalent when targets were pronounced by a L1 or a L2 speaker. This suggests that bottom-up activation in sub-phonemic, non-target language representations does not spread strongly enough to influence target language recognition, and also that there is a very limited role for top-down connections of language nodes with lexical representations within that language. It may therefore be concluded that such information is not strong enough to modulate cross-lingual interactions (bottom-up) and that listeners do not use this information as a cue to activate a non-target language (top-down). The fact that salient language cues (such as information about the native accent of the speaker) do not influence lexical access may be a surprising conclusion. However, this is exactly what has been found in the visual domain, as the recent BIA+ model (Dijkstra & Van Heuven, 2002) does no longer contain such top-down connections, in contrast to its precursor, the BIA model (Dijkstra & Van Heuven, 1998). As such, our results seem to be inconsistent with Ju and Luce (2004). These authors found that Spanish-English bilinguals fixated pictures whose English names were phonologically similar to Spanish auditory targets more frequently than control distracters, but only when the Spanish target words were altered to contain English-appropriate voice onset times. Taken together, it appears that at least this acoustic feature influences cross-lingual lexical activation in bilinguals, although the sub-phonemic differences between the utterances in the different languages of the present study do not. Note that Ju and Luce's findings for Spanish (L1) targets that were not altered with L2 voice onset times are also inconsistent with the interference effects of Marian and colleagues, even though both used bilinguals that had been living in a L2 dominant environment. Ju and Luce argued that this discrepancy might originate from the fact that the stimuli of Marian and colleagues (just as those of the present study, and just as in natural word recognition) contained words starting with a variety of sounds, including nasals (e.g., *marker*) and fricatives (e.g., *fish*), whereas all of the Ju and Luce stimuli began with voiceless stops. Hence, the present findings are only consistent with Ju and Luce's when we reduce the claim that a strong acoustic cue (e.g., voicing, as present in their stimulus manipulation) might inhibit cross-lexicon activation, unlike recognition of materials that diverge on a wider range of acoustic parameters, and are not artificially manipulated on a single language-specific contrast (cfr. Marian and colleagues and the present study).

Note however that we did observe faster word recognition for words spoken by a speaker in her native language than spoken by a speaker in her L2, suggesting that sub-phonemic activation in a non-target language does have *some* influence on recognition, even if it does not annul cross-lingual interactions. At first sight, this observation is inconsistent with the work of Bradlow and Bent (2008). These authors demonstrated that non-native listeners are faster to recognize speech from other non-native speakers with the same L1. A possible explanation for these different results is the fact that the task participants had to perform was different in the Bradlow and Bent (2008) study and the present study. Whereas Bradlow and Bent (2008) investigated accuracy scores on a sentence recognition task, our participants were instructed to make lexical decisions on isolated words. As a consequence, a direct comparison of the results of both studies is not feasible. A plausible explanation for the results of the present study is that the native speaker's pronunciation of the targets provides a better match with stored lexical representations than the utterances of L2 speakers. This way, the threshold for recognition is exceeded faster. Another possibility is that the non-accented speech represents a less noisy signal than the accented speech, and that the accented speech either requires the extra step of talker normalization, or that the accented speech represents a poorer signal-to-noise ratio. At least, this interaction effect between the target listening language and the speaker's native language proves that the different types of speaker utterances indeed contained language-specific acoustic information. As such, this constitutes a successful manipulation check of sub-phonemic differences between languages.

At a theoretical level, the present L1 and L2 effects show great similarity with the findings in the bilingual visual word recognition literature. In that domain, it has also recently been shown that participants do not use language cues to decide which lexicon to access. For example, when reading words in a sentence context, this context is not used as a cue to restrict lexical access (Duyck et al., 2007; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche et al., 2009; Van Assche et al., 2011; Van Hell & de Groot, 2008). Additionally, Thierry and Wu (2007) investigated whether differences in script between the Chinese and English language can restrict lexical access. These authors demonstrated by means of brain potentials that English words were automatically and unconsciously translated into their Chinese counterpart when Chinese-English bilinguals made semantic relatedness decisions about English words, indicating that even script is not used as a cue to guide lexical access. However, the current findings are more surprising than the analogues in the visual domain, given that auditory, but not visual word presentation contains information about the language to which the word belongs, and given the fact that bilinguals are able to determine a word's language membership on the basis of just the initial phonemes (Grosjean, 1988).

Taken together, this study supports a view of bilingual lexical access that is highly nonselective. Whereas the results in bilingual visual word recognition have demonstrated that L1 or L2 readers do not restrict lexical access to one lexicon by using cues such as sentence context (e.g., Van Assche et al., 2009), or even script (Thierry & Wu, 2007), the results of the present study extend these findings to the auditory domain in bilingual word recognition. Apparently, bilinguals do not use speech-specific cues or sub-phonemic differences between speakers to restrict lexical access to the currently relevant lexicon when listening in L2, nor in L1.

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 520-529.

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

Blumenfeld, H., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, *22*, 633-660.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707-729.

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

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1029-1040.

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

Dijkstra, T., Moscoso del Pradio Martín, F., Schulpen, B., Schreuder, R., & Baayen, R. H. (2005). A roommate in cream: Morphological family size effects on interlingual homograph recognition. *Language and Cognitive Processes, 20*, 7-41.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition, *Journal of Memory and Language, 42,* 445-465.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189-225), Mahwah, NJ, USA: Lawrence Erlbaum Associates.

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

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1340-1359.

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers, 36,* 488-499.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 663-679.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 27,* 143-165.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25,* 57-70.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12,* 613-656.

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

Graddol, D. (2004). The future of language. *Science, 303,* 1329-1331.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3,* 233–274.

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears - Constraints on bilingual lexical activation. *Psychological Science, 15(5),* 314-318.

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

Léwy, N., & Grosjean, F. (1997). *A computational model of bilingual lexical access.* Manuscript in preparation, Neuchâtel University, Neuchâtel, Switzerland.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35,* 381-390.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing, 19,* 1-36.

Marian, V., Blumenfeld, H., & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research, 37,* 141-170.

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24,* 173-193.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language, 86,* 70-82.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71-102.

Marslen-Wilson, W. D. (1990). Activation, competition and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives,* (pp. 148-172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1-71.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29-63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1-86.

Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.

Norris, D. G., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34,* 191-243.

Schulpen, B., Dijkstra, T., Schriefers, H. J., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1155-1178.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55,* 197-212.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281-284.

Thierry, G., & Wu, J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Science, 104,* 12530-12535.

Van Assche, E., Drieghe, D., Duyck, W. Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading, *Journal of Memory and Language, 64,* 88-107.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science, 20*, 923-927.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780-789.

Van Hell, J. G. & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128,* 431-451.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

Acknowledgments

This research was made possible by the Research Foundation-Flanders (FWO) of which Evelyne Lagrou is a research fellow. We thank Denis Drieghe and the research group of Simon Liversedge (Southampton University) for their support to make the control study with English monolingual participants (Experiment 2) feasible.

Appendix. Interlingual homophones and matched control words in Experiments 1 and 3. Asterisks indicate targets and control words excluded from the analysis for eliciting errors in a translation test by more than 30 % of participants.

L2 HOMOPHONE	phonology	CONTROL	L1 HOMOPHONE	phonology	CONTROL
BAY	[beI]	hug	BIJ	[bEI]	vol
BEAT	[bi:t]	nose	BIET	[bi:t]	kras
BILL	[bIl]	skin	BIL	[bIl]	lap
BONE	[b@Un]	lion	BOON	[bo:n]	kooi
BOSS	[bOs]	rush	BOS	[bOs]	hek
BRIEF	[bri:f]	nerve	BRIEF	[bri:f]	roman
BULL	[bUl]	page	BOEL	[bu:l]	boer
COOK	[kUk]	fast	KOEK	[ku:k]	zeef
COW	[kaU]	bag	KOU	[kAU]	lui
CRATE*	[kreIt]	mouse	KRIJT*	[krEIt]	lepel
DAY	[deI]	boy	DIJ	[dEI]	mus
DOLL	[dOl]	moon	DOL	[dOl]	kam
FATE	[feIt]	pope	FEIT	[fEIt]	trap
HAY*	[heI]	bug	HEI*	[hEI]	das
HEAL	[hi:l]	duck	HIEL	[hi:l]	kous
HOOK	[hUk]	farm	HOEK	[hu:k]	rook
KNOCK	[nOk]	chief	NOK	[nOk]	pit
LAKE	[leIk]	soft	LIJK	[lEIk]	kast
LANE	[leIn]	king	LIJN	[lEIn]	maag
LEAD	[li:d]	risk	LIED	[li:t]	lift
LEAF	[li:f]	wine	LIEF	[li:f]	buik
LEASE*	[li:s]	scarf*	LIES*	[li:s]	saai
LIST	[lIst]	wing	LIST	[lIst]	slak
LOOP*	[lu:p]	fist	LOEP*	[lu:p]	zoen
MAIL	[meIl]	fork	MIJL	[mEIl]	vork
MESS	[mEs]	safe	MES	[mEs]	aap
MOOD	[mu:d]	silk	MOED	[mu:t]	zout
PALE	[peIl]	twin	PIJL	[pEIl]	wolk
PET	[pEt]	fox	PET	[pEt]	put
PLANE	[pleIn]	towel	PLEIN	[plEIn]	plooi
PRAISE	[preIz]	poison	PRIJS	[prEIs]	straf
PRAY	[preI]	claw	PREI	[prEI]	wilg
PROOF	[pru:f]	widow	PROEF	[pru:f]	draad
QUICK	[kwIk]	chair	KWIK	[kwIk]	klad
RAISE	[reIz]	elbow	REIS	[rEIs]	dame
RAY*	[reI]	jar	RIJ*	[rEI]	jas
ROOM	[ru:m]	wife	ROEM	[ru:m]	mand
SLIM	[slIm]	deaf	SLIM	[slIm]	spin
STREAM	[stri:m]	pillow	STRIEM	[stri:m]	fazant
TAIL	[teIl]	meat	TEIL	[tEIl]	deeg
THIGH*	[TaI]	devil	TAAI*	[ta:j]	tang
TRACK	[tr&k]	taste	TREK	[trEk]	muur
VET	[vEt]	spy	VET	[vEt]	kip
WAY	[weI]	old	WEI	[wEI]	mop

Footnotes

1. These participants grew up in the former Soviet Union, but immigrated to the United States in their early teens, and were students at a top-tier American University at the time of testing. Only two of the participants stated that Russian was their preferred language at the time of testing, five stated that English was their preferred language, and five had no preference between Russian and English.

2. Note that the most straightforward evidence for both L1-L2 and L2-L1 interference was observed in the study by Marian et al. (2003). In this study the size of the interference effect was equivalent in both L2 and L1. Although the fixation time difference between targets and interlingual distracters was also the same in both L2 and L1 in the study by Spivey and Marian (1999), the L2 results demonstrated that participants fixated as much on interlingual distracter items as on unrelated distracter items. According to the authors this asymmetry across the two languages reflects a general tendency to scan the entire display before fixating on the target when instructions are presented in L2.

3. Although French is in fact the second language of children raised in Flanders, we consider it here as the third language because our participants are much more proficient in English.

4. We reported these measures because the native and non-native speaker differed in pronunciation duration (see further). When latency analyses were based on reaction times measured from (auditory) target onset, the same pattern of results was obtained.

5. To investigate the effect of word frequency on the homophone effect, we ran an ANOVA (including the data of Experiments 1A and 1B) with target type (interlingual homophone vs. control) and word frequency (low vs. high) as the independent variables and RTs as the dependent variable. This analysis revealed a significant interaction between target type and word frequency $F1(1,64) = 28.54$, $p < .001$; $F2(1,72) = 15.57$, $p < .001$, indicating that the homophone effect was larger for homophones with a low frequent English meaning.

6. We opted for a control experiment in another country because a control experiment with these homophones in Dutch monolinguals is by definition not possible. According to Graddol (2004) in the Netherlands nearly 80 % of the population claims fluency in English. We suspect that this number is comparable in the Dutch-speaking community of Belgium. For our study, this implies that the entire population has at least some knowledge of English, except specific groups that would also differ on several other related variables.

7. To investigate the effect of word frequency on the homophone effect, we ran an ANOVA (including the data of Experiments 3A and 3B) with target type (interlingual homophone vs. control) and word frequency (low vs. high) as the independent variables and RTs as the dependent variable. This analysis revealed a significant interaction between target type and word frequency $F1(1,57) = 4.88$, $p < .05$; $F2(1,74) = 7.63$, $p < .01$, indicating that the homophone effect was larger for homophones with a low frequent Dutch meaning.

Figure Captions

Figure 1. Graphical presentation of RTs on homophones and matched control words in Experiment 1, when targets were pronounced by a native Dutch speaker (1A) and a native English speaker (1B). The vertical bars represent the 95 % confidence interval.

Figure 2. Graphical presentation of RTs on homophones and matched control words in Experiment 2, when targets were pronounced by a native Dutch speaker (2A) and a native English speaker (2B). The vertical bars represent the 95 % confidence interval.

Figure 3. Graphical presentation of RTs on homophones and matched control words in Experiment 3, when targets were pronounced by a native Dutch speaker (3A) and a native English speaker (3B). The vertical bars represent the 95 % confidence interval.

Figure 1

Graphical presentation of RTs on homophones and matched control words in Experiment 1, when targets were pronounced by a native Dutch speaker (1A) and a native English speaker (1B). The vertical bars represent the 95 % confidence interval.

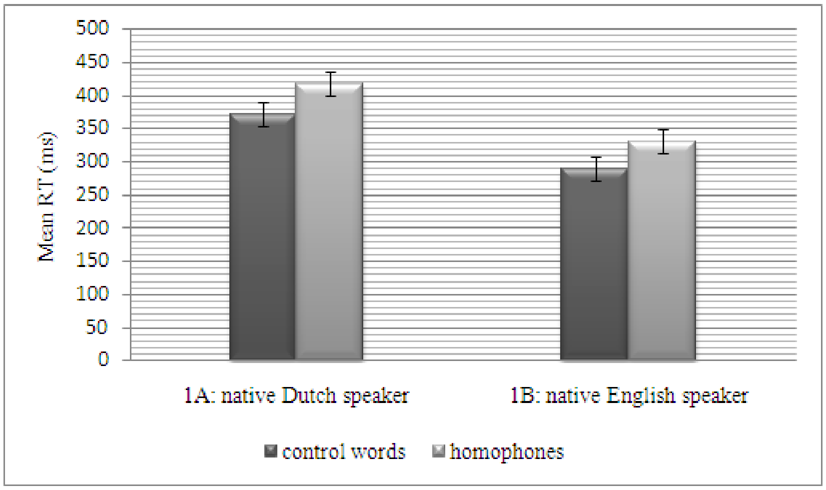

Figure 2

Graphical presentation of RTs on homophones and matched control words in Experiment 2, when targets were pronounced by a native Dutch speaker (2A) and a native English speaker (2B). The vertical bars represent the 95 % confidence interval.

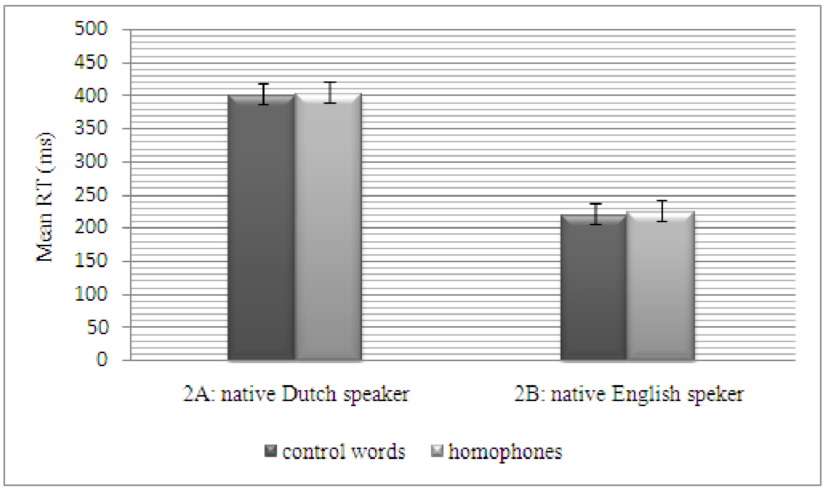

Figure 3

Graphical presentation of RTs on homophones and matched control words in Experiment 3, when targets were pronounced by a native Dutch speaker (3A) and a native English speaker (3B). The vertical bars represent the 95 % confidence interval.

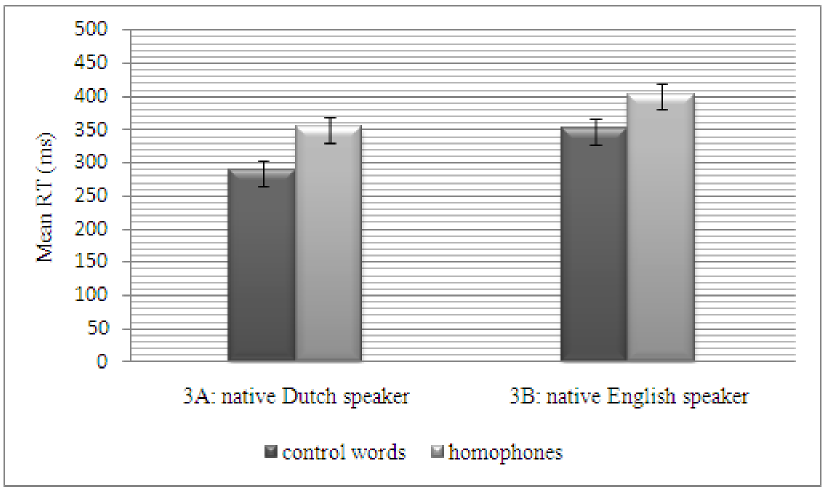

Table 1

Self-reported rating (7-point Likert Scale) of L1, L2, and L3 Proficiency (Experiments 1A, 1B, 3A, and 3B).

Language	Skill	Experiment 1A	Experiment 1B	Experiment 3A	Experiment 3B
	Writing	5.85 (0.74)	6.31 (0.58)	5.97 (0.93)	6.31 (0.60)
	Speaking	6.00 (0.70)	6.46 (0.61)	6.03 (0.97)	6.38 (0.82)
L1 (Dutch)	Reading	6.00 (0.89)	6.57 (0.56)	6.44 (0.56)	6.55 (0.57)
	Understanding	6.85 (0.36)	6.54 (0.56)	6.38 (0.58)	6.55 (0.51)
	General Proficiency	5.94 (0.55)	6.37 (0.55)	6.28 (0.58)	6.45 (0.63)
	Writing	4.65 (0.98)	4.89 (0.80)	4.97 (0.78)	5.00 (0.93)
	Speaking	4.74 (1.02)	5.14 (0.91)	5.38 (0.75)	5.38 (0.86)
L2 (English)	Reading	5.09 (0.97)	5.71 (0.83)	5.72 (0.85)	5.59 (0.98)
	Understanding	5.38 (0.78)	5.57 (0.78)	5.69 (0.90)	5.66 (0.77)
	General Proficiency	4.91 (0.79)	5.20 (0.83)	5.34 (0.70)	5.34 (0.72)
L3 (French)	General Proficiency	4.03 (1.14)	4.29 (1.10)	4.06 (1.05)	4.48 (0.99)

Table 2

Mean RTs and Effect (in Milliseconds), and Accuracy (% Errors) as a function of word type (both RT and Accuracy) for Experiment 1A and 1B.

	Controls		Homophones		Effect	
Speaker	RT	% Errors	RT	% Errors	RT	% Errors
Experiment 1A: Native Dutch speaker	372	11.15	417	20.95	45	9.80
Experiment 1B: Native English speaker	290	10.57	331	15.10	41	4.53

Table 3

Mean target durations (in ms) in Experiment 1 and 3 when targets were pronounced by the Dutch or the English native speaker as a function of word type. Standard deviations are presented between brackets.

	Native Dutch speaker		Native English speaker	
	controls	homophones	controls	homophones
Experiment 1: L2 lexical decision task	437 (80)	449 (70)	606 (94)	599 (100)
Experiment 3: L1 lexical decision task	435 (73)	406 (68)	604 (130)	567 (115)

Table 4

Mean RTs and Effect (in Milliseconds), and Accuracy (% Errors) as a function of word type (both RT and Accuracy) for Experiment 2A and 2B.

	Controls		Homophones		Effect	
Speaker	RT	% Errors	RT	% Errors	RT	% Errors
Experiment 2A: Native Dutch speaker	399	14.12	402	15.69	3	1.57
Experiment 2B: Native English speaker	218	6.58	223	5.87	5	-0.71

Table 5

Mean RTs and Effect (in Milliseconds), and Accuracy (% Errors) as a function of word type (both RT and Accuracy) for Experiment 3A and 3B.

	Controls		Homophones		Effect	
Speaker	RT	% Errors	RT	% Errors	RT	% Errors
Experiment 3A: Native Dutch speaker	289	5.51	354	13.98	65	8.47
Experiment 3B: Native English speaker	352	10.43	404	21.05	52	10.62