Is a FAN Always FUN?

Citation for published version:
Ota, M, Hartsuiker, R & Haywood, SL 2010, 'Is a FAN Always FUN? Phonological and Orthographic Effects in Bilingual Visual Word Recognition' Language and Speech, vol. 53, no. 3, pp. 383-403. DOI: 10.1177/0023830910371462

Digital Object Identifier (DOI):
10.1177/0023830910371462

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Language and Speech

Is a FAN always FUN? Phonological and orthographic effects in bilingual visual word recognition

Mitsuhiko Ota

University of Edinburgh

Robert J. Hartsuiker

Ghent University

Sarah L. Haywood

University of Edinburgh

Running head: Phonology and orthography in bilingual word recognition

Author correspondence to:

Mitsuhiko Ota (mits@ling.ed.ac.uk)

School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK

Acknowledgements:

This study was sponsored by a British Academy grant (SG-33008) awarded to Mitsuhiko Ota and Rob Hartsuiker, an Edinburgh University Development Trust Research Fund grant (EO8679) awarded to Rob Hartsuiker, and a British Academy Postdoctoral Fellowship (PDF/2005/131) awarded to Sarah Haywood. We thank Ellen Simon, Mieke Van Herreweghe, Paul Siakaluk and an anonymous reviewer for their helpful comments on an earlier version of the manuscript.

ABSTRACT

A visual semantic categorization task in English was performed by native English speakers (Experiment 1) and late bilinguals whose first language was Japanese (Experiment 2) or Spanish (Experiment 3). In the critical conditions, the target word was a homophone of a correct category exemplar (e.g., A BODY OF WATER – SEE; cf. SEA) or a word that differed from the correct exemplar by a phonological contrast absent in the bilinguals' first language (e.g., USED FOR COOLING DOWN – FUN; cf. FAN). Homophones elicited more false positive errors and slower processing than spelling controls in all groups. The Japanese-English bilinguals, but not the Spanish-English bilinguals, also displayed 'near-homophone' effects (i.e., homophone-like effects from minimal pairs on nonnative contrasts). We conclude that second-language visual word recognition is influenced by first-language phonology, although the effect is conditioned by the first-language orthographic system. Near-homophone effects can occur when the orthographic systems of the late bilingual's two languages are different in type (e.g., alphabetic vs. non-alphabetic), but may be blocked if the languages use the same writing script (e.g., Roman alphabet).

Keywords: visual word recognition, phonology, orthography, second language, bilinguals

INTRODUCTION

Although identification of written or printed words begins with the visual processing of letter symbols, there is substantial evidence that the phonological information behind the orthographic representations plays a crucial role in the process. Recoding of orthography to phonological information has been shown to occur during or even before readers access the lexical entries of visually presented words. In lexical identification tasks, responses to target words can be primed by prior or immediately subsequent exposure to phonologically similar visual words (e.g., Berent & Perfetti, 1995; Brysbaert, 2001; Drieghe & Brysbaert, 2002; Grainger & Ferrand, 1994; Lukatela & Turvey, 1990; 1994). For example, identification of the target word (e.g., *rate*) is more likely when a pseudo-word briefly presented before the target is a pseudo-homophone (e.g., RAIT) rather than a graphemically similar pseudo-word (e.g., RALT) (Perfetti & Bell, 1991).[1]

Another body of research suggests that phonological information also mediates access to the meanings of visual words. Orthographically presented words have been shown to activate phonological information, which is then used in identifying the words, and hence their meanings. In semantic categorization tasks where participants are asked to decide whether a particular word is a member of a semantic category (e.g., FLOWER), participants are more likely to commit false positive errors for homophones (e.g., ROWS) and pseudo-homophones (e.g., ROWZ) than for a spelling-matched control (e.g., ROBS) (Van Orden, 1987; Van Orden, Johnston, & Hale, 1988; Van Orden, Pennington, & Stone, 1990). This indicates that words that share phonological representations are confusable because identification of visually presented words is dependent on the phonological information readers access in orthographic words. Similarly, in semantic-relatedness judgment tasks, participants are less accurate and slower in rejecting unrelated pairs of words, such as LION-BARE, where one member of the pair is a homophone of a word related to the other member (i.e., BEAR in this example) (Lesch & Pollatsek, 1998; Luo, Johnson, & Gallo, 1998). Further evidence for such phonological mediation comes from priming experiments showing that reaction times in a lexical decision task are reduced by a prime word whose homophone is semantically related to the target, for example BEACH (homophone of BEECH) for TREE (e.g., Lesch & Pollatsek, 1993; Lukatela & Turvey 1991), and also from proofreading and eye movement data (Jared & Rayner, 1999).

All the research mentioned above has been conducted within essentially monolingual populations. In the case of bilingual speakers, particularly late bilinguals who learned one language (the native language, or L1) as a child and the second language (L2) later in life, questions arise as to the extent to which phonological mediation can take place in the visual recognition of L2 words, and whether the relevant phonological information can originate in the L1 in addition to the language of the orthographic words. In Ota, Hartsuiker, and Haywood (2009), we have addressed these questions through a semantic-relatedness judgment task of English words, designed after Luo et al. (1998) and given to three groups of participants: native speakers of English, Japanese, and Arabic. Just as expected, native speakers of English were less accurate and slower in rejecting pairs that contained a word with a homophone related to the other member of the pair (e.g., MOON – SON) in comparison to spelling controls (e.g., MOON – SIN). This was also the case with the nonnative speakers of English, demonstrating that the phonology of L2 words is being processed in L2 visual word recognition too. However, unlike the native English speakers, the nonnative participants also exhibited homophone-like effects in judging pairs that contained a word that differed phonologically from a related word by a segmental contrast missing in their L1. Thus, native speakers of Japanese had relatively more errors and slower response times in rejecting pairs such as KEY – ROCK (ROCK is a 'near-homophone' of LOCK without the /l/-/r/ contrast, which Japanese lacks), and native speakers of Arabic showed comparable effects with pairs such as SAND – PEACH (PEACH is a near-homophone of BEACH without the /p/-/b/ contrast, which Arabic lacks). The results from this study not only show that phonological mediation can take place in bilingual visual word recognition but also that the phonology of the L1, in addition to that of the L2, is involved in the silent reading of L2 words. Late bilingual readers confuse orthographic L2 words representing phonologically distinct lexical items when the relevant sound contrast is absent in their L1. This happens despite the possibility of directly accessing lexical meanings from orthographic representations (the so-called direct access or lexical route; Coltheart, Davelaar, Jonasson, & Besner, 1977, Paap & Noel, 1991, Rubenstein, Lewis, & Rubenstein, 1971).

There are indications, however, that phonological mediation in bilingual visual word recognition is conditioned by the orthographic systems of the two languages involved. For instance, Wang, Koda, and Perfetti (2003; 2004) investigated the effects of different L1 writing systems on L2 reading and concluded that native speakers of a non-alphabetic language (Chinese) relied less on phonological information than native speakers of an alphabetic language (Korean) in reading an alphabetic L2 (English) (see, however, the criticisms raised by Yamada, 2004). Kim and Davis (2003), on the other hand, found no homophone priming in lexical decision or semantic categorization performed by Korean(L1)-English(L2) bilinguals. If the L1 and the L2 share a script, the grapheme-phoneme conversion (GPC) rules of the two languages also seem to affect each other. Thus, bilinguals experience crosslinguistic effects when processing interlingual homographs (words with the same spelling but a different meaning across languages) (Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Haigh & Jared, 2007; Kerkhofs, Dijkstra, Chwilla, & de Bruijn, 2006; Lemhöfer & Dijkstra, 2004). For instance, a visually presented nonword in bilinguals' L1 can prime a homophone in their L2 (Brysbaert, Van Dyck, & Van de Poel, 1999). An example for Dutch-English bilinguals is *bleem* and *blame*: the nonword *bleem* is homophonous to the English word *blame* only when read according to Dutch GPC rules. Van Wijnendaele and Brysbaert (2002) demonstrated that such crosslinguistic priming occurs not only from L1 to L2, but also from L2 to L1. These studies suggest that in order to understand when and which of the two languages of the bilingual can affect L2 visual word recognition, we need to take into consideration the orthographic differences between the languages involved.

In sum, previous research on bilingual visual word recognition has shown that lexical access can be mediated by the L1 phonology under certain conditions, but it is still not clear to what extent this effect arises independent of tasks, phonological contrasts, and the writing systems of the L1 and L2. The purpose of this study was to address two aspects of this issue. First, we were interested in finding out whether the near-homophone effect that was observed in the semantic-relatedness judgment task of Ota et al. (2009) is robust enough to be replicated in a different experimental paradigm and in other nonnative contrasts. Our second aim was to explore the role of L1 orthography in L2 phonological mediation. Specifically, we set out to test whether phonological meditation by L1 phonology in L2 visual word recognition can be observed when the L1 shares a writing script with the L2. In Ota et al. (2009), we compared an alphabetic L1 (Arabic) with a non-alphabetic L1 (Japanese), and demonstrated that in both cases, phonological mediation from the L1 occurs in L2 visual word recognition, contrary to predictions following Wang et al. (2003, 2004). However, neither of the two L1s investigated had the same writing script as the L2 English (i.e., the Roman alphabet). The effects of L1 GPC rules on phonological mediation in L2 word recognition were, therefore, largely left unexplored.

To this end, we ran three experiments using a semantic category decision task designed after Van Orden (1987). In this task, participants are asked to judge whether a particular word matches a semantic category or definition. A genuine homophone effect induces more false positives when the target word is a homophone of a category-matched word (e.g., ‘A FLOWER’ – ‘ROWS’) than when the target is a spelling control word (e.g., ‘A FLOWER’ – ‘ROBS’). A near-homophone effect would induce more false positives when the target word differs from a category-matched word by a contrast lacking in the L1 (e.g., ‘A FASTENING DEVICE’ – ‘ROCK’ for native speakers of Japanese) than when the target is a spelling control (e.g., ‘A FASTENING DEVICE’ – ‘SOCK’). Participants were native speakers of English (Experiment 1), late Japanese-English bilinguals with Japanese as the L1 (Experiment 2), and late Spanish-English bilinguals with Spanish as the L1 (Experiment 3). In order to test near-homophone effects, we included English minimal pairs on /æ/-/ʌ/, /b/-/v/, and /l/-/r/. Japanese lacks all of these contrasts, and Spanish lacks all but the last.

As in Ota et al. (2009), participants were screened for their ability to auditorily identify the critical phonemes in English. The motivation behind this was twofold. First, we wanted to ensure that the two late bilingual groups were comparable in their

L2 phonological abilities. Second, we intended to exclude participants who could not identify these English phonemes at all. By eliminating such individuals, we were focusing on the effects of L1 phonology on the L2 lexicon rather than the phonetic perception of L2 sounds. Although constructing distinct phonological representations in the L2 lexicon usually requires reliable perception of the relevant L2 contrasts, recent work on L2 phonology has shown that perception and lexical representation do not always entail each other (Cutler, Weber, & Otake, 2006; Escudero, Hayes-Harb, & Mitterer, 2008; Pallier, Colomé, & Sebastián-Gallés, 2001; Weber & Cutler, 2004). Most importantly, indeterminate lexical coding of L2 sounds is not necessarily attributable to misperception (Ota et al., 2009; Pallier et al., 2001). L2 contrasts lacking in the L1 may be perceived but not be encoded reliably in the L2 lexicon, rendering words that differ by such contrasts functionally homophones. In Pallier et al. (2001), for example, Spanish-dominant bilingual speakers of Spanish and Catalan (but not their Catalan-dominant counterparts) showed repetition priming across Catalan minimal-pair words that differ in a phonological contrast lacking in Spanish (e.g., /e/-/ɛ/, /s/-/z/). The effect was observed even though the Spanish-dominant bilinguals performed just as well as the native Catalan speakers in a lexical decision task that involved exactly the same contrasts, indicating that their auditory perception of the relevant sounds was accurate. It is not entirely clear why a sound contrast that is perceivable by L2 speakers can still fail to fully separate the lexical representations of words that differ by that contrast. One possibility is that traces of acoustic exemplars that can be discriminated acoustically may converge during memory consolidation into lexical representations. Alternatively, L2 sounds that do not contrast in the L1 may be coded in the mental lexicon with a single representational category that corresponds to the two L2 sounds (Pallier et al., 2001). Either way, non-L1 contrasts may have overlapping phonological representations in the L2 lexicon, such that activation of one member of the pair will also activate the other member. It was this type of representational ambiguity in the bilingual lexicon that we aimed to address.

Previous studies using semantic categorization tasks have shown consistent homophone effects in accuracy but somewhat mixed results in latency (cf. Coltheart, Patterson, & Leavy, 1994; Jared & Seidenberg, 1991; Van Orden, 1987; Van Orden et al., 1988). For this reason, our main predictions focused on error patterns. In Experiment 1, we predicted that native speakers of English would make more false positive errors judging homophones in comparison to spelling controls, but they would not show any near-homophone effects from the minimal pairs. In Experiment 2, we predicted that Japanese-English bilinguals would make relatively more false positive errors in both homophones and minimal pairs involving all three English phonological contrasts. For Experiment 3, in which Spanish-English bilinguals were tested, we needed to factor into our predictions the fact that Spanish and English share a writing script (the Roman alphabet). If native speakers of Spanish reading L2 English words do not access Spanish GPC rules, the process of phonological mediation should be the same as the Japanese-English bilinguals, and we predict more false positive errors in homophones and minimal pairs involving the two phonological contrasts missing in Spanish (i.e., /æ/-/ʌ/ and /b/-/v/, but not /l/-/r/). If, on the other hand, Spanish-English bilinguals access Spanish GPC rules (as well as English ones) while processing English visual words, we may not find near-homophone effects in English minimal pairs involving the /æ/-/ʌ/ contrast, such as FAN and FUN. This is because in Spanish orthography the letters <A> and <U> are consistently mapped onto the phonemes /a/ and /u/, respectively. Therefore, any potential confusion between English words that involve the non-Spanish contrast /æ/-/ʌ/ may be prevented by the simultaneous activation of the native contrast /a/-/u/. The situation is different for the /b/-/v/ pairs because the orthographic contrast between <B> and <V> does not translate into a phonological contrast in Spanish; both letters are mapped onto the same phoneme /b/. Thus, we still expect the Spanish-English bilinguals to show near-homophone effects for English minimal pairs that involve the /b/-/v/ contrast even if Spanish GPC rules are accessed during English visual word recognition.

EXPERIMENT 1: NATIVE SPEAKERS OF ENGLISH

Each experiment involved a phoneme identification screening session and (for those participants passing the screening criteria) an experimental session, consisting of the semantic categorization task, followed by an off-line screening test of lexical knowledge. The lexical knowledge task was added to allay concerns that any attested homophone effects are due to inaccurate knowledge of the spelling of the test items (Starr & Fleming, 2001). Appendices B and C describe the method of the screening tests.

Method

Participants. Twenty native speakers of English (4 males and 16 females) took part in the experiment. They were all members of the University of Edinburgh community and were paid for their participation.

Materials. Thirty-two word pairs served as experimental stimuli, which consisted of 8 minimal pairs for each of the three phonological contrasts as well as 8 pairs of homophones. All minimal pairs contained the single letters <A>, <U>, <B>, <V>,

<L>, and <R> corresponding to the phonemes /æ/, /ʌ/, /b/, /v/, /l/, and /r/, respectively. These phoneme-grapheme pairs have a fairly high rate of correspondence in English, ranging from 86% (for /ʌ/-<U>) to 99% (for /v/-<V>) in type count (Hanna, Hanna, Hodges, & Rudorf, 1966). Hereafter, we will refer to the minimal pairs on each of the three phonological contrasts /æ/-/ʌ/, /b/-/v/, and /l/-/r/ as the A-U pairs (e.g., FAN-FUN), B-V pairs (e.g., BET-VET), and L-R pairs (e.g., LOCK-ROCK), respectively.

A minimally different spelling control was coupled to each pair, with the constraints that the control differed in only a single grapheme from either member of the experimental pair and that its phonological difference from each member of the pair would not involve a contrast missing in Japanese or Spanish.[2] For example, FAN-FUN was given FIN as a spelling control. We used separate spelling controls for each member of the pairs BRAKE-BREAK and BOAT-VOTE in order to compensate for the large difference in orthography between the pair members.[3] See Appendix A for a complete list of experimental and control items. The mean log frequency counts for the experimental words and spelling controls (based on the English wordform frequency list of the CELEX lexical database; Baayen, Piepenbrock, & Van Rijn, 1993) were 1.4 vs. 1.3 for real homophones, 1.3 vs. 1.1 for A-U, 1.0 vs. 1.1 for B-V, and 1.5 vs. 1.4 for L-R. The differences between these means did not approach statistical significance [t< 1 for homophones, B-V, and L-R; $t(15) = 1.33$, $p = .21$ for A-U]. The experimental words were divided into two lists matched in frequency so that a given participant would see only one member of each homophone or minimal pair. Each list was presented to half the participants, and participants were randomly assigned to the lists.

For each triplet of critical items (i.e., homophone/minimal pair and its spelling control), two short category definitions were written, one corresponding to each pair member. For example, for the triplet (FAN-FUN; FIN) we constructed the definitions 'USED FOR COOLING DOWN' and 'SOMETHING ENJOYABLE'. The definitions were then coupled to the opposite member of the homophone/minimal pair as well as the control word. In this case, half the participants saw 'USED FOR COOLING DOWN' (the definition intended for FAN) coupled to the words FUN and FIN (See Table 1). The other half saw 'SOMETHING ENJOYABLE' coupled to FAN and FIN. Thus, each participant saw the same category definition twice, once with the foil (the wrong member of the homophone/minimal pair) and once with the spelling control, and the correct response was 'no' for both.

INSERT TABLE 1 ABOUT HERE

In addition to these critical items, each stimulus list also contained 128 filler items. The filler materials consisted of 64 semantic definitions, each appearing twice but with different words. For 64 definition-word combinations, the correct response was 'yes' on both occasions (e.g., 'A BIRD' – 'DUCK', 'SWAN'), for 32 definition-word combinations it was 'yes' on the first occasion and 'no' on the second occasion (e.g., 'A TYPE OF TV PROGRAMME' – 'NEWS', 'CORN'), and for the remaining 32 combinations it was 'no' on the first, but 'yes' on the second occasion (e.g., 'AN ACCESSORY – 'MASS', 'RING'). Since each list contained 64 critical definition-word combinations, all for which the correct response was 'no', exactly half of the complete set of critical and filler items a given participant saw required a 'yes' response.

The two stimulus lists had the same pseudo-random order of critical and filler materials so that a given critical item in List A would occur in the same position as the corresponding item in List B. Furthermore, each list was divided in half, so that all 96 definitions were presented once before any definition was repeated. Each half was completely balanced with respect to the number of items in each experimental condition and with respect to the number of items requiring a 'yes' or 'no' response.

Procedure. Each trial began with a fixation point presented in the center of the screen for 1000 ms, followed by a definition. The definition remained on the screen for 1500 ms. Subsequently, a word was presented for 300 ms, followed by a mask (a row of #-signs, as many as the number of letters in the stimulus word). The participants responded by pushing the right ('yes') or left ('no') button of a button box. The word was presented in an 18 point courier new font. The session began with 44 practice trials. The first five trials of the experiment proper were fillers. The experiment was broken into four quarters and participants were given an opportunity to take a self-timed break between each quarter.

Data analysis. We analyzed the proportion of errors in each condition (experimental, control) for each phonological/orthographic contrast (homophones, A-U, B-V, and L-R) separately. For each comparison, we constructed generalized linear mixed effects models with condition as a fixed effect, and with subject and item (i.e., definition) as crossed random effects (Baayen, Davidson, & Bates, 2008). Because error data are categorical, we used the logistic link function in these analyses (Jaeger, 2008). The models were implemented using the *lme4* library (Bates & Sarkar, 2007) in R (R Development Core Team, 2006).

We conducted a lexical knowledge test after the semantic categorization task (see Appendix C for the methodological details of this test). If a participant responded 'I don't know' or made a mistake on any combination of a test word (either an experimental or control item) and a definition in the lexical knowledge test, we excluded the participant's response to that word and also the item matching that word in the semantic categorization task. For example, if a participant's response to 'SOMETHING ENJOYABLE - FUN' in the lexical knowledge task was 'no' or 'I don't know', we removed that person's semantic categorization task responses for FUN and for the spelling control FIN.

Two control items in the B-V condition proved to be problematic and were excluded from the analysis, along with the corresponding experimental items in all experiments. First, we realized only in retrospect that the control word CAN has the sense 'indicating permission' (this is the second sense for the verb in the Oxford Advanced Learner's Dictionary, Fourth Edition) which is semantically related to our definition of BAN: 'FORBID OR PROHIBIT'. Second, in the lexical knowledge test, 20% of the native speakers rejected the definition 'A PARKING ATTENDANT' for VALET (presumably because Edinburgh 'parking attendants' monitor for parking violations rather than park cars for customers).

Although our main interest is in false positive errors, we also analyzed reaction time data, following the same procedural steps taken for the analysis of error rates. Here we obtained *p*-values by way of Monte Carlo Markov Chain simulations, based on 10,000 samples (Baayen et al., 2008). Additionally, we now excluded trials on which participants made an error and trials with latencies that were more than 2.5 Standard Deviations above the mean.

Results and Discussion

Not surprisingly, the native speakers of English scored at ceiling on the phoneme identification task. Their mean correct score was 100% for /æ/-/ʌ/, 99% for /b/-/v/, and 99% for /l/-/r/. However, they were not at ceiling on the lexical knowledge test: There were 11 false rejections (involving the words *suck*, *flash*, *flesh*, *boat*, *ton*, *melt*, and *right*). The corresponding observations in the semantic categorization task (as well as their matched observations) were excluded from the error and reaction time analyses. These accounted for 1.8% of the data.

After exclusions, there remained 42 errors (3.5%) in 1,178 responses. As predicted, errors were more common in the homophone condition than in the corresponding spelling control condition (see Table 2). There was a significant effect of condition (coefficient = 1.82, SE = 0.62, Z = 2.92, $p < .01$). As expected, none of the minimal pair contrasts showed a significant difference between the experimental items and spelling controls: A-U: coefficient = 1.07, SE = 0.77, Z = 1.39, $p = .16$; B-V: coefficient = 0.0003, SE = 1.007, Z = 0.0003, $p = 1$; L-R: coefficient = 0.74, SE = 1.25, Z = 0.59, $p = .55$.

INSERT TABLE 2 ABOUT HERE

The bottom panel of Table 2 shows the mean response latency for each contrast, with exclusions of observations that were errors (3.5%) or outliers (1.8%). The resulting analyses were based on 1,115 observations. Participants were significantly slower at correctly rejecting homophones than matched control items (coefficient = 55.48, SE = 18.13, t = 3.06, $p < .01$). As expected there was no effect for the A-U contrast (coefficient = 32.25, SE = 18.92, t = 1.70, $p = .09$) or for the B-V contrast (coefficient = -22.09, SE = 21.49, t = -1.03, $p = .31$). Unexpectedly, experimental items in the L-R contrast were rejected more slowly than their controls (coefficient = 38.91, SE = 18.31, t = 2.13, $p < .05$).

The accuracy results in Experiment 1 replicated the finding of Van Orden (1987). More false positive errors were elicited by homophones (e.g., SON paired with the definition of SUN) than by spelling controls (e.g., SIN with the definition of SUN). Furthermore, no difference was found between minimal pairs (e.g., FAN-FUN) and their spelling controls (e.g., FIN), indicating that the homophone effect does not extend to minimal pairs. These outcomes confirm our predictions.

By and large, the response latencies were also as predicted, with longer reaction times induced by homophones but not by minimal pairs. One exception to this general pattern was the L-R contrast, where a significant difference was found between the experimental items and the spelling controls. This may be a spurious outcome, as no such effect was exhibited by the native English speakers' L-R performance in the semantic-relatedness judgment task of Ota et al. (2009). It is of course possible that certain phonological contrasts, such as /l/-/r/, can slow down, if not cause more errors in, semantic category decisions. A variety of memory retention and speech error evidence shows that /l/ and /r/ are one of the most similar sounding pairs in English (Wicklegren, 1965; Stemberger, 1991; Frisch, Pierrehumbert, & Broe, 2004), and the proximity of the two sounds might cause homophone-like latency effects in semantic categorization. The issue will benefit from further investigation.

EXPERIMENT 2: JAPANESE-ENGLISH BILINGUALS

The same experiment as Experiment 1 was conducted with Japanese-English bilinguals. For this population, we predicted that the homophone effects should also be observed in the minimal pair items. That is, Japanese-English bilinguals should produce relatively large false positive error rates not only for homophones but also for words containing A-U, B-V, and L-R, because the phoneme pairs associated with these letters are not contrastive in Japanese.

Method

Participants. Thirty-one native speakers of Japanese with knowledge of English took part in the experiment. They were paid for their participation. Twenty participants reached the criterion for the phonological identification test (a minimum of 11/16 correct for each phonological contrast) and were invited to take part in the main experiment. The 20 participants in the main experiment (15 females and 5 males) were aged 27 years on average (range 20 - 35). They were all members of the Edinburgh University community. Indicators of their proficiency level in English are summarized in Table 3.

INSERT TABLE 3 ABOUT HERE

Materials, procedure, and data analysis. See Experiment 1.

Results and Discussion

The Japanese participants, although scoring better than chance, were not completely at ceiling on the phonological identification task. Their mean correct score was 86% for /æ/-/ʌ/, 93% for /b/-/v/, and 86% for /l/-/r/. They also selected the option 'I don't know' or produced errors several times in the lexical knowledge task. This led to the exclusion of 304 responses (25.3%) from the semantic categorization data.

After exclusions, there remained 165 errors (18.4%) in 896 responses. As predicted, errors were more common in the experimental condition than in the corresponding spelling control condition in all critical contrasts (Table 4). The comparison with the spelling control condition was significant for the homophones (coefficient = 2.59, SE = 0.51, Z = 5.11, $p < .001$), the A-U contrast (coefficient = 1.49, SE = 0.38, Z = 3.94, $p < .001$), the B-V contrast (coefficient = 1.46, SE = 0.51, Z = 2.86, $p < .01$), and the L-R contrast (coefficient = 0.76, SE = 0.36, Z = 2.11, $p < .05$).

INSERT TABLE 4 ABOUT HERE

The bottom panel of Table 4 shows the mean response latency for each contrast. The reaction time analysis excluded observations that were errors (18.4%) or outliers (2.7%), resulting in 707 remaining observations. The experimental items were rejected slower than spelling controls for each contrast. The effect on latencies was significant for the homophones (coefficient = 147.36, SE = 48.19, $t = 3.06$, $p < .01$), the A-U contrast (coefficient = 176.41, SE = 63.48, $t = 2.78$, $p < .01$), and the B-V contrast (coefficient = 134.47, SE = 60.32, $t = 2.23$, $p < .05$). However, the effect was not significant for the L-R contrast (coefficient = 62.47, SE = 58.13, $t = 1.08$, $p = .28$).

The relatively higher error rate in the homophone condition shows that processing of nonnative written words is also phonologically mediated just as in native visual word recognition. Furthermore, the error rates in the A-U, B-V, and L-R conditions support the prediction that the lack of the /æ/-/ʌ/, /b/-/v/, and /l/-/r/ phonemic contrasts in Japanese causes homophone-like effects for minimal pairs on these contrasts.

The latency results showed that semantic category decision was slowed down by homophones and minimal pairs, although with the exception of the L-R contrast. The lack of latency effects in the L-R pairs is surprising at first glance. However, this statistical null result may simply be due to lack of power. Because of the fairly large number of errors in the lexical knowledge test and the main task, the analysis for this condition was based on only 188 observations. Note that the reaction times show a difference of almost 100 ms in the predicted direction (Table 4).

EXPERIMENT 3: SPANISH-ENGLISH BILINGUALS

Experiment 3 was conducted with Spanish-English bilinguals. As with the Japanese-English bilinguals, we predicted that the homophone effects should be observed among the real homophone items.

As for the minimal pairs, different outcomes were anticipated depending on whether the late bilinguals accessed Spanish GPC rules as well as English ones during the task, a possibility due to the shared writing script in the two languages. The phonological contrast /æ/ vs. /ʌ/ (as in /fæn/ vs. /fʌn/), which is missing in Spanish, is represented by the orthographic contrast <A> vs. <U> (e.g., FAN vs. FUN) in the English words used in the experiment. But in Spanish orthography, <A> and <U> are associated with the Spanish phonemes /a/ and /u/, respectively. Therefore, if the Spanish-English bilinguals activated Spanish GPC rules (in addition to English GPC rules) while reading the English words, they would have access to distinctive Spanish phonological representations (e.g., /fan/ vs. /fun/), and they would not confuse orthographic English words that differ by <A> and <U>. If, however, no Spanish GPC rules were involved in processing English A-U pairs, we would expect near-homophone effects to emerge, as the English phonological contrast associated to the orthographic pair (/æ/ - /ʌ/) is not available in Spanish.

The contrast between /b/ and /v/ (as in /bɛt/ vs. /vɛt/) is also missing in Spanish. This phonological distinction is represented by the contrast between <B> and <V> in the English words in the experiment (e.g., BET vs. VET). However, unlike in the case of <A> and <U>, <B> and <V> are not associated with different phonemes in Spanish GPC. They are both mapped onto the single phoneme /b/, as evident in the pronunciation of *baca* (/baka/ 'luggage') and *vaca* (/baka/ 'cow'). We thus predict Spanish-English bilinguals to exhibit near-homophone effects for English B-V pairs regardless of their potential access to Spanish GPC rules.

The contrast between /l/ and /r/ exists in Spanish, and in both English and Spanish, the phonemes /l/ and /r/ are typically associated with the letters <L> and <R>, respectively. So we do not expect Spanish-English bilinguals to show near-homophone effects for L-R pairs.

In sum, our predictions for this experiment were as follows. There would be significantly more false positives for real homophones and B-V words than for their spelling controls. There would not be more errors produced for L-R words in comparison to their spelling controls. There would be significantly more false positive errors for A-U words than for the corresponding spelling controls, if the Spanish-English bilinguals had no access to Spanish GPC rules during the task, but no difference would be obtained if the bilinguals had access to Spanish GPC rules.

Method

Participants. Twenty-seven native speakers of Spanish were paid to participate. Twenty participants reached the criterion for the phonological identification test (the same criterion as in Experiment 2) and were invited to take part in the main experiment. The twenty participants in the main experiment (14 females and 6 males) were on average 25 years old (range 19 - 30). They were all members of the Edinburgh University community. Indicators of their proficiency level in English are summarized in Table 3. None of these indicators differed significantly from those of the Japanese-English bilingual participants in Experiment 2, except for self-ratings for L1 and L2 reading skills.

Materials, procedure, and data analysis. See Experiment 1.

Results and Discussion

In the phoneme identification task, the Spanish-English bilinguals had a mean correct score of 92% for /æ/-/ʌ/, 92% for /b/-/v/, and 99% for /l/-/r/. As in the case of the Japanese participants, some Spanish speakers selected 'I don't know' in the lexical knowledge task or made errors on that task. All corresponding responses for these items were excluded from the main task (282 responses; 23.5%). After exclusions, there remained 66 errors (7.2%) in 918 responses. The breakdown of the errors is shown in Table 5. Errors were significantly more common in the homophone condition than in the corresponding spelling control condition (coefficient = 2.41, SE = 0.71, Z = 3.40, $p < .001$). No significant differences were found for the A-U contrast (coefficient = 0.63, SE = 0.59, Z = 1.08, $p = .28$), the B-V contrast (coefficient = -1.13, SE = 0.67, Z = -1.68, $p = .09$), or the L-R contrast (coefficient = -0.42, SE = 0.72, Z = -0.58, $p = .56$).

INSERT TABLE 5 ABOUT HERE

The bottom panel of Table 5 shows the reaction times for the correct responses in each category. The reaction time analysis excluded observations that were errors (7.2%) or outliers (2.2%), resulting in 832 remaining observations. Critical items were rejected significantly slower in the homophone than in the control condition (coefficient = 125.99, SE = 49.45, $t = 2.55$, $p < .05$). No significant differences were found for the A-U contrast (coefficient = -7.10, SE = 46.30, $t = -0.15$, $p = .88$), the B-V contrast (coefficient = 58.68, SE = 48.90, $t = 1.20$, $p = .23$), or the L-R contrast (coefficient = 43.19, SE = 42.74, $t = 1.01$, $p = .31$).

The results are clear-cut. In both accuracy and latency, Spanish-English bilinguals showed homophone effects, but no near-homophone effects. The lack of A-U effects is consistent with the L1 GPC access scenario we described above. The mapping of <A> to /a/ and <U> to /u/ in Spanish orthography could prevent the participants from confusing English orthographic words contrasting in <A> and <U>. This explanation is not applicable to the lack of effects in the B-V condition, however, because Spanish GPC rules map both <B> and <V> onto /b/. A possible account for this case will be discussed below.

GENERAL DISCUSSION

To summarize the results from the three experiments, accuracy data revealed homophone effects in all three groups, while near-homophone effects (i.e., false positive errors on minimal pairs of nonnative contrasts) were observed only in the Japanese-English bilingual group. Latency data largely followed the same pattern. We now discuss the implications of these findings for the two goals of the study.

The first purpose of this study was to test the robustness of near-homophone effects across tasks and contrasts. The outcomes from Experiments 1 and 2 show that the near-homophone effects obtained in the semantic-relatedness decision tasks of Ota et al. (2009) can indeed be replicated in a semantic category decision task and also in more than one nonnative contrast in a language. The Japanese-English bilinguals in Experiment 2 were less accurate and, for the most part, slower in rejecting category foils that involved minimal pairs of the nonnative /æ/-/ʌ/, /b/-/v/, and /l/-/r/ contrasts. In contrast, the native speakers of English in Experiment 1 exhibited such effects only from foils involving real homophones. In other words, when the L1 and L2 use different writing scripts, minimal pairs (involving nonnative contrasts) can induce homophone-like effects in L2 word recognition. Overall, these results add more support to our claim (Ota et al., 2009) that the lexical representations accessed during

L2 visual word processing can be undermined by the lack of relevant phonological contrasts in the L1.

The second purpose of our study was to examine whether near-homophone effects can still be observed when the L1 and the L2 share a writing script. Results from Experiment 3 indicate that the answer to this question is negative. There were two near-homophone sets for the Spanish-English bilinguals (A-U and B-V), but in neither set did we find a difference in accuracy or latency. This is in stark contrast with the near-homophone effects exhibited by Japanese-English bilinguals as well as the Arabic-English bilinguals in Ota et al. (2009), whose L1 writing systems do not involve the Roman alphabet. Note that this difference cannot be ascribed to a difference in phonological proficiency between the Spanish-English and Japanese-English bilinguals groups, whose accuracy scores of /æ/-/ʌ/ and /b/-/v/ identification in the pretest did not differ significantly. Nor is there evidence for difference between the groups in familiarity with lexical items used in the experiments, as the exclusion percentage from the lexical knowledge test (i.e., the proportion of errors and 'don't know' responses) was similar between the Spanish speakers (23.5%) and the Japanese speakers (25.3%). As can be seen in Table 3, the two groups are also comparable with respect to other potential indicators of general L2 proficiency level, such as length of residence in English-speaking countries, proportion of L2 use, and the age at which participants began reading English. One exception is self-rated L2 reading proficiency, which was lower for the Japanese group. But the Japanese speakers also assigned themselves a significantly lower *L1* reading proficiency, indicating that this difference is likely due to cultural, rather than linguistic, differences between the two groups. The most plausible reason for the lack of near-homophone effects in the Spanish-English bilinguals' A-U condition, therefore, is the fact that the L1 GPC rules map the letter <A> and <U> to different phonemes in Spanish: /a/ and /u/. If bilinguals have difficulties in 'turning off' their L1 GPC rules, as research on interlingual homophones suggests, then pairs such as FUN and FAN may also activate Spanish phonological representations (i.e., /fan/ and /fun/) which are different from the English phonological representations (i.e., /fæn/ and /fʌn/), and more importantly, are phonemically distinct in the L1. So even if the Spanish readers' /æ/-/ʌ/ contrast is indeterminate, the /a/-/u/ contrast Spanish can block phonological confusion between FAN and FUN.

The lack of a near-homophone effect in the B-V pairs cannot be explained in the same manner, since the orthographic contrast between <B> and <V> does not reflect a phonemic contrast in Spanish. But there is a plausible orthographic account for this too. Spanish has an extremely shallow orthography with very few cases of one-to-many or many-to-one spelling-sound correspondence. The one noticeable exception to this otherwise consistent system occurs with the letters <B> and <V> as illustrated by the examples of homophones already mentioned: *baca* (/baka/ 'luggage') and *vaca* (/baka/ 'cow'). Since the rare homophones in Spanish are found with these letters, experienced Spanish readers may have a high level of awareness of this orthographic contrast and monitor their performance more carefully in processing visual words containing <B> and <V> in their L2 as well. As a relevant observation, we note that the Spanish-English bilinguals in Experiment 3 produced fewer errors and longer latencies for the experimental items in the B-V condition in relation to their spelling controls. Although not statistically significant, this tendency towards a speed-accuracy tradeoff is consistent with the interpretation that the participants were engaged in a response strategy that compensated for the lack of phonological contrast corresponding to the difference between <B> and <V> in English or Spanish.

The different outcomes in the A-U and B-V conditions between the Spanish group and the Japanese group are, therefore, arguably both attributable to the fact that the Spanish-English bilinguals were processing L2 words orthographically represented in the same writing script as their L1. Taken together with the results from Ota et al. (2009), in which clear near-homophone effects were displayed by native speakers of Japanese and Arabic, the results from the current study indicate that we need to identify different levels of L1-L2 orthographic similarities to understand the exact impact of L1 orthography on the mediation of L1 phonology in L2 visual word recognition.

At one level, L1 phonological mediation is not contingent on a particular *class* of orthographic system. Both Japanese and Arabic use non-Roman-alphabetic writing systems, but while Arabic writing is alphabetic (in the sense that there is correspondence between letters and phonemes), none of the three orthographic systems used in Japanese is alphabetic; the two *kana* systems are fundamentally syllabaries, and the *kanji*, borrowed from Chinese, is largely ideographic. Although phonological information is encoded at some level of Japanese writing, orthography-phonology correspondence at the segmental level is marginal at best. Our Japanese-English data from Ota et al. (2009) and the current study provide evidence that even readers with a non-alphabetic L1 background engage in a considerable amount of phonological recoding in reading an alphabetic L2. This sheds new light on the recent debate between Wang et al. (2003, 2004) and Yamada (2004). According to Wang et al. (2003), native speakers of non-alphabetic languages should depend less on phonological information in identifying English words because their L1 orthography does not actively involve prelexical phonological processing. Although Wang et al.'s (2003) evidence comes from the contrast between Chinese (non-alphabetic) and Korean (alphabetic but non-Roman), the claim should apply to Japanese. While we do not dispute the relative L1 effects on the balance between phonological and orthographic information used in L2 reading, we have reasons to believe that phonological effects are robust enough to be found in L2 reading whether the L1 orthographic system is alphabetic or not — as long as the L1 and L2 scripts are different.

At another level, however, we have seen evidence that L1 phonological mediation is affected by L1 orthography when the L1 and L2 scripts are the same. The Spanish-English bilinguals in our study did not show near-homophone effects in silently processing English words. Although we appealed to different accounts for the A-U and B-V conditions, in both cases, the ultimate reason for the lack of near-homophone effects was attributed to the involvement of the L1 GPC, which could only occur if the L1 and the L2 used identical writing scripts. We therefore predict a systematic difference in L1 phonological mediation between late bilingual speakers of same-script languages (e.g., Turkish-English, German-Spanish, or Uzbek-Russian) and those of different-script languages (e.g., Hebrew-English, Greek-Spanish, or Chinese-Russian).

CONCLUSIONS

In this study, we examined the roles of L1 phonology and orthography in bilingual reading by comparing monolingual readers of English with bilingual readers of English from two different L1 backgrounds: Japanese (a non-Roman alphabetic language) and Spanish (a Roman alphabetic language). The semantic categorization results from the native English speakers and Japanese-English bilinguals generally show that the type of false positive errors elicited by homophones in monolinguals can also be induced by near-homophones in bilinguals, where such near-homophones involve phonological contrasts that are not present in the L1 of the bilingual reader. However, performance of Spanish-English bilinguals indicates that the near-homophone effect may be pre-empted when the L1 and L2 employ the same writing script. Our experiments show that the involvement of L1 phonology in L2 visual word recognition is dependent more on the identity between the L1 and L2 scripts rather than the types of orthographic system (e.g., alphabetic or non-alphabetic) used in the L1 and the L2.

FOOTNOTES

[1] In referring to orthographic words used in experimental work (including the current study), we put the words in the letter-case in which they were presented (e.g., WORD or *word*). Outside such methodological contexts, we put orthographic words in lowercase italics.

[2] The BAG-BUG pair and FLASH-FLUSH pair had spelling controls (BEG and FLESH) that differed from one of the pair members by an /æ/-/ɛ/ contrast, which does not exist in Japanese or Spanish. However, BAG-BEG and FLASH-FLESH were deemed not to be confusable because speakers of these languages are known to assimilate English /æ/ with their /a/ and English /ɛ/ with their /e/ (Flege, Bohn & Jang, 1995; Fox, Flege, & Munro, 1995; Strange et al., 1998).

[3] This orthographic discrepancy could not be resolved for BERRY and VERY, for which the spelling control was only MERRY.

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. Philadelphia: Linguistic Data Consortium.

Bates, D. M., & Sarkar, D. (2007). *lme4: Linear mixed-effects models using S4 classes*, R package version 0.99875-6.

Berent, I., & Perfetti, C. A. (1995). A Rose is a REEZ: The two-cycles model of phonological assembly in reading English. *Psychological Review*, *102*, 146-184.

Brysbaert, M. (2001). Prelexical phonological coding of visual words in Dutch: Automatic after all. *Memory & Cognition*, *29*, 765-773.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 137-148.

Coltheart, V., Davelaar, E., Jonasson, T. V., & Besner, D. (1977). Access to the internal lexicon. In D. Stanislav (Ed.), *Attention and performance*, vol. 6 (pp. 532-555). Hillsdale, NJ: Lawrence Erlbaum.

Coltheart, V., Patterson, K., & Leahy, J. (1994). When a ROWS is a ROSE: Phonological effects in written word comprehension. *The Quarterly Journal of Experimental Psychology*, *47A*, 917-955.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, *34*, 269-284.

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

Drieghe, D. & Brysbaert, M. (2002). Strategic effects in associative priming with words, homophones, and pseudohomophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 951-961.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*, 345-360.

Flege, J. E., Bohn, O.-S., Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437-470.

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, *97*, 2540-2551.

Frisch, S., Pierrehumbert, J., & Broe, M. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, *22*, 179-228.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33*, 218-233.

Hanna, P. R., Hanna, J. S., Hodges, R. E. & Rudorf, E. H. Jr. (1966). *Phoneme-grapheme correspondence as cues to spelling improvement*. Washington, DC: U.S. Office of Education.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 623-644.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Jared, D., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General*, *128*, 219-264.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General*, *120*, 358-394.

Kerkhofs, R., Dijkstra, T, Chwilla, D. J., & de Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs : RT and N400 effects. *Brain Research*, *1068*, 170-183.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, *49*, 484-499.

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

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Phonology: Learning, Memory, and Cognition*, *19*, 285-294.

Lesch, M. F., & Pollatsek, A. (1998). Evidence for the use of assembled phonology in accessing the meaning of printed words. *Journal of Experimental Phonology: Learning, Memory, and Cognition*, *24*, 573-592.

Lukatela, G., & Turvey, M. T. (1990). Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology*, *2*, 325-344.

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 951-966.

Lukatela, G., & Turvey, M. T. (1998). Learning to read in two alphabets. *American Psychologist, 53*, 1057-1072.

Luo, C. R., Johnson, R. A., & Gallo, D. A. (1998). Automatic activation of phonological information in reading: Evidence from the semantic relatedness decision task. *Memory and Cognition*, *26*, 833-843.

Ota, M., Hartsuiker, R. J., & Haywood, S. L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition*, *111*, 263-269.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print and sound: Still a good horse race. *Psychological Research*, *53*, 13-24.

Perfetti, C. A., & Bell, L. C. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, *27*, 59-70.

R development core team (2006). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Starr, M. S., & Fleming, K. K. (2001). A rose by any other name is not the same: The role of orthographic knowledge in homophone confusion errors. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*, 744-760.

Stemberger, J. (1991). Apparent anti-frequency effects in language production: The addition bias and phonological underspecification. *Journal of Memory and Language*, *30*, 161-185.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English voweld by Japanese speakers. *Journal of Phonetics*, *26*, 311-344.

Van Orden, G. C. (1987). A Rows is a Rose: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181-198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 371-386.

Van Orden, G. C., Pennington, B., & Stone, G. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488-522.

Van Wijnedaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*, *28,* 616-627.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and non-alphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, *87*, 129-149.

Wang, M., Koda, K., & Perfetti, C. A. (2004). Language and writing systems are both important in learning to read: A reply to Yamada. *Cognition*, *93*, 133-137.

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

Wicklegren, W. A. (1965). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America*, *39*, 388-398.

Yamada, J. (2004). An L1-script-transfer-effect fallacy: A rejoinder to Wang et al. (2003). *Cognition*, *93*, 127-132.

APPENDIX A. HOMOPHONES AND NEAR-HOMOPHONES USED IN EXPERIMENTS 1-3, WITH SPELLING CONTROLS

(Near-)Homophones		Spelling controls
Homophone Condition		
BRAKE	BREAK	BRAVE / BREAD
CENT	SENT	DENT
CELL	SELL	TELL
MEAT	MEET	MELT
SEA	SEE	SET
SON	SUN	SIN
HEAL	HEEL	HELL
TOE	TOW	TON
A-U Condition		
BAG	BUG	BEG
CAP	CUP	COP
FAN	FUN	FIN
FLASH	FLUSH	FLESH
HAT	HUT	HIT
MAD	MUD	MID
SACK	SUCK	SICK
TRACK	TRUCK	TRICK

APPENDIX A (Continued)

B-V Condition

*BAN	VAN	CAN
*BALLET	VALET	CHALET
BOAT	VOTE	GOAT / NOTE
BIKING	VIKING	HIKING
BENDING	VENDING	MENDING
BERRY	VERY	MERRY
BET	VET	NET
BEST	VEST	TEST

L-R Condition

LATE	RATE	DATE
LUST	RUST	DUST
LAP	RAP	MAP
LICE	RICE	NICE
LIGHT	RIGHT	NIGHT
LAW	RAW	PAW
LOCK	ROCK	SOCK
LOAD	ROAD	TOAD

Note. Unless specified, the spelling control was used in both lists. The two triplets marked with ‘*’ were discarded from all analyses.

APPENDIX B. METHOD OF THE PHONEME IDENTIFICATION SCREENING TASK

The screening task was administered to all participants. Only participants that performed above chance level for all three the contrasts /æ/-/ʌ/, /b/-/v/, and /l/-/r/ were invited to take part in the main experiment. It was a two-alternative forced choice matching task involving auditory and visual nonsense syllables.

Materials

The critical items were recordings of six phonological minimal pairs of nonsense syllables, two pairs for each phonological contrast. The syllables, spoken by a female native speaker of English (S.H.), were recorded on high quality tape in a studio and subsequently digitized. To match the requirement of the experimentation software (E-prime 1.0), the digital files were re-sampled at the rate of 11 kHz.

In two pairs, the syllables had the consonant-vowel structure CV, and in four pairs, the structure CVC(C). Additionally, 16 filler syllables were recorded with similar CV structures. The critical pairs were /pæz/-/pʌz/ and /tæsp/-/tʌsp/ (/æ/-/ʌ/ contrast); /ba/-/va/ and /bɪlp/-/vɪlp/ (/b/-/v/ contrast); /la/-/ra/ and /lɪlp/-/rɪlp/ (/l/-/r/ contrast). Each experimental item consisted of an auditory presented syllable (e.g., /pæz/), followed by two visually presented syllables in block letters, one that matched the auditory syllable (<PAZZ>) and one that matched the other member of the minimal pair (<PUZZ>). In order to camouflage the phonological contrasts of interest, filler items included experimental syllables paired with visually presented syllables differing on another contrast (<BA>-<DA>) and paired with visually presented syllables differing on more than one phonological feature (<LA>-<TA>).

Additionally, pairs of non-experimental syllables were auditorily presented with visual syllables differing in one feature (<SA>-<ZA>) or more features (<NA> – <DA>).

We compiled a master list containing 96 filler trials and 48 experimental trials (16 for each contrast). Within the experimental trials, each auditory syllable was presented four times, twice with the corresponding visual syllable on the left side of the screen and twice with that syllable on the right side of the screen. A separate random order was determined for each participant. The experiment began with 10 practice items.

Procedure

Each trial began with a fixation point, followed by the auditory presentation of the target syllable after 1500 ms. The audio files were presented using high quality headphones and volume was set to a comfortable level. Immediately after the offset of the syllable, the two visually presented syllables appeared, one to the left and to the right of the centre of the fixation point. The participants were instructed to press the left button of a button box if the auditory syllable matched with the visual syllable on the left, and the right button if it matched the visual syllable on the right.

APPENDIX C. METHOD OF THE LEXICAL KNOWLEDGE SCREENING TEST

The lexical knowledge test was an off-line definition/word matching test using the same 64 critical target words each participant saw in the semantic categorization task, but this time accompanied with a correct definition. In addition, 64 fillers were presented along with an incorrect definition. The materials were presented on a sheet of paper and the participant ticked one of three boxes (‘yes’, ‘no’, ‘I don’t know’). Two lists were constructed, which matched the two lists from the semantic categorization task, both in the same pseudo-random order.

TABLE 1

Example of an experimental and a control item, in each of two lists

Condition	List A	List B
Experimental		
Category	Used for cooling down	Something enjoyable
Target	FUN	FAN
Control		
Category	Used for cooling down	Something enjoyable
Target	FIN	FIN

TABLE 2

Mean percentage of errors and mean response latencies (standard error of the mean in parentheses) in each condition of Experiment 1 (native speakers of English)

	Experimental	Control	Difference
Errors (%)			
homophones	11.4 (2.5)	2.5 (1.3)	8.9
A-U	5.2 (1.8)	2.0 (1.1)	3.2
B-V	1.7 (1.2)	1.7 (1.2)	0.0
L-R	1.9 (1.1)	1.3 (0.9)	0.6
Response latencies (ms)			
homophones	697 (18)	648 (14)	49
A-U	685 (16)	659 (17)	26
B-V	636 (15)	655 (21)	-19
L-R	672 (16)	637 (15)	35

TABLE 3

Proficiency data for the participants in Experiment 2 (Japanese-English bilinguals) and Experiment 3 (Spanish-English bilinguals)

	Japanese			Spanish		
	M	SD	range	M	SD	range
Length Residence[ns]	2;3	2;0	0;1-6;1	2;2	2;0	0;1-8;2
L2 use ≥ L1 use (N)[ns]	15			16		
Age Reading English[ns]	11;2	2;5	5 – 15	11;1	3;5	5-18
Self-rated L2 reading proficiency**	7.3	1.3	4–9	8.5	1.1	6-10
Self-rated L1 reading proficiency*	9.3	1.2	5–10	9.8	0.4	9–10

Note. '*' denotes a significant differences between the groups at the $p < .05$ level, '**' denotes significance at the $p < .01$ level, *ns* denotes not significant on independent-samples *t*-tests (Chi-squared test for L2 use ≥ L1 use). Length Residence = length of residence in UK or other English-speaking country (years; months); L2 use ≥ L1 use (N) = number of participants who use English as much or more than their L1 on a daily basis; Age Reading English = age participant started reading English (years; months); Self-rated L2 and L1 reading proficiency = self-rated proficiency in reading L2 (L1) on a 10-point Likert scale ranging from "not at all" (1) to "perfect" (10).

TABLE 4

Mean percentage of errors and mean response latencies (standard error of the mean in parentheses) in each condition of Experiment 2 (Japanese-English bilinguals)

	Experimental	Control	Difference
Errors			
homophones	35.1 (4.6)	5.4 (2.2)	29.7
A-U	29.4 (4.1)	10.3 (2.7)	19.1
B-V	21.7 (4.3)	6.5 (2.6)	15.2
L-R	23.5 (3.9)	13.4 (3.1)	10.1
Response latencies			
homophones	1121 (48)	989 (34)	132
A-U	1271 (61)	1108 (44)	163
B-V	1180 (56)	1065 (43)	115
L-R	1167 (53)	1071 (49)	96

TABLE 5

Mean percentage of errors and mean response latencies (standard error of the mean in parentheses) in each condition of Experiment 3 (Spanish-English bilinguals)

	Experimental	Control	Difference
Errors			
homophones	17.7 (3.5)	2.5 (1.4)	15.2
A-U	9.5 (2.7)	6.0 (2.2)	3.5
B-V	4.1 (2.0)	10.2 (3.1)	-6.1
L-R	3.2 (1.6)	4.8 (1.9)	-1.6
Response latencies			
homophones	1148 (49)	1028 (38)	120
A-U	1085 (39)	1088 (40)	-3
B-V	1080 (43)	1010 (44)	70
L-R	1060 (35)	1013 (38)	47