Bilingualism: Language and Cognition 12 (1), 2009, 23–39 © 2008 Cambridge University Press doi:10.1017/S1366728908003465
First published online 14 August 2008

Phonological inconsistency in word naming: Determinants of the interference effect between languages

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

Dutch–English participants named words and nonwords with a between-language phonologically inconsistent rime, e.g., GREED and PREED, and control words with a language-typical rime, e.g., GROAN, in a monolingual stimulus list or in a mixed list containing Dutch words. Inconsistent items had longer latencies and more errors than typical items in the mixed lists but not in the pure list. The consistency effect depended on word frequency, but not on language membership, lexicality, or instruction. Instruction did affect the relative speed and number of errors in the two languages. The consistency effect is the consequence of the simultaneous activation of two sublexical codes in the bilinguals' two languages and its size depends on the activation rate of the associated lexical representations (high-frequency words versus low-frequency words and nonwords) and on the decision criteria that monitor the response conflict at the decision level: the timing for responding (time criterion) in each language depends on the composition of the stimulus list and the likelihood of responses in either language.

Every Dutch–English bilingual knows that the English word TOOL sounds like /tɔl/ when pronounced in Dutch and that the Dutch word GEEL ("yellow", /ɣel/) sounds like /giːl/ or /dʒiːl/ when pronounced in English. Thus, bilinguals can apply spelling-to-sound conversions of one language to letters strings that are words in the other language. The current study investigates the role of AUTOMATIC PHONOLOGICAL RECODING in the other language on target word naming in the native language (L1) or the second language (L2). In particular, we will examine the effects of lexicality, dominance of the target language, instructions, and stimulus list composition on the size of the between-language phonological interference effect and discuss them in relation to the mechanisms for executive control that have been suggested in the literature on bilingual visual word recognition.

In the bilingual domain, several studies have indicated that the phonology of the non-target language affects word recognition in the target language (e.g., Nas, 1983; Beauvillain and Grainger, 1987; Doctor and Klein, 1992; Dijkstra, Grainger, and Van Heuven, 1999; Jared and Kroll, 2001; Jared and Szucs, 2002). In the masked priming paradigm, studies have shown that the prelexical phonological recoding of a consciously inaccessible letter string is an automatic process in both languages of a bilingual (Brysbaert, Van Dyck and Van de Poel, 1999; Van Wijnendaele and Brysbaert, 2001). In the Brysbaert et al. study (1999), Dutch–French participants identified French (L2) target words (e.g., SOURD /suːʁ/) in a perceptual identification task. The monolingual stimulus list contained only L2 target words and the bilingual participants were unaware of the primes, which were presented for 42 ms. The primes were nonwords that were homophonic to the L2 targets when the Dutch (L1) spelling-to-sound rules were applied (e.g., "soer" /su.r/). Recognition rates were significantly better on the pseudohomophone prime trials than on the graphemic control trials. The same stimulus materials did not elicit a phonological priming effect when tested in a group of French monolinguals. The authors concluded that L1 (Dutch) spelling-to-sound rules had been automatically activated even though participants were only consciously aware of the L2 words in the experiment.

Using the same stimulus materials, Van Wijnendaele and Brysbaert (2001) observed the same effects with French–Dutch bilinguals. These participants identified the French target words (L1) better following a pseudohomophone according to L2 grapheme–phoneme correspondences than following a graphemic control. The priming effect fell in the same range as the effect in the Brysbaert et al. study. The authors used these comparable effect sizes to conclude that grapheme–phoneme conversion rules for L1 and L2 are activated automatically and simultaneously during visual word recognition in bilinguals.

Studies using different tasks, like word naming and lexical decision, have produced results that are compatible with this idea of automatic phonological recoding, but also suggest that the effect of this process can vary in size depending on experiment- or language-related factors (Duyck, 2005; Jared and Kroll, 2001). In particular, language dominance (L1 vs. L2) and the composition

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

of the stimulus list (monolingual vs. mixed language blocks) appear to be relevant factors. In a masked priming study, Duyck (2005) observed an asymmetry between L1 and L2 pseudohomophone primes. Whereas L1 pseudohomophones facilitated lexical decisions on associatively related L2 targets (e.g., "greis", a pseudohomophone of Dutch "grijs"/"grey" facilitated responses on BLACK), L2 homophones had no effect on associatively related L1 targets. Given the evidence favoring automatic phonological recoding in L1 and L2 in the Van Wijnendaele et al. (2001) study and his own finding that translation priming on L2 words is stronger when the prime is a high-frequency L2 homophone to the target's L1 translation (*hook*/*hoek* "corner"), Duyck concludes that L2 phonological recoding is automatic but that phonological representations are probably less accessible in L2 than L1 (i.e., the temporal delay assumption).

Jared and Kroll (2001) showed that the effect of prelexical phonological recoding in the non-target language is also determined by list composition. Rather than measuring this effect by means of the processing advantage from masked between-language pseudohomophones, they used phonological interference effects in word naming as a diagnostic. The authors' critical targets were English words with French enemies (e.g., STALE; the rime -*ale* is pronounced /al/ in French), which they presented for naming to French–English and English–French bilinguals. One of their theoretical interests concerned the recent use of French on the naming latencies for the critical items. With that purpose in mind they used a three-blocks design: a block of English target items, an intermittent block of French filler items and a second block of English target items. The French–English bilinguals (naming in L2, Experiment 4) named the English words with French enemies more slowly than English words without enemies (e.g., SPURT), but only after the intermittent block of French words (although the interaction between block and word type was not significant). They did, however, show a significant error effect in both the first and the second block: words with French rime enemies elicited more errors than words without enemies. The English–French bilinguals (naming in L1, Experiment 1), also showed a larger phonological effect on word naming latencies in the second English block than in the first, but in contrast to the French–English bilinguals, they did not show a significant effect on error rates in either block.

The finding that the pronunciation of a target word is affected by previous reading in the non-target language does not discredit the idea that automatic phonological recoding in the non-target language takes place. It does show that list composition can modify its impact on the target response measure. Two accounts have been proposed to explain this finding. According to Jared and Kroll, the French words in the intermittent block raised the activation levels of the phonological representations of all French words in the lexicon, causing more inhibition from these representations on the target language lexical candidates. In the first block, activation levels in French were too low to affect naming latencies in English. This account is in accordance with the Language Mode Hypothesis (Grosjean, 1998) that bilinguals can either be in a monolingual or a bilingual mode depending on the environment, the task demands and their expectations (see also the Bilingual Interactive Activation (BIA) model of bilingual visual processing by Dijkstra and Van Heuven, 1998).

An alternative account, proposed by Dijkstra and Van Heuven (2002) in the new BIA model (BIA+), asserts that list composition does not affect the accessibility of lexical representations in the non-target language but instead the participants' reliance on these representations in the decision stage of the task. In the BIA+ model responses are selected by means of dynamic decision criteria, i.e., selection can rely more or less on some types of information depending on their relevance in the experiment. The effect of stimulus composition can thus be explained: in a monolingual context it is safe to ignore activated representations in the non-target language, because responses are limited to one language only. In contrast, in a mixed language context representations from both languages have to be taken into account, because responses in both languages are required.

It is interesting to note that the effect of list composition on non-target language interference (and the same discussion about its interpretation) has also been observed when the between-language competition arises at the lexical level rather than the sublexical one, i.e., with interlingual homographs (see Dijkstra and Van Heuven, 2002; Dijkstra 2005 for an overview). For example, in both language-specific lexical decision tasks (respond YES to the words of one language and NO to the words of the other language and nonwords; e.g., Dijkstra and Van Heuven, 1998) and in language-specific word naming (Smits, Martensen, Dijkstra, and Sandra, 2006) reaction times to interlingual homographs were longer than those to monolingual control words in mixed-language stimulus lists but not in monolingual lists (see also De Groot, Delmaar and Lupker, 2000; Dijkstra, De Bruijn, Timmermans and Ten Brinke, 2000; Dijkstra, Timmermans and Schriefers, 2000; Dijkstra, Van Jaarsveld and Ten Brinke, 1998; Von Studnitz and Green, 2002).

The work with interlingual homographs points to another factor that is of interest for our study of the phonological interference effect: task demands. The inhibition for these words in a language-specific lexical decision task turned into facilitation in a generalized lexical decision task (Dijkstra and Van Heuven, 1998)

(YES to words of either language and NO to nonwords). This raises the possibility that the instructions on how to deal with the experimental item types can also modify the effect of phonological interference.

The goal of the current study is to systematically study a number of factors that could modify the interference effect caused by bilingual phonological recoding in the non-target language. We adopted Jared and Kroll's rationale and used naming speed and errors as measures of competition. Word targets were unique to one language, but were pronounceable with a different pronunciation in the other language (e.g., SPOON, which is pronounced /spɔn/ in Dutch). We will focus on the impact of the following four factors: (a) stimulus list composition, (b) the lexical status of the target items (to distinguish between prelexical and lexical phonology), (c) language dominance, and (d) task demands.

Stimulus list composition

The target items were tested in two different stimulus lists. The pure list contained only English words and English-like nonwords, whereas the mixed list also contained Dutch words and Dutch-like nonwords. The use of two stimulus lists serves two purposes. First, we wanted to replicate Jared and Kroll's (2001) finding that between-language phonological interference depends on a recent encounter with the non-target language. If so, we should find larger interference effects in our mixed condition. Second, we wanted to create a condition that should elicit maximal phonological interference in order to test the effects of the three other factors under study. If a monolingual stimulus list indeed causes an expectancy-based top–down suppression of the other-language lexical representations, language mixing should remove this suppression. Given a random presentation, the language of the upcoming responses is unpredictable and hence participants cannot actively suppress lexical activation or phonological recoding in one language. Thus, the mixed list offers the ideal context for studying the role of the three following factors.

Lexicality

We included both English and Dutch words and English-like and Dutch-like nonwords. Nonwords were used to test whether phonological competition at the sublexical level can be resolved on the basis of lexical activation. For words, the correct pronunciation can be retrieved lexically and this information may be used to quickly resolve a phonological conflict that arises at the level of grapheme–phoneme conversion. In contrast, a pronunciation conflict for nonwords cannot be lexically resolved. Therefore, nonwords may suffer more phonological interference than words. If both words and nonwords show the same degree of interference, this suggests that the articulatory codes needed for responding are mainly based on sublexical information.

Language dominance

The size of the phonological interference effect could differ between L1 and L2 target words. One might expect a larger effect on L2 targets, either because the sublexical generation of an L1 pronunciation is faster or because lexical access to the phonological representation of similar L1 words (and the resulting competition with the L2 target word) is easier. Brysbaert and colleagues' finding (Brysbaert et al., 1999; Van Wijnendaele and Brysbaert, 2001) that L1 and L2 pseudohomophones prime equally well suggests that bilinguals engage in phonological recoding in both their languages simultaneously. Still, even if both codes are generated equally fast, their disturbance on the pronunciation of the target will depend on the speed with which the target's lexical representation is accessed. If lexical access for L2 words is on average longer than for L1 words, a larger phonological interference effect is expected for L2 words. This would be in line with the temporal delay hypothesis and with Duyck's (2005) finding that L1 pseudohomophones prime associatively related L2 words more strongly than vice versa. Because Duyck used a monolingual list, it seems reasonable to expect an even stronger phonological interference effect on L2 targets in the mixed list of the current study, which offers optimal interference opportunity, but only if phonological interference in word naming is primarily lexical rather than sublexical. Jared and Kroll (2001) reported a similar asymmetry when using the interference paradigm adopted in the present study: larger interference effects for L2 word naming than for L1 word naming. However, comparison between their participant groups is difficult as their language backgrounds were different. Moreover, comparisons between studies are complicated by differences between techniques, languages, the participants' proficiency levels, and so on. Therefore, we presented our participants in the mixed list with both L1 and L2 words, allowing for a within-participants comparison.

Instruction (task demands)

In Green's (1998) Inhibitory Control model of bilingual processing and in the BIA+ model (Dijkstra and Van Heuven 2002), a different set of instructions is reflected in a different task schema and/or a different set of decision criteria. To test whether and how decision processes affect phonological interference, we manipulated the task schema by varying the instructions for the nonwords (the strict list condition: pronounce in English vs. the lenient mixed condition: pronounce as you like). As a mixed stimulus list does not allow for the suppression of lexical representations in either language, any differences in the phonological interference effect for words due to differences in nonword instructions should be explained as the result of changes in the decision criteria.

Method

Participants

Participants were Dutch–English bilingual students in the Language Department at the University of Antwerp or staff members. Students received credits or were paid for participation. Staff members volunteered. There were 33 participants in the pure list condition. Their mean age was 20.0 and their mean number of years of experience with English was 7.6. There were 32 participants in the strict mixed condition, with a mean age of 20.2 and a mean number of years experience of 6.5. There were 28 participants in the lenient mixed condition. Their mean age was 19.10, their mean number of years experience with English 7.8. All participants contributed only once. All participants received English instruction in school from the second or third year of secondary education (approximately 14 years of age) until the last year (at approximately 18 years of age). Furthermore, in Flanders, the Dutch speaking part of Belgium, most English language television programs and movies are not dubbed, but subtitled, and English music is very common, so all participants had quite some experience with hearing English from a very young age onwards up until the moment of testing.

Materials

From the Celex-database (Baayen, Piepenbrock and Van Rijn, 1993) all monosyllabic English and Dutch word forms were extracted. All within- and between-language homographs and all words with a word form frequency of 0 were removed.

There were three groups of English words: two groups of test items and one group of filler items. Forty-two English words were selected that had a within-language consistent rime. Twenty-one of these had a rime that was inconsistent between English and Dutch, e.g., SEED, pronounced as /**set**/ in Dutch (inconsistent items). A second set of 21 items had rimes that are orthographically impossible in monosyllabic Dutch word forms, e.g., SIDE (typical items). These were matched to the inconsistent items, on an item-by-item basis for frequency and letter length, and groupwise for first phoneme, complexity of the onset and occurrence frequency of the rime in English. In addition we selected 21 English words with between-language consistent rimes as a filler group, e.g., SKIP, pronounced as /skɪp/ in Dutch. We inserted these fillers to vary the consistency of interlingual rimes in the experimental list, as the test items had either between-language inconsistent rimes or language-specific rimes.

There were also three groups of 21 Dutch words (two groups of test items and one group of filler items) selected according to the same criteria. An example of a between-language inconsistent Dutch test word is SLAG (/slɑx/, meaning "slap"), an example of a Dutch test word with a monolingual rime is SLIJK (/slɛik/, meaning "mud"), and an example of a between-language consistent filler word is STOF (/stɒf/, meaning "dust" or "fabric").

There were two groups of nonwords. In one group, the items were created by recombining the onsets and the rimes of a subset of the inconsistent English test words, e.g., SOOF, which would be pronounced as /sɔf/ in Dutch. This resulted in 16 inconsistent nonwords. We added another five nonwords with between-language inconsistent rimes, to obtain a total of 21 inconsistent nonwords. The inconsistent nonwords were matched groupwise to the inconsistent English words for onsets and rime frequency. The second group of nonwords had between-language consistent rimes, e.g., SELM, which would be pronounced /sɛlm/ in Dutch.

In total, there were 63 English words, 63 Dutch words, and 42 nonwords. For the mean numbers of letters and frequencies per item type, see Table 1. All the test items and their characteristics are listed in the appendix.

Three list conditions were created. The pure list condition consisted of all English words and nonwords. The two mixed list conditions were created by adding the Dutch words. In the Strict Mixed Condition, participants were instructed to read all nonwords in English. In the Lenient Mixed Condition, participants were told they could choose and vary the language of pronunciation for the nonwords. In total, the Pure List Condition consisted of 105 trials and the mixed list conditions of 168 trials each. Three list versions of each condition were created by randomizing the order of items.

Procedure

Participants sat in front of a computer screen at a distance of approximately 50 cm. Presentation of the trials and the registration of reaction times was monitored by DMDX (see Forster and Forster, 2003). Before starting the experiment, the participants read the instructions on a sheet of paper. All instructions were in English. The instructions included an example of an English word (HORSE) and a nonword (BLIFT). Participants were instructed to read the presented words out loud as fast as possible with as few errors as possible. In the Pure List Condition and the Strict Mixed List Condition, it was stressed that participants had to read the nonwords as if they were English words. In the Lenient Mixed List Condition, participants were told explicitly that they

Table 1. *Number of letters, word form frequency, log frequency, and rime frequency per item type.*

English words				
Rime	word frequency	log frequency	rime frequency	number of letters
Typical	38.86	1.19	12.52	4.57
Inconsistent	37.52	1.18	10.29	4.48
Dutch words				
Rime	word frequency	log frequency	rime frequency	number of letters
Typical	36.95	1.07	13.09	4.24
Inconsistent	38.86	1.1	16.43	4.1
Nonwords				
Rime	word frequency	log frequency	rime frequency	number of letters
Inconsistent	—	—	10.33	4.33

could choose the pronunciation for the nonwords. Each experimental list was preceded by a practice list of 20 items. For the Pure List Condition, the practice list contained 12 English words and 8 nonwords. For the Mixed List Conditions, it contained 8 English words, 7 Dutch words, and 5 nonwords.

Each trial began with an attention marker (+) presented in the center of a black screen for 500 milliseconds (ms), preceded by a pause of 500 ms. Then the test item was presented in the center of the screen in a white Courier 24 bold letter type. The next trial started after the pronunciation of the item or after 1500 ms if no response was given. The responses were marked on a score form by the experiment supervisor. Five response types were distinguished: English pronunciations, Dutch pronunciations, mixed pronunciations, pronunciations that were indistinguishable between English and Dutch, and pronunciations that did not fit into any of the aforementioned categories.

Results

Data cleaning

The data were cleaned per List Condition (Pure List, Strict Mixed List, and Lenient Mixed List). Errors were pronunciations in the wrong language, i.e., Dutch responses to English words, e.g., /sidə/ for SIDE; English responses to Dutch words, e.g., /blɪʒ/ for BLIJ (/blɛi/); responses that did not fit either English or Dutch spelling-to-sound conversion rules (SSC-rules), e.g., /snɛik/ for SNEAK; and hesitations, stutters, etc. Dutch responses to nonwords, e.g., /spɔm/ for SPOOM, were labeled as errors in the Pure List and the Strict Mixed List, and as a separate category in the Lenient Mixed List. This enabled us to compare between the reaction times (RTs) of English responses to nonwords in the Lenient Mixed Condition and those in the Strict Mixed Condition. Because responses in the wrong language formed the vast majority of errors, we analyzed their proportions separately from the other errors.

Reaction times for error responses and all RTs under 150 ms were removed before determining outliers for the reaction time analysis. Participant means and item means were calculated per item type. All individual RTs that were more than 2.56 standard deviations (SDs) slower or faster than both their participant and item mean were removed.

Participants were removed whose mean RT to typically English or Dutch items was more than three SDs higher or lower than the overall mean for these word types in the same List Condition. This resulted in the removal of two participants from the Strict Mixed Condition (remaining $n = 30$). No participants were removed from the other two List Conditions.

Items whose mean RT deviated more than three SDs from the overall condition mean in a particular list were removed from all lists. This led to the removal of the items FLEET (English test word with an inconsistent rime) and POOK (Dutch test word with an inconsistent rime).

In all, 10% of data were removed from the pure list, 16% from the strict mixed list, and 20 % from the lenient mixed list.

In the following sections, we will first present an analysis of the English test words in the different list conditions. Second, we will compare the English test words to the Dutch test words. Third, we will analyze the consistency effect for nonwords.

English words

Overall analysis

For an overview of RTs and errors, see Table 2. An analysis of variance including the within-participants factor Consistency (inconsistent versus typical items), and the between-participants factor List Condition was conducted

Table 2. *Reaction times and number of wrong-language responses for words, per item type.*

	English words		Dutch words	
	inconsistent	typical	inconsistent	typical
RTs				
Pure	494	505	—	—
Strict Mixed	537	532	558	532
Lenient Mixed	568	541	546	522
Errors				
Pure	0	0		
Strict Mixed	3	0	11	0
Lenient Mixed	8	2	10	0

Strict Mixed: Mixed list with mandatory English pronunciation for the nonwords.
Lenient Mixed: Mixed list with choice of pronunciation for nonwords.
NOTE: The reaction times to the English words are the reaction times of the English responses. The reaction times to the Dutch items are the reaction times of the Dutch responses. The number of wrong-language responses indicates the number of times a Dutch response was given to an English word and the number of English responses to a Dutch word.

on the RTs of correct responses to English words. The analysis showed that the main effect of Consistency was not significant [F1 (1,88) = 1.74, p = .19; F2 < 1]. The main effect of List Condition was significant [F1 (2,88) = 3.15, p < .05; F2 (2,38) = 52.15, p < .001]. The interaction between List Condition and Consistency was also significant [F1 (2,88) = 4.88, p < .05; F2 (2,38) = 5.41, p < .01]. Separate analyses, involving only two list conditions, showed that the consistency effect in the Pure List was the reverse of the effect in the Mixed Lists, both for the Lenient Mixed list [F1 (1,59) = 8.55, p < .01; F2 (1,39) = 11.07, p < .001] and in the participant analysis for the Strict Mixed list [F1 (1,61) = 5.63, p < .05; F2 (1,39) = 2.25, p = .14]. The effect of Consistency did not differ between the two Mixed Lists [F1 (1,56) = 1.58, p = .21; F2 (1,39) = 3.27, p = .08].

As the consistency effect differed between the pure and the mixed lists, we analyzed these lists separately. We collapsed across the two mixed lists, because the above analyses indicated no difference in the size of the consistency effect between those. In the PURE LIST, inconsistent English words were named faster than typical English words, but this effect was not significant in the item analysis [F1 (1,32) = 10.48, p < .01; F2 < 1]. In the MIXED LISTS, inconsistent words were named more slowly than typical words, but again this difference was not significant in the item analysis [F1 (1,57) = 4.39, p < .05; F2 < 1].

The same analyses were performed on the proportion of wrong-language responses to English words. Inconsistent words elicited significantly more wrong-language responses than typical words [F1 (1,58) = 52.89, p < .001; F2 (1,39) = 7.52, p < .01]. There was a significant effect of List Condition [F1 (2,88) = 26.88, p < .001; F2 (2,38) = 9.00, p < .005] and a significant interaction between Consistency and List Condition [F1 (2,88) = 11.80, p < .001; F2 (2,38) = 3.06, p = .06]. Separate analyses, involving only two list conditions, showed that the effect of Consistency was smaller in the Pure List than in the Mixed Lists, both the Lenient Mixed List [F1 (1,59) = 28.42, p < .001; F2 (1,39) = 4.52, p < .05], and the Strict Mixed List [F1 (1,61) = 9.72, p < .005; F2 (1,39) = 5.90, p < .05]. The consistency effect was the same in the two mixed lists [F1 (1,56) = 3.18, p = .08; F2 (1,39) = 1.71, p = .19].

As the above analyses showed a difference between the pure and the mixed lists, we analyzed these separately. We collapsed across the two mixed lists, because they did not differ in the size of the consistency effect. In the PURE LIST, inconsistent English words caused more wrong-language responses than typical English words, but this difference was not significant [F1 (1,32) = 3.20, p = .08; F2 (1,39) = 1.99, p = .17]. In the MIXED LISTS, inconsistent words caused significantly more wrong-language responses than typical words [F1 (1,57) = 43.55, p < .001; F2 (1,40) = 7.41, p < .05].

Frequency effects

High-frequency words might be subject to less interference from an inconsistent phonological code in the non-target language than low-frequency words. In order to assess this possibility, we performed a frequency-determined median split on the typical and inconsistent English words. However, a direct comparison between the corresponding subsets (above-median versus below-median) of these word types was unwarranted, because they were not matched on onset and rime frequency (typical and inconsistent words had been matched on these factors on a groupwise rather than a pairwise basis). Accordingly, we assessed the effect of consistency in these subsets indirectly by comparing the effect of language mixing on the two word types. In order to do so, we subtracted an item's mean naming latency in the pure list from its mean naming latency in the mixed list (either strict or lenient, depending in the analysis). Interference from the phonology of the non-target language will surface in the form of larger difference scores for the inconsistent words than for the typical words. By performing an analysis on these difference scores with the factor frequency (above or below median), we can assess whether phonological interference is stronger for low-frequency words than for high-frequency words. For an

Table 3. *Differences in reaction times between the mixed and the pure lists for low and high frequency English words.*

	Frequency	Inconsistent	Typical	Consistency effect
RT Strict Mixed – RT pure list	LF	64	19	45
	HF	33	58	−25
RT Lenient Mixed – RT pure list	LF	83	29	54
	HF	64	45	19

Strict Mixed: Mixed list with mandatory English pronunciation for the nonwords.
Lenient Mixed: Mixed list with choice of pronunciation for nonwords.

overview of the difference scores per frequency category, see Table 3.

An item analysis on the difference scores including the within-items factor List Condition (lenient or strict), and the between-items factors Consistency (inconsistent or typical) and Frequency (high frequency or low frequency) was conducted. The analysis shows that List Condition had a significant effect: the difference scores were larger for the Lenient Mixed Condition than for the Strict Mixed Condition [F2 (1,37) = 11, $p < .005$]. Consistency had a significant effect as well [F2 (1,37) = 10, $p < .005$]: inconsistent words had larger difference scores than typical words. Frequency did not have a significant effect [F2 < 1]. There was, however, a significant interaction between Consistency and Frequency [F2 (1,37) = 7, $p < .01$]: the consistency effect was larger for low-frequency words than for high-frequency ones. The other interactions were not significant.

English and Dutch words

For an overview of RTs and errors, see Table 2. An analysis of variance including the within-participant factors Language (English word or Dutch word) and Consistency (inconsistent item or typical Dutch/English item) and the between-participants factor List Condition (Strict Mixed List or Lenient Mixed List) was conducted on the RTs of correct responses to English and Dutch words. The effect of Language was not significant [F1 < 1, F2<1]. The Consistency effect was significant: typical English or Dutch words were named faster than inconsistent words [F1 (1,56) = 12,49, $p < .005$; F2 (1,78) = 4.34, $p < .05$]. There was a significant interaction between List Condition and Language [F1 (1,56) = 4.52, $p < .05$; F2 (1,78) = 10.41, $p < .01$]: English and Dutch word naming was equally fast in the Strict Mixed List, but English word naming was slower than Dutch word naming in the Lenient Mixed List. The other interactions were not significant.

The same analysis was conducted on the proportion wrong-language responses to English and Dutch words. Dutch words elicited more wrong-language responses than English words, but this main effect of Language was only significant in the participant analysis [F1 (1,56) = 7.80; F2 < 1]. Consistency had a significant effect: Inconsistent items had more wrong-language responses than typical items [F1 (1,56) = 148.16, $p < .005$; F2 (1,78) = 13.73, $p < .001$]. List Condition had a significant effect [F1 (1,56) = 5.75, $p < .05$; F2 (1,78) = 8.58, $p < .005$]: There were more wrong-language responses in the Lenient Mixed List than in the Strict Mixed List. There was a significant interaction between Language and List Condition [F1 (1,56) = 7.77, $p < .001$; F2 (1,78) = 15.38, $p < .001$]: in the Lenient Mixed List, English words elicited more wrong-language responses than Dutch words, whereas the reverse was true for the Strict Mixed List. The interaction between Language and Consistency was significant in the participant analysis, but not in the item analysis [F1 (1,56) = 23.94, $p < .01$; F2 (1,78) = 2.43, $p = .12$]: Dutch words had a larger consistency effect than English words. The other interactions were not significant.

Nonword analysis

Is the effect of phonological interference the same for words and nonwords? Our analysis of the word data showed that inconsistent English words suffered more from language mixing than typical English words, indicating that the RT increase for inconsistent words in the mixed list reflects (at least in part) phonological interference.

In an additional analysis we assessed the presence of phonological interference in nonwords by using the effect of language mixing on typical English words as a baseline. As the latter effect cannot be due to phonological interference, a larger effect of language mixing on inconsistent nonwords would reveal the presence of phonological interference. Even though the two item types in this comparison differ with respect to two factors (phonological consistency and lexicality) the analysis is legitimate, if the latency difference between words and nonwords is constant across lists. That is, if lexicality does not affect language mixing. This was indeed the case as was shown by an analysis with the factors Lexicality (inconsistent nonword or inconsistent word) and List

Table 4. *Reaction times and number of Dutch responses for nonwords.*

RTs	RTs	Number of Dutch responses
Pure	515	7
Strict Mixed	580	13
Lenient Mixed	595	43

Strict Mixed: Mixed list with mandatory English pronunciation for the nonwords.
Lenient Mixed: Mixed list with choice of pronunciation for nonwords.
NOTE: The reaction times to the nonwords are the reaction times of the English responses. The number of wrong-language responses indicates the number of times a Dutch response was given to nonwords.

Condition (Pure versus Strict Mixed or Lenient Mixed). The interaction was never significant [Pure versus Lenient Mixed List: F1 (1,59) = 1; p = .25; F2 < 1; Pure versus Strict Mixed List: F1 (1,61) = 4, p = .07; F2 (1,39) = 1; p = .28]. As such, a comparison of the language mixing effect between the inconsistent nonwords and the typical English words will only reflect the effect of phonological consistency and not the effect of lexicality.

The data for this analysis were the differences obtained by subtracting the mean RT in the Pure List from the mean RT in a Mixed List (as in the analysis of the frequency factor; for the difference scores, see Table 5). In the analysis including the within-item factor List Condition (Lenient or Strict) and the between items factor Consistency (inconsistent nonword or typically English word) we found indeed that the difference scores for inconsistent nonwords were larger than the difference scores for typical English words [F2 (1,40) = 4, p < .05], indicating a phonological interference effect for the nonwords. List Condition did not have a significant effect [F2 (1,40), p = .24]. The interaction was not significant [F2 < 2].

General discussion

The goal of the present paper was to study which mechanisms modify the phonological interference effect in word naming. We studied four factors: stimulus list composition, lexicality, language dominance, and nonword naming instruction. By manipulating stimulus list composition we attempted to replicate Jared and Kroll's (2001) observation that the effect of phonological interference depends on list composition. The manipulation of lexicality made it possible to assess the relative contributions of sublexical and lexical phonological codes. The factor language dominance addressed the issue whether sublexical phonological codes are generated simultaneously in L1 and L2. Finally, nonword naming instruction allowed us to assess the impact of a decision stage factor on the magnitude of phonological interference.

The between-language phonological interference effect was tested by contrasting two types of items: items of which the rime had conflicting pronunciations in the two languages of the bilinguals (e.g., GREED, pronounced /ɣret/ in Dutch) and items whose rimes occurred in only one language (e.g., GROAN, *-oan* does not exist as a rime in Dutch). Dutch–English bilinguals named English words and nonwords in a monolingual list or in a mixed language list with Dutch words, whose rimes were also either between-language inconsistent or typical for Dutch. We used two sets of instructions for the mixed language lists: nonwords had to be named in English (the Strict Mixed Condition) or participants could choose and vary the language of pronunciation for the nonwords (the Lenient Mixed Condition).

In the MIXED LANGUAGE LISTS we found a reliable consistency effect for both English and Dutch words. Items with conflicting rimes between languages had slower naming latencies and elicited more errors than items with language-typical rimes. The consistency effect was equally large for nonwords and words. Within the set of English words, the effect was significantly smaller for the higher-frequency words than for the lower-frequency words (the interaction between frequency and consistency could not be assessed for Dutch words, as the statistical analysis required a comparison with the pure list). Although varying the instructions with respect to nonword naming did not change the consistency effect, it did have a general effect on the mean reaction times and error rates for items in the two languages. In the Strict

Table 5. *Differences in reaction times between the mixed lists and the pure list for inconsistent nonwords and typically English words.*

	Inconsistent nonwords	Typically English words	Consistency effect
RT Strict Mixed – RT pure list	60	28	32
RT Lenient Mixed – RT pure list	69	37	32

Strict Mixed: Mixed list with mandatory English pronunciation for the nonwords.
Lenient Mixed: Mixed list with choice of pronunciation for nonwords.

Mixed Condition, word naming was equally fast in English and in Dutch, and Dutch words elicited more errors than English words. In the Lenient Mixed Condition, word naming was faster in Dutch than in English and there was no difference in error rates. In the MONOLINGUAL LIST there was no phonological interference effect. Inconsistent words were even named somewhat faster than typical words (an effect that was not significant in the item analysis) and their error rates did not differ.

The first factor that was manipulated in the current study was STIMULUS LIST COMPOSITION. In the mixed lists we observed a between-language phonological interference effect, which confirms that sublexical phonology can be activated simultaneously for the two languages of bilinguals (Brysbaert et al., 1999, and Van Wijnendaele and Brysbaert, 2001). However, our data do not allow us to confirm or reject the hypothesis that simultaneous phonological activation in both languages is automatic and therefore occurs in all situations, because we observed no interference effect for the inconsistent items in the pure list.

The observed differences between the mixed lists and the pure list in this study are parallel to the differences observed by Jared and Kroll (2001). They observed a consistency effect on English targets with between-language inconsistent rimes when they were named after the participants had named a French filler block, and a smaller (or no) such effect when they were named before the French filler block. Jared and Kroll explained this difference in terms of suppression of lexical representations in the non-target language, which takes place as long as participants encounter only words from a single language, but disappears when recent experience leads them to expect words from the other language as well. This account could be extended to the suppression of non-target language sublexical phonology in a pure list through inhibitory connections from the lexical to the sublexical level. However, an interpretation along these lines would contradict that activation of sublexical phonology in the non-target language is automatic, as was proposed by Brysbaert et al. (1999) and Van Wijnendaele and Brysbaert (2001).

There is a way to reconcile the notion of automatic phonological recoding in the two languages of a bilingual with the absence of a between-language phonological interference effect in word naming in a pure list. Dijkstra and Van Heuven (2002) argued in their BIA+ model that top–down suppression of all lexical representations in one language on the basis of expectations is not possible and that the outcomes of the sublexical and lexical activation processes are evaluated at a decision level according to dynamic decision criteria, such as language membership. Depending on these criteria, a response selection problem may or may not arise. According to this view, words with between-language inconsistent rimes would cause a selection problem in a mixed list but not in a pure list. In our pure list only English responses are required. Therefore, even though sublexical phonological recoding or lexical activation would also lead to Dutch response candidates, the latter could be safely ignored and hence not delay the selection process. In the mixed language lists, however, the response language for the next trial is not predictable. As such, on each trial the process of response selection has to consider all response options. As automatic phonological recoding in L1 and L2 will yield conflicting phonological codes for inconsistent words but not for typical English or Dutch words, the selection process will be delayed and more prone to errors in the case of inconsistent words. This account rejects the idea of top–down suppression of the non-target language, while offering a way to integrate the effect of list composition with the hypothesis that sublexical phonological codes are activated automatically and simultaneously in the bilingual's two languages during visual word recognition (Brysbaert et al., 1999; Van Wijnendaele and Brysbaert, 2001). In all fairness, our effect of list composition does not by itself reject the suppression hypothesis, even though an interpretation of our data in terms of automatic phonological activation and list-dependent decision criteria is plausible and offers a coherent account of apparently contradictory findings in previous research.

The second factor we examined was the LEXICALITY of our items. The presence of a lexical representation for words (hence a stored pronunciation) could make it easier to resolve the conflict between two sublexical phonological codes. However, there was no overall difference between words and nonwords when we compared the effect of language mixing (pure vs. mixed list) on the consistency effect. On closer inspection, it became clear that the magnitude of the consistency effect does not so much depend on the presence of a lexical representation but on its accessibility: lower-frequency inconsistent words suffered significantly more from language mixing than higher-frequency words. This interaction between frequency and consistency is in line with findings in the monolingual domain, where larger phonological interference effects for low-frequency words than for high-frequency words have also been observed (Seidenberg, Waters, Barnes, and Tanenhaus, 1984; Waters and Seidenberg, 1985; Taraban and McClelland, 1987).

The finding that an effect of phonological interference particularly affects low-frequency words and nonwords suggests a sublexical locus of the effect. In such an account, inconsistent nonwords and inconsistent English words are recoded into two sublexical phonological codes, according to L1 and L2 grapheme–phoneme mappings. When the item is a high-frequency word in the target language, the quick accessibility of its stored pronunciation will make it possible to resolve the selection problem very fast or even prevent it from taking place. In contrast, when the item is a low-frequency word in the target language, lexical access will be slow (or impossible in the case of nonwords), providing opportunity for a response conflict between the two phonological codes. Note that this dual-route account does not deny a role for lexical phonology in the naming of low-frequency words. The error percentages indicate that participants rely on the stored pronunciation of these words to avoid making a non-target language response, but the reaction times show that this process is too slow to prevent a response conflict. Accordingly, the different effect of phonological consistency on high-frequency and low-frequency words reflects the relative speed with which their prelexical and lexical phonological codes become available, rather than whether they become available at all.

Even though nonwords have no lexical representation, they too can generate phonological information through sublexical recoding and through lexical activation. The lexical route will produce partial activation of L1 and L2 orthographic neighbors, while at the same time the L1 and L2 sublexical codes cause a response conflict. The way in which this conflict is resolved depends on the instructions for nonword pronunciation. When participants are free to choose the language for nonword pronunciation (Lenient Mixed Condition), they may choose the pronunciation that is most strongly supported by the partially activated words in the lexicon. However, when they must pronounce nonwords according to the pronunciation rules for a particular language (Strict Mixed Condition), they have to select the appropriate phonological code on the basis of language information. This information can be derived from the analogy between the sublexical phonological codes and the phonological representations of the partially activated neighbors (see Lemhöfer and Dijkstra, 2004). For example, the nonword SPOOM will activate the orthographic and phonological representations of English ROOM, BLOOM, GROOM, and so on, and by analogy /spu:m/ will be "tagged" as the English response for SPOOM.

Our adoption of a dual-route account to explain how lexicality and frequency shape the size of the phonological interference effect strongly resembles Rastle and Brysbaert's (2006) conclusions concerning the role of automatic phonological recoding in L1 and especially the way in which they modeled this process. On the basis of an exhaustive meta-analysis review of the literature on masked pseudohomophone priming (in English) and two novel experiments, these authors showed that masked pseudohomophone priming is a small but reliable effect, reflecting the existence of a fast sublexical process of automatic phonological recoding in L1. Importantly, the most successful simulation of their findings within the DRC framework (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001) set out from the assumption that a word's phonological representation in the mental lexicon is activated through two independent input "channels": a sublexical phonological code and a connection to the word's orthographic representation (which mediates lexical access). Thus participants in a lexical-decision task can discriminate between true words like *brain* and pseudohomophones like *brane* on the basis of activation levels, which will be higher in the case of a real word due to the orthographic route. Rastle and Brysbaert's proposal that response selection in a lexical decision task hinges on the joint operation of sublexical and lexical phonology comes close to our own claim that the occurrence of between-language phonological interference in word naming depends on the relative speed with which an item's stored pronunciation and its sublexical phonological code become available.

The third factor under study in the mixed language lists was the LANGUAGE MEMBERSHIP of the test words. We observed equally large effects of phonological consistency on L1 and L2 words. This is in accordance with the findings of Van Wijnendaele and Brysbaert (2001). They found that masked nonword primes facilitated the perceptual identification of L1 target words when L2 grapheme–phoneme conversion rules made the primes homophonic to their targets. This effect of automatic phonological recoding in L2 was just as large as the effect observed earlier by Brysbaert et al. (1999), who used the same materials but presented them to participants with the reverse language dominance. For these participants, the targets were L2 words and the critical primes were pseudohomophones according to L1 grapheme–phoneme correspondences. The authors concluded from these studies that there is automatic and simultaneous activation of L1 and L2 spelling-to-sound rules. Our finding that phonological consistency has the same effect on L1 and L2 words supports this claim.

The last factor under study was the effect of INSTRUCTIONS. As was mentioned above, participants were instructed to name the nonwords in English in one mixed list condition (the Strict Condition) and told they could choose and vary the language of pronunciation for the nonwords in the other mixed list condition (the Lenient Condition). Although the consistency effect was the same across the mixed list conditions, the relationship between the two languages on overall response speed and error rates changed. This is evident from the naming latencies on the inconsistent words, but also on the typical words. In the Strict Mixed List, English word responses were equally fast as Dutch word responses, but in the Lenient Mixed List English word responses were significantly slower than Dutch word responses.

It is implausible that this effect on word naming speed is due to the suppression of L2 representations in the Lenient Condition. Mixing words in an unpredictable sequence of L1 and L2 targets creates circumstances that make language suppression counterproductive. Rather, we suggest that the timing of responses is determined on the basis of a separate time criterion for each language. These criteria reflect the participants' confidence in making responses in that language. As the effect of nonword instruction indicates, the criteria are sensitive to the probability with which a response in the language must be made. In the Strict Condition, the requirement to name the nonwords in English implies a 2:1 ratio of English to Dutch responses. In the Lenient Condition, the freedom to select a response language for nonwords implies a much lower ratio (almost 1:1 on average, see Table 4). As an English response was more probable in the Strict Condition, participants will have been more confident to select the English pronunciation and, accordingly, have adopted an earlier time criterion for English responses compared to participants in the Lenient Condition. This is reflected in a different relationship between the English and Dutch word naming speeds in the two mixed conditions.

The notion of a time criterion can also account for another effect: the finding that typical English words are named more slowly in the mixed lists than in the pure list. It appears that when one item type in L2 causes a processing delay (i.e., inconsistent L2 words) participants become more cautious to name the "easy" L2 items as well (i.e., typical L2 words). This suggests that one word type in a language determines participants' overall confidence for making responses in that language, i.e., the time criterion for that language. This fits in with the literature on the effect of mixing item types (e.g., Lupker, Brown, and Colombo, 1997; Lupker, Kinoshita, Coltheart, and Tamsen, 2003). For instance, Lupker et al. (2003) found that pictures were named more slowly when participants had to name words in the same list than when they were presented in a pure list and that this mixing cost increased as a function of word difficulty (a higher cost when the presented words had more letters, lower frequencies, and multiple syllables). It seems as though the time criterion for picture responding moved in accordance with the word naming latencies. These findings can be transferred to our own experiment: delaying the naming response to inconsistent L2 words by adding L1 words had a "spill-over" effect on the phonologically unproblematic, typical L2 words, which suggests the involvement of a time criterion.

Conclusions

The effects of our four factors on word naming supports the following verbal description of how bilinguals process words with between-language phonologically inconsistent rimes. The finding of phonological interference when a stored pronunciation takes relatively long to access (low-frequency words) or does not exist (nonwords) reveals the existence of a sublexical process of grapheme–phoneme mapping, which recodes the orthographic string into a phonological code. Equally-sized interference effects on the naming of L1 and L2 words indicate that this phonological recoding occurs simultaneously in the two languages of relatively proficient bilinguals such as the participants in the current study. When this process yields two conflicting phonological codes (between-language inconsistent rimes), it creates the potential for a response conflict in a naming task, which will manifest itself in the form of slower responses and more errors. Two factors can prevent or quickly resolve this conflict: one at the lexical level and one at the decision level. When the target is a high-frequency word the fast retrieval of its stored pronunciation will enable the response mechanism to quickly select the appropriate pronunciation. When the task schema at the decision stage excludes responses on the basis of language information (pure list), the presence of conflicting phonological codes will not cause a response conflict either. Hence, we conclude that the current study is compatible with the hypothesis of automatic phonological recoding (see Rastle & Brysbaert, 2006) and with the notion that such recoding occurs in parallel in the two languages of bilinguals (Brysbaert et al. 1999, Van Wijnendaele and Brysbaert, 2001). Moreover, the concept of dynamic decision criteria from the BIA+ framework (Dijkstra & Van Heuven, 2002) allows us to integrate this conclusion with a seemingly contradictory effect: our replication of Jared and Kroll's finding (2001) that the between-language phonological interference effect depends on stimulus list composition.

Appendix: The test items

Table A1. *Number of letters, frequency, and rime consistency measures of the English test items.*

	Number of letters	Frequency per million	Log frequency	Rime consistency within language	Rime frequency within language	Dutch–English rime consistency
INCONSISTENT ENGLISH WORDS						
gland	5	2	0.30	0.92	12	0.46
groom	5	4	0.60	1.00	8	0.40
bleed	5	4	0.60	1.00	20	0.47
tweed	5	5	0.70	1.00	20	0.47
slam	4	5	0.70	1.00	22	0.47
reef	4	7	0.85	1.00	1	0.08
broom	5	7	0.85	1.00	8	0.40
bald	4	9	0.95	1.00	1	0.00
greed	5	9	0.95	1.00	20	0.47
spoon	5	12	1.08	1.00	9	0.40
fleet	5	15	1.18	1.00	10	0.32
tool	4	16	1.20	0.88	8	0.64
stir	4	18	1.26	1.00	3	1.00
dirt	4	21	1.32	1.00	4	0.60
skirt	5	21	1.32	1.00	4	0.60
seed	4	29	1.46	1.00	20	0.47
proof	5	32	1.51	1.00	7	0.30
salt	4	44	1.64	0.67	3	0.40
speed	5	79	1.90	1.00	20	0.47
lie	3	87	1.94	0.86	7	0.33
feel	4	362	2.56	1.00	9	0.45
Average	4.48	37.52	1.18	0.97	10.29	0.44
TYPICAL ENGLISH WORDS						
groan	5	2	0.30	1.00	3	1.00
graze	5	4	0.60	1.00	11	1.00
bloke	5	5	0.70	1.00	12	1.00
sneak	5	5	0.70	0.86	14	0.77
bluff	5	5	0.70	1.00	14	0.92
probe	5	7	0.85	1.00	3	1.00
rack	4	8	0.90	1.00	23	0.95
glare	5	8	0.90	0.94	18	0.88
beam	4	9	0.95	1.00	10	1.00
spine	5	12	1.08	1.00	21	1.00
toast	5	15	1.18	1.00	4	1.00
float	5	17	1.23	1.00	10	1.00
steam	5	18	1.26	1.00	10	1.00
tray	4	21	1.32	1.00	29	1.00
lawn	4	22	1.34	1.00	10	1.00
sale	4	34	1.53	1.00	15	1.00
dawn	4	35	1.54	1.00	10	1.00
fuel	4	44	1.64	1.00	2	1.00
scale	5	71	1.85	1.00	15	1.00
seat	4	81	1.91	0.82	17	0.82
side	4	393	2.59	1.00	12	1.00
Average	4.57	38.86	1.19	0.98	12.52	0.97

Table A2. *Number of letters, frequency, and rime consistency measures of the Dutch test items.*

	Number of letters	Frequency per million	Log frequency	Rime consistency within language	Rime frequency within language	Dutch–English rime consistency
INCONSISTENT DUTCH WORDS						
pook "poker"	4	1	0.00	1.00	14	0.40
ree "deer"	3	1	0.00	1.00	24	0.49
goor "filthy"	4	2	0.30	1.00	21	0.68
leem "clay, loam"	4	2	0.30	1.00	8	0.67
sleet "wear"	5	2	0.30	0.96	23	0.64
snee "cut"	4	3	0.48	1.00	24	0.49
slee "sled"	4	3	0.48	1.00	24	0.49
pand "building"	4	8	0.90	1.00	11	0.46
koor "choir"	4	10	1.11	1.00	21	0.68
kalk "limestone"	4	11	1.04	1.00	9	0.43
hees "hoarse"	4	11	1.04	1.00	20	0.53
flag "vlag"	4	18	1.26	0.83	12	0.25
vat "barrel"	3	20	1.23	0.89	26	0.46
bloot "naked"	5	24	1.38	1.00	23	0.70
kern "core"	4	35	1.54	1.00	1	0.25
blad "leaf"	4	51	1.71	1.00	11	0.39
slag "battle, slap"	4	54	1.73	0.83	12	0.25
zand "sand"	4	56	1.75	1.00	11	0.46
vlees "meat"	5	81	1.99	1.00	20	0.53
bleek "pale"	5	174	2.24	1.00	19	0.57
stad "city"	4	249	2.40	1.00	11	0.39
Average	4.10	38.86	1.10	0.98	16.43	0.49
TYPICAL DUTCH WORDS						
pij "habit"	3	1	0.00	1.00	17	1.00
vlo "flea"	3	1	0.00	1.00	11	0.83
huig "uvula"	4	1	0.00	1.00	6	1.00
geul "ditch"	4	2	0.30	1.00	5	1.00
snaar "string"	5	2	0.30	1.00	21	1.00
spijs "food"	5	2	0.30	1.00	11	1.00
slijk "mud"	5	4	0.60	1.00	13	1.00
koek "biscuit"	4	7	0.85	1.00	10	1.00
zoen "kiss"	4	10	1.00	1.00	10	1.00
stoet "parade"	5	11	1.00	1.00	17	1.00
kaal "bald"	4	13	1.11	1.00	25	0.95
raar "weird"	4	16	1.20	1.00	21	1.00
voer "forage"	4	19	1.28	1.00	15	0.92
bruin "brown"	5	26	1.41	1.00	8	1.00
keus "choice"	4	30	1.48	1.00	6	1.00
lijf "body"	4	51	1.64	1.00	10	1.00
broek "pants"	5	56	1.75	1.00	10	1.00
steun "support"	5	57	1.76	1.00	7	1.00
blij "happy"	4	101	1.98	1.00	17	1.00
pijn "pain"	4	149	2.27	1.00	18	1.00
vrij "free"	4	217	2.34	1.00	17	1.00
Average	4.24	36.95	1.07	1.00	13.10	0.99

Table A3. *Number of letters, frequency, and rime consistency measures of the nonword test items.*

	Number of letters	Rime consistency within language	Rime frequency within language	Dutch–English rime consistency
Inconsistent nonwords				
rald	4	1.00	1	0.00
speef	5	1.00	1	0.08
bir	3	1.00	3	1.00
slild	5	0.67	3	0.40
sirt	4	1.00	4	0.60
lirt	4	1.00	4	0.60
stie	4	0.86	7	0.33
soof	4	1.00	7	0.30
brool	5	0.88	8	0.64
floom	5	1.00	8	0.40
spoom	5	1.00	8	0.40
gloon	5	1.00	9	0.40
deek	4	1.00	9	0.38
greel	5	1.00	9	0.45
skand	5	0.92	12	0.46
gree	4	0.95	20	0.46
preed	5	1.00	20	0.47
teed	4	1.00	20	0.47
blee	4	0.95	20	0.46
fam	3	1.00	22	0.47
twam	4	1.00	22	0.47
Average	4.33	0.96	10.33	0.44

Table A4. *Number of letters, frequency, and rime consistency measures of the Dutch and English filler items.*

	Number of letters	Frequency per million	Log frequency	Rime consistency within language	Rime frequency within language	Dutch–English rime consistency
CONSISTENT ENGLISH WORDS						
desk	5	87	1.94	0.00	0	0.00
fist	4	19	1.28	0.90	10	0.94
flint	5	13	1.11	0.92	12	0.95
brink	5	6	0.78	1.00	19	0.96
priest	6	34	1.53	0.00	0	1.00
brisk	5	9	0.95	1.00	3	1.00
left	4	573	2.76	1.00	6	1.00
soil	4	46	1.66	1.00	8	1.00
boil	4	21	1.32	1.00	8	1.00
spoil	5	9	0.95	1.00	8	1.00
toil	4	3	0.48	1.00	8	1.00
song	4	33	1.52	1.00	9	1.00
rim	3	8	0.90	1.00	13	1.00
grim	4	16	1.20	1.00	13	1.00
skim	4	4	0.60	1.00	13	1.00
sting	5	4	0.60	1.00	17	1.00
twin	4	16	1.20	1.00	18	1.00
grin	4	12	1.08	1.00	18	1.00
skin	4	96	1.98	1.00	18	1.00
skip	4	5	0.70	1.00	23	1.00
grit	4	3	0.48	1.00	25	1.00
Average	4.33	48.43	1.19	0.90	11.86	0.95
CONSISTENT DUTCH WORDS						
plomp "chubby"	5	1	0.00	1.00	13	1.00
vrek "miser"	4	1	0.00	1.00	22	1.00
gesp "buckle"	4	2	0.30	1.00	4	1.00
lomp "rude"	4	2	0.30	1.00	13	1.00
spil "pivot"	4	2	0.30	1.00	18	1.00
stip "dot"	4	2	0.30	1.00	25	1.00
stipt "punctual"	5	4	0.60	1.00	14	1.00
pret "fun"	4	8	0.90	0.95	19	0.96
kont "backside"	4	10	1.00	1.00	11	0.78
heks "witch"	4	11	1.04	1.00	2	1.00
romp "torso"	4	12	1.08	1.00	13	1.00
vlot "raft"	4	16	1.20	1.00	28	1.00
vel "skin"	3	21	1.32	1.00	20	1.00
kist "chest"	4	29	1.46	1.00	16	0.94
bril "glasses"	4	32	1.51	1.00	18	1.00
bron "source"	4	42	1.62	1.00	12	0.83
zon "sun"	3	43	1.63	1.00	12	0.83
stof "dust"	4	71	1.85	1.00	15	0.93
stel "couple"	4	84	1.92	1.00	20	1.00
blik "can"	4	175	2.24	1.00	31	1.00
vol "full"	3	293	2.47	1.00	25	1.00
Average	3.95	41.00	1.10	1.00	16.71	0.97

Table A5. *Number of letters, frequency, and rime consistency measures of the nonword filler items.*

	Number of letters	Rime consistency within language	Rime frequency within language	Dutch–English rime consistency
CONSISTENT NONWORDS				
fom	3	0.83	6	0.94
spom	4	0.83	6	0.94
skint	5	0.92	12	0.95
grink	5	1.00	19	0.96
prink	5	1.00	19	0.96
relk	4	1.00	1	1.00
selm	4	1.00	1	1.00
bilk	4	1.00	3	1.00
slem	4	1.00	3	1.00
sept	4	1.00	4	1.00
lelf	4	1.00	4	1.00
brift	5	1.00	8	1.00
floil	5	1.00	8	1.00
tift	4	1.00	8	1.00
glong	5	1.00	9	1.00
grelt	5	1.00	9	1.00
dilt	4	1.00	10	1.00
stin	4	1.00	18	1.00
blop	4	1.00	22	1.00
twop	4	1.00	22	1.00
spip	4	1.00	23	1.00
Average	4.29	0.98	10.24	0.99

References

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

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

Brysbaert, M., Van Dyck, G. & Van de Poel, M. (1999). *Journal of Experimental Psychology: Human Perception and Performance*, 25 (1), 137–148.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.

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

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

Dijkstra, T., De Bruijn, E., Timmermans, M. & Ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition*, 3 (1), 69–78.

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

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

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

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 (3), 175–197.

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

Doctor, E. A. & Klein, D. (1992). Phonological processing in bilingual word recognition. In R. J. Harris (ed.), *Cognitive processes in bilinguals*, pp. 237–252. Amsterdam: Elsevier.

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 (6), 1340–1359.

Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers*, 35, 116–124.

Green, D. W. (1998). Mental control of the bilingual lexico–semantic system. *Bilingualism: Language and Cognition*, 1 (2), 67–81.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1 (2), 131–149.

Jared, D. & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, 44, 2–31.

Jared, D. & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition*, 5 (3), 225–239.

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, 4, 533–550.

Lupker, S. J., Brown, P. & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 570–590.

Lupker, S. J., Kinoshita, S., Coltheart, M. & Taylor, T. E. (2003). Mixing costs and mixing benefits in naming words, pictures, and sums. *Journal of Memory and Language*, 49, 556–575.

Nas, G. (1983). Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior*, 22 (5), 526–534.

Rastle, K. & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, 53, 97–145.

Seidenberg, M. S., Waters, G., Barnes, M. A. & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, 23, 383–404.

Smits, E., Martensen, H., Dijkstra, T. & Sandra, D. (2006). Naming interlingual homographs: Variable competition and the role of the decision system. *Bilingualism: Language and Cognition*, 9 (3), 281–297.

Taraban, R. & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language*, 26, 603–631.

Van Wijnendaele, I. & Brysbaert, M. (2001). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*, 28 (3), 616–627.

Von Studnitz, R. E. & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, 5 (1), 1–23.

Waters, G. & Seidenberg, M. S. (1985). Spelling–sound effects in reading: Time-course and decision criteria. *Memory and Cognition*, 13 (6), 557–572.

Received December 15, 2006
Revision received April 26, 2007
Accepted May 15, 2007