John Benjamins Publishing Company

When MOOD rhymes with ROAD

Dynamics of phonological coding in bilingual visual word perception

Martin van Leerdam, Anna M. T. Bosman and
Annette M. B. de Groot

University of Amsterdam / Radboud University Nijmegen / University of
Amsterdam

Three experiments investigated whether perception of a spelling-to-sound
inconsistent word such as MOOD involves coding of inappropriate phonology
caused by knowledge of enemy neighbors (e.g., BLOOD) in non-native speakers.
In a new bimodal matching task, Dutch-English bilinguals judged the corre-
spondence between a printed English word and a speech segment that was or
was not the printed word's rime. Evidence for coding of inappropriate phonol-
ogy was obtained with trials in which the speech segment was derived from an
English enemy neighbor. In such trials, error rates increased significantly relative
to control trials. This effect was also found when speech segments were derived
from Dutch enemy neighbors, which suggests inappropriate coding of cross-
language phonology. These findings are consistent with a strong phonological
theory of word perception (Frost, 1998), in which phonological coding is essen-
tially a language non-selective process.

Keywords: visual word perception, phonology, spelling-to-sound consistency,
second language, bilingualism

Does MOOD rhyme with ROAD? On the face of it, the answer is that it does not, be-
cause according to the spelling-to-sound correspondence rules of English orthog-
raphy, the rimes of MOOD and ROAD have dissimilar pronunciations. However, as
we will show, in Dutch-English bilinguals the visual processing of MOOD may ac-
tually elicit phonology appropriate to ROAD. The fact that for the bilingual reader
spelling relates to sound differently across languages allows for the possibility that
contextually inappropriate phonology from the native language emerges jointly
with correct phonology during second-language word perception. Returning to

the example, MOOD may elicit inadvertent phonology rhyming with ROAD because this particular phonology is concordant with knowledge of spelling-to-sound relations in Dutch orthography. Native speakers of Dutch have learned, to asymptotic degrees, that the rime -OOD is pronounced as in the Dutch word LOOD (meaning 'lead') and many other words with this ending — which happens to match the rime of ROAD.

Mapping form to multiple functions

Cognitive systems deal with inconsistent experience by accommodating manifold relations between surface forms and multiple cognitive functions (i.e., percepts and actions, cf. MacWhinney, 2005; Van Orden, Pennington, & Stone, 1990). In the process of mapping form to function, consequently, all previously associated functions are potential candidates. Multiple interpretations of the same form may cause ambiguity in perception and action, which must be resolved. Ambiguity arises if the system considers both contextually appropriate and inappropriate interpretations. In dynamic systems theory, this system property is known as *multistability* (e.g., Van Orden, Jansen op de Haar, & Bosman, 1997). Multistability refers to the coexistence of several final states, or equilibrium points, which are neither completely stable nor totally unstable. Due to the structure of the underlying basins of attraction, multistable systems are very sensitive to perturbations: Even a small perturbation can cause a prompt transition from one system state to another. Therefore, it is possible for an appropriate interpretation to reach a final, asymptotic state and then to be perturbed to settle into a contextually inappropriate interpretation.

In this study we take a close look at one of the most intricate workings of the human mind: The process of reading printed words. We focus on the relations between the spellings and pronunciations of words, and investigate experimentally how manifold relations affect the process of visual word perception. An example of a spelling that has multiple pronunciations is the rime -OOD, which is pronounced differently in the words MOOD and BLOOD. In the process of mapping spelling to sound, multiple phonological candidates give rise to ambiguity. From a dynamic systems perspective, this ambiguity may reflect multistability, the coexistence of both appropriate and inappropriate phonological states. If the mapping of spelling to sound behaves as a multistable process, a perturbation may cause the system to jump from appropriate to inappropriate phonology.

The present study introduces a new and, as we shall see, effective experimental paradigm with which this feature of multistability may be examined. Our first objective is to show that processing of spelling-to-sound ambiguous words involves

coding of inappropriate phonology in non-native speakers. The way we accomplish this goal is to establish the experimental conditions (i.e., the perturbation) in which, in a within-language situation, the reading system is prompted to exchange perception of an appropriate pronunciation for a potential but inappropriate pronunciation. If such an exchange can be realised, this would show that coding of inappropriate phonology is part of the perception of spelling-to-sound ambiguous words. The second objective of this study is to show that for spellings that relate to sounds differently across languages, inappropriate phonology is coded across languages as well. Finally, a more general objective is to provide a highly sensitive new methodology that may be of interest to other researchers who aim to study cognitive systems exhibiting multistable behavior resulting from ambiguity in perception and action.

In this study, we investigated how spelling-to-sound ambiguity affects reading performance. Ambiguity at the word-body level arises when the same spelling body (e.g., -OOD) has more than one possible pronunciation. Following the seminal work of Glushko (1979; see also Jared, McRae, & Seidenberg, 1990), words like MOOD are traditionally defined as spelling-to-sound *inconsistent* because their spelling body maps onto more than one pronunciation, and words like MOON are defined as spelling-to-sound *consistent* because their spelling body has a single pronunciation. Words that share a spelling body with other words are called *neighbors*. They are classified as *friends* if their spelling bodies are pronounced the same way and *enemies* if they are pronounced differently. Neighbors of an inconsistent word always include one or several enemies, but some words have more enemies than others. This implies that spelling-to-sound inconsistency is a matter of degree (Jared et al., 1990; Ziegler, Stone, & Jacobs, 1997; see also Holden, 2002). To quantify the degree of inconsistency, counts of number and frequency of friends and enemies can be converted into consistency ratios (e.g., Ziegler et al., 1997). A consistency ratio greater than .5 indicates that a word has more and stronger friends than enemies (i.e., it contains a *typical mapping*) and a consistency ratio smaller than .5 indicates that a word has more and stronger enemies than friends (i.e., it contains an *atypical mapping*).

Phonological coding in visual word perception

Phonological coding, the process of mapping spelling to sound, is a central and primary constituent of visual word perception. This assertion lies at the heart of the strong phonological theory of word perception (Frost, 1998; see also Carello, Turvey, & Lukatela, 1992; Van Orden et al., 1990). According to the strong phonological theory, phonological coding is a *mandatory* process, which means

that a phonological structure is automatically generated in the process of word perception even though the explicit pronunciation of the phonological structure is not required and may sometimes even hinder task performance. Furthermore, phonological coding is an early and primary source of constraint on word perception. Phonological structures arise very rapidly and mediate the process of word perception by acting as a coherent frame for other ongoing processes.

A central and primary role for phonology is consistent with the phonological coherence hypothesis advanced by Van Orden and his colleagues (Van Orden & Goldinger, 1994; Van Orden et al., 1990). Rooted in dynamic systems theory, the phonological coherence hypothesis is typically expressed in terms of adaptive resonances within a triangular framework consisting of a level of orthographic processing units, a level of phonologic processing units, and a level of semantic processing units (see Bosman & Van Orden, 1997; Van Orden et al., 1990, Van Orden et al., 1997). In this resonance framework, a bidirectional connective matrix links the processing units of all levels into a fully interdependent recurrent network. During word perception, a pattern of activation over orthographic processing units flows forward to create a pattern over all associated phonological and semantic processing units. In turn, these processing units feed activation (i.e., top-down expectations) back to the orthographic units, transforming phonological and semantic patterns back into an orthographic form. *Resonance* is achieved when feedforward and feedback sources of interactive activation are mutually reinforcing. Within limits, the dynamic flow of activation is self-perpetuating, integrating the separate sources of orthographic, phonological, and semantic information into a coherent perceptual experience (see Gottlob, Goldinger, Stone, & Van Orden, 1999).

As opposed to the strong phonological theory of Frost (1998), traditional theories of word reading do not assign a central and primary role to phonological coding. In the earlier years, the empirical literature was dominated by the dual route theory (Coltheart, 1978). This theory postulates that two independent procedures are required for converting print to speech. In the major, fast procedure word pronunciation is directly addressed in a mental lexicon and in a secondary, slow procedure it is assembled through the application of grapheme-to-phoneme correspondence rules (i.e., the delayed phonology hypothesis). The two routes towards pronunciation are specially tailored to handle different kinds of words. The addressed (or lexical) route is appropriate for pronouncing both regular words (e.g., MINT) and irregular words (e.g., PINT), whereas the assembled (or nonlexical) route is suitable only for pronouncing (low-frequency) regular words and pseudowords. The dual route cascaded (DRC) model proposed by Coltheart and his colleagues (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) essentially implements the classic dual route theory. In this model, a phonological representation is formed either indirectly through a

relatively slow rule-based system of grapheme-to-phoneme correspondence rules (i.e., assembled phonology via the nonlexical route) or directly via a relatively fast lexical look-up. However, Seidenberg and McClelland (1989) demonstrated that a single system, a parallel distributed connectionist network, can be successfully trained to assemble phonology both for (low-frequency) regular words and for irregular ones, suggesting that the direct lexical route is in fact redundant.

The role of phonological coding in bilingual visual word perception

The cognitive system's capability to accommodate the faculty of language per se is impressive, but now consider the cognitive system of the bilingual, which harbours *two* languages. For decades, the nature of this coexistence and its implications for processing has been a source of dispute among linguists and cognitive psychologists. One of the major questions being posed is whether the processing of a word in one of the bilingual's languages is influenced by knowledge of words from the other language. Recent insights hold that under many circumstances bilinguals cannot suppress the non-target language. In a nutshell, a large number of psycholinguistic studies have supported the theoretical position that bilingualism entails a single integrated language system that processes words from the two languages basically *language non-selectively* (e.g., Bijeljac-Babic, Biardeau, & Grainger, 1997; De Groot, Delmaar, & Lupker, 2000; Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra & Van Hell, 2003; Dijkstra & Van Heuven, 2002; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Van Hell & Dijkstra, 2002). According to Grosjean (1997, 2001), the degree in which bilingual word processing proceeds language non-selectively may be a function of various factors. One such factor concerns aspects of *stimulus-list composition*, such as whether words from one language or from both languages are included in the experiment. This may affect the language mode of the participant, which, in turn, determines the extent to which the non-target language is co-activated with the target language during task performance (i.e., the language mode hypothesis, see also Dijkstra & Van Hell, 2003; Dijkstra & Van Heuven, 2002).

Research on bilingual word perception has yielded strong evidence that processing words in, especially, the second language (L2) proceeds essentially language non-selectively. In the bilingual studies referred to above, visual word perception was looked upon as a process of mapping spelling directly onto meaning. However, because research in the monolingual domain has convincingly demonstrated that phonological coding is a central and primary constituent of visual word perception, current models of bilingual word perception should take this process into account.

If the written forms of a bilingual's two languages both exploit the alphabetic principle, it is possible that spellings relate to sounds differently across languages. This situation in fact holds for Dutch-English bilinguals. Building on the notion of language non-selective processing, this state of affairs gives rise to the question whether bilinguals engage knowledge of spelling-to-sound relations from one or both languages when reading in either one of them. Specifically, the question of interest is whether, in case of bilingual processing of inconsistent English (L2) words, a phonological structure driven by knowledge of *Dutch* spelling-to-sound relations is initially coded as well. If it is established that for Dutch bilingual readers a spelling body such as -ood correlates very strongly with its pronunciation in Dutch, in fact more strongly than with any other English pronunciation, we would expect this to be the case.

The above considerations lead us to the more general question whether phonological coding is as fundamental to bilingual visual word perception as it is to monolingual visual word perception. If we were to observe language non-selective spelling-to-sound effects in a bilingual reading task, this would appear to be the case.

To date, only relatively few studies have investigated the role of phonology in bilingual visual word perception (e.g., Brysbaert, Van Dyck, & Van der Poel, 1999; Dijkstra, Grainger, & Van Heuven, 1999; Duyck, Diependaele, Drieghe, & Brysbaert, 2004; Gollan, Forster, & Frost, 1997; Jared & Kroll, 2001; Nas, 1983; Thierry & Wu, 2004; Tzelgov, Henik, Sneg, & Baruch, 1996; Van Wijnendaele & Brysbaert, 2002). To illustrate, in an ingenious experiment Nas (1983) provided an early demonstration of cross-language effects of spelling-to-sound knowledge. In this experiment, Dutch-English bilinguals performed an English lexical decision task ("is the letter string an English word?"). The critical manipulation concerned the type of nonwords that participants had to reject. Specifically, half of the nonwords were so called interlingual pseudohomophones (e.g., snay), letter strings that do not occur in English or Dutch, but that sound like Dutch words (i.e., *snee*, meaning 'cut') if processed as English words. Nas observed that interlingual pseudohomophones were rejected slower than control nonwords, and they were also more often misclassified as words. Apparently, reading English involved English-based phonological coding of letter strings which, in turn, gave rise to Dutch meaning activation. Apart from demonstrating the occurrence of mandatory phonological coding in visual word processing, this study showed that L2 reading may activate word knowledge (i.e., phonological structures) of the first language (L1) even if it hinders performance.

As a final illustration, Brysbaert et al. (1999) investigated phonological coding in L2 word processing. They asked Dutch-French bilinguals to identify briefly presented French target words. Each target word was preceded by a briefly pre-

sented, masked L1 Dutch word or nonword prime. Brysbaert et al. observed that with primes homophonic to the target participants recognised more French target words than with graphemic controls. Importantly, the homophonic primes were only homophonic with the French target word according to *Dutch* spelling-to-sound correspondence rules. These phonological priming effects thus indicate that word processing in the L2 is sensitive to L1 phonological information.

To conclude, the above findings support the hypothesis that phonology plays a leading role in bilingual visual word perception. They show, firstly, that processing words in the L2 involves mandatory phonological coding according to non-native intralingual spelling-to-sound knowledge, and, secondly, that the processing of L2 words can at the same time deploy phonological codes arising from L1 intralingual spelling-to-sound knowledge.

Simultaneous cross-language phonological coding. Even though phonological coding in L2 word perception has been demonstrated in the studies discussed so far, it has not yet been established that processing an L2 word involves the *simultaneous* coding of multiple, cross-language phonological structures in a single perceptual event. To address the specific question whether word perception in bilinguals involves the simultaneous coding of cross-language phonology, Jared and Kroll (2001) conducted a word-naming study with French-English bilinguals. The technique they used to reveal such coding was to examine the influence of *interlingual* enemy neighbors. Specifically, Jared and Kroll wondered whether L2 naming of an English word is slowed by spelling-to-sound knowledge of L1 French enemy words, analogously to the monolingual case, where naming of a word such as PINT is slowed by spelling-to-sound knowledge of English enemies such as MINT. Contrary to their expectation, Jared and Kroll (2001, Experiment 3) found little evidence that such is the case. In a subsequent experiment, they used French-English bilinguals who were relatively more proficient in their L1 French. With this new group of participants, they observed an effect of L1 French enemies on L2 English word naming. This interlingual consistency effect increased when participants first named a block of French words. Jared and Kroll also tested French-English bilinguals reading in their L1 and found little evidence of an effect of L2 English enemies on L1 French word naming unless they had recently spoken in their L2. Interestingly, the results of Jared and Kroll also showed that the intralingual consistency effect observed in L2 English word naming was larger than the interlingual consistency effect. That is, the impact of the relatively weak L2 English on English word naming turned out to be larger than the impact of the relatively strong French L1.

In sum, the interlingual consistency effects obtained by Jared and Kroll (2001) suggest that under some circumstances processing of an L2 word involves simultaneous, mandatory coding of multiple, *cross-language* phonological structures.

The present study

The present investigation explores the phonological coding processes of Dutch-English bilinguals reading English (L2) words. Our starting point is the hypothesis that phonology plays a central and primary role in visual word perception. This raises the question whether L2 word perception in bilinguals, just as L1 word perception, involves mandatory phonological coding. Furthermore, the hypothesis that L2 word perception proceeds language non-selectively gives rise to the question whether Dutch-English bilinguals might engage spelling-to-sound knowledge from both languages simultaneously when perceiving L2 words. In this study, cross-language phonological coding refers to inadvertent native Dutch phonology emerging simultaneously with appropriate (and inappropriate) non-native English phonology.

Our purpose is to study the process of phonological coding before and after it is constrained by global coherence of orthographic-phonologic-semantic activation dynamics — a transient integration of spelling, sound, and meaning information that can be utilised for word identification or to launch word pronunciation. In this view it is assumed that word perception is a *continuous* process, a process that does not stop when the system reaches a state of equilibrium (e.g., Goldinger, Azuma, Abramson, & Jain, 1997). One reason we adopt this dynamic systems approach is that it may be difficult if not impossible to detect phonological coding processes in tasks that demand a "cognitive moment of identification" (Perfetti & Tan, 1998). For example, in word naming (the task used by Jared & Kroll, 2001) or in reading for meaning, global processes of word perception (e.g., phonologic-semantic dynamics) may operate highly efficiently, so much so that they conceal the more local processes, the ones originating from the early phases of word perception. Consequently, a bilingual reader may manage to read an L2 word both fast and correctly, and show no or weak evidence of cross-language interference despite the fact that the non-target language is active (cf. Jared & Kroll, 2001, Experiments 3 and 4).

The print-to-speech correspondence task. To accommodate our requirement that experimental observations reflect processes originating from the early phases of word perception, we devised a new bimodal reading task. In this so-called *print-to-speech correspondence task*, two stimuli are presented simultaneously to the visual and auditory modalities. One of the stimuli is a visually presented printed word (e.g., MOOD), and the other is an auditorily presented speech segment, which may or may not be the word's spoken rime (i.e., the phonological body). The printed word is presented for approximately 200 ms, which, according to estimates of Rayner and Pollatsek (1989) is sufficient to launch word pronunciation or word identification.

The task to be performed by the participant is to judge whether the printed word and the spoken rime are congruent with one another. For example, on a *match trial,* the participant may be presented with the English word MOOD accompanied by the spoken rime of the word MOOD. In this case, the two stimuli are congruent with one another and a *yes* response would be appropriate. If however, on a *no-match trial,* the word MOOD is accompanied by an unrelated spoken rime, for example the spoken rime of the word BRIDE, the two stimuli are not congruent with one another and a *no* response is required. Spelling-to-sound ambiguity of the printed words may influence performance on both match trials and no-match trials. Accurate task performance may be elicited by a word with a consistent spelling-to-sound mapping (CM, e.g., MOON), because orthographic-phonologic dynamics cohere quickly in consistent words. Performance on a word with an inconsistent but typical spelling-to-sound mapping (TM, e.g., MOOD) may be more error-prone. For such a word, orthographic-phonologic dynamics cohere slower because of competition between local resonances. Finally, for a word with an inconsistent and atypical spelling-to-sound mapping (AM, e.g., BLOOD) performance may be even worse. Here, the competition between local resonances needs even more time to be resolved, because the relative self-consistency of inappropriate spelling-to-sound associations is lower for words with atypical mappings.

In view of our goals, one feature of the print-to-speech correspondence task is of particular interest, namely the possibility that a printed inconsistent word (e.g., MOOD) is accompanied by the spoken rime of one of its *enemies* (e.g., BLOOD). In such a "catch trial", the appropriate phonological structure that emerges from spelling is obviously not congruent with the spoken rime. However, we may expect inappropriate phonological structures, that have been inhibited in the course of word processing, to be restored at full strength if they are probed by external form-similar codings. If such an external probe consists of a spoken rime presented in the print-to-speech correspondence task, we may be able to detect inappropriate phonological structures that were launched at the start of word presentation and continued to lie dormant. In terms of multistability, an appropriate orthographic-phonologic resonance may reach a final, asymptotic state and then be perturbed to settle into an inappropriate orthographic-phonologic resonance.

To summarize, the specific question whether processing of an inconsistent word involves coding of inappropriate phonology is addressed by comparing performance on *catch trials* (e.g., printed word MOOD; spoken rime of BLOOD) and *no-match trials* (e.g., printed word MOOD; spoken rime of BRIDE). It is expected that performance on catch trials is worse than performance on no-match trials. This prediction can also be stated as: catch > no-match (we will refer to this difference as the Trial Type effect). This effect is predicted, because for no-match trials inappropriate phonological codings are not fostered by spoken rimes whereas in catch

trials this is expected to occur. Furthermore, the Trial Type effect should be larger for words with atypical mappings than for words with typical mappings, because inappropriate phonologic codings are stronger for AM-words such as BLOOD than for TM-words such as MOOD. This prediction for the interaction effect of Trial Type and Word Type can also be stated as: Trial Type effect (AM) > Trial Type effect (TM).

Unique contribution of this study. The present study is an effort to contribute to our understanding of monolingual and bilingual visual word perception. Guided by the general principle of manifold form-function relations and conceived within the resonance framework of Van Orden and Goldinger (1994), it brings together two central theoretical issues in reading research: The role of phonology in monolingual visual word perception and the language non-selective view of bilingual word processing. Building on related work (Brysbaert et al., 1999; Dijkstra et al., 1999; Jared & Kroll, 2001), this study aims to add new insights in the use of spelling-to-sound knowledge in the processing of L2 words. Foremost, it extends the work of Jared and Kroll by exploring the exact dynamics of phonological coding that may be obscured in word-naming responses. In their study, Jared and Kroll relied on the word-naming task and had to infer the process of inappropriate cross-language phonological coding *indirectly* from neighborhood effects. The unique feature of the present study is that, using the idea of multistability, we aim to detect the competing inappropriate phonological coding originating in the initial stages of word perception *on the fly*. Our new experimental paradigm creates in a simple but effective way the circumstances wherein the reading system is perturbed to exchange appropriate phonology for possible but inappropriate phonology. Such an exchange is only possible if coding of inappropriate phonology actually arises in the perception of spelling-to-sound ambiguous words.

General method

Participants

A total of 200 Dutch first-year psychology students from the University of Amsterdam, The Netherlands, participated in Experiments 1–3 for either course credit or a small financial compensation. The participants were Dutch-English bilinguals, with Dutch as their native language and English as their strongest foreign language. All were fairly fluent in English: They had learned it at school for about 3–4 hours a week, starting at the age of ten and continuing until the end of secondary school. The Dutch education system emphasizes the productive use of new languages, which implies that the participants also acquired knowledge of pronunciations. Furthermore, their university education in psychology required them to

read mainly in English. Each participant took part in one of a total of three experiments, according to when he or she arrived at the laboratory. None participated in more than one experiment. A total of 60 students participated in Experiment 1, 80 in Experiment 2, and 60 in Experiment 3.

Materials

Selection of printed word stimuli. The printed word stimuli used in Experiments 1–3 consisted of 120 monosyllabic English words. These were extracted from the linguistic database of Ziegler et al. (1997). Two experimental word lists were created. One list consisted of 60 English words that have Dutch neighbors (e.g., MOOD), and the other consisted of 60 English words that do not have Dutch neighbors (e.g., SAID). Homophones and interlingual homographs were avoided in these lists, and so were words that were expected to be unfamiliar to the Dutch participants. A word was excluded if it was semantically or phonologically unfamiliar to the first author or to any of five Dutch students from the same population as the ones tested in this study.

In both experimental word lists three word types were contrasted that differed in the degree of (in)consistency of spelling-to-sound mappings. In each list 20 words had *consistent* spelling-to-sound mappings, 20 words had *typical* spelling-to-sound mappings, and 20 words had *atypical* spelling-to-sound mappings. The consistent words (CM) had spelling bodies that mapped onto a single phonological body (e.g., MOON). The typical and atypical inconsistent words (TM and AM) had spelling bodies that mapped onto more than one phonological body. The latter word types differed in the number and frequency of friends and enemies. Typical inconsistent words (e.g., MOOD) had a larger number of friends than enemies. Furthermore, the frequency of the friends was generally higher than the frequency of the enemies. In contrast, atypical inconsistent words (e.g., BLOOD) had a larger number of enemies than friends and the frequency of the enemies was generally higher than of the friends. In order to ensure that the two groups of inconsistent words were similar in visual form, both contained the same set of 20 spelling bodies (e.g., -OOD in MOOD–BLOOD; -AID in SAID–PAID). The complete list of word stimuli is provided in Appendix. Summary statistics for the relevant variables for each experimental word list are presented in Table 1. Word frequency estimates were taken from the English corpus type lexicon of the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993; Burnage, 1990). Bigram frequency estimates were collected by means of the computer program LexStat (van Heuven, 2000) and using the Kucera and Francis (1967) database. Familiarity and imageability ratings were obtained from the Medical Research Council (MRC) psycholinguistic database (Coltheart, 1981).

Table 1. Characteristics of the English Printed Words Used in Experiments 1 and 2. (CM = Consistent Spelling-to-Sound Mappings, TM = Typical Spelling-to-Sound Mappings, AT = Atypical Spelling-to-Sound Mappings, NOF = Number of Friends, FOF = Frequency of Friends)

	No Dutch Neighbors			With Dutch Neighbors		
	CM	TM	AM	CM	TM	AM
Mean number of letters	4.7	4.8	4.8	4.1	4.0	4.1
Mean log frequency	1.9	1.8	1.8	2.0	2.0	2.0
Mean log bigram frequency	5.9	6.0	5.9	6.1	6.1	6.0
Mean familiarity	5.7	5.6	5.6	5.4	5.6	5.7
Mean imagability	5.0	4.8	4.8	4.2	4.9	4.7
Mean NOF	5.6	5.6	1.4	4.6	4.5	1.9
Mean log summed FOF	2.6	2.7	2.0	2.8	2.8	2.1
Mean consistency ratio	1.0	0.8	0.2	1.0	0.8	0.2

Coupling Printed Words to Spoken Rimes. The sound stimuli used in Experiments 1–3 consisted of speech segments (i.e., spoken rimes) with a vowel-consonant (VC) structure (e.g., -ood). These were obtained by having a female native British-English speaker read aloud lists of words and pseudowords. She was instructed to pronounce the words in Received Standard English. The resulting pronunciations were all recorded and subsequently edited such that of each letter string's spoken form only the *rime* remained.

Three types of trials were created. *Match trials* consisted of a printed word and a spoken rime that were congruent with each other (and hence required a *yes* response). This congruency was accomplished by using the actual rime of the printed word's spoken form. *No-match trials* and *catch trials* consisted of a printed word and a spoken rime that were not congruent with each other (and hence required a *no* response). For no-match trials this incongruency was accomplished by using the rime of an unrelated spoken word. This unrelated word shared only the coda (i.e., final consonant) with the target word. An example of such a trial is when the target word mood is accompanied by the spoken rime of bride. Finally, incongruency was accomplished for catch trials by using the rime of an *enemy* of the target word. Likewise, this enemy word shared only the coda with the target word. An example of such a trial is when the target word mood is accompanied by the spoken rime of blood.

Statistical data analysis. Null hypothesis significance tests were augmented with 95% confidence intervals for principal mean differences (e.g., Kirk, 1995; Loftus & Masson, 1994; Masson & Loftus, 2003; Maxwell & Delaney, 1990). In the statistical

analyses performed in this study, error variance was estimated for a Latin square design (see Maxwell & Delaney, 1990; Myers & Well, 1995; Pollatsek & Well, 1995). To ensure correspondence between the appropriate ANOVA and a confidence interval, the same estimate of error variance was used for both. For all sets of pairwise comparisons, the Bonferroni procedure was used, for the null hypothesis tests as well as for the confidence intervals (see Kirk, 1995; Maxwell & Delaney, 1990). This involved restricting the per-comparison alpha level and the construction of simultaneous 95% confidence intervals (95% SCI, see Maxwell & Delaney, 1990).

Experiment 1

The present study's first objective was to examine whether the process of English (L2) word perception in Dutch-English bilinguals involves mandatory, *intralingual* phonological coding. As pointed out before, we used the print-to-speech correspondence task to address the question whether perception of an inconsistent word is influenced by spelling-to-sound knowledge of English enemy words or, stated differently, whether processing of such a word involves competition between appropriate and inappropriate phonological codings. We may conclude that such competition occurs if we can demonstrate that processing an inconsistent word involves coding of inappropriate phonology. For this purpose, the use of catch trials is of crucial importance. As explained earlier, catch trials may enable us to detect local phonological codings that were launched at the start of word presentation but that have been inhibited in the course of word processing. Thus if, in a catch trial, MOOD is accompanied by the spoken rime of BLOOD, and if inappropriate phonology is indeed part of the initial conditions of word perception, this spoken rime may restore the degraded, inappropriate coding to such a degree that participants may find it difficult, or are actually unable, to perceive a mismatch. That is, the spoken rime may put the degraded, inappropriate coding full-blown back into competition. Consequently, participants may occasionally react with an incorrect *yes* response (i.e., a false-positive), thus indicating that they perceived MOOD's phonology to rhyme with the rime of BLOOD.

Method

Participants and materials. A group of 60 Dutch-English bilinguals served as participants. They were presented with the printed English words and spoken rimes described in the General Method section.

Procedure. The printed word stimuli were displayed in lowercase letters in the centre of the computer screen of an Apple Macintosh PowerPC 4400/200, using a standard Macintosh font (Geneva, size 14). Presentation of the printed words was synchronised with the refresh cycle of the screen. The speech segments (i.e., the spoken rimes) were presented at a comfortable level through a set of headphones (SONY MDR-V100). Error rates were collected by means of two button boxes interfaced with the serial ports of the computer. Stimulus presentation and data recording were controlled by the computer program fLexi, an experiment generator developed at the Department of Psychology of the University of Amsterdam.

Participants were tested individually in a quiet and normally lit room. They were seated at approximately 50 cm in front of the computer screen and were given verbal and written instructions, followed by a block of practise trials. None of the practise words contained spelling bodies occurring in the experimental words. An experimental session involved a series of experimental trial blocks. The first two trial blocks were preceded by a block of 30 (Experiments 1 and 2) or 10 (Experiment 3) practise trials. The order of trials was randomised for each participant and for each trial block.

Each trial started with a fixation region that was displayed in the centre of the computer screen. The fixation region consisted of two horizontal dashes extending approximately 2.1 cm. It remained on the screen for approximately 500 ms and was then replaced by a printed word. The printed word was displayed for approximately 200 ms and was in turn replaced by a pattern mask that remained on the screen for another 1.8 sec, or until the participant responded. Sound stimuli were presented simultaneously with the onset of the printed word stimuli. Participants were instructed to decide as quickly and accurately as possible whether print and speech were congruent with one another by pressing either the *yes* or *no* button, using the index fingers. They all used the index finger of the dominant hand for *yes* responses. After the response, immediate visual feedback was provided. The feedback was displayed for approximately 750 ms. If the participant pressed the correct button the phrase "correctly responded" appeared; if not, the word "wrong" appeared. If the participant failed to respond within 2000 ms after onset of the printed word, the trial was aborted and the word "slowly…" appeared. Irrespective of the participant's response, the total duration of a trial was approximately 2500 ms. The next trial was initiated after an interstimulus interval of approximately 1000 ms.

Experimental design. Five groups of mismatch trials were created representing specific combinations of Trial Type (catch trial vs. no-match trial) and Word Type (AM vs. TM vs. CM). These five combinations were: no-match trial AM, no-match trial TM, no-match trial CM, catch trial AM, and catch trial TM. In this experiment and the following ones, the spelling body of a typical inconsistent word

Table 2. Overview of No-Match Trials and Catch Trials Within Four Trial Blocks A, B, C, and D Following a Single Latin Square

Latin Square Sequence	Primary Test		Replication	
	Position 1	Position 2	Position 3	Position 4
A B C D	PAID No-Match	SAID Catch	PAID Catch	SAID No-Match
B A D C	SAID Catch	PAID No-Match	SAID No-Match	PAID Catch
C D A B	PAID Catch	SAID No-Match	PAID No-Match	SAID Catch
D C B A	SAID No-Match	PAID Catch	SAID Catch	PAID No-Match

(e.g., SAID) also appeared in an atypical inconsistent word (e.g., PAID). To prevent intralist-priming effects of spelling bodies, the two words containing the same spelling body were presented in separate blocks of trials. In Experiments 1 and 2 inconsistent words were paired with a spoken rime to create a no-match trial, but also with a spoken rime to create a catch trial. Consequently, a spelling body appeared in four blocks of trials, A, B, C, and D. For example, in Block A the word PAID appeared together with the spoken rime of PLEAD to create a no-match trial. In Block B, the neighbor SAID appeared together with the spoken rime of PAID to create a catch trial. In Block C, the word PAID appeared again, together with the spoken rime of SAID, to create a catch trial. Finally, in Block D the word SAID appeared again, together with the spoken rime of PLEAD, to create a no-match trial. In Experiments 1 and 2, each participant was presented with each of the four trial blocks. Hence, for each participant data was obtained for each possible combination of Trial Type and Word Type. The temporal order of all four trial blocks was counterbalanced across four different participant groups according to a single Latin square. Participants were randomly assigned to the four rows of the square (see Table 2).

Words used in Blocks A and B also appeared in Blocks C and D, but, for mismatch trials, they were paired with different spoken rimes. Therefore, data obtained from performance on Blocks C and D can be conceived as a (within-participants) replication of data obtained from performance on Blocks A and B, and visa versa. Because presenting the same printed words may affect the data in complicated ways, we analysed the data separately for the blocks presented in the first two temporal positions (Positions 1 and 2, *primary test*), and the blocks presented in the second two temporal positions (Positions 3 and 4, *replication*). The data obtained from Positions 1 and 2 were considered the most valid, because the trial blocks in these positions contained unique words, not presented before. The two separate

blocks of trials containing the same spelling bodies (Blocks A and B for participant groups 1 and 2, and Blocks C and D for participant groups 3 and 4) were administered in two separate experimental sessions conducted one week apart, a precautionary procedure also used by Jared (1997). The other two trial blocks, administered in Positions 3 and 4, immediately followed the trial block administered in Position 2. Thus, different stimulus pairs were processed by different participants. This potential source of variance can be isolated and removed from the estimate of error variance, which may improve the efficiency of the design. Procedures for this are provided by Kirk (1995) and Myers and Well (1995; see also Pollatsek & Well, 1995). In the present case they simply involved adding participant groups (i.e., the four rows of the Latin square) as a between-subjects variable in an ANOVA, resulting in a treatments × participants(group) error term.

In Experiment 1, the list of English words with Dutch neighbors (e.g., MOOD) was used for match trials and the list of English words without Dutch neighbors (e.g., SAID) was used for no-match and catch trials. In Experiments 2 and 3 this was the other way around. In all three experiments match trials were treated as filler trials, meaning that only the data of the no-match and catch trials will be reported.

Results

For each participant, percentages of false-positive errors were calculated within the five groups of trials, separately for the first two blocks (i.e., primary test) and
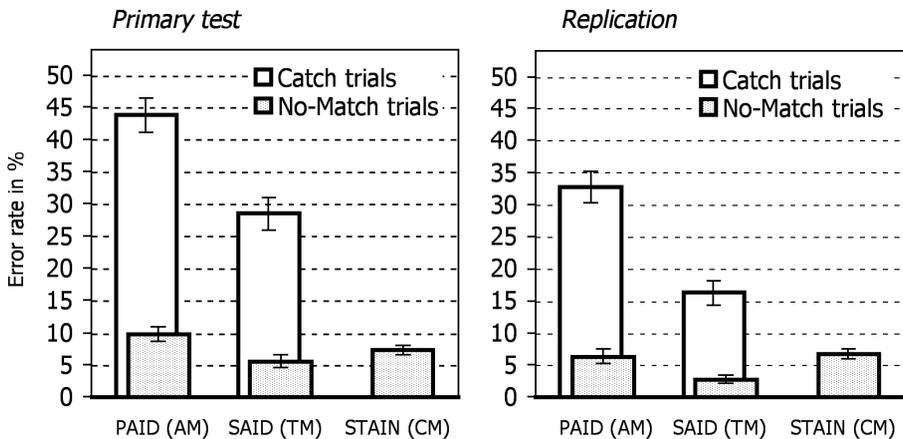
Figure 1. Mean percentages of false-positives as a function of Trial Type (catch trials vs. no-match trials) and Word Type (AM vs. TM vs. CM) for primary test (left panel) and replication (right panel) in Experiment 1. Error bars represent the standard error of the mean.

for the second two blocks (i.e., replication). Response latencies were also collected. However, because the results of the response latencies generally converged with those of the error rates and to save space, they are not reported below. A participant produced an error in a mismatch trial when he or she pressed the *yes* button when presented with a printed word and a spoken rime that were in fact incongruent with one another. Figure 1 shows the mean percentages of false-positives for the catch trials and no-match trials, both for the primary test and the replication. These mean percentages were collapsed over trial blocks and participant groups. The error bars represent the standard error of the mean.

Recall that catch trials should induce more errors than no-match trials (i.e., the Trial Type effect). This was tested with the contrast: catch > no-match. Furthermore, a Trial Type by Word Type interaction effect was expected, captured by the contrast Trial Type effect (AM) > Trial Type effect (TM). Table 3 presents interval estimates of these contrasts. It turned out that Trial Type was indeed associated with different numbers of false-positives. Figure 1 shows that error rates for catch trials reached strikingly high levels, up to 44% for atypical inconsistent words. Overall, participants made more errors on catch trials than on no-match trials, $F(1, 56) = 234.53$, $MSE = 299.85$, $p < .001$. As Figure 1 shows, this Trial Type effect was larger in the primary test than in the replication, with a difference of 8.4 percentage points (95% CI 4.9 to 11.9). The Trial Type effect was also larger for atypical inconsistent words (AM) than for typical inconsistent words (TM). The

Table 3. Planned Comparisons (with 95% Confidence Intervals) of No-Match Trials and Catch Trials for Words with Atypical Spelling-to-Sound Mappings (AM) and Words with Typical Spelling-to-Sound Mappings (TM) for Experiments 1–3

	Overall		AM		TM		Difference	
	Catch > No-match		Catch > No-match		Catch > No-match			
Error rates								
Exp 1 (primary test)	28.4	(24.5, 32.4)	33.8	(27.3, 40.4)	23.0	(17.9, 28.1)	10.8	(4.5, 17.2)
Exp 1 (replication)	20.0	(16.8, 23.2)	26.3	(21.2, 31.5)	13.7	(9.8, 17.6)	12.7	(8.1, 17.2)
Exp 2 (primary test)	26.8	(24.0, 29.6)	33.9	(29.5, 38.2)	19.8	(16.2, 23.3)	14.1	(10.1, 18.2)
Exp 2 (replication)	19.6	(16.9, 22.3)	26.5	(22.2, 30.8)	12.8	(9.4, 16.1)	13.8	(9.9, 17.6)
Exp 3	20.9	(18.4, 23.3)	38.5	(32.8, 44.2)	12.3	(7.8, 16.9)	26.2	(18.8, 33.6)

Trial Type effect (AM) > Trial Type effect (TM) contrast was statistically significant, $F(1, 56) = 11.62$, $MSE = 302.98$, $p < .01$.

The above results will be discussed simultaneously with those of Experiment 2. To anticipate, the results of both experiments strongly suggest that in Dutch-English bilinguals the processing of an inconsistent L2 English word indeed involves inadvertent coding of inappropriate phonology resulting from spelling-to-sound knowledge of enemy neighbors in L2.
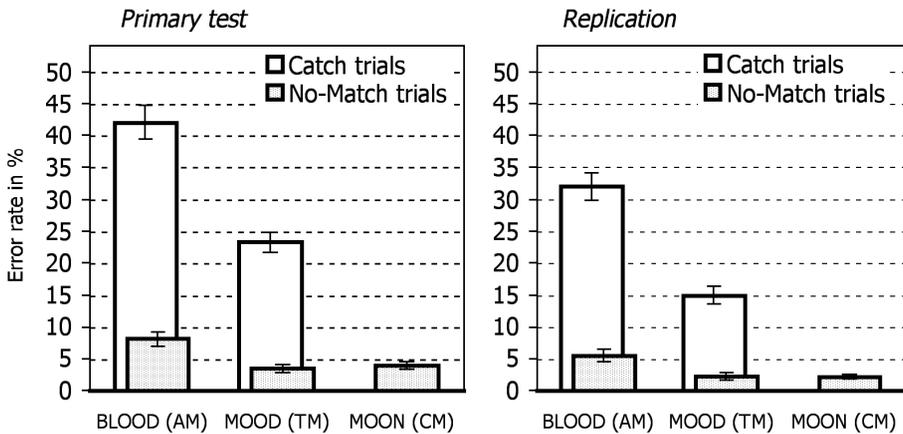
Experiment 2

In Experiment 2 and also in Experiment 3, the list of English words without Dutch neighbors (e.g., SAID) was used for the match trials whereas the list of English words with Dutch neighbors (e.g., MOOD) was used for no-match trials and catch trials. In the previous experiment, this was the other way around. Except for this change, Experiment 2 was identical to Experiment 1. Thus, the crucial difference with Experiment 1 is that the present experiment used English words that have Dutch enemy neighbor words such as LOOD (in addition to English enemy neighbors). As explained in the Introduction section, this particular type of English words is suitable for studying cross-language spelling-to-sound consistency effects. With this type of words, Experiment 2 explored the impact of knowledge of *English* enemy words on English word perception. In Experiment 3, the same words were used to study the impact of knowledge of *Dutch* enemy words on English word perception.

Method

Participants, materials, and procedure. A group of 80 Dutch-English bilinguals served as participants. They were presented with the printed English words and the spoken rimes described in the General Method section. The experimental design and procedure were identical to those used in Experiment 1.

Results

As in Experiment 1, error rates for catch trials reached stunningly high levels of up to 42% for atypical inconsistent words. Overall, participants made more errors on catch trials than on no-match trials, $F(1, 76) = 424.88$, $MSE = 203.02$, $p < .001$. As Figure 2 suggests, this Trial Type effect was again larger in the primary test than in the replication (see Table 3), with a difference of 7.2 percentage points (95% CI 3.9 to 10.5). Furthermore, Figure 2 suggests that the Trial Type effect was again larger

Figure 2. Mean percentages of false-positives as a function of Trial Type (catch trials vs. no-match trials) and Word Type (AM vs. TM vs. CM) for primary test (left panel) and replication (right panel) in Experiment 2. Error bars represent the standard error of the mean.

for atypical inconsistent words (AM) than for typical inconsistent words (TM), with a statistically significant Trial Type effect (AM) > Trial Type effect (TM) contrast, $F(1, 76) = 48.40$, $MSE = 164.90$, $p < .001$.[1]

Discussion

Experiments 1 and 2 provided evidence for the hypothesis that perception of an inconsistent word involves inadvertent coding of inappropriate phonology in L2. For Dutch-English bilinguals, rejecting a catch trial consisting of an inconsistent printed word (e.g., MOOD) and a spoken rime derived from an enemy of the word (e.g., BLOOD) appears to demand exceptional effort. Participants frequently responded with false-positives, thus indicating that they, for instance, perceived MOOD's phonology to rhyme with the rime of BLOOD. The observed effects of Trial Type indicate that, in catch trials, spoken rimes probe the degraded, inappropriate phonological codings to such a degree that competition between appropriate and inappropriate codings is resumed at full strength. In terms of multistable perception, this probe may cause the system to jump from an appropriate to an inappropriate phonological coding. Moreover, the Trial Type effect was considerably larger for words with atypical mappings than for words with typical mappings. That is, in a catch trial, BLOOD's phonology is more readily perceived to rhyme with that of MOOD than in the reversed case. It therefore seems that degraded, inappropriate phonological codings corresponding to highly self-consistent mappings are more readily restored than codings corresponding to less self-consistent mappings.

Taken together, the finding that rejecting a catch trial takes extraordinary effort strongly indicates that, in Dutch-English bilinguals, processing of an inconsistent English word involves inadvertent coding of an inappropriate phonological structure, just as it has been shown to occur in monolingual English speakers. This supports the view that perception of inconsistent English words by Dutch-English bilinguals involves mandatory phonological coding, resulting in multiple, competing phonological structures.

Experiment 3

As we have seen so far, under specific conditions MOOD may be perceived to rhyme with BLOOD. But is it also true that, due to manifold *cross-language* spelling-to-sound relations, MOOD may be perceived to rhyme with the Dutch word LOOD (which rhymes with ROAD)? This question follows from the second objective of this study, which seeks to investigate whether the process of L2 English word perception involves coding of inappropriate phonology due to knowledge of L1 *Dutch* enemy neighbors. This research goal thus tests the possibility that L2 word perception in bilinguals proceeds essentially language non-selectively.

Evidence for inadvertent coding of cross-language phonology may be observed in performance on catch trials, such as when MOOD is accompanied by a spoken rime derived from a Dutch enemy neighbor (e.g., the Dutch word LOOD). Thus, if coding of inappropriate Dutch phonology is actually part of the initial conditions of the perception of MOOD, the spoken rime of LOOD may restore the degraded, inappropriate coding to such a degree that participants may find it difficult to, or may even be unable to, perceive a mismatch. Note that in contrast to the previous experiments, Experiment 3 also included catch trials for consistent words such as MOON.

The general predictions are the same as those examined in the previous experiments. In addition to the earlier comparisons, Experiment 3 explored the effect of stimulus-list composition on performing the print-to-speech correspondence task. Studies of Dijkstra et al. (1998), De Groot et al. (2000), and Jared and Kroll (2001) have shown that stimulus-list composition affects the degree in which L2 words are processed language non-selectively (but see Dijkstra & Van Hell, 2003; Van Hell & Dijkstra, 2002;). In order to examine the effect of stimulus-list composition in Experiment 3, two experimental conditions were compared. In the Dutch-fillers condition the stimulus materials were mixed with an additional 25% Dutch filler trials and in the (neutral) English-fillers condition they were mixed with an additional 25% English filler trials. Adding Dutch trials to the stimulus set was expected to affect the relative activation of the non-target language. Processing

Dutch words should boost the Dutch language system, which may then enhance (inappropriate) phonological coding of English words according to Dutch spelling-to-sound knowledge. This, in turn, may lead to more false-positive errors for catch trials. However, because it is assumed that in the initial conditions of word perception all phonological structures that have previously been associated with a spelling body are launched, irrespective of the relative dominance of the non-target language system, we might as well obtain a null effect of stimulus-list composition. Essentially, this reasoning corresponds with the view of Dijkstra and Van Hell (2003), that word processing is always language non-selective.

Method

Participants. Sixty Dutch-English bilinguals served as participants.

Materials. The same printed word stimuli were used as in Experiment 2. For Experiment 3, English word stimuli were selected that contained spelling bodies which have distinct pronunciations in English and Dutch. The spelling body *-ood*, for example, is pronounced as rhyming with ROAD in Dutch, as in the Dutch word LOOD. These types of words were required to create "Dutch" catch trials that consist of a printed English word and a spoken rime derived from a *Dutch* enemy neighbor. Due to restrictions of stimulus selection, only half of the atypical inconsistent words, typical inconsistent words, and consistent words were suitable for this purpose. For Experiment 3, these specific words were used to create catch trials (with spoken rimes derived from either English or Dutch enemy neighbors), and the remaining inconsistent words were used to create no-match trials. Another difference with the previous experiments was that in Experiment 3 inconsistent words were paired with a single spoken rime. Thus, an inconsistent word was used to create either a no-match trial or a catch trial. Consequently, because printed words and spoken rimes now formed exclusive pairs, in contrast to Experiments 1 and 2 there was no need for using more than two trial blocks. In Experiment 3, the 40 inconsistent word stimuli were paired with only one of the available sound stimuli, thus creating either a no-match trial or a catch trial. Of both the group of 20 words with typical mappings and the group of 20 words with atypical mappings, 10 served to create no-match trials and 10 served to create catch trials. Note that the two groups of words had the same spelling bodies.

In creating the catch trials, words of which the spelling body had a pronunciation similar to the corresponding Dutch word were discarded. Table 4 shows the statistics for the relevant variables, separately for the word stimuli allocated to no-match trials and for those allocated to catch trials. In this new arrangement of word stimuli the balance on the set of stimulus characteristics was preserved. However,

Table 4. Characteristics of the English Printed Words Used in Experiment 3. (CM = Consistent Spelling-to-Sound Mappings, TM = Typical Spelling-to-Sound Mappings, AT = Atypical Spelling-to-Sound Mappings, NOF = Number of Friends, FOF = Frequency of Friends)

	No-match trials			Catch trials		
	CM	TM	AM	CM	TM	AM
Mean number of letters	4.3	3.7	3.9	3.9	4.2	4.2
Mean log frequency	2.0	1.9	1.8	1.9	2.1	2.2
Mean log bigram frequency	6.2	6.0	6.0	6.0	6.1	6.0
Mean familiarity	5.3	5.5	5.6	5.5	5.7	5.7
Mean imagability	4.0	5.4	4.5	4.3	4.4	4.9
Mean NOF	6.0	6.3	0.9	3.1	2.7	2.8
Mean log summed FOF	3.0	2.7	1.8	2.6	2.8	2.3
Mean consistency ratio	1.0	0.8	0.2	1.0	0.7	0.3

it turned out that, for catch trials, the differences in number and frequency of English friends and enemies between the groups of word stimuli representing the three types of words were slightly altered. The corresponding consistency ratios were reduced but not totally eliminated.

Experimental design and procedure. As in Experiments 1 and 2, the spelling body of an English word with a typical mapping (e.g., MOOD) also appeared in an English word with an atypical mapping (e.g., BLOOD). To prevent intralist-priming effects of spelling bodies, the two words containing the same spelling body were presented in two separate blocks of trials. The two separate blocks of trials containing the same spelling bodies were administered in one experimental session, interrupted by a short break. In the mismatch trials, each inconsistent word was either paired with a spoken rime to create a no-match trial, or with a spoken rime to create a catch trial. The two trial blocks A and B contained equal numbers of words from all three word types. In each block, there were 10 catch trials and 10 no-match trials for inconsistent words, and 10 no-match trials for consistent words. Each participant was presented with each of the two trial blocks. Hence, for each participant data was obtained for each available combination of Trial Type and Word Type, using every word only once. The temporal order of the two trial blocks was counterbalanced across two different participant groups according to a single Latin square. Participants were randomly assigned to the rows of the square.

Experiment 3 used the same basic design as the previous experiments. However, for half of the participants the stimulus materials were mixed with (25%) Dutch filler trials (Dutch-fillers condition) and for the other half they were mixed

with (25%) English filler trials (English-fillers condition). Therefore, in the analyses Filler Type was treated as a between-subjects variable. The filler trials consisted of equal numbers of match trials and no-match trials. For the Dutch-fillers condition an additional set of Dutch trials was created to serve as practise trials. Participants in this condition were informed that the experimental materials not only consisted of English words but also of Dutch words.

Results

Figure 3 shows that error rates for catch trials again reached strikingly high levels, up to 50% for atypical inconsistent words. Overall, participants made more errors on catch trials than on no-match trials, $F(1, 58) = 290.36$, $MSE = 45.08$, $p < .001$. Furthermore, Figure 3 shows that the Trial Type effect was larger for the AM-words than for the other word types (TM and CM). There was no evidence of a Trial Type by Filler Type interaction effect ($F < 1$). The Trial Type effect (AM) > Trial Type effect (TM) contrast gave a difference of 26.2 percentage points, with a 95% SCI of 18.8 to 33.6, $F(1, 58) = 75.94$, $MSE = 270.49$, $p < .001$. Similarly, the Trial Type effect (AM) > Trial Type effect (CM) contrast gave a difference of 26.7 percentage points with a 95% SCI of 18.4 to 35.0, $F(1, 58) = 62.70$, $MSE = 340.23$, $p < .001$, whereas the Trial Type effect (TM) > Trial Type effect (CM) gave a 0.5 difference in percentage points, with a 95% SCI of -4.4 to 5.4, $F(1, 58) = 0.06$, $MSE = 118.19$, $p = .802$.

Analyses evaluating the effect of stimulus-list composition. As can be seen in Figure 3, the two groups of participants receiving different types of fillers produced nearly
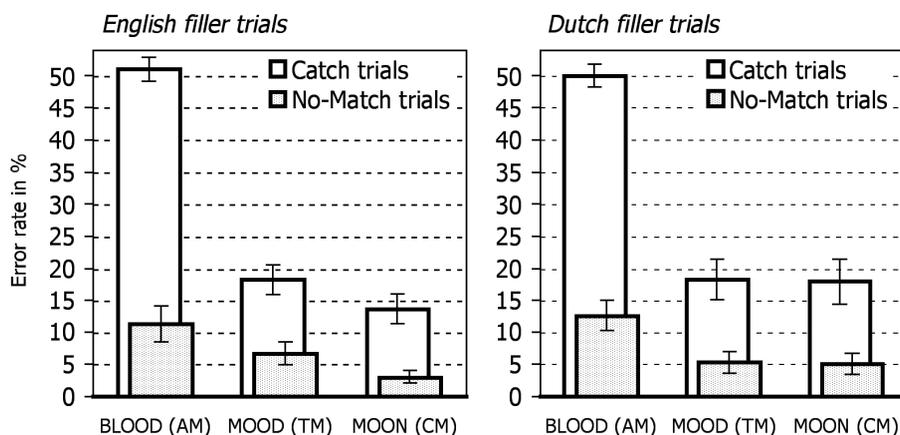
Figure 3. Mean percentages of false-positives as a function of Trial Type (catch trials vs. no-match trials) and Word Type (AM vs. TM vs. CM) for the English-fillers condition (left panel) and the Dutch-fillers condition (right panel) in Experiment 3. Error bars represent the standard error of the mean.

identical results. Overall, participants in the Dutch-fillers condition did not produce significantly more false-positive errors on mismatch trials than participants in the English-fillers condition. The difference in group means was 0.9 percentage points (95% CI -3.1 to 4.9) and not statistically significant ($F < 1$). Furthermore, contrasting no-match-trial and catch-trial performance, the Trial Type effect was not significantly larger in the Dutch-fillers condition than in the English-fillers condition, with a statistically non-significant group difference of 0.4 percentage points ($F < 1$, 95% CI -4.5 to 5.4).

Discussion

Experiment 3 extends the findings of Experiments 1 and 2 by demonstrating the occurrence of *interlingual* spelling-to-sound consistency effects in Dutch-English bilinguals. It was found that spelling-to-sound knowledge of *Dutch* enemy neighbors affected the ability to perceive a mismatch between an English printed word and a spoken rime. As in the previous experiments, the results obtained with the print-to-speech correspondence task were compelling: Rejecting a catch trial that consisted of a printed inconsistent word (e.g., MOOD) and a spoken rime that was derived from a Dutch enemy of the word (e.g., LOOD) appeared to demand a remarkably large effort. Participants responded very frequently with false-positives, thus indicating that they, for instance, perceived MOOD's phonology to rhyme with the rime of the Dutch word LOOD. This finding strongly supports the hypothesis that multistable perception of an interlingually inconsistent word such as MOOD involves inadvertent cross-language coding of inappropriate Dutch phonology.

A large Trial Type effect with this new type of spoken rimes was expected for Dutch-English bilinguals. For these participants, spelling-to-sound knowledge of Dutch words is likely to be stronger (i.e., more self-consistent) than spelling-to-sound knowledge of English words. Thus, a highly self-consistent Dutch phonological coding that has been inhibited for an atypical inconsistent word such as BLOOD is restored quite instantly by a fostering sound stimulus. Consequently, in a catch trial, a degraded, inappropriate Dutch phonological coding is pulled readily into competition again, which hinders perception of a mismatch between print and sound. In sum, in agreement with the notion of a leading role of phonology in bilingual visual word perception, the Trial Type effects observed in Experiment 3 support a language non-selective view of bilingual word perception, in which the perception of an inconsistent word involves simultaneous phonological coding in both of the bilingual's languages.[2]

The results of Experiment 3 support the conclusion reached by Brysbaert et al. (1999) and Jared and Kroll (2001) that L2 reading not only engages spelling-to-sound knowledge of the target language, but also that of the non-target

language, L1. Moreover, they extend our current knowledge of the influence of cross-language enemy neighbors on word processing such as gathered by Jared and Kroll (2001). These investigators observed that the impact of enemy spelling-to-sound knowledge on L2 word naming was larger if it concerned enemies of the relatively weak L2 than if it concerned enemies of the relatively strong L1. Even when the non-target L1 was boosted by a previous trial block, its impact remained relatively small. What this finding seems to indicate is that inadvertent coding of inappropriate phonology due to knowledge of L1 enemy neighbors is relatively insubstantial. In contrast, the present experiment suggests that inadvertent coding of inappropriate L1 Dutch phonology is substantial. From the principle of self-consistency this finding makes sense, because spelling-to-sound knowledge in the L1 is stronger than in the L2. Nevertheless, even though it may be expected that inadvertent coding of L1 phonology occurs more firmly than of L2 phonology, the fact remains that for L2 word-naming performance the impact of L1 enemy spelling-to-sound knowledge was rather small in Jared and Kroll's (2001) study.

How can this deviant pattern of results be understood? The key to the solution may be the possibility that inadvertent coding of inappropriate L1 phonology takes place in the initial conditions of word perception, a phase that may be revealed by the print-to-speech correspondence task but not by the word-naming task. In this initial phase, appropriate and inappropriate phonological structures emerge in proportion to their statistical dominance. However, incorrect codings are rapidly constrained by semantic feedback as the system moves towards an appropriate phonological structure. Because in the word-naming task, although present initially, phonological codings relevant to words from the non-target language are inhibited, cross-language competition between local orthographic-phonologic resonances is resolved quickly, resulting in relatively unobstructed word-naming performance.

Finally, Experiment 3 incorporated a stimulus-list composition variable (i.e., Filler Type). It was expected that adding Dutch filler trials to the stimulus set would increase the relative dominance of the non-target language, L1 Dutch. Yet, a comparison of the Dutch-fillers and English-fillers conditions yielded no evidence of enhanced (inappropriate) phonological coding according to Dutch spelling-to-sound knowledge. This result is clearly not in accordance with observations of marked stimulus-list composition effects such as reported in Dijkstra et al. (1998). Again, this discrepancy may be understood by assuming that the print-to-speech correspondence task reveals the ballistic nature of phonological coding, whereas the lexical decision task used in these other studies reveals processes of global-level linguistic coding.

General discussion

In the mental lexicon, English words reside among neighbors, some of which are friends and others are enemies. Our first objective in this study was to assess the impact of English enemy neighbors on the visual perception of English words by non-native speakers, using a new and, as evidenced by the data, highly sensitive bimodal task. Specific support for the assumption that word perception involves coding of inappropriate phonology came from the class of catch trials where an inconsistent printed English word was accompanied by a spoken rime derived from an enemy neighbor. In all experiments, responding to these trials appeared extraordinary difficult, suggesting that printed words were perceived to rhyme with an enemy. The fact that, for example, MOOD can be probed to be perceived as rhyming with BLOOD strongly suggests that perception of an inconsistent word is multistable, involving inadvertent coding of intralingual enemy phonology. This finding is an unequivocal demonstration of the mandatory nature of phonological coding. In terms of the resonance framework, manifold relations between spelling and sound (i.e., phonological ambiguity) imply multistable local orthographic-phonologic resonances that are resolved through successive cycles of cooperative and competitive interactive activation (Van Orden & Goldinger, 1994).

These findings are clearly not in accordance with dual route theory, which emphasizes a non-phonological process for word reading and in which the phonological route to the reading response is of secondary importance. Although its successor, the DRC model of Coltheart and his colleagues (Coltheart et al., 1993; 2001), assigns a more central function to phonological computation, the way it is currently parameterized is not consistent with the fast and mandatory nature of phonological coding. To accommodate the fact that phonological coding plays a primary role in reading, Lukatela, Eaton, Lee, Carello, and Turvey (2002) suggest a change in the DRC model's parameter settings. Specifically, they propose to impose a delay on the start of processing along the lexical route and to assign higher weights to activation along the nonlexical route. Such a modification would have the effect that assembled phonology via the nonlexical route precedes addressed phonology via the lexical route. However, changing the parameter settings of the model this way would imply a rather drastic break with the classic dual route theory, because it effectively rejects the delayed phonology hypothesis (see also Van Orden et al., 1990, for discussion).

Interlingual phonological coding in bilingual word perception

For the second objective of this study, Experiment 3 introduced cross-language Dutch enemy neighbors to investigate the impact of manifold *interlingual* spelling-

to-sound relations, thus seeking support for the hypothesis that L2 word perception in bilinguals involves coding of inappropriate L1 phonology. Once again, such evidence was obtained from the catch trials. A spectacularly large number of false-positive responses occurred: Dutch-English bilinguals frequently pressed the "yes" button when a word like MOOD was presented jointly with the Dutch pronunciation of the printed word's spelling body (which in Dutch sounds like OAD as in ROAD). This finding indicates that perception of an inconsistent word includes coding of cross-language enemy phonology and, therefore, that phonological coding in printed word perception proceeds language non-selectively. In terms of dynamic systems theory, cross-language phonological coding appears to be based on multistable, interlingual spelling-to-sound dynamics. Note however that the present experiments only tested L2 word perception. Plausibly, in L1 word perception by bilinguals inappropriate L2 phonology has a relatively small impact (e.g., Jared & Kroll, 2001, Experiments 1 and 2; Haigh & Jared, 2007).

The strong phonological theory of Frost (1998) offers a coherent account of the primary role of phonology in visual word perception. At the core of this account rests the assumption that phonological assembly is a mandatory process. The present findings extend the available empirical evidence on phonological coding in monolingual and bilingual word reading, yielding unequivocal evidence of the mandatory nature of phonological coding by demonstrating that bilingual word processing may actually initiate *and* preserve enemy phonology arising from knowledge of cross-language spelling-to-sound knowledge. In contrast, coding of cross-language enemy phonology can not be easily explained by the traditional dual-route model. A more detailed discussion of the present results within the dual-route framework is beyond the scope of this paper (cf. Van Wijnendaele & Brysbaert, 2002).

Interlingual phonological coding and language mode

A key experimental factor that may influence the degree in which bilingual word processing proceeds language non-selectively concerns the composition of the stimulus list, that is, whether words from just one or from both languages are presented to the participant. A couple of studies employing the common naming and lexical-decision tasks have shown an effect of this manipulation, with a larger influence of the non-target language when the stimulus set contains a number of stimulus words from this language than when only words from the target language are presented to the participants (e.g., Jared & Kroll, 2001). The reason presumably is that including words from the non-target language boosts this language's activation level. Accordingly, adding Dutch filler trials to the stimulus set in Experiment 3 was expected to increase the activation level of non-target Dutch, resulting in

a larger Trial Type effect in the Dutch-fillers condition than in the English-fillers condition. However, no effect of the stimulus-list manipulation was obtained, suggesting that, unlike naming and lexical decision, the print-to-speech correspondence task is insensitive to this variable.

As explained previously, this apparent task effect can be accounted for by assuming that performance on the print-to-speech correspondence task reflects processes of phonological coding in the initial conditions of word perception, the ballistic nature of which precludes external influences such as an effect of language mode. It follows that phonological coding in bilingual word perception proceeds in a language non-selective fashion (cf. Brysbaert et al., 1999; Dijkstra & Van Hell, 2003), even when the stimulus-list composition favors a monolingual language mode. This suggests that the degree in which bilingual word processing exploits word knowledge of both the languages varies according to the time course of the form-function dynamics, with language non-selective processing occurring predominantly during an early stage of mandatory (local-level) phonological coding.

In conclusion, according to our findings, phonological coding in L2 word perception in bilinguals involves language non-selective processing, which seems unaltered by the language mode the bilingual reader is in. These findings are consistent with a general view on word processing, which holds that phonological assembly is mandatory, whereas the use of lexical knowledge may be subject to strategic control (Drieghe & Brysbaert, 2002; see also Frost, 1998).

Authors' note

Notes

1. In this study, statistical analyses were performed on participants' false-positive error rates. Analyses over item means are not appropriate here, because the stimuli were matched across conditions and the selected word items consisted of a non-random and exhaustive selection from the item population (cf. Dijkstra et al., 1999; Jared & Kroll, 2001; see Clark, 1973; Raaijmakers, 2003; Raaijmakers, Schrijnemakers, & Gremmen, 1999). However, to examine the possibility

that only one or a few items were responsible for the Trial Type effects, an analysis of items was carried out for the catch trial data of Experiment 2. We used the data of this experiment because its design is more basic than that of Experiment 3, and item means are based on the responses of a large number of participants. Repeated-measures ANOVAs with items as the random variable (i.e., F_2 analyses) were performed to examine the statistical significance of the catch > no-match contrast for atypical and typical inconsistent words. If only a few items underlie these Trial Type effects, this would be associated with relatively high error variance and non-significant F-ratios. The catch > no-match contrast involved comparing the item mean of a catch trial against the corresponding item mean of a no-match trial. Both for atypical and typical inconsistent words the Trial Type effect was statistically significant, $F(1, 19) = 19.70$, $MSE = 462.61$, $p < .001$ and $F(1, 19) = 22.35$, $MSE = 118.17$, $p < .001$. For atypical and typical inconsistent words, 17 and 18 out of 20 word pairs, respectively, showed a Trial Type effect. In conclusion, the item analysis shows that most word pairs contribute to the overall Trial Type effects.

2. We realize that our account of the Trial Type effect would be even more convincing if we had tested a control group of monolingual English speakers and found them not to exhibit this effect. After all, one could imagine the effect was caused by inadvertent characteristics of the materials rather than being the result of knowing Dutch. One indication that such an account is implausible is that in Experiments 1 and 2 Dutch-English bilinguals similarly produced many false-positives in catch trials when the sound stimuli were based on English enemy neighbors. Yet, the catch trials' sound stimuli in Experiment 3, based on Dutch enemy neighbors, were constructed according to exactly the same procedures as the catch trials' sound stimuli in Experiments 1 and 2. One could furthermore want to argue that, for one reason or another, in Experiment 3 the sound stimuli used for catch trials were more difficult to distinguish from the phonological codings of the corresponding English target words than the sound stimuli used for no-match trials. A reason could be that the catch trials' sound stimuli happened to sound like English words that match the corresponding visually presented words. Such, however, was clearly not the case. After all, the sound stimuli were derived from recordings of Dutch words, produced according to Dutch spelling-to-sound conversion rules. In conclusion, it appears that neither differences in matching procedures of the stimulus materials in Experiments 1 and 2 on the one hand and Experiment 3 on the other hand, nor inappropriate matching of the stimulus materials within Experiment 3, can provide a plausible alternative account for the large trial-type effect in Experiment 3.

References

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

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

Bosman, A. M. T., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C. A. Perfetti, L. Rieben, & M. Fayol, (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 173–194). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Burnage, G. (1990). CELEX: A guide for users. Nijmegen, The Netherlands: SSN.

Carello, C, Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 211–226). Amsterdam: North-Holland.

Clark, H. H. (1973). The language-as-fixed-effect-fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic Press.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589–608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (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 nonselective access to bilingual memory. *The Quarterly Journal of Experimental Psychology, 53A,* 397–428.

Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (1999). Recognizing 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 Hell, J.G. (2003). Testing the language mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism, 6*, 2–16.

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

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

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.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The size of the cross-lingual masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology, 51,* 1–9.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71–99.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674–691.

Goldinger, S. D., Azuma, T., Abramson, M., & Jain, P. (1997). Open wide and say "blay!" Attentional dynamics of delayed naming. *Journal of Memory and Language, 37,* 190–216.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122–1139.

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Van Orden, G. C. (1999). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 561–574.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Mahwah, NJ: Erlbaum.

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Oxford, England: Basil Blackwell.

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

Holden, J. G. (2002). Fractal characteristics of response time variability. *Ecological Psychology,14*, 53–86.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36,* 505–529.

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., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language*, *29*, 687–715.

Kirk, R. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd.). Pacific Grove, CA: Brooks/Cole.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review, 1*, 476–490.

Lukatela, G., Eaton, T., Lee, C., Carello, C., & Turvey, M. T. (2002). Equal homophonic priming for words and pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 3–21.

MacWhinney, B. (2005) A Unified Model of Language Acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Pacific Grove, CA: Brooks/Cole.

Myers, J. L., & Well, A. D. (1995). *Research design and statistical analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in visual Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 101–118.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 785–794.

Raaijmakers, J. G. W. (2003). A further look at the "Language-as-Fixed-Effect Fallacy". *Canadian Journal of Experimental Psychology, 57*, 141–151.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The Language-as-Fixed-Effect Fallacy": Common misconceptions and alternative solutions. *Journal of Memory and language, 41*, 416–426.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, N. J.: Prentice-Hall.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *Neuroreport, 15,* 1555–1558.

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional reading via the phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 336–339.

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

Van Heuven, W. J. B. (2000). *Visual word recognition in monolingual and bilingual readers: Experiments and computational modelling*. Unpublished doctoral dissertation, University of Nijmegen, The Netherlands.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1269–1291.

Van Orden, G. C., Jansen op de Haar, M. A., & Bosman, A. M. T. (1997). Complex dynamic systems also predict dissociations but they do not reduce to autonomous components. *Cognitive Neuropsychology, 14,* 131–165.

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

Van Wijnendaele, 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.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What's the pronunciation of _OUGH and the spelling for /u/? A database for computing feedforward and feedback inconsistency in English. *Behavior Research Methods, Instruments, and Computers, 29,* 600–618.

Appendix

Printed Word Stimuli Used in Experiments 1–3 (CM = Consistent Spelling-to-Sound Mappings, TM = Typical Spelling-to-Sound Mappings, AT = Atypical Spelling-to-Sound Mappings).

English words without Dutch neighbors			English words with Dutch neighbors		
CM	TM	AM	CM	TM	AM
stain	said	paid	moon	mood	blood
force	tall	shall	she	camp	swamp
world	trash	wash	then	scan	swan
course	catch	watch	bet	bar	war
point	thread	plead	theft	yard	ward
bird	leak	steak	such	barn	warn
girl	dear	bear	hold	gasp	wasp
with	heard	beard	hook	sat	what
fix	feast	breast	rich	where	here
prince	cheat	great	cup	twist	christ
wait	weight	height	them	give	five
cape	grey	key	spent	two	go
teach	load	broad	pact	come	home
ground	north	worth	charge	son	won
grass	both	cloth	walk	none	stone
bluff	through	though	seem	work	fork
lung	mouth	youth	sane	nose	lose
spy	brown	known	shift	ghost	lost
peace	pull	dull	nude	hour	four
stick	brush	bush	jerk	love	move

Author's address

Martin van Leerdam
University of Amsterdam
Department of Psychology
Psychonomy Section
Roetersstraat 15
1018 WB Amsterdam
The Netherlands

m.van.leerdam@planet.nl