Language and Cognitive Processes

ISSN: 0169-0965 (Print) 1464-0732 (Online) Journal homepage: https://www.tandfonline.com/loi/plcp20

The locus of the orthographic consistency effect in auditory word recognition

Paulo Ventura , José Morais , Chotiga Pattamadilok & Régine Kolinsky

To cite this article: Paulo Ventura , José Morais , Chotiga Pattamadilok & Régine Kolinsky (2004) The locus of the orthographic consistency effect in auditory word recognition, Language and Cognitive Processes, 19:1, 57-95, DOI: 10.1080/01690960344000134

To link to this article: https://doi.org/10.1080/01690960344000134

Published online: 03 Jun 2010.

LANGUAGE AND COGNITIVE PROCESSES, 2004, *19* (1), 57–95

The locus of the orthographic consistency effect in auditory word recognition

Paulo Ventura
Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Portugal

José Morais
Research Unit in Cognitive Neurosciences (UNeSCog), Université Libre de Bruxelles, Belgium

Chotiga Pattamadilok and Régine Kolinsky
Research Unit in Cognitive Neurosciences (UNeSCog), Université Libre de Bruxelles, Belgium
Fonds National de la Recherche Scientifique, Belgium

In Experiments 1–2, we replicated with two different Portuguese materials the consistency effect observed for French by Ziegler and Ferrand (1998). Words with rimes that can be spelled in two different ways (inconsistent) produced longer auditory lexical decision latencies and more errors than did consistent words. In Experiment 3, which used shadowing, no effect of orthographic consistency was found. This task difference could reflect the confinement of orthographic influences to either decisional or lexical

Correspondence should be addressed to Paulo Ventura, Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa, Alameda da Universidade, 1649–013 Lisboa, Portugal. Email: paulo.ventura@fpce.ul.pt

The authors wish to express their gratitude to Kathleen Rastle, Johannes Ziegler, and Dennis Norris who made very stimulating comments on a previous version of this paper. Preparation of this article was partly supported by the Centro de Psicologia Clínica e Experimental—Desenvolvimento, Cognição e Personalidade of the Universidade de Lisboa, Portugal and by the Human Frontier Science Program RGP 53/2002 "An interdisciplinary approach to the problem of language and music specificity". The third and last authors are Research Fellow and Research Associate, respectively, of the Belgian Fonds National de la Recherche Scientifique (F.N.R.S.). We would like to thank Mireille Cluytens for helpful assistance during material recording, and are grateful to Sandra Fernandes and José Querido for their help in running the experiments.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960344000134

processes. In Experiment 4, we tried to untangle these two interpretations by comparing two situations in which a shadowing response was made contingent upon either a lexical or a phonemic criterion. A significant effect of orthographic consistency was observed only in lexically contingent shadowing. We thus argue that lexical but not sublexical processes are affected by orthographic consistency.

INTRODUCTION

Models of spoken word recognition typically assume that speech is processed without any modulation by orthographic knowledge. This is the case of NAM (Luce, Pisoni, & Goldinger, 1990), Cohort (Marslen-Wilson, 1987), TRACE (McClelland & Elman, 1986), ShortList (Norris, 1994), and Merge (Norris, McQueen, & Cutler, 2000). Yet, several observations suggest that orthographic knowledge influences speech processing.

The first striking demonstration of the intervention of orthographic knowledge on a phonological task involved rhyme judgement. Seidenberg and Tanenhaus (1979) had listeners monitor a short list of spoken words for a word that rhymed with a prior cue word. Rhyme decisions were faster when the cue and rhyming words were orthographically similar (e.g., PIE-TIE) than when they were not (RYE-TIE). This response pattern was still observed when rhyme production frequencies were equated for orthographic similar and dissimilar rhymes (Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981). Seidenberg and Tanenhaus (1979) also showed that the time taken to decide that two spoken words rhyme is shorter when their spellings are similar (TOAST-ROAST) than when they are not (TOAST-GHOST), and that the opposite effect holds for negative decisions (which are faster for LEAF-REF than for LEAF-DEAF). In a similar vein, in auditory syllable monitoring, detection times depend on the spelling similarity between the syllable target and the carrier word (Taft & Hambly, 1985).

Since this work, evidence gathered by other authors also suggested that spoken word recognition is influenced by orthography. It is important to distinguish, in the literature, situations in which orthographic information is explicitly provided from those in which it is not.

Printed redundant cues (Frost, 1991; Frost & Katz, 1989; Frost, Repp, & Katz, 1988) or primes (Dijkstra, Frauenfelder, & Schreuder, 1993) facilitate spoken language recognition. Conversely, encoding a spoken word prime (e.g., BEAD) affects colour naming of orthographically similar visual words (BREAD), in comparison to both unrelated (SKILL) and phonetic control (FEED) primes (Tanenhaus, Flanigan, & Seidenberg, 1980). It has also been shown that searching for a phoneme target like /v/ is quite difficult when the written carrier is incongruent with the target's spelling (as in the word OF, cf. Schneider & Healy, 1993).

However, it is doubtful that these cross-modal situations reflect normally activated spoken word processing routes. In fact, Kouider and Dupoux (2001) observed cross-modal repetition priming effects only when the masked visual prime (spelling the auditory target) was consciously perceived by the participants. When prime duration was sufficiently short so as to avoid the involvement of conscious strategies, no cross-modal priming effect emerged, while within-modal (visual) priming was observed whatever the prime duration.

Nonetheless, orthographic influences can be observed in situations in which orthographic information is not presented explicitly. Within these purely auditory studies, one should further distinguish between those examining metalinguistic abilities and those examining speech recognition processes.

Indeed, alphabetic knowledge radically alters the way listeners consciously segment speech, presumably by providing them with an abstract model for analysing speech sounds (Olson, 1996). The study of both illiterate adults (Morais, Bertelson, Cary, & Alegria, 1986; Morais, Cary, Alegria, & Bertelson, 1979) and non-alphabetic readers (de Gelder, Vroomen, & Bertelson, 1993; Read, Zhang, Nie, & Ding, 1986) clearly illustrates this point. In fact, both populations are unable to deal explicitly with phonemes in tasks like phoneme deletion or addition.

Several developmental metalinguistic studies have shown a bidirectional coupling between orthography and phonology rapidly established from the onset of written language acquisition. For instance, children count an additional phoneme in [pitʃ] compared to [ritʃ] due to the number of letters in the corresponding spellings (Ehri & Wilce, 1979). Orthographic strategies were also observed in Japanese children in various metaphonological tasks (Mann, 1986; Spagnoletti, Morais, Alegria, & Dominicy, 1989). The influence of letter names both on children's ability to connect print and speech (Treiman, Tincoff, & Richmond-Welty, 1996) and on sound judgements (Treiman & Cassar, 1997) provides another example of orthographic influence. As a consequence of this coupling between phonology and orthography, the impact of orthographic knowledge on metalinguistic performance increases as children become more fluent readers/writers (Ehri & Wilce, 1980; Treiman & Cassar, 1997; Zecker, 1991).

Hence, not surprisingly, similar orthographic effects are observed in adult readers' metalinguistic behaviour. Adults are for example influenced by the number of letters in the word spelling (Moats, 1994), by the precise manner in which their alphabet represents phonology (Ben-Dror, Frost, & Bentin, 1995), and by the orthographic representations of their non-alphabetic first acquired writing system (Nakamura, Kolinsky, Spagnoletti, & Morais, 1998).

Recently, Ventura, Kolinsky, Brito-Mendes, and Morais (2001) showed that the explicit and intentional operations called forth by blending two CVC monosyllabic words (cf. Treiman, 1983, 1986) into a new CVC word are also influenced by orthography. Participants preferred the C/VC blend (onset vs. rime[1]) for the words spelled with a final consonant (orthographic monosyllables, e.g., BAR, ''bar''-MEL, ''honey''), but the CV/C blend (onset plus peak vs. coda) for the words spelled with a final mute ''e'' (orthographic disyllables, e.g., CURE, ''heal''-PELE, ''skin''). This response pattern was observed even when instructions emphasised the importance of the stimuli ''sounds'' and when rimes were identical (e.g., PAR –''pair'' vs. PARE –''stop''). There was no substantial reduction of the orthographic influence when crossing the acoustic portion of the rime between stimuli allowed control of possible acoustic-phonetic effects. The whole pattern of results shows that phonological judgements about the structure of the syllable are shaped by the links between phonology and orthography.

The findings described so far were obtained with metalinguistic tasks that reflect rather late, post-perceptual processes. Some studies tried to explore the influence of orthography through on-line measures of word recognition.

Jakimik, Cole, and Rudnicky (1985; cf. also Slowiaczek, Soltano, Wieting, & Bishop, 2003; Experiment 1) examined whether lexical decisions about spoken words are influenced by the spelling of an immediately preceding auditory prime. They found a facilitatory effect for monosyllabic word targets preceded by bisyllabic word primes when the common syllable was both phonologically and orthographically similar (e.g., MESSAGE-MESS), but not when words were only related by either sound (DEFINITE-DEAF) or spelling (LEGISLATE-LEG). Yet, the authors acknowledged that their results probably reflect an orthographic strategy rather than usual recognition processes. As a matter of fact, polysyllabic words were followed by repeated first syllables 30% to 50% of the time, and some of the participants claimed to have tried to anticipate items deliberately.

Slowiaczek et al. (2003) tested paired primes and targets with the same number of syllables to reduce the tendency to use the relationship between polysyllabic primes and monosyllabic targets. Participants engaged in either a lexical decision (Experiments 3 and 4) or a shadowing (Experiments 2, 5, and 6) task with a single-trial (Experiments 2, 3, 5, and 6) or (as Jakimik et al., 1985) subsequent-trial (Experiment 4) priming procedure. In Experiments 2 to 4, the authors found a facilitatory effect

[1] Throughout the paper, we use the term *rime* to refer to the phonological rime and the term *body* to refer to the orthographic rime.

when the initial syllables shared orthography (e.g., RATIO-RATIFY), but not when they shared phonology alone (e.g., NUISANCE-NOODLE) or both phonology and orthography (e.g., FUNNEL-FUNNY). However, as acknowledged by the authors, in these experiments, primes and targets in the orthography condition shared not only graphemes (3.35, on average) but also phonemes (1.67, on average). In addition, the high proportion of related trials (50%) may have fostered strategic processing. Unfortunately, Slowiaczek et al. (2003) dealt with these two problems separately. In Experiment 5, in which they observed an influence of the number of shared graphemes (one or three) between primes and targets in a situation where the number of shared phonemes was held constant at one, the proportion of related trials was maintained at 50%. In Experiment 6, in which they found a facilitatory effect both when the initial syllables shared orthography alone and when they shared phonology and orthography, the proportion of related trials was reduced to 20%, but the material was the same as the one of Experiments 2 to 4. Thus, the way Slowiaczek et al. chose to deal with the two above mentioned problems makes the interpretation of their results difficult.

Other studies showed an influence of orthography at the phoneme level. In a French phoneme monitoring study, Frauenfelder, Segui, and Dijkstra (1990) observed slower detection times for the phoneme /k/ than for /p/ and /t/ in words compared with pseudo-words. The authors attributed this effect to the variety of orthographic representations for /k/ as opposed to /p/ or /t/ (in French, the phoneme /k/ can be spelled in many different ways: ‘‘c’’, ‘‘cc’’, ‘‘ch’’, ‘‘cch’’, ‘‘k’’, ‘‘ck’’, ‘‘cu’’, ‘‘q’’, ‘‘qu’’, ‘‘cqu’’). A similar consistency effect was observed by Dijkstra, Roelofs, and Fieuws (1995) in a Dutch study where participants monitored for phonemes (e.g., /k/) with either a primary spelling (‘‘k’’, as in PAPRIKA) or a secondary spelling (‘‘c’’, as in REPLICA).

Yet, phoneme monitoring probably involves a strong metalinguistic component. Indeed, it usually yields very low scores in illiterate adults. For example, in Morais et al. (1986), illiterates performed at only 37% correct when they had to detect the phoneme /k/. Importantly, however, illiterate adults are not different from literate ones in all phonemic-based tasks or phenomena: no effect of literacy was observed for categorical perception or audio-visual linguistic integration nor for the occurrence of feature blending or migration errors (Castro, 1993; Kolinsky & Morais, 1993; Morais, Castro, Scliar-Cabral, Kolinsky, & Content, 1987; Morais & Kolinsky, 1994, 1995; see also Castro, Vicente, Morais, Kolinsky, & Cluytens, 1995). The two last phenomena provide evidence for automatic extraction of intra-syllabic phonetic information by illiterates, despite the fact that these people are unable to analyze explicitly speech into phonemes. Besides, illiterate adults are perfectly able to discriminate minimal pairs (Adrián, Alegria, & Morais, 1995; Scliar-Cabral, Morais, Nepomuceno, & Kolinsky, 1997). This is actually not surprising, since even babies demonstrate consonant discrimination abilities (Eimas, Siqueland, Jusczyk, & Vigorito, 1971).

There is thus a clear dissociation between early, perceptual, spoken word recognition processes and late, post-perceptual, metalinguistic processes in terms of the presence or absence of literacy effects (cf. Morais & Kolinsky, 2001, for an overview). Hence, the fact that illiterate adults are quite poor at phoneme monitoring strongly suggests a metalinguistic component in this task.

Without attributing the orthographic effect they obtained in phoneme monitoring to metalinguistic representations, Dijkstra et al. (1995) considered that it was mediated by lexical access. As a matter of fact, this effect was greater for words for which lexical access had presumably taken place because their uniqueness point (henceforth, *UP*[2]) preceded the /k/ (as in PAPRIKA vs. REPLICA), and was not clear for other cases (e.g., KABOUTER vs. CABARET).

Hallé, Chéreau, and Segui (2000) provided another demonstration of the influence of orthography at the phoneme level, and interpreted their results as emerging from the lexical access mechanism itself. The authors used words in which voice assimilation alters the pronunciation of a consonant letter. For instance, the French word ABSURDE is pronounced /apsyrd/, although the canonical pronunciation of ''b'' is /b/. In a phoneme detection task, French listeners exhibited a strong tendency to detect /b/ rather than /p/ in such words. In another experiment, Hallé et al. used a phonemic gating task. The first gate presented the initial portion of the word until the point of maximum spectral stability in the vowel preceding the labial stop. There were a total of eight gates, whose duration increased by 40-ms steps. Up to gate 5, /p/ responses outnumbered /b/ responses. From gate 6 onward the pattern of responses gradually reversed until /b/ responses became dominant at gate 8. This led the authors to conclude that the orthographic code causing the interference effect emerged from cohort activation mechanisms. For example, before sufficient acoustic information is presented to guess ABSURDE but as soon as /s/ is heard, the cohort of words would reduce to words that all share the ''b'' spelling for /p/.

There are however several problems with this interpretation. First, as Hallé et al. pointed out, there are more words in French beginning with ''abs'' than with ''abt''. Given this difference, the cohort activation account predicts that reporting /b/ should be more likely for the words with ''s'' than with ''t''. No such difference was found. Second, Hallé and colleagues

[2] UP is defined as the point from word onset at which only one candidate remains in the cohort of possible words (e.g., Marslen-Wilson & Welsh, 1978)

acknowledged that orthographic and morphophonemic codes were almost perfectly congruent in their materials. For example, the word ABSURDE also has an underlying morphophonemic code /absyrd/ because it contains the prefix {ab-}. Thus, for most of the stimuli used in the Hallé et al. study, the morphophonemic code could also have interfered with the phonological code. In addition, even if one admits that the interference effect was orthographic, a cohort mediated mechanism can hardly account for orthographic effects occurring at the end of words and involving higher grain size (rime-body) units.

Such effects have been observed by Ziegler and Ferrand (1998), who reported the strongest evidence so far available of an impact of orthographic knowledge on spoken word recognition. Indeed, their manipulation did not require the target's competing spellings or pronunciations to be introduced before or at the same time as the target word, nor did the target word need to be matched to its competitor. These authors found that inconsistent French words with rimes that can be spelled in multiple ways (e.g., the French word NOM is inconsistent because its rime could be spelled, -ON, -ONC, -ONG, etc.) produced longer auditory lexical decision latencies and more errors than did words with rimes that are spelled only one way.

One potential problem with Ziegler and Ferrand's (1998) study is that it relied on the comparison between different words. It is thus possible that the inconsistent words used in their study were not only more ambiguous in terms of having multiple spellings but also in terms of their phonetic composition. In fact, most consistent items had a CVC structure (e.g., DUNE /dyn/) while some inconsistent words had an open CV structure with a more complex nasal sound (e.g., DAIM /dE$/). However, a new auditory lexical decision study (Ziegler, Ferrand, & Montant, unpublished data) excludes the possibility that performance differences are only due to phonetic complexity. It showed performance differences between French words that have exactly the same rime (e.g., /E$/, as in NAIN and DAIM) but differ in the probability with which their phonology maps into their spelling. Words with the dominant spelling (AIN) led to better performance than words with subdominant spelling (AIM), even if performance on consistent words was better than performance on either of these two inconsistent word types. Since both types of inconsistent words had the same rime, the spelling probability effect, and by extension most probably the consistency effect, must reflect the activation of orthographic information during auditory word recognition.

Ziegler and Ferrand (1998) interpreted their consistency effect assuming a bi-directional flow of activation, namely a bi-directional coupling between orthography and phonology that would be functional in both visual and auditory word recognition. According to them, phonology is automatically activated during reading, and orthography is automatically activated during speech processing. This idea is clearly stated in the title of their article, according to which "orthography shapes the *perception* of speech" (our italics). Thus, once reading is acquired, the influence of orthography would pervade the whole speech recognition system. In particular, Ziegler and Ferrand (1998) refer to models in which this coupling between orthography and phonology does not occur, or not only at the word level, but also between various grain-size *sublexical* units (e.g., Stone, Vanhoy, & Van Orden, 1997; Van Orden & Goldinger, 1994; Van Orden, Jansen op de Haar, & Bosman, 1997; Ziegler, Montant, & Jacobs, 1997), for example between phonemes and letters (or graphemes), and/or between bodies and rimes.

We share with Ziegler and Ferrand (1998) the idea of a coupling between orthography and phonology. Like them, we believe that orthographic influences on speech processing deserve close attention. However, we suggested that such influences are not pervasive but, on the contrary, restricted to some aspects of the recognition process (Ventura et al., 2001). In particular, the consistency effect might stem from post-access, lexical or decisional components. Indeed, the lexical decision situation used by Ziegler and Ferrand (1998) does not only imply some lexically based analysis[3] but, since it includes a decisional component, is also often considered to involve substantial post-lexical processing (e.g., Goldinger, 1996; Jakimik et al., 1985; Slowiaczek & Hamburger, 1992). Thus, the consistency effect observed in lexical decision could stem either from lexical competition between the several orthographic word representations activated by an inconsistently spelled rime, or from interference at the time the decision has to be taken.

The aim of the present work was to examine these two hypotheses and thus to determine the locus of the consistency effect in speech processing.

More precisely, one of the interpretations of the locus of the consistency effect is that it occurs relatively late in the processing route. This interpretation relies on a broad distinction between two successive stages of processing, perceptual and post-perceptual (see Kolinsky, 1998; Morais & Kolinsky, 1994; Morais, Kolinsky, Ventura, & Cluytens, 1997; Ventura et al., 2001). At the first stage, operations of segmentation of the acoustic representation of the speech signal are carried out in a modular way (in Fodor's 1983 sense of modularity), hence independently of literacy-induced knowledge. The phonological output of that stage does not

[3] This claim does not imply that full lexical access is always necessary to respond in a lexical decision task. Several authors actually agree on the notion that a familiarity mechanism might allow fast non-lexically based decisions, without unique item identification (e.g., Balota & Chumbley, 1984).

however reach consciousness without transformation: it is re-elaborated by attentional processes at a second, post-perceptual, stage, the output of which is conscious recognition. Under that view, the influence of orthography is confined to late processes taking place at the post-perceptual stage or further on in the processing route, probably at the decisional stage.

Nevertheless, the idea of distinct stages or levels of processing has a long history in the field of spoken word recognition, and a number of authors made a critical distinction not so much in terms of two successive (early vs. late) stages of processing but rather in terms of lexical versus sublexical representations. The lexical level includes representations that correspond to words and is often characterised as consisting of competitive processes among these lexical representations (e.g., Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994). The sublexical level includes independent representational entities corresponding to components of spoken words or pseudo-words. Although sublexical and lexical representations may become activated at similar time scales, there is evidence that these two levels of representation have dissociable effects on processing (e.g., Vitevitch & Luce, 1998; 1999; see also Luce & Large, 2001).

Thus, another interpretation of the locus of the consistency effect observed by Ziegler and Ferrand (1998) is that it only occurs at the lexical level. This view seems compatible with the fact that no consistency effect was observed for pseudo-words in their study. Yet, one could still argue, as these authors did, that pseudo-words are rejected by a "time out" mechanism when no single word unit has passed the activation threshold before a certain time threshold. If it were the case, the absence of a consistency effect for pseudo-words would not counter their view of ubiquitous coupling between orthography and phonology.

The present paper includes four experiments aimed at contrasting more directly Ziegler and Ferrand's view with the hypothesis that the influence of orthography is confined to late (probably decisional) and/or lexical processes.

The first step in our study was to try to replicate in Portuguese the consistency effect initially reported by Ziegler and Ferrand (1998) for French. This was done to ensure that the effect could be disclosed in Portuguese. Indeed, given that the Portuguese orthographic system is shallower than the French system, the degree of rime-body inconsistency is weaker in Portuguese than in French. In Experiment 1, participants were auditorily presented with consistent and inconsistent words and pseudo-words, and instructed to decide as quickly as possible whether each stimulus was or not a real Portuguese word.

The second step in our study was to evaluate the hypothesis that the consistency effect reflects a general mechanism that pervades all stages of speech processing. If it were the case, we would expect consistency to influence performance not only in lexical decision, but also in other on-line auditory tasks. In Experiments 2 and 3, we thus compared the consistency effect in lexical decision and in shadowing. In this task, the listener has to repeat as rapidly as possible any spoken sequence, be it a word or a pseudo-word. We used in both tasks a new material that was specifically designed to control for phonetic biases linked to a vocal response.

Theoretically, a shadowing response requires only a precise analysis of the phonetic properties of the stimulus word (or pseudo-word) to build an articulatory plan, and does not rely on any binary choice decision. As such, it is considered as biasing participants towards the sublexical level and/or as requiring less decisional processes than lexical decision (e.g., Balota & Chumbley, 1984; Jakimik, Cole, & Rudnicky, 1985; Norris, 1986; Vitevitch & Luce, 1999).

This does not imply however that shadowing necessarily bypasses lexical access. Indeed, shadowing latencies are somewhat affected by word frequency (Radeau & Morais, 1990), by neighbourhood density (Luce & Pisoni, 1998; see also Vitevitch, 2002 for the effect of onset density) and by the location of the UP (Radeau & Morais, 1990), but to a lesser extent than lexical decision.

For example, a direct task comparison found frequency effects to be significantly stronger as well as more reliable in lexical decision than in shadowing (Radeau, Morais, & Segui, 1995). The reliability of frequency effects in shadowing is also questioned by Connine, Mullennix, Shernoff, and Yelen's (1990) observation of a familiarity but not of a word frequency effect. Thus, lexical access does not seem as mandatory in shadowing as in lexical decision. Coherent with this view, effects of word frequency and of lexical status were observed only in distant shadowers, not in close ones (Marslen-Wilson, 1985). As regards UP effects, direct comparison between shadowing and lexical decision is not possible. Indeed, lexical decision is inappropriate for testing UP effects, since a word can in principle become a non-word at any moment after the UP (e.g., by modification of one phoneme), so that the listener is supposed to wait for the stimulus offset before deciding. Instead, Radeau, Mousty, and Bertelson (1989) used another binary choice decision task, namely gender decision, on the same set of stimuli with early or late UPs used by Radeau and Morais (1990). They observed strong UP effects representing 30% to 53% of the time interval between word onset and UP, while the effect observed in shadowing was smaller (8%) and statistically unreliable (Radeau & Morais, 1990). This task difference might be linked to a different involvement of the lexicon, or of strategic processes. As a matter of fact, UP effects involve a strategic component, as shown both by blocking effects (the 53% effect mentioned above was obtained with blocked presentation; with mixed presentation, Radeau et al., 1989, reported effects ranging from 30% to 37%) and by the fact that UP effects are only observed with rather slow speaking rates and not at normal ones (Radeau, Morais, Mousty, & Bertelson, 2000).

The whole set of results is thus consistent with the view that late, strategic processes and/or lexical representations are less involved in shadowing than in lexical decision[4]. If the consistency effect observed by Ziegler and Ferrand (1998) reflected a general mechanism, orthographic consistency would be expected to influence lexical decision and shadowing performances to a similar extent. On the contrary, if the influence of orthography were restricted to later and/or lexical stages of spoken word recognition, then orthographic consistency would be expected to affect lexical decision far more than shadowing. In addition, we also predicted that the effect of consistency would be observed with words, but not with pseudo-words, even in the shadowing task.

The third and last step in our study was to evaluate whether the consistency effect stems from lexical or decisional components. For that purpose, in Experiment 4 we compared two shadowing situations in which the shadowing response was made contingent upon either a lexical or a phonemic criterion.

EXPERIMENT 1: DOES ORTHOGRAPHIC CONSISTENCY AFFECT LEXICAL DECISION PERFORMANCE IN PORTUGUESE?

As already mentioned, the degree of rime-body inconsistency is weaker in Portuguese than in French. It was thus worth trying to replicate in Portuguese the consistency effect initially reported by Ziegler and Ferrand (1998) for French.

Using a lexical decision task, we presented participants with a material similar to the one used by Ventura et al. (2001). *Consistent words* included *consistent rimes* that can be spelled in only one way, namely either with or without a final mute "e". For example, in Portuguese the rime /av/ can only be spelled with a final mute "e" (e.g., CAVE, "basement") and the rime /er/ can only be spelled with a final C (e.g., LER, "read"). In contrast,

[4] Note that an opposite opinion about the requirements of the shadowing task has been stated by Slowiaczek (1994) and by Liu, Bates, Powell, and Wulfeck (1997), who claimed that shadowing a word can only be done with full lexical analysis. These authors based their assumption on the observation of semantic priming effects in shadowing. However, task requirements are not easily inferred on the basis of priming: the semantic links present in the material can induce specific strategies which are not necessarily at work when no such link is introduced.

inconsistent words included *inconsistent rimes* that can be spelled in two different ways. For example, the rime /ɛl/ can be spelled either with a final mute ‘‘e’’ (e.g., PELE, ‘‘skin’’) or with a final C (e.g., MEL, ‘‘honey’’).

Method

Participants. In all the experiments of this study, participants were undergraduates from the University of Lisbon and obtained an extra credit in an introductory psychology course for their participation. All were native speakers of Portuguese and none reported an hearing impairment. In Experiment 1, they were 21, aged 18–34 years (median: 19 years).

Stimuli. The stimulus set consisted of 40 monosyllabic CVC Portuguese words and 40 monosyllabic CVC pseudo-words listed in Appendix 1. Half of the words were consistent and the other half were inconsistent. Ten different rimes were used in the consistent set (/av/, /ɛd/, /ɛm/, /er/, /ɛv/, /ɔd/, /ɔm/, /or/, /ɔt/, /um/) and eight different rimes were used in the inconsistent set (/al/, /ar/, /ɛl/, /ɛr/, /il/, /ir/, /ɔl/, /ul/). The pseudo-word set was constructed by replacing the initial phoneme of each word. Thus each pseudo-word ended with the same rime of either a consistent or an inconsistent word (henceforth, consistent vs. inconsistent pseudo-words, respectively). All pseudo-words were pronounceable.

Characteristics of the words appear in Table 1. As can be seen, since all words were CVC monosyllables and had a UP of 3 phonemes, they all became discriminable when the final consonant was heard.

As already mentioned, rimes of inconsistent words had always two possible spellings (either CVC or CVCV). Each word was also selected on the basis of its consistency ratio (CR), which is the sum of words with the same rime and body relative to the sum of words with the same rime. The CR thus reflects the degree of rime/body (in)consistency, and varies between 0 and 1. It is, by definition, 1 for consistent words. A CR of 0.5 indicates an equal number of words with each spelling (CVC and CVCV)

TABLE 1
Characteristics of the words used in Experiment 1

Variables	*Consistent*	*Inconsistent*
Length (No. phonemes)	3	3
Uniqueness point	3	3
No. rime spellings	1	2
Consistency ratio	1	0.39
No. phonological neighbours	9.5	12.35
Frequency (5-point scale)	3.7	3.56
Mean duration (in ms)	621	586
(minimum–maximum)	(439–828)	(375–730)

while a CR greater than 0.5 indicates a larger number of words with the same spelling, and a CR smaller than 0.5 indicates a larger number of words with different spelling, as was the case, on average, for inconsistent words.

Pseudo-word selection was also done on the basis of a statistical analysis of consistency. An independent sample of 24 participants was asked to write down 107 pronounceable pseudo-words constructed by adding an initial phoneme to the 18 different rimes used in our study. They performed the task individually in a quiet room, pseudo-words being presented via the loudspeakers of a computer. The order of the pseudo-words for each participant was randomised. For each pseudo-word we first calculated the ratio of the sum of CVCV spellings relative to the sum of both CVC and CVCV spellings. We subtracted this ratio from 0.5 and the absolute value of the result was used as a CR for pseudo-words, which varies between 0 (maximum inconsistency) and 0.5 (maximum consistency). For inconsistent and consistent pseudo-words, mean CR was 0.09 and 0.35, respectively, $t(38) = 10.84$, $p < .0001$.

A standard objective frequency count does not exist in Portuguese, but large correlations have often been reported between subjective and objective (e.g., Kucera & Francis, 1967; Thorndike & Lorge, 1944) word frequency (e.g., Balota, Pilotti, & Cortese, 2001; Carroll, 1971; Galbraith & Underwood, 1973; Gordon, 1985; Segui, Mehler, Frauenfelder, & Morton, 1982; Shapiro, 1969; Tryk, 1968; Whalen & Zsíga, 1994). Hence, subjective frequency was estimated by an independent group of 70 participants who rated 250 monosyllabic CVC words on a 5-point scale, ranging from least (1) to most (5) frequent. The consistent and inconsistent words used did not differ significantly on subjective frequency, $t(38) = 0.52$, $p > .10$.

Durations of the words and pseudo-words were measured for each condition. Mean duration was slightly shorter for inconsistent than for consistent words, but not significantly so, $t(38) = 1.01$, $p > .10$. It was 677 ms (range: 455–881) for consistent pseudo-words and 615 ms (range: 412–768) for inconsistent pseudo-words. Although the last difference fails to reach the conventional level of statistical significance, $t(38) = 1.81$, $p = .08$, we used ANCOVAs in the item analyses reported below, with pseudo-word duration as covariate.

There was a significant difference in the number of phonological neighbours—defined as the number of words that can be obtained by changing one phoneme—between consistent and inconsistent words, $t(38) = 3.22$, $p < .005$. Because of this difference, item analyses contrasting performance to consistent and inconsistent words were performed using neighbourhood size as covariate.

All items were recorded in a soundproof room on a digital audio tape recorder (Sony DTC-55ES) by a male Portuguese native speaker. Stimuli were then transferred to a computer at a sampling rate of 22.05 kHz and 16-bit conversion using the Cool Edit sound editor. The inter-trial interval was 2 s.

Procedure. Participants were tested individually in a sound-attenuated room. The order of stimulus presentation was randomised. Presentation and timing were controlled by E-prime 1.0 (Schneider, Eschman, & Zuccolotto, 2002a, 2002b) running on a 366 MHz Celeron PC. Stimuli were presented auditorily at a comfortable level through Sennheiser HD 600 headphones. Participants were required to perform a lexical decision task on each stimulus by pressing as quickly and accurately as possible one of two response buttons of a button box (PST SRB 200A) that was placed in front of them. Both right-handed subjects and left-handed ones (two participants) were required to respond with their preferred hand. The experimental session was preceded by eight practice trials. During the experimental session, no feedback was provided.

Results and discussion

Reaction times (RTs) were measured from stimulus onset to response onset. Incorrect responses were removed and analysed separately. RTs longer than 1670 ms or shorter than 355 ms (2.5% of the data) were removed from the RT analysis. This criterion was based on the examination of the overall distribution of RTs for correct responses (mean = 1011 ms, SD = 263) and was set close to mean ± 2.5 SD.

Mean RTs and error rates for consistent and inconsistent words and pseudo-words are presented in Table 2. Planned comparisons were performed on correct mean RTs and error rates averaged by subjects (t_1) and by items (t_2).

Latency analyses revealed that participants were faster by 52 ms, on average, on lexical decisions to consistent than to inconsistent words, $t_1(20)$

TABLE 2
Mean reaction times (ms), standard deviation for correct responses and percentage of errors for consistent and inconsistent words and pseudo-words (Experiment 1: Lexical decision)

		Consistent	*Inconsistent*
Words	RT (ms)	882	934
	SD	94	96
	Errors (%)	9.5	19.1
Pseudo-words	RT (ms)	1090	1058
	SD	163	141
	Errors (%)	11.5	11.2

$= 3.91, p < .001$; $t_2(38) = 2.09, p < .05$. On the contrary, they were faster by 32 ms, on average, on lexical decisions to inconsistent than to consistent pseudo-words, but this difference, as we will see below, could be accounted for by covariate effects and is not statistically significant in the item analysis, $t_1(20) = 3.43, p < .005$; $t_2(38) = 1.0, p > .10$.

Error rate analyses show that participants made fewer errors (by 9.6%, on average) on lexical decisions to consistent than to inconsistent words, although this difference falls short of statistical significance in the item analysis, $t_1(20) = 4.6, p < .0005$; $t_2(38) = 1.86, p = .07$. Error rates did not vary significantly between consistent and inconsistent pseudo-words, $t_1(20) = 0.11, p > .10$; $t_2(38) = 0.01, p > .10$.

In our word stimulus set, inconsistent words had more neighbours than consistent words and tended to be slightly shorter (by 35 ms). Because these factors could affect lexical decision performance, separate covariate analyses on the item data were performed. With either number of neighbours or word duration as covariate, the consistency effect for words was still reliable when considering latencies, $F(1, 37) = 7.31, p < .01$, $MSe = 13,562.7$; $F(1, 37) = 12.38, p < .001$, $MSe = 8471.2$, respectively, while for errors it still falls short of statistical significance, $F(1, 37) = 3.69$, $p = .06$, $MSe = 0.03$; $F(1, 37) = 3.62$, $p = .06$, $MSe = 0.03$. In addition, since inconsistent pseudo-words were significantly shorter (by 62 ms) than consistent pseudo-words, analyses on the item data were also performed using duration as covariate. These show that, for pseudo-words, there was no reliable effect of orthographic consistency on either latencies or errors (both Fs < 1).

In sum, we replicated the consistency effect in auditory lexical decision initially reported by Ziegler and Ferrand (1998) for French. Inconsistent Portuguese words produced longer auditory lexical decision times and more errors than did words with rimes that are spelled only one way. This result is particularly striking when we consider the degree of rime-body inconsistency in the two languages. Indeed, in our material inconsistent words had rimes that can be spelled in only two ways, while Ziegler and Ferrand's ones had rimes that can be spelled in multiple ways (about six spellings, on average). The consistency effect thus appears to be quite robust.

In the next two experiments, we evaluated the hypothesis that the consistency effect reflects a general mechanism that pervades all stages of spoken word recognition. As discussed in the Introduction, if it were the case, we would expect consistency to influence performance not only in lexical decision, but also in other on-line auditory tasks like shadowing. In Experiments 2 and 3, we thus compared the consistency effect in these two tasks. A new material was presented to different groups of participants who were instructed, in Experiment 2, to decide as quickly as possible whether the auditorily presented stimulus was or was not a real Portuguese word and, in Experiment 3, to pronounce each item as rapidly as possible.

EXPERIMENT 2: LEXICAL DECISION-MATCHED MATERIAL

In Ziegler and Ferrand's (1998) study, the stimulus categories were not matched for initial phonemes. Such a matching is, in principle, not necessary for a lexical decision task. In order to keep the Portuguese replication as close as possible to the original French study, it was also the case for the material in Experiment 1.

Since the next step was to make a direct comparison of the influence of orthographic consistency in lexical decision (this experiment) and in shadowing (Experiment 3), we used in both experiments a new material specifically designed to control for phonetic biases linked to the vocal response required in Experiment 3. Current research suggests that the first two phonemes need to be controlled across conditions (Kessler, Treiman, & Mullennix, 2002; Rastle & Davis, 2002). We thus matched the two initial phonemes across the experimental manipulations. Care was also taken to match stimulus duration across conditions more adequately than in Experiment 1. We also considered several types of inconsistent words, as described below.

Method

Participants. A fresh group of 46 students, aged 18–34 years (median: 19 years) was tested.

Stimuli. The stimulus set consisted of 40 monosyllabic CVC Portuguese words and 40 monosyllabic CVC pseudo-words. Half of the words were orthographically consistent and the other half inconsistent. Consistent words included rimes that can be spelled in only one way. In contrast, inconsistent words included rimes that can be spelled in two different ways. For half of the inconsistent words, the inconsistency was the same as the one used in Experiment 1, i.e., the VC rime can be spelled either with a final mute ''e'' or with a final C. For the other half of the inconsistent words, the rime can be spelled with two different orthographic consonantal codas, as /ɔs/, which can be spelled ''SS'' (e.g., POSSE, ''hold'') or ''Ç'' (e.g., COÇE, ''scratch'').

Each word in the consistent set was matched with a word in the inconsistent set as regards the first two phonemes (CV). For example, the consistent word BATE, ''hit'', shares the first two phonemes, /ba/, with the inconsistent word BAR, ''bar''. The last phoneme of the initially matched words differed. Thus, although BATE and BAR share the two initial phonemes, they have different rimes (the consistent /at/ rime and the inconsistent /ar/ rime, respectively).

Seventeen different rimes were used in the consistent set (/at/, /af/, /av/, /ag/, /ap/, /aʀ/, /ɔv/, /ɔm/, /ɔd/, /ɔt/, /ɔʀ/, /ɛv/, /ɛm/, /ɛk/, /ɛd/, /ed/, /um/) and thirteen in the inconsistent set (/ar/, /al/, /as/, /ɔr/, /ɔl/, /ɔʃ/, /ɔs/, /ɔz/, /ɛʃ/, /ɛr/, /ɛl/, /ɛʃ/, /ur/). The pseudo-word set was constructed by replacing the initial phoneme of each word. Thus each pseudo-word ended with the same rime as either a consistent or an inconsistent word. In addition, each pseudo-word in the consistent set was matched with a pseudo-word in the inconsistent set as regards the first two phonemes (CV). For example, for the initially matched word pair BATE-BAR, the corresponding pseudo-words were /jat/ and /jar/. All pseudo-words were pronounceable. All the items are listed in Appendix 2.

Characteristics of the words appear in Table 3. As in Experiment 1, since all words were CVC monosyllables and had a UP of 3 phonemes, they all became discriminable when the final consonant was heard.

As in Experiment 1, inconsistent words had always two possible spellings, and were selected using the CR.

Pseudo-words were also selected on the basis of a statistical analysis of orthographic consistency. Besides the CR norms available for the pseudo-words belonging to the initial pool (see Experiment 1) that we used here too, we collected additional CR norms for 62 new pronounceable pseudo-words. A fresh sample of 24 participants was asked to write down these pseudo-words. Procedure was the same as in the pre-test described in Experiment 1. The CR for inconsistent and consistent pseudo-words was 0.11 and 0.47, respectively, $t(38) = 18.99$, $p < .0001$.

Subjective word frequency norms for words (see Experiment 1) were similar for consistent and inconsistent words, $t(38) = 1.3$, $p > .10$.

Contrary to what happened in Experiment 1, stimulus duration was now carefully controlled so as to be unconfounded with orthographic consistency. There was thus no significant stimulus duration difference

TABLE 3
Characteristics of the words used in Experiments 2 to 4

Variables	*Consistent*	*Inconsistent*
Length (No. phonemes)	3	3
Uniqueness point	3	3
No. rime spellings	1	2
Consistency ratio	1	0.5
No. phonological neighbours	11.35	14.2
Frequency (5-point scale)	3.48	3.1
Mean duration (in ms)	604	607
(minimum–maximum)	(448–716)	(409–737)

between consistent and inconsistent words, $t(38) = 0.01$, $p > .10$. Similarly, there was no significant difference in stimulus duration, $t(38) = 0.21$, $p > .10$, between consistent pseudo-words (mean duration: 595 ms; range: 446–714) and inconsistent ones (mean duration: 600 ms; range: 396–740).

In spite of the efforts deployed in stimulus selection, there was still a difference in the number of phonological neighbours between consistent and inconsistent words, $t(38) = 2.2$, $p < .05$. Item analyses were thus run using neighbourhood size as covariate.

All items were recorded in a soundproof room on a digital audio tape recorder (Aiwa HDS1) by a male Portuguese native speaker and then transferred to a computer as in Experiment 1.

Procedure. Presentation and timing were controlled by E-prime 1.0 (Schneider et al., 2002a, 2002b) running on a 1.7 MHz Pentium PC. The stimuli were presented as in Experiment 1.

Results and discussion

RTs were measured from stimulus onset to response onset. Incorrect responses were removed and analysed separately. As in Experiment 1, correct RTs longer or shorter than mean (950 ms) ± 2.5 SD (SD = 228) were removed (2% of the data).

Mean RTs and error rates for consistent and inconsistent words and pseudo-words are presented in Table 4. Planned comparisons were performed on correct mean RTs and error rates averaged by subjects (t_1) and by items (t_2).

Latency analyses revealed that participants were faster by 46 ms, on average, on lexical decisions to consistent words than to inconsistent words, although this difference falls short of statistical significance in the item analysis, $t_1(45) = 5.95$, $p < .0001$; $t_2(38) = 1.83$, $p = .07$. Latency did

TABLE 4
Mean reaction times (ms), standard deviation for correct responses and percentage of errors for consistent and inconsistent words and pseudo-words (Experiment 2: Lexical decision)

		Consistent	*Inconsistent*
Words	RT (ms)	870	916
	SD	76	93
	Errors (%)	12.2	29
Pseudo-words	RT (ms)	972	976
	SD	106	103
	Errors (%)	11.1	8.2

not vary significantly between consistent and inconsistent pseudo-words, $t_1(45) = 0.56, p > .10$; $t_2(38) = 0.03, p > .10$.

Error rate analyses showed that participants made fewer errors (by 16.8%, on average) on lexical decisions to consistent words than to inconsistent words, $t_1(45) = 10.52, p < .0001$; $t_2(38) = 2.82, p < .01$. On the contrary, they made fewer errors (by 2.9%, on average) on lexical decisions to inconsistent than to consistent pseudo-words. This difference is however not statistically significant in the item analysis $t_1(45) = 2.32$, $p < .05$; $t_2(38) = 0.72, p > .10$.

Since inconsistent words had more neighbours than consistent ones, analyses on the item data were also run using number of neighbours as covariate. They showed that the consistency effect for words was still reliable for both latencies, $F(1, 37) = 4.24, p < .05, MSe = 12{,}767.7$, and errors, $F(1, 37) = 7.21, p < .01, MSe = 0.04$.

In sum, we replicated the consistency effect in auditory lexical decision reported in Experiment 1, this time using a better controlled material and several types of orthographic inconsistencies. Indeed, some rimes were inconsistent as regards the presence or absence of a final mute "e", while others (e.g., /ɔs/) were spelled with different orthographic C (e.g., POSSE vs. COÇE).

Inconsistent Portuguese words produced longer auditory lexical decision times and more errors than did words with rimes that are spelled only one way. We also observed an unexpected accuracy advantage for inconsistent pseudo-words, but this effect was not statistically reliable.

In the following experiment, we presented the same material as in Experiment 2 to a fresh sample of participants, but now with instructions to pronounce each item as rapidly and as accurately as possible.

EXPERIMENT 3: SHADOWING

As discussed in the Introduction, the observation of a consistency effect that would be similar in shadowing and in lexical decision could be predicted by a strong interactive view of spoken word perception. The absence or attenuation of the orthographic consistency effect in the shadowing task is predicted by a view assuming orthographic influences to be confined to late and/or lexical processing.

We also predicted that the effect of consistency would be observed with words, but not with pseudo-words, even in the shadowing task. The absence of a consistency effect for pseudo-words can be explained by a "time-out mechanism" in lexical decision, but not in shadowing. As discussed by Ziegler and Ferrand (1998), some highly interactive models might even predict that in naming tasks there should be a stronger consistency effect for pseudo-words than for words, because the former should pattern more like low- than like high-frequency words. This would occur if one assumes that the greater amount of learning for high-frequency words reinforces spelling-to-sound mappings at the biggest grain size (i.e., the word level). Inconsistency at this level is smaller than inconsistency at the subword level. Thus, the more stable word-level grain sizes of high-frequency words should help the inconsistent words to overcome competition at subword grain sizes more efficiently.

Method

Participants. A fresh group of 46 students, aged 18–36 years (median: 19 years) was tested.

Stimuli and procedure. Stimuli and conditions of presentation were the same as in Experiment 2.

Contrary to Experiments 1 and 2, in which participants performed lexical decision, in Experiment 3 they were asked to shadow the items, i.e., to pronounce each item as rapidly and accurately as possible. The vocal response triggered a voice key connected to a button box (PST SRB 200A). Naming latencies were measured from stimulus onset to the triggering of the voice key by the participant's response. The session was tape-recorded for naming errors collection.

Results

Less than 1% of naming latencies were not collected due to voice key failures. Incorrect naming responses were removed and analysed separately. As in the former experiments, correct RTs longer or shorter than mean (764 ms) ± 2.5 SD (SD = 165) were removed (1% of the data). Mean RTs and error rates for consistent and inconsistent words and pseudo-words are presented in Table 5. Planned comparisons were

TABLE 5
Mean reaction times (ms), standard deviation for correct responses and percentage of errors for consistent and inconsistent words and pseudo-words (Experiment 3: Shadowing)

		Consistent	*Inconsistent*
Words	RT (ms)	752	760
	SD	104	108
	Errors (%)	2	1.7
Pseudo-words	RT (ms)	767	765
	SD	110	111
	Errors (%)	8.8	6

performed on correct mean RTs and error rates averaged by subjects (t_1) and by items (t_2).

Consistent words were not pronounced significantly faster, $t_1(45) = 1.7$; $t_2(38) = 0.043$, $p > .10$ in both cases, and did not elicit less errors, $t_1(45) = 0.48$; $t_2(38) = 0.62$, $p > .10$ in both cases, than inconsistent words.

There was no RT effect of orthographic consistency for pseudo-words either, $t_1(45) = 0.52$; $t_2(38) = 0.07$, $p > .10$ in both cases. Error rate analyses showed that participants made fewer errors (by 2.8%, on average) on shadowing responses to inconsistent than to consistent pseudo-words, but this difference falls far from significance in the item analysis, $t_1(45) = 2.33$, $p < .05$; $t_2(38) = 0.82$, $p > .10$.

Analyses of the word item data using number of neighbours as covariate confirmed that there was no reliable consistency effect on either latencies or errors (both *F*s < 1).

Additional analyses and between-tasks comparison of Experiments 2 and 3

In order to compare directly the influence of orthographic consistency on lexical decision and shadowing, we estimated the RT consistency effect size for Experiments 2 and 3 using the necessary adjustment procedures for repeated measures (see e.g., Dunlap, Cortina, Vaslow, & Burke, 1996; Johnson & Eagly, 2000). The null result in Experiment 3 suggests a failure to replicate the consistency effect obtained in Experiment 2. In fact, while the effect size point estimate of the contrast between consistent and inconsistent words was $d = 1.01$ for Experiment 2 (lexical decision), it was $d = 0.26$ for Experiment 3 (shadowing). Estimating the 5% one-sided confidence interval for the effect size obtained in Experiment 2 (the procedure recommended by e.g., Cohen, 1994, and Schmidt, 1996) we found a lower limit of $d = 0.38$. The difference found in Experiment 3 is thus clearly out of the range of the confidence interval for the effect size of Experiment 2. This conclusion holds true for accuracy data, since the contrast between consistent and inconsistent words in Experiment 3 (effect size point estimate: $d = 0.07$) is out of the range of the 5% one-sided confidence interval of the effect size estimate in Experiment 2 ($d = 1.55$, lower limit: $d = 0.86$). Hence, it is highly unlikely that we were dealing with the same effect in both experiments, and we can be fairly confident that there was no consistency effect hidden in the shadowing data.

However, considering together the results of the two experiments, it appeared that responses were faster for shadowing than for lexical decision, $t_1(90) = 8.53$, $p < .0001$; $t_2(79) = 22.8$, $p < .0001$. It could thus be argued that consistency did not affect shadowing only because responses were faster in this task. This interpretation predicts that the size of the consistency effect should be correlated with response speed. In order to test it, we calculated for each participant the size of the consistency effect (the difference between latencies to consistent and inconsistent words) and the global response latency for the whole material (words and pseudo-words). There was no significant correlation between size of the consistency effect and global latency ($r = .24$, $p > .10$, for lexical decision; $r = .11$, $p > .10$, for shadowing). This allows us to rule out the possibility that the lack of consistency effect in shadowing would be due to the overall faster response rate.

Nevertheless, it could still be argued that the difference between our consistent and inconsistent words was not strong enough for the consistency effect to emerge in shadowing. Indeed, in our stimulus set, inconsistent rimes could be spelled in only two ways, while Ziegler and Ferrand (1998) used inconsistent rimes that could be spelled in multiple ways (about six spellings, on average). However, we can discard this argument on the basis of the following analysis.

Although all our inconsistent words had rimes that can be spelled in only two ways, they had different CR (consistency ratio) values. As already explained, the CR is the sum of words with the same rime and spelling relative to the sum of words with the same rime, and thus reflects the degree of rime/body (in)consistency. We divided the inconsistent words used in Experiment 3 into two sets, using the median CR (0.50) as the cut-off point: below the median CR (0.36 CR, on the average) and above the median CR (0.64 CR, on the average), the CR difference between the two sets being statistically significant, $t(18) = 5.46$, $p < .0001$. Below and above median CR inconsistent words remained controlled for stimulus duration, $t(18) = 0.48$, $p > .10$, and number of neighbours, 13 and 15.4, respectively, $t(18) = 1.23$, $p > .10$, but differed on word frequency, 3.5 vs. 2.7, respectively, $t(18) = 2.3$, $p < .05$. Given the definition of the CR, our below-the-median CR set was similar to Ziegler et al.'s (unpublished) subdominant spelling inconsistent words, i.e., they were inconsistent words presenting an infrequent spelling of their phonology, while our above-the-median CR set was similar to their dominant spelling inconsistent words. If our failure to find a consistency effect in shadowing were due to the weakness of rime/body inconsistency in our material, then one might expect the effect to occur at least for the below-median CR set. However, the item analyses performed with word frequency as covariate showed almost identical shadowing latencies for both types of inconsistent words (760.8 ms and 762.1 ms, respectively; $F < 1$) and post-hoc tests (Tukey HSD) indicate that consistent words were not pronounced faster than any type of inconsistent words. Thus, even inconsistent words that present an infrequent spelling of their phonology did not lead to a performance decrease.

In short, we can be fairly confident that, contrary to lexical decision, there is no significant orthographic consistency effect in shadowing. This task difference was predicted on the basis of the assumption that the two tasks tap different stages of processing. In particular, we assumed that lexical decision involves substantially more post-access, lexical and/or decisional components than shadowing, which relies on earlier processing stages and/or on sublexical representations.

In order to check whether the present data support this interpretation, we examined in both tasks the effect of a lexical variable, namely word frequency. Indeed, if lexical decision performance involves more lexical processing and/or more decision processes (cf. Balota & Chumbley, 1984), it ought to be more dependent on word frequency than shadowing performance is. The correlation between word RT and word frequency was computed separately for Experiments 2 and 3. There was a significant negative correlation between word RT and word frequency for lexical decision ($r = -.56, p < .0001$), but not for shadowing ($r = -.15, p > .10$). We also checked whether fewer participants were affected by word frequency in shadowing than in lexical decision. To this aim, we computed for each participant the correlation between word RT and word frequency. A remarkable dissociation was observed between the two tasks. In lexical decision, 29 out of the 46 participants gave evidence of significant ($p < .05$) negative correlations between RT and word frequency, while in shadowing this was the case for only four out of the 46 participants. The difference between these two distributions is significant, $\chi^2(1) = 29.53$, $p < .0001$, further supporting our assumption that shadowing was accomplished through lesser lexical processing than lexical decision.

Discussion

The shadowing task was used to determine if the effect of multiple rime spelling is circumscribed to lexical and/or post-lexical decision stages of spoken word processing or, alternatively, if it pervades earlier stages leading to the activation of word phonological forms.

In accordance with the idea that the influence of orthography is restricted to late, decisional and/or lexical stages of speech processing, and in contrast to what we observed in lexical decision, we found no evidence of an orthographic consistency effect in the shadowing task. Even inconsistent words that present an infrequent spelling of their phonology did not lead to a performance decrease. This result seems difficult to reconcile with the view that our manipulation of consistency was not strong enough for the consistency effect to emerge in shadowing. Additional analyses confirmed that the absence of an orthographic consistency effect is not due to shadowing responses being faster than lexical decision.

In addition, we did not observe any detrimental effect of orthographic inconsistency on pseudo-word processing. Contrary to what was the case for lexical decision, no ‘‘time-out mechanism’’ might account for such a response pattern.

Taken together, the results of Experiments 2 and 3 thus clearly show that the consistency effect does not reflect a general coupling between phonology and orthography at all speech processing stages.

The notion that lexical representations and/or decisional processes are less involved (or at least, less mandatory) in shadowing than in lexical decision is supported by two observations. One is that frequency was significantly correlated to response latencies only in lexical decision. The other is that many more participants showed a significant correlation between response latencies and word frequency in lexical decision than in shadowing. These results, together with those reported in the Introduction on the comparison between the two tasks, converge on the notion that lexical representations affect lexical decision to a greater extent than shadowing.

This does not imply that the lexical/sublexical distinction is the critical one in the occurrence or not of the consistency effect. In the following experiment, we tried to evaluate whether the consistency effect stems from lexical or decisional components. For that purpose, we compared two situations in which a shadowing response was made contingent upon either a lexical or a phonemic criterion. More precisely, participants were presented with the same stimuli as in Experiments 2 and 3, but now with instructions either to shadow the items only if they were words (*lexically contingent shadowing*) or to shadow the items only if they were preceded by their first phoneme (*phonemically contingent shadowing*).

While both tasks still require to repeat the spoken sequence as rapidly as possible, as in the simple shadowing situation used in Experiment 3, they now involve an additional decisional component, one lexical (*lexically contingent shadowing*) and the other sublexical (*phonemically contingent shadowing*).[5] If the orthographic consistency effect were arising at a decisional stage of processing, then consistency would be expected to

[5] The tasks used in Experiment 4 might be considered as new variants of the ‘‘cued shadowing’’ situation analysed by Bates and Liu (1996). However, we preferred not to use this terminology because most of the cued shadowing tasks aimed at examining priming effects and hence used peripheral cues to a target word, for example its position within a word pair (e.g., Radeau, Morais, & Dewier, 1989) or a voice shift (e.g., Liu et al., 1997). As the name we chose suggests, the task we used is actually closer to the *lexically contingent naming task* (also called *Go No-go naming*) used in written word identification studies (e.g., Andrews & Heathcote, 2001), in which participants are instructed to read the item aloud only if it is a word.

influence performance to a similar extent in both tasks. On the contrary, if the influence of orthography were restricted to lexical representations of spoken words, then the consistency effect would be expected to be stronger in lexically than in phonemically contingent shadowing.

EXPERIMENT 4: LEXICALLY VS. PHONEMICALLY CONTINGENT SHADOWING

In the present experiment, one group of participants received instructions to shadow the items only if they were words. An orthographic consistency effect was expected on the basis of the idea that it depends on either lexical activation or the presence of a decisional component in the task. Additionally, the demonstration of a consistency effect in this task would allow us to rule out the possibility that, in shadowing, output processes related to the generation and the execution of articulatory motor programmes would mask perceptual effects.

Both the auditory lexical decision task used in Experiments 1 and 2 and the lexically contingent shadowing task rely on lexical representations and have a decisional component. We used a second task, in which participants were instructed to pronounce an item only if it was preceded by its first phoneme. Thus, the decisional component of this contingent shadowing situation was sublexical. If the consistency effect were arising at a decisional stage, we should observe it in this task as well. On the contrary, if the influence of orthography was restricted to lexical representations of spoken words, the consistency effect should disappear.

Method

Participants. A fresh group of 70 students (36 and 34, respectively, for lexically and phonemically contingent shadowing), aged 18–44 years (median: 20 years) was tested.

Stimuli. Stimuli were the same as in Experiments 2 and 3. Target phonemes for phonemically contingent shadowing were recorded in a soundproof room on a digital audio tape recorder (Aiwa HDS1) by a female native speaker of Portuguese and then transferred to a computer as in Experiments 1 and 2.

Procedure. Stimuli for lexically contingent shadowing were presented as in Experiment 3. Contrary to Experiment 3, in which participants were asked to shadow all the items, they were instructed to shadow an item only if it was a Portuguese word.

Each trial in phonemically contingent shadowing began with the auditory presentation of a target phoneme. After a 500 ms delay, an item (either a word or a pseudo-word) was presented. Participants were asked to shadow this item as rapidly and accurately as possible only when it started with the target phoneme. Each participant received all items twice, once with the first phoneme as target and once with a different phoneme as target. The inter-trial interval was 2 s.

Presentation and timing were controlled and naming latencies collected as in Experiment 3.

Results

Lexically contingent shadowing. Less than 0.5% of naming latencies were not collected due to voice key failures. Failures to shadow a target word (misses) and incorrect naming responses were removed and analysed separately. Correct word RTs longer or shorter than mean (1010 ms) $\pm$ 2.5 SD (SD = 258) were removed (3% of the data). Mean RTs and error rates for consistent and inconsistent words are presented in the upper part of Table 6. Planned comparisons were performed on correct mean RTs and error rates averaged by subjects (t_1) and by items (t_2).

Consistent words were pronounced faster than inconsistent words (by 59.5 ms, on the average), although the difference falls short of statistical significance in the item analysis, $t_1(35) = 5.3$, $p < .0001$; $t_2(38) = 1.8$, $p = .07$. Error rates did not vary between consistent and inconsistent words, $t_1(35) = 1.02$; $t_2(38) = 1.17$, $p > .10$ in both cases.

Analyses of the word item data using number of neighbours as covariate still showed the consistency effect to fall short of statistical significance for

TABLE 6
Mean reaction times (ms), standard deviation for correct responses and percentage of errors for consistent and inconsistent words (lexically contingent shadowing and phonemically contingent shadowing) and pseudo-words (phonemically contingent shadowing)

			Consistent	*Inconsistent*
Lexically contingent shadowing				
	Words	RT (ms)	967	1026
		SD	105	127
		Errors (%)	0.6	1.2
Phonemically contingent shadowing				
	Words	RT (ms)	760	758
		SD	120	120
		Errors (%)	2.2	3.2
	Pseudo-words	RT (ms)	758	749
		SD	127	124
		Errors (%)	4.9	6.5

latencies, $F(1, 37) = 3.37$, $p = .07$, $MSe = 15{,}486.2$, and to be not reliable for errors, $F(1, 37) = 1.33$, $p > .10$, $MSe = 0.001$.

These results seem similar to those observed when participants had to perform a lexical decision task on the same material, as was the case in Experiment 2. Direct comparison between the two experiments showed that the contrast between consistent and inconsistent words in the present task (RT effect size point estimate: $d = 0.89$) is within the range of the 5% one-sided confidence interval of the RT effect size estimated in Experiment 2 ($d = 1.01$, lower limit: $d = 0.38$). Hence, the effect is likely to be the same in both experiments.

Phonemically contingent shadowing. Less than 1.5% of naming latencies were not collected due to voice key failures. Errors for target present trials (failures to shadow an item and incorrect naming responses) were removed and analysed separately. Correct target present RTs longer or shorter than mean (768 ms) $\pm$ 2.5 SD (SD = 191) were removed (1.7% of the data). Mean RTs and error rates for consistent and inconsistent words and pseudo-words are presented in the lower part of Table 6. Planned comparisons were performed on correct mean RTs and error rates averaged by subjects (t_1) and by items (t_2).

As can be seen in Table 6, the RT difference between consistent and inconsistent words was very small and not significant, $t_1(33) = 0.3$; $t_2(38) = 0.18$, $p > .10$ in both cases. Coherently, consistent words did not elicit fewer errors, $t_1(33) = 1.09$; $t_2(38) = 0.98$, $p > .10$ in both cases, than inconsistent words. Neither did orthographic consistency affect pseudo-word processing, RT: $t_1(33) = 1.43$; $t_2(38) = 0.33$, $p > .10$ in both cases; error rate: $t_1(33) = 1.24$; $t_2(38) = 0.65$, $p > .10$ in both cases.

Analyses of the word item data using number of neighbours as covariate confirmed that there was no reliable consistency effect on either latencies, $F(1, 37) < 1$, or errors, $F(1, 37) = 2.15$, $p > .10$, $MSe = 0.002$.

These results seem similar to those observed when participants had to perform a simple shadowing task on the same material. Direct comparison between the two experiments showed that in the present task the RT contrast between consistent and inconsistent words (RT effect size point estimate: $d = -0.05$) is within the range of the 5% one-sided confidence interval of the RT effect size estimated in Experiment 3 ($d = 0.26$, lower limit: $d = -0.33$).

Additional analyses and between-tasks comparison. Direct comparison of the two tasks showed the RT consistency effect size for phonemically contingent shadowing ($d = -0.05$) to be clearly out of the range of the 5% one-sided confidence interval of the effect size of lexically contingent shadowing ($d = 0.89$; lower limit: $d = 0.20$). Hence, it is highly unlikely that we were dealing with the same effect in both tasks, and we can be fairly confident that there was no consistency effect hidden in the phonemically contingent shadowing data. This mirrors the conclusion of the similar comparison made between lexical decision (Experiment 2) and simple shadowing (Experiment 3).

However, when considering the two tasks together, it appeared that responses were faster for phonemically than for lexically contingent shadowing, $t_1(68) = 8.65, p < .0001$; $t_2(39) = 17.36, p < .0001$, as was the case when we compared the simple shadowing situation (Experiment 3) to lexical decision (Experiment 2). It could thus be argued here too that the consistency effect did not come out in the phonemically contingent shadowing task only because responses were faster in this situation. If it were the case, the size of the consistency effect should be correlated with response speed. However, while for lexically contingent shadowing the correlation between the size of the consistency effect and global latency is significant ($r = .34, p < .05$), we found no significant correlation for phonemically contingent shadowing ($r = .01, p > .10$). We thus performed a covariance analysis with task (lexically vs. phonemically contingent shadowing) as the independent variable, consistency effect (RT inconsistent—RT consistent) as the dependent measure and global latency for words as covariate. There was still a reliable difference between the two tasks as regards the size of the consistency effect, $F(1, 67) = 3.92, p < .05$, $MSe = 2963.9$. This allows us to rule out the possibility that the absence of a consistency effect in the phonemically contingent shadowing task would be due to the overall faster response rate.

We can thus be fairly confident that, contrary to what we observed in the lexically contingent shadowing task, there was no significant orthographic consistency effect in the phonemically contingent shadowing situation. Both tasks include a decisional component. Thus, the between-task difference appears to be linked to the fact that, while phonemically contingent shadowing biases participants to sublexical representations, only lexically contingent shadowing necessarily involves lexical representations.

On the basis of this interpretation, we should observe lexically contingent shadowing performance to be more dependent than phonemically contingent shadowing performance on lexical variables like word frequency. This was the case: the correlation between word RT and word frequency was significant only for lexically contingent shadowing ($r = -.49, p < .001$), not for phonemically contingent shadowing ($r = -.16, p > .10$). In addition, fewer participants were affected by word frequency in phonemically than in lexically contingent shadowing. Indeed, in lexically contingent shadowing, 16 out of the 36 participants gave evidence of significant ($p < .05$) negative correlation between RT and word frequency, while in phonemically contingent shadowing this was the case for only 6 out of the 34 participants. The difference between these two distributions is significant, $\chi^2(1) = 5.73$, $p < .05$, supporting our assumption that phonemically contingent shadowing was accomplished through lesser lexical processing than lexically contingent shadowing. This mirrors the conclusion of the similar comparison made between lexical decision (Experiment 2) and simple shadowing (Experiment 3).

Discussion

Taking together the results of Experiments 2 and 3 (lexical decision vs. shadowing) and those of Experiment 4 (lexically vs. phonemically contingent shadowing), we can conclude that the decisional component of the task is not critical in determining the emergence of the consistency effect. Rather, orthographic consistency seems to affect performance only when the task relies on lexical representations.

The demonstration of a consistency effect in the lexically contingent task also runs clearly against the notion that shadowing is by itself not sensitive enough to disclose the consistency effect.

GENERAL DISCUSSION

Contrary to what might be expected from current models of spoken word recognition, the present study converges with a growing body of experimental evidence, exhaustively reviewed in the Introduction, which demonstrates that the recognition of spoken words is modulated by orthographic knowledge.

In Experiment 1, we replicated in Portuguese the consistency effect initially reported by Ziegler and Ferrand (1998) for French words. Inconsistent Portuguese words including rimes that can be spelled in two different ways produced both significantly longer auditory lexical decision times and significantly more errors than did consistent words including rimes that can be spelled only one way. The orthographic consistency effect was replicated in Experiment 2, using a new material for which the two initial phonemes were matched across experimental manipulations.

As we argued in the Introduction, if the influence of orthographic knowledge occurred at all stages of speech processing, including those concerned with the extraction of phonetic and phonological information from the acoustic signal, then the consistency effect should be observed not only for words but also for pseudo-words. Yet, when differences were observed for pseudo-words in the present study, they reflected an unexpected advantage for inconsistent pseudo-words and were statistically unreliable. Thus, like Ziegler and Ferrand (1998), we did not observe any advantage of consistent over inconsistent pseudo-words. Nevertheless, if in lexical decision pseudo-words were rejected by a ‘‘time out’’ mechanism, the absence of a consistency effect on pseudo-word processing in this task would not counter Ziegler and Ferrand’s view.

In our study we had resort to stronger tests of the idea that orthographic knowledge pervades all stages of speech processing. They consisted in manipulating the decisional and lexical demands of the task and in looking for the occurrence of a consistency effect in several tasks using the same material that elicited such an effect in lexical decision. If speech processing were modulated by orthographic knowledge from its very early stages and/or at sublexical levels of processing, the consistency effect should also affect performance in tasks that do not necessarily rely on lexical representations or that do not involve a strong decisional component.

In Experiment 3, we required the participants to immediately repeat the speech stimulus just heard. This shadowing task does not involve a binary decision on whether the stimulus is or is not a word and can be assumed to require much less lexical processing than lexical decision does. The last assumption is based on evidence, reviewed in the Introduction, concerning the much weaker role of word frequency, neighbourhood density and location of the UP in shadowing compared with binary choice decision tasks like lexical or gender decision.

In addition, in shadowing no ‘‘time-out mechanism’’ might account for the absence of a consistency effect on pseudo-word processing. On the contrary, highly interactive models might predict an even stronger consistency effect for pseudo-words than for words, because the more stable word-level grain sizes would help the inconsistent words to overcome competition at subword grain sizes more efficiently (Ziegler & Ferrand, 1998).

In such a shadowing situation, we did not find any effect of orthographic consistency on either response latency or error rate. This was true not only for pseudo-words but also for words. Consistent with the view that orthography only modulates late and/or lexical stages of word recognition, we found that the size of the effect on either response latency or accuracy observed in Experiment 3 (shadowing) is clearly not within the confidence interval of the size of the effect in Experiment 2 (lexical decision).

As said above, our prediction that, contrary to lexical decision, shadowing would not elicit a consistency effect was based on previous data indicating that this task involves less lexical and/or decisional processing than lexical decision. It was therefore important to find empirical support for this assumption in the present study, too. Consistent with this hypothesis, we found many more participants to display a word frequency effect, i.e., shorter response latencies with increasing word frequency, in the lexical decision task than in the shadowing one.

So far, the present results supported the notion that the influence of orthographic knowledge on spoken word recognition is confined to late, decisional and/or lexical stages of processing. Experiment 4 was designed to evaluate whether the consistency effect stems from lexical or decisional components. For that purpose, we compared two situations in which a shadowing response was made contingent upon either a lexical or a phonemic criterion. In the lexically contingent shadowing situation, participants were instructed to shadow the items only if they were words. In the phonemically contingent shadowing situation, they were instructed to shadow the items only if they were preceded by their first phoneme. No effect of orthographic consistency was observed in the latter situation, while a significant effect was observed in lexically contingent shadowing. In addition, the effect size estimated in phonemically contingent shadowing was clearly not within the confidence interval of the effect size estimated in lexically contingent shadowing.

These results mirror those obtained when comparing the simple shadowing situation (Experiment 3) to the lexical decision task (Experiment 2). Similarly, we also observed word frequency to affect performance far more in lexically than in phonemically contingent shadowing.

Taken together, the two sets of data thus lead to converging results when comparing, on the one hand, simple shadowing to lexical decision, and, on the other hand, phonemically to lexically contingent shadowing. In addition, we argued that these between-task differences cannot be accounted for by an effect of overall response speed. Indeed, we noticed that in both sets of data the consistency effect emerged only when responses were relatively slow, which was the case for simple shadowing in comparison with lexical decision, and for lexically in comparison to phonemically contingent shadowing. However, no significant correlation was observed between the size of the consistency effect and global latency for simple shadowing and lexical decision (Experiments 2 and 3), and there still was a reliable effect size difference between lexically and phonemically contingent shadowing when we partialled out global latency. It thus seems hard to believe that the consistency effect did not come out in the simple and phonemically contingent shadowing tasks only because responses were much faster in these two situations.

Our results thus support the notion that the influence of orthographic knowledge is confined to lexical stages of speech processing. The comparison between the two contingent shadowing tasks clearly shows that by itself the decisional component of the task is not critical in determining the emergence of the consistency effect. This does not imply however that the lexical/sublexical distinction is critical for the occurrence or not of the consistency effect. Indeed, both the auditory lexical decision task and the lexically contingent shadowing task involve a decisional as well as a lexical component. In other words, we lack a decision-free situation in which we could observe the consistency effect while biasing participants towards lexical representations. We are thus allowed to conclude that lexical activation is a *necessary* condition of the occurrence of the orthographic consistency effect. Whether it is a *sufficient* condition cannot be determined on the basis of the present study. It remains possible that only a lexically based decision (an inherent component of the lexically contingent task used in Experiment 4) can elicit the consistency effect.

The present conclusions, though valid for Portuguese, are not necessarily so for other languages, since the orthographic inconsistency we exploited is restricted to only two possible spellings. By contrast, the rime inconsistencies considered by Ziegler and Ferrand (1998) for the French language concerned, on average, about six different spellings. We thus cannot exclude the possibility that for languages in which orthographic inconsistencies are stronger than in Portuguese the influence of orthography in spoken word recognition extends to early and/or prelexical processing stages. Although our own study produced results that make this hypothesis unlikely (see the Additional analyses and between-tasks comparison of Experiments 2 and 3), more work is needed to check whether similar results would be obtained in deeper orthographies, like the French one.

While our conclusion is obviously restricted to speech processing, it is interesting to note that a similar debate is taking place as regards written word recognition. Indeed, while several written word recognition models assume reciprocal connections between orthographic and phonological units at various grain sizes (e.g., Stone et al., 1997; Van Orden & Goldinger, 1994; Van Orden et al., 1997) some authors proposed instead a restrictive interactivity account in which interactions are limited to lexical processing levels (i.e., to whole word orthographic representations) and do not concern the sublexical and prelexical level (e.g., Peereman, Content, & Bonin, 1998). However, the situation may be different in the case of written word recognition. Given that alphabetically written words are learned on the basis of their sub-lexical correspondences with the phonological segments of the spoken words, phonological representations may persist playing a role even at the early and/or prelexical processing stages when the recognition of written words becomes automatic.

On the contrary, when the child begins to acquire literacy, the reorganisation of his/her basic processes of speech perception under the influence of a still growing, necessarily imperfect and unstable, body of orthographic knowledge would introduce undesirable sources of error in the speech system. Such a view makes plausible the notion that, at least for orthographic systems that do not exhibit a high degree of phonographic inconsistency, the orthographic influence is limited to lexical (or, possibly, both lexical and decisional) processing stages.

Nevertheless, since lexical influences of orthographic knowledge seem to impair spoken word recognition, as shown by the consistency effects studied in the present paper, it remains to be understood why there should be orthographic effects at all. Although this problem goes far beyond the scope of the present paper, we would like to suggest that orthographic representations might help mnemonic representations to become more stable and precise. As a matter of fact, several studies on illiterate adults show that they are very poor at repeating pseudo-words (e.g., Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998; Morais & Mousty, 1992) and have very low verbal spans (e.g., Morais et al., 1986), clearly inferior to their visuo-spatial span (see Morais & Kolinsky, 2001). Lack of automatic activation of orthographic representations provides a potential explanation of illiterates' verbal span inferiority. If this interpretation were correct, orthographic influences on spoken word recognition would provide far more benefits than costs to a literate mind.

Manuscript received June 2002
Revised manuscript received June 2003

REFERENCES

Adrián, J.A., Alegria, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology*, *30*, 329–353.

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 514–540.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency on the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Balota, D.A., Pilotti, M., & Cortese, M.J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition*, *29*, 639–647.

Bates, E., & Liu, H. (1996). Cued shadowing. *Language and Cognitive Processes*, *11*, 577–582.

Ben-Dror, I., Frost, R., & Bentin, S. (1995). Orthographic representation and phonemic segmentation in skilled readers: A cross-language comparison. *Psychological Science*, *6*, 176–181.

Carroll, J.B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior*, *10*, 722–729.

Castro, S.-L. (1993). *Alfabetização e percepção da fala*. [Alphabetization and spoken word perception] Porto: Portugal: Instituto Nacional de Investigação Cientifica.

Castro, S.-L., Vicente, S., Morais, J., Kolinsky, R., & Cluytens, M. (1995). Segmental representation of Portuguese in 5- and 6- year olds: Evidence from dichotic listening. In I. Hub Faria & J.Freitas (Eds.), *Studies on the acquisition of Portuguese. Proceedings of the First Lisbon Meeting on Child Language* (pp. 1–16). Lisboa: Colibri.

Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, *121*, 1053–1063.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003.

Connine, C.M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1084–1096.

de Gelder, B., Vroomen, J., & Bertelson, P. (1993). The effects of alphabetic-reading competence on language representation in bilingual Chinese subjects. *Psychological Research*, *55*, 315–321.

Dijkstra, T., Frauenfelder, U.H., & Schreuder, R. (1993). Bidirectional grapheme-phoneme activation in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 931–950.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology*, *49*, 264–271.

Donnenwerth-Nolan, S., Tanenhaus, M.K., & Seidenberg, M.S. (1981). Multiple code activation in word recognition: Evidence from rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 170–180.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., & Burke, M.J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*, 170–177.

Ehri, L.C., & Wilce, L.S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, *71*, 26–40.

Ehri, L.C., & Wilce, L.S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, *1*, 371–385.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Frauenfelder, U.H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 77–91.

Frost, R. (1991). Phonetic recoding of print and its effect on the detection of concurrent speech in amplitude-modulated noise. *Cognition*, *39*, 195–214.

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word processing. *Memory and Cognition*, *17*, 302–311.

Frost, R., Repp, B.H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory and Language*, *27*, 741–755.

Galbraith, R.C., & Underwood, B.J. (1973). Perceived frequency of concrete and abstract words. *Memory and Cognition*, *1*, 56–60.

Goldinger, S.D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, *11*, 559–567.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language*, *24*, 631–645.

Hallé, P.A., Chéreau, C., & Segui, J. (2000). Where is the /b/ in ''absurde'' [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, *43*, 618–639.

Jakimik, J., Cole, R.A., & Rudnicky, A.I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, *24*, 165–178.

Johnson, B.T., & Eagly, A.H. (2000). Quantitative synthesis of social psychological research. In H.T. Reis & C.M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 496–528). Cambridge: Cambridge University Press.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, *47*, 145–171.

Kolinsky, R. (1998). Spoken word recognition: A stage-processing approach to language differences. *European Journal of Cognitive Psychology*, *10*, 1–40.

Kolinsky, R., & Morais, J. (1993). Intermediate representations in spoken word recognition: A cross-linguistic study of word illusions. In *Proceedings of the '93 Eurospeech Conference* (pp. 731–734). Berlin.

Kouider, S., & Dupoux, E. (2001). A functional disconnection between spoken and visual word recognition: Evidence from unconscious priming. *Cognition*, *82*, B35–B49.

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

Liu, H., Bates, E., Powell, T., & Wulfeck, B. (1997). Single-word shadowing and the study of lexical access. *Applied Psycholinguistics*, *18*, 157–180.

Luce, P.A., & Large, N.R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*, 565–581.

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

Luce, P.A., Pisoni, D.B., & Goldinger, S.D. (1990). Similarity neighborhoods of spoken words. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122–147). Cambridge, MA: MIT Press.

Mann, V. (1986). Phonological awareness: The role of reading experience. *Cognition*, *24*, 65–92.

Marslen-Wilson, W.D. (1985). Speech shadowing and speech comprehension. *Speech Communication*, *4*, 55–73.

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

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

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

Moats, L.C. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia*, *44*, 81–102.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech analysis. *Cognition*, *24*, 45–64.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323–331.

Morais, J., Castro, S.-L., Scliar-Cabral, L., Kolinsky, R., & Content, A. (1987). The effects of literacy on the recognition of dichotic words. *Quarterly Journal of Experimental Psychology*, *39 A*, 451–465.

Morais, J., & Kolinsky, R. (1994). Perception and awareness in phonological processing: The case of the phoneme. *Cognition*, *50*, 287–297.

Morais, J., & Kolinsky, R. (1995). The consequences of phonemic awareness. In B. de Gelder, & J. Morais (Eds.), *Speech and reading: Comparative approaches* (pp. 317–337). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Morais, J., & Kolinsky, R. (2001). The literate mind and the universal mind. In E. Dupoux (Ed.), *Language, brain and cognitive development: Essays in honor of Jacques Mehler* (pp.

463–480). Cambridge, MA: MIT Press. [Also in E. Dupoux (Ed.), *Les langages du cerveau.* Ed. Odile Jacob, 2002.]

Morais, J., Kolinsky, R., Ventura, P., & Cluytens, M. (1997). Levels of processing in the phonological segmentation of speech. *Language and Cognitive Processes*, *12*, 35–39.

Morais, J., & Mousty, P. (1992). The causes of phonemic awareness. In J. Alegria, D. Holender, J. Junça de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 193–211). Amsterdam: Elsevier.

Nakamura, M., Kolinsky, R., Spagnoletti, C., & Morais, J. (1998). Phonemic awareness in alphabetically literate Japanese adults: The influence of the first acquired writing system. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *17*, 417–450.

Norris, D. (1986). Word recognition: context effect without priming. *Cognition*, *22*, 93–136.

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

Norris, D., McQueen, J.M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*, 299–370.

Olson, D.R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, *60*, 83–104.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language*, *39*, 151–174.

Radeau, M., & Morais, J. (1990). The uniqueness point effect in the shadowing of spoken words. *Speech Communication*, *9*, 155–164.

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory and Cognition, 17, 525–535.*

Radeau, M., Morais, J., Mousty, P., & Bertelson, P. (2000). The effect of speaking rate on the role of the uniqueness point in spoken word recognition. *Journal of Memory and Language*, *42*, 406–422.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1297–1311.

Radeau, M., Mousty, P., & Bertelson, P. (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, *51*, 123–128.

Rastle, K., & Davis, M.H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 307–314.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic reading. *Cognition*, *24*, 31–44.

Schmidt, F.L. (1996). Statistical significance testing and cumulative knowledge in Psychology: Implications for training of researchers. *Psychological Methods*, *1*, 115–129.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime user's guide.* Pittsburgh: Psychology Software Tools Inc.

Schneider, V.I., & Healy, A.F. (1993). Detecting phonemes and letters in text: Interactions between different types and levels of processes. *Memory and Cognition*, *21*, 739–751.

Scliar-Cabral, L., Morais, J., Nepomuceno, L., & Kolinsky, R. (1997). The awareness of phonemes: So close-so far away. *International Journal of Psycholinguistics*, *13*, 211–240.

Seidenberg, M.S., & Tanenhaus, M.K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 546–554.

Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia*, *20*, 615–627.

Shapiro, B.J. (1969). The subjective estimation of relative word frequency. *Journal of Verbal Learning and Verbal Behavior*, *8*, 248–251.

Slowiaczek, L. (1994). Semantic priming in a single-word shadowing task. *American Journal of Psychology*, *107*, 245–260.

Slowiaczek, L.M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1239–1250.

Slowiaczek, L.M., Soltano, E.G., Wieting, S.J., & Bishop, K.L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology*, *56 A*, 233–262.

Spagnoletti, C., Morais, J., Alegria, J., & Dominicy, M. (1989). Metaphonological abilities of Japanese children. *Reading and Writing*, *2*, 221–244.

Stone, G.O., Vanhoy, M.D. & Van Orden, G.C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, *36*, 337–359.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, *24*, 320–335.

Tanenhaus, M.K., Flanigan, H.P., & Seidenberg, M.S. (1980). Orthographic and phonological activation in auditory and visual word recognition. *Memory and Cognition*, *8*, 513–520.

Thorndike, E.L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Bureau of Publications, Teachers College, Columbia University.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, *15*, 49–74.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, *25*, 476–491.

Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology*, *33*, 771–780.

Treiman, R., Tincoff, R., & Richmond-Welty, E.D. (1996). Letter names help children to connect print and speech. *Developmental Psychology*, *32*, 505–514.

Tryk, H.E. (1968). Subjective scaling of word frequency. *American Journal of Psychology*, *81*, 170–177.

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. (1997). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology*, *14*, 131–165.

Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language and Cognitive Processes*, *16*, 393–418.

Vitevitch, M.S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 270–278.

Vitevitch, M.S., & Luce, P.A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*, 325–329.

Vitevitch, M.S., & Luce, P.A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*, 374–408.

Whalen, D.H., & Zsiga, E.C. (1994). Subjective familiarity of English word/name homophones. *Behavior Research Methods, Instruments, and Computers*, *26*, 402–408.

Zecker, S.G. (1991). The orthographic code: Developmental trends in reading-disabled and normally-achieving children. *Annals of Dyslexia*, *41*, 178–192.

Ziegler, J.C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory recognition. *Psychonomic Bulletin and Review*, *5*, 683–689.
Ziegler, J.C., Ferrand, L., & Montant, M. (2002). Interdependence of written and spoken language processing. *Unpublished manuscript*.
Ziegler, J.C., Montant, M., & Jacobs, A.M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language*, *37*, 533–554.

APPENDIX 1

Material of Experiment 1

Words		*Pseudo-words*	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*
/kor/	/sal/	/bor/	/bal/
/dor/	/mal/	/gor/	/nal/
/por/	/tal/	/zor/	/pal/
/ver/	/val/	/ger/	/dal/
/ter/	/lar/	/ber/	/tar/
/ler/	/mar/	/ner/	/ʒar/
/kav/	/gar/	/dav/	/zar/
/nav/	/rir/	/pav/	/gir/
/mɛd/	/fɛl/	/lɛd/	/dɛl/
/pɛd/	/mɛl/	/tɛd/	/rɛl/
/lɛm/	/ʒɛl/	/bɛm/	/gɛl/
/tɛm/	/pɛl/	/pɛm/	/tɛl/
/dɛv/	/kɛr/	/sɛv/	/sɛr/
/nɛv/	/fɛr/	/fɛv/	/mɛr/
/bɔd/	/ʒɛr/	/ʒɔd/	/vɛr/
/pɔd/	/vil/	/kɔd/	/sil/
/tɔm/	/sɔl/	/rɔm/	/ʒɔl/
/pɔt/	/mɔl/	/zɔt/	/ʃɔl/
/kum/	/sul/	/dum/	/kul/
/lum/	/bul/	/pum/	/vul/

APPENDIX 2

Material of Experiments 2 to 4

Words		*Pseudo-words*	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*
/bat/	/bar/	/ʒat/	/ʒar/
/ʃɔv/	/ʃɔr/	/pɔv/	/pɔr/
/kɔm/	/kɔl/	/ʒɔm/	/ʒɔl/
/dɛv/	/dɛʃ/	/mɛv/	/mɛʃ/
/fum/	/fur/	/sum/	/sur/
/gaf/	/gar/	/ʃaf/	/ʃar/
/ʒɛm/	/ʒɛr/	/sɛm/	/sɛr/
/pɔd/	/pɔz/	/ʒɔd/	/ʒɔz/
/pɔt/	/pɔs/	/fɔt/	/fɔs/
/ʀɔd/	/ʀɔs/	/mɔd/	/mɔs/
/ʀed/	/ʀeʃ/	/ʃed/	/ʃeʃ/
/sɛk/	/sɛl/	/ʀɛk/	/ʀɛl/
/kav/	/kal/	/pav/	/pal/
/mɔʀ/	/mɔl/	/nɔʀ/	/nɔl/
/mɔt/	/mɔʃ/	/gɔt/	/gɔʃ/
/nɔv/	/nɔʃ/	/tɔv/	/tɔʃ/
/pag/	/pas/	/ʀag/	/ʀas/
/pɛd/	/pɛʃ/	/kɛd/	/kɛʃ/
/tap/	/tal/	/zap/	/zal/
/vaʀ/	/val/	/ʒaʀ/	/ʒal/