European Journal of Cognitive Psychology

ISSN: 0954-1446 (Print) 1464-0635 (Online) Journal homepage: http://www.tandfonline.com/loi/pecp20

Reading strategies in orthographies of intermediate depth are flexible: Modulation of length effects in Portuguese

César F. Lima & São Luís Castro

To cite this article: César F. Lima & São Luís Castro (2010) Reading strategies in orthographies of intermediate depth are flexible: Modulation of length effects in Portuguese, European Journal of Cognitive Psychology, 22:2, 190-215, DOI: 10.1080/09541440902750145

To link to this article: http://dx.doi.org/10.1080/09541440902750145

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
2010, 22 (2), 190–215

Reading strategies in orthographies of intermediate depth are flexible: Modulation of length effects in Portuguese

César F. Lima and São Luís Castro
Speech Lab of the Faculty of Psychology and Education, University of Porto, Portugal

This paper examines the role of grapheme–phoneme conversion for skilled reading in an orthography of intermediate depth, Portuguese. The effects of word length in number of letters were determined in two studies. Mixed lists of five- and six-letter words and nonwords were presented to young adults in lexical decision and reading aloud tasks in the first study; in the second one, the length range was increased from four to six letters and an extra condition was added where words and nonwords were presented in separate, or blocked, lists. Reaction times were larger for longer words and nonwords in lexical decision, and in reading aloud mixed lists, but no effect of length was observed when reading words in blocked lists. The effect of word length is thus modulated by list composition. This is evidence that grapheme–phoneme conversion is not as predominant for phonological recoding in intermediate orthographies as it is in shallow ones, and suggests that skilled reading in those orthographies is highly responsive to tasks conditions because readers may switch from smaller segment-by-segment decoding to larger unit or lexicon-related processing.

Keywords: Grapheme–phoneme conversion; Length effect; Orthography; Portuguese; Skilled reading.

A central question in the study of reading is how the orthographic properties of a language influence the neurocognitive processes involved in visual word recognition. Are reading processes largely universal, or are they dependent on how orthography represents the speech stream? Orthographies differ in the

---

Correspondence should be addressed to São Luís Castro, Laboratório de Fala, Faculdade de Psicologia e de Ciências da Educação, Universidade do Porto, Rua Doutor Manuel Pereira da Silva, P 4200–392 Porto, Portugal. E-mail: slcastro@fpce.up.pt

The support from the University of Porto through the grant “Programa de Investigação Científica na Pré-Graduação 2006 UP/CGD” is gratefully acknowledged. César Lima holds a doctoral fellowship from the Portuguese Foundation for Science and Technology (SFRH/BD/39306/2007). We thank Francesca Peressotti for helpful discussions, and also Kathleen Rastle and one anonymous reviewer for their comments on an earlier version of the manuscript.

http://www.psypress.com/ecp DOI: 10.1080/09541440902750145

consistency and complexity of the mappings between letters and sounds (Liberman, Liberman, Mattingly, & Shankweiler, 1980). In shallow orthographies, such as those of Greek, Italian, or Spanish, the orthographic and phonemic codes are isomorphic, i.e., the correspondences between graphemes and phonemes are regular and unambiguous. By contrast, in deep orthographies such as English the relation between graphemes and sounds is not transparent: the same grapheme may represent different sounds in different contexts, and the same sound can be spelled with different letters or groups of letters (e.g., Frost, 1994; Frost, Katz, & Bentin, 1987). How these differences in orthography may impact on visual word recognition has been recently analysed in cross-language neuroimaging and behavioural studies. Paulesu and colleagues (2000) had English and Italian adults read aloud similar words and nonwords in their native languages. Italian readers had shorter reaction times for both types of stimuli than English readers, and their brain activations were greater in the left superior temporal regions that are associated with phoneme processing. English readers, in contrast, had a larger advantage of words over nonwords than Italian readers, and their activations were greater in the left posterior inferior temporal gyrus and in the anterior inferior frontal gyrus, areas that are associated with word retrieval. In a behavioural cross-language study, Seymour, Aro, Erskine, and the COST Action A8 network (2003) compared reading acquisition in 13 European orthographies. The rate of progress in English was found to be much slower than in shallow orthographies. Whereas children learning Finnish, German, Greek, Italian, and Spanish scored close to 100% accuracy in word and nonword reading at the end of the first grade, children learning English remained at 34% correct for words and 29% correct for nonwords.

Cross-language differences in reading were first taken into account in the orthographic depth hypothesis (Frost et al., 1987; Katz & Feldman, 1983). This hypothesis is based on the dual-route theoretical framework, according to which there are two basic pathways for reading: a phonological route in which letter strings are segmented, and then serially converted into sounds using grapheme–phoneme correspondence rules; and a lexical route that maps directly whole-word orthographic configurations to word phonology by accessing word knowledge stored in the mental lexicon. The orthographic depth hypothesis postulates that readers adapt their reliance on one or another of these routes depending on the orthographic properties of the language. In shallow orthographies the phonological route would be preferred since the mappings between letters and sounds are relatively direct and consistent; in deep orthographies readers would be encouraged to use the direct lexical/orthographic route because grapheme–phoneme correspondences are often equivocal (e.g., Frost, 1994). Although this view was until recently the most prominent framework for interpreting cross-language differences in reading, its fundamental prediction that the phonological effects are reduced in deep orthographies has been challenged (for a review, see Frost, 1998). Rayner, Sereno, Lesch, and Pollatsek (1995), for example, observed strong phonological priming effects in an eye movement paradigm in English, thus demonstrating that there is an early automatic activation of phonological codes when reading a deep orthography.

Recently, the psycholinguistic grain size theory (Goswami & Ziegler, 2006; Ziegler & Goswami, 2005, 2006) has built upon the orthographic depth hypothesis. Instead of a dichotomous opposition between reading routes (lexical phonology or word-as-a-whole vs. grapheme–phoneme conversion), it assumes a continuous differentiation of the phonological units involved in reading. According to this perspective, cross-language variation reflects differences in the phonological recoding strategies that are developed in response to orthography. (“Phonological recoding” is used to refer to the mapping from orthography to phonology; e.g., Share, 1995). From a developmental point of view, children can succeed in reading shallow orthographies by relying almost exclusively on phonological recoding at the grapheme–phoneme level, because grapheme/phoneme correspondences are simple and direct. In deep orthographies, however, using these small units of phonological recoding would likely result in mispronunciations and thus children are led into converting print to speech by using larger chunks, like patterns of letters, rhymes, syllables, or even whole words. Learning to read in nonshallow orthographies would involve phonological processing of differently sized units, or “grain sizes”, and would require multiple phonological recoding strategies instead of grapheme–phoneme conversion only. This would explain why learning to read in English proceeds more slowly than in more transparent orthographies, as Seymour and colleagues (2003) have shown. It could also explain why the characteristics of the reading lists modulate the performance of English children, but not of German ones: It is probably due to the cost of switching between smaller and larger units of processing that English children performed worst when the lists included two types of nonwords, those that could be read by grapheme–phoneme conversion and those that required larger units of phonological recoding, whereas German children’s performance was unaffected by list composition (Goswami, Ziegler, Dalton, & Schneider, 2003).

The psycholinguistic grain size theory proposes that skilled reading cannot be dissociated from its developmental history: it is learning to use different phonological recoding units as function of orthography that leads to differences in word processing by adults (Ziegler & Goswami, 2005). In shallow orthographies skilled readers are expected to rely almost exclusively on small units of phonological recoding (grapheme–phoneme conversion), whereas in deep orthographies readers are expected to use larger units (e.g., whole word; bodies and rhymes; syllables) or multiple units varying in size. Empirical support in favour of this proposal comes from a study comparing

English and German adults reading identical words and nonwords in their own languages (Ziegler, Perry, Jacobs, & Braun, 2001). The effect of length manipulated in number of letters was used as a marker for grapheme–phoneme conversion, and body-rhyme effects were used as a marker for conversion of larger units. Length effects were stronger in German than in English, and the reverse occurred for body-rhyme effects. This pattern of results shows that readers of shallow orthographies (such as German) rely predominantly on the conversion of smaller rather than larger units, while readers of deep orthographies (such as English) rely comparatively more on larger rather than smaller units. A similar pattern was obtained in a comparison of Arabic, a deep orthography, with French, that is relatively less deep. Using event-related brain potentials, Simon, Bernard, Lalonde, and Rebai (2006) were able to show that adult skilled readers in French displayed a specific component that is taken to mark the use of grapheme–phoneme conversion, the N320 (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999), but Arabic readers did not. These results suggest that skilled reading in French is more dependent on grapheme–phoneme conversion than is skilled reading in Arabic.

Length effects are a good behavioural signature of the recoding of small units. If reading is achieved by converting graphemes into phonemes, then reaction times should increase as a function of the number of letters in words and nonwords since there are more chunks to be processed. Several studies examining length effects have supported the hypothesis that skilled reading in shallow orthographies is achieved by decoding grapheme–phoneme correspondences. For example, in Italian, Peressotti and Mulatti (2005) compared five- and six-letter long words in lexical decision and reading aloud tasks, and found that longer words took more time to be processed than shorter ones. Thus, one letter is enough to elicit significant length effects in a shallow orthography. Another study with Italian (Spinelli et al., 2005) compared children of different ages without dyslexia, children with dyslexia, and young adults in reading words ranging from three to eight letters. Although length effects were stronger in children, particularly in those with dyslexia, they also occurred in adults (with five- to eight-letter long words). The ubiquity of length effects in Italian is corroborated by results demonstrating that the number of letters is a significant predictor of reading latencies in adults (Bates, Burani, D'Amico, & Barca, 2001). In another shallow orthography, Spanish, a study aiming to identify which variables determine reading latency in skilled readers (Cuetos & Barbón, 2006) revealed that length, measured in number of letters and syllables, was one of the two best predictors of reading latencies (the other was subjective age of acquisition).

Findings about length effects in deep orthographies unveil a different picture. For example, Weekes (1997) observed that reaction times are as fast for six-letter as for three-letter words when reading aloud in English, a result suggesting that units larger than the grapheme/phoneme are used. The notion that words are processed in parallel through the lexical route is embedded in computational models of reading such as the dual route cascaded model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). This model predicts an interaction between length and lexicality, in that length is expected to affect nonword—but not word—processing. The absence of length effects has indeed been reported for lexical decision in English (Baayen, Feldman, & Schreuder, 2006; Balota, Cortese, Sergent-Marshall, Sipeler, & Yap, 2004; but see New, Ferrand, Pallier, & Brysbaert, 2006), a finding that corroborates that reading in deep orthographies is not strictly based on grapheme–phoneme conversion. Note, however, that length effects have been found in reading aloud (Baayen et al., 2006; Balota et al., 2004); these effects emerged in the context of powerful experimental designs, in which the number of items tested and the length range were extremely large. On the whole, then, length effects are clear and ubiquitous in shallow orthographies, showing that grapheme–phoneme conversion enjoys a dominant status; and length has residual or no impact in deep orthographies, probably because readers use larger units of phonological recoding or resort to multiple phonological recoding strategies.

Given that most studies about cross-language differences in reading are focused on orthographies in opposing extremes, much less is known about what happens in intermediate regions of orthographic depth. Results from reading aloud in French, not as deep as English but far from transparent (e.g., Ziegler, Jacobs, & Stone, 1996), suggest that reliance on grapheme–phoneme conversion is modulated by task conditions. Content and Peereman (2003) observed that, in comparison with lists composed exclusively by words, when the reading lists included nonwords, thus presumably engaging smaller units of processing, the effects of length and regularity were enhanced, whereas the effect of frequency was attenuated. We study here another orthography of intermediate depth, European Portuguese. Several orthographic and phonetic properties concur to characterise European Portuguese as an orthography of intermediate depth (Castro & Gomes, 2000; Fernandes, Ventura, Querido, & Morais, 2008; Seymour et al., 2003). There are only five vowel letters for 14 vocalic phonemes, so vocalic grapheme–phoneme correspondences are not trivial. Syllabic division adds another source of difficulty: although CV syllables are predominant in Portuguese (Gomes, 2001; Vigário & Falé, 1994), syllable boundaries in speech are often not clear due to vocalic reduction. Unstressed syllables, notably in word final position, are reduced in fluent speech to the point of being acoustically inexistent (Azevedo, 2005; Delgado-Martins, 2002; Mateus & Andrade, 2000). In words ending with the letters <e> or <o>, the final vowel of the canonical CV syllable is suppressed in speech such that, for example, words such as "vale", *valley*, are orthographically disyllabic and phonologically monosyllabic, /val/ (more examples are: "pele", *skin*, /pɛl/; "ministro", *minister*, /mniʃtr/). Thus, when reading in European Portuguese some vowels letters are not pronounced, and the syllabification of written words is not commensurate with fluent speech. Consonant grapheme–phoneme correspondences also exhibit some complexity, including digraphs (e.g., <lh> corresponds to /λ/, <ch> to /ʃ/), contextual regularities (e.g., <s> corresponds to /ʃ/ at the end of words or syllables, to /s/ at the beginning of words, and to /z/ in intervocalic position), and cases of arbitrary matching (e.g., <x> is read as /ʃ/ in "taxa", *fee*, as /z/ in "exacto", *exact*, and as /ks/ in "taxi", *cab*). Research on reading acquisition and developmental dyslexia in European Portuguese is consistent with this view. In the European cross-language study (Seymour et al., 2003), Portuguese children were far from the near ceiling performance obtained for shallow orthographies and clustered with French around 80% correct. In a study on developmental dyslexia in Portuguese, Sucena, Castro, and Seymour (in press) observed that Portuguese children present a pattern different from what has been described for shallow orthographies (where disfluency affects latencies, not accuracy) and deep orthographies (where accuracy is impaired). Accuracy for reading simple words with 1-to-1 grapheme–phoneme correspondences was not impaired in dyslexic 9-year-old children, but it was substantially impaired for complex words. It appears that dyslexia in Portuguese does not preclude the development of decoding abilities at the level of normal beginning readers (as for shallow orthographies), but the development of the orthographic lexicon is impaired such that not only reaction times but also accuracy is affected.

The goal of this paper is to examine how the intermediate position of Portuguese in the transparency–opacity continuum influences skilled reading. Is phonological recoding predominantly based on grapheme–phoneme conversion, or do larger reading units play a role? How flexible is the recruitment of different units of phonological recoding? In order to address these questions we conducted two studies using reading aloud and lexical decision tasks. The effect of length measured in number of letters was taken as marker of phonological recoding at the grapheme/phoneme level. For nonwords, we predicted robust length effects since the absence of lexical representations biases towards the use of smaller units. For words, if in Portuguese the optimal phonological recoding strategy is the grapheme/phoneme (as in shallow orthographies), length effects should be consistently observed. If, however, the intermediate orthography of Portuguese forces readers to use larger units of phonological recoding, length should have minimal or no impact in word reading. The flexibility of reading strategies will be addressed in the second study by manipulating list composition (words and nonwords mixed, or blocked). By comparing our results with those obtained for Italian (Peressotti & Mulatti, 2005) and for Spanish (Cuetos & Barbón, 2006), we aim to clarify whether skilled reading in an intermediate orthography conforms to the pattern observed for transparent orthographies, or differs from it. If the results for Portuguese are different from shallow orthographies, this would support the hypothesis that the basic cognitive processes involved in visual word recognition differ across languages in a graded manner, that is, even when these are not situated in opposing extremes of the transparency–opacity continuum.

STUDY 1

In this study we determined the effects of length in reading aloud and lexical decision. Two groups of college students read mixed lists of words and nonwords varying in length (five and six letters) and frequency (high and low). One group performed the reading aloud task, and the other the lexical decision task. This design and conditions were chosen so that we were able to compare our results with Peressotti and Mulatti's (2005) for Italian.

Methods

Participants. Sixty students at the University of Porto volunteered to take part in this study (mean age = 23.2 years, $SD = 7.6$); 31 performed the reading aloud task, and 29 the lexical decision task. All were native speakers of European Portuguese and skilled readers with no known reading disorders, and had normal or corrected-to-normal vision.

Materials. The stimuli consisted of 200 items, 100 Portuguese words and 100 pronounceable nonwords. All were orthographic disyllables that started with a consonant and ended with a vowel. Words were selected from Porlex, a lexical database for European Portuguese (Gomes & Castro, 2003). Five- and six-letter long words, 50 each, were chosen so that half in each group would be of high frequency ($M = 858.6$ counts per million) and half of low frequency ($M = 46.6$) while keeping neighbourhood density similar across conditions (see Table 1). Frequency values were taken from the Corlex frequency database (Bacelar do Nascimento et al., 2007). All words, including low frequency ones, were relatively common, and similar in syllable structure and grapheme–phoneme consistency (see Appendix 1 for the complete list of stimuli). Nonwords were derived by replacing the fourth letter of each word. Thus, they also fell into four groups according to length and frequency (baseword frequency). Words and nonwords were pseudorandomised and divided into two blocks of trials, each block including words and nonwords (mixed blocks). The presentation order of the blocks was counterbalanced across subjects.

TABLE 1
Characteristics (mean values) of words used as stimuli in Study 1

	High frequency		*Low frequency*	
	5 letters	*6 letters*	*5 letters*	*6 letters*
Written frequency	844	874	48	46
Orthographic neighbours	3.12	2.84	2.72	2.56
Number of phonemes	4.92	5.2	4.68	5.36

Procedure. Participants were tested individually in one experimental session lasting about 15 minutes. The presentation of the stimuli and recording of response reaction times (RTs) were controlled from SuperLab V4.0 (Abboud, Schultz, & Zeitlin, 2006) running on a Macintosh PowerBook G4 computer. In the reading aloud experiment, vocal RTs were collected by a microphone placed in front of the participants' lips and connected to a Cedrus SV-1 voice-key. In the lexical decision task, RTs were collected using the keys "<" and "-" with the labels "Sim" ("yes"), and "Não" ("no"). Items were presented in lowercase Gill Sans characters, size 110, on the centre of the computer screen. Participants were told that the letter strings appearing on the screen would be real words or invented ones. In the reading aloud task, they were instructed to read them aloud as quickly and accurately as possible; in the lexical decision task, they were instructed to respond whether the letter strings were real words or not by pressing the Yes or No keys, as quickly and accurately as possible. The position of the Yes and No keys in the keyboard was counterbalanced across participants. In order to familiarise them with the task, the session started with 20 practice trials after which the first block was presented; a short pause was made before the second block. RTs were measured from the appearance of the item on the screen until the participant began to respond, and errors were recorded by the experimenter.

Results

Reading aloud

Mean RTs in milliseconds (ms) and error rates in percentages are presented in Table 2. Errors corresponded to 2.2% of the responses. Voice-key misfirings and outliers (RTs below 200 ms, above 2000 ms, or deviating more than 3 *SD* from the grand mean) were discarded from RT analyses (2.5% of correct responses). Words and nonwords were analysed separately and submitted to analyses of variance (ANOVAS) by subjects (F_1) and by items (F_2). For the F_1 analyses, frequency (high and low; baseword frequency for nonwords) and length (five and six letters) were treated as within-subjects

TABLE 2
Mean latencies by condition (in ms) for reading aloud and lexical decision experiments in Study 1

	Reading aloud			
	Words		*Nonwords*	
Frequency	*5 letters*	*6 letters*	*5 letters*	*6 letters*
High	662 (0.1)	677 (0.3)	705 (2.9)	725 (4.4)
Low	681 (0.8)	691 (1.3)	699 (2.6)	731 (5.2)
	Lexical decision			
High	661 (1.9)	662 (1.7)	896 (7.5)	952 (8.2)
Low	809 (15.7)	826 (17.0)	900 (6.1)	1008 (9.8)

Values in parentheses represent error rates in percentages.

factors. For the F_2 analyses, frequency and length were treated as between-subjects factors. A three-way ANOVA with the factors lexicality, frequency, and length was also carried out to examine the effect of lexicality and possible interactions.

Words. Latencies were smaller for shorter words (671 ms) than for longer ones (684 ms), and for high frequency words (669 ms) than for low frequency ones (686 ms), as indicated by the main effects of length and frequency, both significant in the F_1 analyses: respectively, $F_1(1, 30) = 24.03$, $p < .0001$, $\eta^2_p = .45$; $F_2(1, 96) = 2.84$, $p = .09$, $\eta^2_p = .03$; and $F_1(1, 30) = 27.63$, $p < .0001$, $\eta^2_p = .48$; $F_2(1, 96) = 2.78$, $p = .09$, $\eta^2_p = .03$. The interaction was not significant, $F\text{s} < 1$. The analysis of errors only revealed an effect of frequency: high frequency words were read more accurately (0.2%) than low frequency ones (1%), $F_1(1, 30) = 11.9$, $p < .01$, $\eta^2_p = .28$; $F_2(1, 96) = 7.90$, $p < .01$, $\eta^2_p = .08$.

Nonwords. Length determined nonword latencies: Five-letter nonwords were read faster (702 ms) than six-letter nonwords (728 ms), $F_1(1, 30) = 30.02$, $p < .0001$, $\eta^2_p = .5$; $F_2(1, 96) = 10.31$, $p < .01$, $\eta^2_p = .08$. The baseword frequency effect was not significant, $F\text{s} < 1$. The ANOVA on errors had similar results: fewer errors were made for shorter (2.7%) than for longer nonwords (4.8%), $F_1(1, 30) = 5.52$, $p < .05$, $\eta^2_p = .16$; $F_2(1, 96) = 4.75$, $p < .05$, $\eta^2_p = .05$, and no other effects reached significance.

The ANOVAs including lexicality showed an advantage of words (677 ms; 0.6%) over nonwords (715 ms; 3.8%), both in latency, $F_1(1, 30) = 73.63$, $p < .0001$, $\eta^2_p = .71$; $F_2(1, 192) = 35.83$, $p < .0001$, $\eta^2_p = .16$, and in error data, $F_1(1, 30) = 40.18$, $p < .0001$, $\eta^2_p = .57$; $F_2(1, 192) = 40.34$, $p < .0001$, $\eta^2_p = .17$.

The Lexicality × Length interaction was marginally significant, since the effect of length was slightly larger for nonwords (effect: 26 ms; 2.1%) than for words (effect: 13 ms; 0.35%); latencies, $F_1(1, 30) = 8.56$, $p < .01$, $\eta_p^2 = .22$; $F_2 < 1$; errors: $F_1(1, 30) = 3.78$, $p = .06$, $\eta_p^2 = .11$; $F_2(1, 192) = 3.03$, $p = .08$, $\eta_p^2 = .02$.

Lexical decision

RTs and error rates for lexical decision can be seen in Table 2. Errors amounted to 8.3%. RTs due to equipment failure or outliers were discarded from further analyses (1.3% of correct responses).

Words. Even though low frequency shorter words were responded to faster (809 ms) than longer words of the same frequency (826 ms), neither the effect of length nor the interaction with frequency reached significance, Fs < 1. There was an advantage of high frequency words (662 ms) over low frequency ones (818 ms), $F_1(1, 28) = 67.86$, $p < .0001$, $\eta_p^2 = .71$; $F_2(1, 96) = 97.73$, $p < .0001$, $\eta_p^2 = .52$. The analysis on errors yielded similar results: there was only an effect of frequency, with a clear advantage of high frequency (1.8%) over low frequency words (16.3%), $F_1(1, 28) = 117.99$, $p < .0001$, $\eta_p^2 = .81$; $F_2(1, 96) = 25.22$, $p < .0001$, $\eta_p^2 = .21$.

Nonwords. In the latency data, the effect of length was significant, $F_1(1, 28) = 22.75$, $p < .0001$, $\eta_p^2 = .45$; $F_2(1, 96) = 23.66$, $p < .0001$, $\eta_p^2 = .2$, as was the effect of baseword frequency, $F_1(1, 28) = 11.88$, $p < .01$, $\eta_p^2 = .3$; $F_2(1, 96) = 2.94$, $p = .09$, $\eta_p^2 = .03$: shorter nonwords were processed faster (898 ms) than longer nonwords (980 ms) and there was an advantage of nonwords with high baseword frequency (924 ms vs. 954 ms for low baseword frequency words). Moreover, the Length × Baseword frequency interaction was significant: The advantage of shorter items was larger for low-frequency nonwords (effect: 108 ms) than for high-frequency ones (effect: 55 ms), $F_1(1, 28) = 9.10$, $p < .01$, $\eta_p^2 = .25$; $F_2(1, 96) = 3.85$, $p = .05$, $\eta_p^2 = .04$. The analysis of errors showed that accuracy was better for shorter (6.8%) than for longer nonwords (9%), an effect that was significant only in the by subjects analysis, $F_1(1, 28) = 7.12$, $p < .02$, $\eta_p^2 = .20$; $F_2 < 1$.

The global ANOVAs showed that words were responded to faster (740 ms) than nonwords (939 ms), and that the advantage of shorter items was significant for nonwords (effect: 82 ms) but not for words (effect: 9 ms), as can be deduced from the main effect of lexicality, $F_1(1, 28) = 49.06$, $p < .0001$, $\eta_p^2 = .64$; $F_2(1, 192) = 279.62$, $p < .0001$, $\eta_p^2 = .59$, and from the Lexicality × Length interaction, $F_1(1, 28) = 11.06$, $p < .01$, $\eta_p^2 = .28$; $F_2(1, 192) = 8.06$, $p < .01$, $\eta_p^2 = .04$. These effects were not significant in the analysis of errors, Fs < 1.

Discussion

In this study we analysed if a difference of one-letter is enough to elicit length effects in skilled reading in Portuguese. For nonwords, as predicted, it is: five-letter items were processed faster than six-letter ones, both in reading aloud and in lexical decision. This result reflects the use of smaller reading units for nonword processing and corroborates the pattern that has been described for languages with different orthographic depths (e.g., Peressotti & Mulatti, 2005; Weekes, 1997; Ziegler et al., 2001). The frequency effect for words, the lexicality effect, and the interactions of length with lexicality, and of length with baseword frequency, which are predicted by reading models such as the dual route cascaded model (Coltheart et al., 2001), were also observed.

Concerning the key question of whether a one-letter difference in length impacted on word processing, the results depended on the task. The advantage of shorter over longer words was clear in reading aloud but it did not reach statistical significance in lexical decision[1]. Therefore, it is likely that Portuguese readers relied on grapheme–phoneme conversion when reading aloud, as do readers of transparent orthographies (e.g., Cuetos & Barbón, 2006). However, this was not so in lexical decision. The absence of a significant length effect and the occurrence of a baseword frequency effect indicate that lexical knowledge played a dominant role. Frequency, a lexicon-related variable, was influential even when the items were nonwords. This is probably due to the fact that nonwords were derived by only changing one letter of the original words. In Spanish, frequency effects for nonwords have also been observed by Perea, Rosa, and Gómez (2005) when nonwords were similar to words.

Differently from results obtained with a similar design for Italian (Peressotti & Mulatti, 2005), where length effects were found in reading aloud and in lexical decision, our results indicate that in Portuguese length effects are less robust. This might be so because in an orthography of intermediate depth the utilisation of smaller units of phonological recoding is not as predominant as it is in shallow orthographies. Skilled readers would rely on grapheme–phoneme conversion if phonological processing is promoted, as when reading words and nonwords in succession (it is possible to read a word or nonword quickly without profound lexical processing), but they would rely on larger units of recoding when the task biases towards

---

[1] Because longer length is associated with increased occurrence of digraphs and with less mute final <e>s in the stimuli, we performed post hoc analyses to determine whether the pattern of length effects would persist if these two variables were controlled for. We repeated the item analyses considering digraphs and the existence of final <e> as covariates, and we generally replicated the same pattern of results: Length influenced word latencies in reading aloud, ANCOVA, $F_2(1, 94) = 2.98$, $p = .08$, but not in lexical decision, ANCOVA, $F_2 < 1$.

the lexicon. Before coming to this conclusion, though, the following cautionary remarks must be considered. In the experiments conducted in Study 1, words and nonwords in different conditions were not matched for the initial grapheme/phoneme. However, differential effects of onset consistency were observed in reading aloud and lexical decision (Balota et al., 2004), and initial phoneme effects in reading aloud have also been reported (Rastle, Croot, Harrington, & Coltheart, 2005). So, the pattern of results obtained in Study 1 could be due, at least in part, to uncontrolled differences in word onset across experimental conditions. Even though analyses with a subgroup of items matched for initial grapheme and phoneme ($n = 104$) yielded similar results (length effects in reading aloud but not in lexical decision), we conducted a second study where this potential source of artefact was better controlled for.

STUDY 2

The second study had three main goals. The first was to replicate the results obtained of the first study with stimuli matched for the initial grapheme and phoneme. The second was to analyse length effects with a larger difference range, from four to six letters. This manipulation was introduced to maximise the odds of obtaining significant length effects in lexical decision, because a one-letter difference may suffice to elicit task-independent length effects in shallow orthographies, but a larger range may be needed to observe the same effect in intermediate orthographies. The third goal was to explore the flexibility of reading strategies in Portuguese by determining the effects of length in a third condition, reading aloud pure lists, that is, blocked presentation of words and of nonwords. Mixed lists composed of words and nonwords together may encourage the use of grapheme–phoneme conversion and/or nonlexical processing, as Balota et al. (2004) have suggested. Blocked lists containing only words, by contrast, may encourage a stronger reliance on lexical knowledge and on larger units of phonological recoding. A plausible hypothesis is that if grapheme–phoneme conversion is the optimal unit of processing, it will be used always irrespective of list composition; however, if there are multiple units of processing, readers will recruit them flexibly in response to task conditions. Indeed, in French the length effect was increased in a mixed condition of reading aloud as compared to a blocked one, indicating that different units of recoding were used depending on context (Content & Peereman, 1992). Therefore, if the correspondence between graphemes and phonemes is the optimal unit of phonological recoding in Portuguese, it should be used with both mixed and blocked lists, as has been consistently described for shallow orthographies like Italian (Bates et al., 2001; Peressotti & Mulatti, 2005; Spinelli et al., 2005) and

Spanish (Cuetos & Barbón, 2006). If, however, grapheme–phoneme correspondence is not the preferential unit of phonological recoding, and Portuguese readers adopt flexible reading strategies to better respond to task conditions, a differential impact of length depending on presentation context, mixed versus blocked lists, would occur.

Another question of interest regards the effect of lexicality. Kinoshita, Lupker, and Rastle (2004) have shown that, in English, the lexicality effect is larger when the fillers are exception words (blocked lists) than when they are nonwords (mixed lists). The same pattern was recently obtained in Italian, a shallow orthography (Pagliuca, Arduino, Barca, & Burani, 2007). Analysing the modulation of the lexicality effect by list composition in Portuguese will allow us to determine whether this strategic effect also occurs in an orthography of intermediately depth and, if yes, to corroborate its generality across languages.

Methods

Participants. Eighty-two Psychology students from the University of Porto took part in this study for course credits (mean age = 21.01 years, *SD* = 3.9); 35 performed reading aloud with mixed lists, 25 with blocked lists, and 22 performed the lexical decision task.

Materials. Three hundred items, 150 Portuguese words and 150 pronounceable nonwords, were used as stimuli (see Appendix 2). Words were orthographic bisyllables that fell into six groups varying in length (four-, five-, and six-letters long) and frequency (high and low; *M* = 757.8 and *M* = 46.8 counts per million, respectively). The number of orthographic neighbours was kept similar across conditions and high and low frequency words had a similar number of phonemes (see Table 3). Words in the different conditions were matched for initial grapheme and phoneme. Nonwords were formed by changing one letter of the second syllable of the original word, while maintaining syllable structure and consonant or vowel identity (consonants replaced by consonants, vowels by vowels). For reading aloud

TABLE 3
Characteristics (mean values) of words used as stimuli in Study 2

	High frequency			*Low frequency*		
	4 letters	*5 letters*	*6 letters*	*4 letters*	*5 letters*	*6 letters*
Written frequency	764	815	694	49	47	43
Orthographic neighbours	3.1	3.1	2.6	3.2	3	2.3
Number of phonemes	4	4.5	5.3	4	4.7	5.6

with mixed lists and lexical decision, items were pseudorandomised and divided into two blocks each including words and nonwords. In reading aloud with blocked lists, items were divided into two blocks, one containing words and the other nonwords. Within each block, items were pseudo-randomised.

Procedure. The procedure was similar to that of Study 1 and the experimental session lasted about 20 minutes. Participants reading blocked lists were forewarned about the type of list they would be asked to read next.

Results and discussion

Reading aloud: Mixed and blocked lists

Mean RTs and error rates are presented in Table 4. In the mixed lists, the error rate was 2.2%; voice-key misfirings and outliers accounted for 5.3% of correct responses. In the blocked lists, the error rate was 1.8%, voice-key misfirings and outliers were 4% of correct responses. ANOVAS were carried out separately for words and nonwords with list composition (mixed and blocked), frequency, and length as factors. Frequency and length were treated as within-subjects factors in the F_1 analyses, and as between-subjects factors in the F_2 analyses, and the reverse with list composition, mutatis mutandi. A four-way ANOVA including lexicality was also calculated.

TABLE 4
Mean latencies by condition (in ms) for reading aloud with mixed lists, lexical decision, and reading aloud with blocked lists (Study 2)

	Reading aloud mixed lists					
	Words			*Nonwords*		
Frequency	*4 letters*	*5 letters*	*6 letters*	*4 letters*	*5 letters*	*6 letters*
High	587 (0)	592 (0.1)	594 (0.8)	635 (3.1)	635 (4.3)	651 (5.5)
Low	601 (2.0)	601 (1.2)	613 (1.6)	621 (1.5)	626 (2.1)	651 (5.7)
	Reading aloud blocked lists					
High	570 (0.4)	569 (0.3)	561 (0.3)	627 (1.3)	627 (4)	643 (6.7)
Low	577 (0.7)	575 (0.5)	581 (0.7)	613 (0.7)	627 (1.1)	644 (4.2)
	Lexical decision					
High	639 (1.1)	649 (0.9)	692 (1.5)	895 (5.7)	882 (4.7)	943 (5.2)
Low	769 (18.5)	762 (13.7)	795 (13.1)	891 (3.1)	895 (5.7)	941 (7.7)

Values in parentheses represent error rates in percentages.

Words. High frequency words were read faster (579 ms) than low frequency words (591 ms), irrespective of list composition, $F_1(1, 58) = 35.67$, $p < .0001$, $\eta_p^2 = .38$; $F_2(1, 144) = 9.15$, $p < .003$, $\eta_p^2 = .06$; Fs < 1 for the Frequency × List composition interaction. The main effect of length was not significant, Fs < 1, but there was an interaction between list composition and length in the by-subjects analysis, $F_1(2, 116) = 3.36$, $p < .04$, $\eta_p^2 = .05$; $F_2(2, 144) = 1.53$, $p = .2$, $\eta_p^2 = .02$. The number of letters determined reading latencies when the list composition was mixed (effect: 9 ms) but not when the lists were blocked (effect: −2 ms), as illustrated in Figure 1. Separate ANOVAS for mixed and blocked lists revealed a significant effect of length for mixed lists in the by-subjects analysis, $F_1(2, 68) = 4.75$, $p < .02$, $\eta_p^2 = .12$; $F_2 < 1$; post hoc Tukey tests showed that this effect was significant between four- and six-letter words, $p < .02$, marginal between five- and six-letter words, $p = .09$, and nonsignificant between four- and five-letters, $p > .05$. For blocked lists, the effect of length was not significant, Fs < 1. Regarding accuracy, high frequency words elicited less errors (0.3%) than low frequency ones (1.1%), $F_1(1, 58) = 11.64$, $p < .002$, $\eta_p^2 = .17$; $F_2(1, 144) = 10.95$, $p < .002$, $\eta_p^2 = .07$. No other effects or interactions reached significance, Fs < 1.

Nonwords. There was an advantage of shorter (624 ms) over longer nonwords (647 ms), that was confirmed by a significant effect of length, $F_1(2, 116) = 21.2$, $p < .0001$, $\eta_p^2 = .27$; $F_2(2, 144) = 7.43$, $p < .0001$, $\eta_p^2 = .09$. Tukey tests revealed differences between the lengths four and five, $p > .05$, five and six, $p < .0001$, and four and six, $p < .0003$. Nonwords derived from low frequency words were read slightly faster (630 ms) than those derived from high frequency words (636 ms), an effect significant only in the

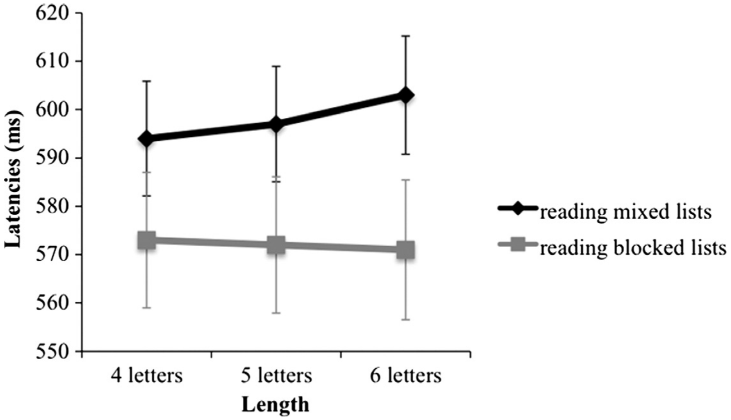

Figure 1. Word reading latencies as function of list composition: words and nonwords (mixed lists) or words only (blocked lists).

by-subjects analysis, F_1 (1, 58) = 9.3, $p < .004$, $\eta^2 = .14$; $F_2 < 1$. No other effects were significant, $F\text{s} < 1$. The error analysis yielded similar results: shorter items elicited less errors (1.6%) than longer ones (5.4%; significant differences between four and five, $p < .05$, and five and six, $p < .002$), and low baseword frequency nonwords less errors (2.5%) than high baseword frequency ones (4.1%), both effects significant in the F_1 analyses: length, $F_1(2, 116) = 32.31$, $p < .0001$, $\eta_p^2 = .36$; baseword frequency, $F_1(1, 58) = 18.14$, $p < .0001$, $\eta_p^2 = .23$; and marginally significant in F_2 analyses, $F_2(2, 144) = 6.97$, $p < .02$, $\eta_p^2 = .09$, and $F_2(1, 144) = 3.1$, $p = .08$, respectively.

The ANOVAs including lexicality confirmed a robust advantage of words (585 ms; 0.7%) over nonwords (613 ms; 3.3%) for RTs, $F_1(1, 58) = 92.73$, $p < .0001$, $\eta_p^2 = .62$; $F_2(1, 288) = 187.0$, $p < .0001$, $\eta_p^2 = .39$, and for errors, $F_1(1, 58) = 72.3$, $p < .0001$, $\eta_p^2 = .55$; $F_2(1, 288) = 30.86$, $p < .0001$, $\eta_p^2 = .1$. The Lexicality × Length interaction was also significant in the F_1 analyses: Length influenced nonwords (effect: 23 ms; 3.8%) but not words (effect: 3 ms; 0.1%); for latencies, $F_1(2, 116) = 13.81$, $p < .00001$, $\eta_p^2 = .19$; $F_2 < 1$, for errors $F_1(2, 116) = 22.14$, $p < .0001$, $\eta_p^2 = .28$; $F_2 < 1$. Importantly, an interaction between list composition and lexicality was obtained: The lexicality effect was larger when the lists were blocked (effect = 58 ms) than when they were mixed (effect = 38 ms): latencies, $F_1(1, 58) = 3.91$, $p = .05$, $\eta_p^2 = .06$; $F_2(1, 288) = 20.7$, $p < .0001$, $\eta_p^2 = .07$; errors, $F\text{s} < 1$.

Overall, the results in the reading aloud task revealed that the effects of word length and of lexicality are modulated by list composition. Regarding the effects of length for mixed lists, the observed advantage of shorter items replicates the pattern found in Study 1. It indicates that grapheme–phoneme conversion is a preferential strategy used by Portuguese skilled readers when reading aloud words and nonwords in mixed lists. However, the effect of length disappeared when the lists were blocked, a clear evidence that in this condition readers were using larger units of processing. Moreover, the effect of lexicality was larger in the blocked condition than in the mixed one. This result extends to Portuguese similar findings obtained for English (Kinoshita et al., 2004) and Italian (Pagliuca et al., 2007).

Lexical decision

As in Study 1, more errors were observed in lexical decision (7%) than in reading aloud (2% overall). Only 1.5% of correct responses were outliers, which were discarded from further RT analyses.

Words. The effect of word length was significant in the analyses of latencies. Shorter words were processed faster (704 ms) than longer ones (743 ms), $F_1(2, 42) = 14.76$, $p < .0001$, $\eta_p^2 = .41$; $F_2(2, 146) = 3.12$, $p < .05$, $\eta_p^2 = .04$. Post hoc Tukey tests showed that this effect was significant between four- and six-letter words, $p < .0002$, and five- and six-letter words, $p < .0003$, but not between four- and five-letter words. An effect of frequency was also obtained: high frequency words were responded to faster (660 ms) than low frequency ones (776 ms), $F_1(1, 21) = 132.66$, $p < .0001$, $\eta^2_p = .86$; $F_2(1, 146) = 73.46$, $p < .0001$, $\eta^2_p = .33$. The Frequency × Length interaction was not significant, Fs < 1. The analysis of errors showed a significant effect of length in the F_1 analysis, $F_1(2, 42) = 5.3$, $p < .008$, $\eta^2_p = .2$; $F_2 < 1$, that should be interpreted in the context of a Frequency × Length interaction, $F_1(2, 42) = 4.62$, $p < .052$, $\eta^2_p = .18$; $F_2 < 1$. For low frequency words, four-letter words elicited slightly more errors (18.3%) than six-letter words (13.5%). The main effect of frequency was also significant: high frequency words elicited fewer errors (1.1%) than low frequency ones (15.2%), $F_1(1, 21) = 72.7$, $p < .0001$, $\eta^2_p = .78$; $F_2(1, 146) = 53.22$, $p < .0001$, $\eta^2_p = .27$.

Nonwords. In the ANOVA on latencies, only the effect of length was significant. Shorter items were processed more rapidly (908 ms) than longer ones (957 ms), $F_1(2, 42) = 10.19$, $p < .0003$, $\eta^2_p = .33$; $F_2(2, 146) = 3.37$, $p < .04$, $\eta^2_p = .04$; differences were significant between four- and six-letter nonwords, $p < .003$, and five- and six-letter nonwords, $p < .001$, *ns* for four- versus five-letter nonwords. Regarding errors, there was only a significant Frequency × Length interaction in the F_1 analysis, $F_1(2, 42) = 4.66$, $p < .02$, $\eta^2_p = .18$; $F_2 < 1$: For low baseword frequency nonwords, shorter items elicited fewer errors (3.1%) than longer items (7.7%), but for high baseword frequency nonwords no such advantage occurred (corresponding values are 5.7% and 5.2%, respectively).

In the global ANOVAs a significant effect of lexicality was obtained only in the analysis of latencies, $F_1(1, 21) = 35.63$, $p < .0001$, $\eta^2_p = .63$; $F_2(1, 288) = 287.05$, $p < .00001$, $\eta^2_p = .5$; for errors, Fs < 1, confirming that words were responded to faster (718 ms) than nonwords (908 ms).

The effect of length was stronger here than in the lexical decision experiment in Study 1. This difference is probably due to the fact that a larger length range was examined, and also to improved control of the materials. Here, we adopted stricter criteria for selecting the words (matched for initial grapheme and phoneme) and for deriving nonwords (maintaining syllable structure and broad phonetic class of the changed segments).

Summing up the results of Study 2, an effect of length was observed for lexical decision and for reading aloud words and nonwords in mixed lists, but not for reading words in pure lists.[2] These results suggest that, in

---

[2] As for Study 1, we conducted item analyses considering digraphs and final <e> as covariates, and the same pattern of results was obtained: length determined latencies in reading aloud words in mixed lists, ANCOVA, $F_2(2, 142) = 2.85$, $p < .05$, and in lexical decision, ANCOVA, $F_2(2, 142) = 3.3$, $p < .04$, but not when the lists were blocked, ANCOVA, $F_2 < 1$.

Portuguese, the use of grapheme–phoneme conversion is important enough to be manifest in different tasks, notably lexical decision, which largely depends on lexical processing. However, the absence of length effects for words in the context of blocked lists indicates that, differently from shallow orthographies (Bates et al., 2001; Cuetos & Barbón, 2006; Peressotti & Mulatti, 2005; Spinelli et al., 2005), grapheme–phoneme conversion is not the optimal phonological recoding unit. The observed pattern shows that in Portuguese, an intermediately consistent orthography, length effects are modulated by task conditions. When the lists included words and nonwords, promoting the phonological recoding at the smallest linguistic unit, Portuguese readers relied on grapheme–phoneme conversion; when words were presented in a purely lexical environment of words with words only, larger units of phonological recoding were preferred. Another noteworthy result was the modulation of the lexicality effect by list composition (see General Discussion).

GENERAL DISCUSSION

The goal of the experiments reported in this paper was to determine whether smaller reading units such as grapheme–phoneme conversion are a predominant phonological recoding strategy for skilled reading in an orthography of intermediate depth, European Portuguese, as they are in transparent orthographies like Italian or Spanish. We used the effects of length, measured in number of letters, as a behavioural signature of grapheme–phoneme conversion. For nonwords, robust length effects were predicted since nonlexical items can easily be read by converting graphemes into phonemes. For words, we had hypothesised that if grapheme–phoneme conversion strategies were indeed preferential, length effects should be ubiquitous as they have been shown to be in shallow orthographies; however, if this phonological recoding unit is not optimal, then length effects should be less prevalent and responsive to task conditions, thus reflecting the use of larger recoding units or of multiple strategies. We will now review the main findings, and then focus on the modulation of length effects.

The main empirical findings can be summarised as follows. For nonwords, as predicted, a robust effect of length was obtained in lexical decision and in reading aloud irrespective of task conditions (mixed and blocked lists). In reading aloud words, however, the impact of length was dependent on task conditions. In Studies 1 and 2, shorter words were responded to faster than longer ones when the lists were mixed, that is, in the context of words and nonwords in arbitrary succession. However, when the words were presented in the context of words only (Study 2, blocked lists), length did not affect reading latency nor accuracy. Regarding lexical decision, although in the first study the advantage of five- over six-letter items did not reach statistical significance, in the second study length effects became apparent in the range of four- to six-letter items. Effects of word frequency and lexicality were also obtained, replicating for Portuguese findings that have been reported for various languages (e.g., Juphard, Carbonnel, & Valdois, 2004; Pagliuca et al., 2007; Valdois et al., 2006). The interaction between lexicality and length, predicted by models such as the dual route cascaded model (Coltheart et al., 2001) or the CDP+ model of reading aloud (Perry, Ziegler, & Zorzi, 2007), was also observed. Finally, the advantage of words over nonwords was larger when reading blocked lists rather than mixed lists. Thus, the lexicality effect in reading was, like the length effect, responsive to task conditions. The modulation of the lexicality effect by list composition extends to an orthography of intermediate depth a result that was obtained for English, a deep orthography (Kinoshita et al., 2004), and for Italian, a shallow orthography (Pagliuca et al., 2007). Altogether, this is evidence that the modulation of the lexicality effect by subtle task conditions is a general, non language-specific, property of the reading mechanism.

Modulation of length effects by list composition

The fact that the effect of length was dependent on task conditions shows that reading in European Portuguese differs from reading in shallow orthographies. Findings from skilled reading in a close related Romance language, Spanish (Cuetos & Barbón, 2006), and also in Italian (Peressotti & Mulatti, 2005), have revealed effects of length in lexical decision and in reading aloud mixed or blocked lists. In these studies, length effects were obtained using stimuli and length ranges that were similar to ours (four- to six-letters in Spanish, five- to six-letters in Italian). The robustness of length effects in these languages is evidence that grapheme–phoneme conversion is the preferred unit of phonological recoding in reading shallow orthographies. In Portuguese, however, length effects were not ubiquitous. Although the number of letters affected visual word recognition in lexical decision and in reading aloud mixed lists, thus revealing reliance on smaller reading units, when words were presented for reading in blocked lists the effect of length disappeared, thus indicating reliance on larger units of phonological recoding. This pattern suggests flexible use of different recoding units according to task demands. When the task encourages the codification of small units, as in the case of mixed lists since nonwords cannot be read without grapheme–phoneme conversion and it is not possible to predict whether the next stimulus is a word or nonword, reading is achieved by grapheme/phoneme conversion. When the task encourages lexical processing, as in the case of word-only lists, grapheme–phoneme conversion no longer drives the reading process. Skilled reading in Portuguese thus involves using different units of phonological recoding, which are flexibly recruited in response to task demands (mixed vs. blocked lists).

The observed flexibility of reading strategies is consistent with the pattern obtained for skilled reading in French (Content & Peereman, 1992). Differently from shallow orthographies, then, grapheme–phoneme conversion is not the optimal unit of phonological recoding for skilled reading in orthographies of intermediate depth, like French and European Portuguese. In line with the results on reading acquisition (Seymour et al., 2003) and on developmental dyslexia (Sucena et al., in press), the present findings extend to skilled reading the evidence that the cognitive processes involved in reading in Portuguese differ from shallow orthographies, and are consistent with the view that Portuguese is indeed an orthography in an intermediate position on the transparency–opacity continuum.

At the behavioural level, the contingency of word length effects may be more apparent in orthographies of intermediate depth because reading is not strongly reliant on grapheme–phoneme conversion. Ziegler and colleagues (2001) refer that the cross-language differences in reading are mainly related to the size of dominant units, the number of different sizes, and the flexibility in using different sizes (p. 383). Our results support the hypothesis that in orthographies of intermediate depth various phonological recoding units are operational and readers are flexible in switching between them in response to task demands.

Implications for studies about length effects

Our results also shed some light on the puzzling scenario that characterises findings about length effects. New et al. (2006) mention how difficult it is to fit them into a clear interpretative pattern. On the basis of the present results, we suggest that some of the variability can be explained by the fact that length effects are both task dependent and language dependent. Length effects will vary depending on whether the task encourages the codification of smaller or larger linguistic units, and also as a function of how direct and unequivocal is the mapping between graphemes and phonemes. Our results demonstrate how a change in list composition can lead to different results, which may appear contradictory at first but that can be understood by taking into consideration the two parameters mentioned. Combining an experimental paradigm that biases towards the lexicon (words-only lists) with an intermediately consistent orthography (Portuguese) results in the absence of length effects; if the experimental paradigm biases towards grapheme–phoneme conversion (nonwords mixed with words), length effects emerge even in an intermediate orthography because grapheme–conversion is more reliable than it is in deep orthographies.

Besides task and orthographic depth, a third aspect that should be considered is the power of the experimental design. Studies with English have uncovered length effects in reading aloud with thousands of items varying between two and seven letters (Baayen et al., 2006; Balota et al., 2004). Indeed, small phonological recoding units at the grapheme–phoneme level are also used in deep orthographies. However, since they are ineffective for successful reading, only a powerful experimental design can pick up their effects. In shallow orthographies length effects are observable with much less items and in a smaller range of length differences; this by itself shows that grapheme–phoneme conversion is a prevalent reading device in these orthographies.

CONCLUSION

In a set of experiments in an orthography of intermediate depth, Portuguese, it was shown that length effects in skilled word reading are modulated by task conditions. Shorter words were processed more rapidly than longer ones in lexical decision and in reading aloud tasks with mixed lists, but length had no impact in reading aloud words when these were presented in the context of words-only. These results suggest that Portuguese skilled readers use grapheme–phoneme conversion strategies when they are useful to succeed in the task, but rely on larger units of phonological recoding when the context of task biases towards lexical knowledge. Therefore, skilled reading in Portuguese requires a flexible and strategic use of different units of phonological recoding. Comparing our results with those reported for shallow orthographies, it can be concluded that the preferential units of phonological recoding in skilled reading vary across orthographies even when they are not in opposing extremes of the orthographic depth continuum.

Original manuscript received March 2008
Revised manuscript received November 2009
First published online July 2009

REFERENCES

Abboud, H., Schultz, W., & Zeitlin, V. (2006). *SuperLab* (Ver. 4.0) [Computer software]. San Pedro, CA: Cedrus Corporation.

Azevedo, M. A. (2005). *Portuguese: A linguistic introduction*. Cambridge, UK: Cambridge University Press.

Baayen, R., Feldman, L., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *55*, 290–313.

Bacelar do Nascimento, M. F., Casteleiro, J. M., Marques, M. L. G., Barreto, F., Amaro, R., & Veloso, R. (2007). *Corlex: Léxico multifuncional computorizado do Português contemporâneo* [lmcpc_dec.txt]. Retrieved from Centro de Linguística da Universidade de Lisboa Web site: http://www.clul.ul.pt/sectores/linguistica_de_corpus/projecto_lmcpc.php

Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316.

Bates, E., Burani, C., D'Amico, S., & Barca, S. (2001). Word reading and picture naming in Italian. *Memory and Cognition*, *29*(7), 986–999.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*, 235–260.

Castro, S. L., & Gomes, I. (2000). *Dificuldades de aprendizagem da língua materna* [*Language difficulties in children*]. Lisboa, Portugal: Universidade Aberta.

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

Content, A., & Peereman, R. (1992). Single and multiple process models of print to speech conversion. In J. Alegria, D. Holender, J. Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 213–236). Amsterdam: North-Holland.

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology*, *18*(3), 415–436.

Delgado-Martins, M. R. (2002). *Fonética do Português: Trinta anos de investigação*. Lisboa, Portugal: Editorial Caminho.

Fernandes, S., Ventura, P., Querido, L., & Morais, J. (2008). Reading and spelling acquisition in European Portuguese: A preliminary study. *Reading and Writing, 21*, 805–821.

Frost, R. (1994). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 116–129.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*(1), 71–99.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 104–115.

Gomes, I. (2001). *Ler e escrever em Português Europeu* [*Reading and writing in European Portuguese*]. Unpublished doctoral dissertation, University of Porto, Portugal.

Gomes, I., & Castro, S. L. (2003). Porlex, a lexical database in European Portuguese. *Psychologica*, *32*, 91–108.

Goswami, U., & Ziegler, J. (2006). Fluency, phonology and morphology: A response to the commentaries on becoming literate in different languages. *Developmental Science*, *9*(5), 451–453.

Goswami, U., Ziegler, J., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics*, *24*, 235–247.

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition*, *55*, 332–340.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 157–166.

Kinoshita, S., Lupker, S., & Rastle, K. (2004). Modulation of regularity and lexicality effects in reading aloud. *Memory and Cognition*, *32*(8), 1255–1264.

Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137–153). Baltimore: University Park Press.

Mateus, M. H., & Andrade, E. (2000). *The phonology of Portuguese*. Oxford, UK: Oxford University Press.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin and Review*, *13*(1), 45–52.

Pagliuca, G., Arduino, L., Barca, L., & Burani, C. (2007). Fully transparent orthography, yet lexical reading aloud: The lexicality effect in Italian. *Language and Cognitive Processes, 23*, 422–433.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S., et al. (2000). A cultural effect on brain function. *Nature Neuroscience*, *3*(1), 91–96.

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception and Psychophysics*, *67*(2), 301–314.

Peressotti, F., & Mulatti, C. (2005). Reading pseudowords: Base-word frequency effect and length effect [Abstract]. In B. Hommel, G. Band, W. La Heij, & G. Wolters (Eds.), *Proceedings of the XIVth meeting of the European Society for Cognitive Psychology* (p. 99). Leiden, The Netherlands: Department of Cognitive Psychology, Leiden University.

Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273–315.

Rastle, K., Croot, K., Harrington, J., & Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 1083–1095.

Rayner, K., Sereno, S., Lesch, M., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science*, *6*(1), 26–32.

Seymour, P. H. K., Aro, M., Erskine, J., & the COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Share, D. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Simon, G., Bernard, C., Lalonde, R., & Rebai, M. (2006). Orthographic transparency and grapheme–phoneme conversion: An ERP study in Arabic and French readers. *Brain Research*, *1104*(1), 141–152.

Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology*, *27*(2), 217–235.

Sucena, A., Castro, S. L., & Seymour, P. (in press). Developmental dyslexia in an orthography of intermediate depth: The case of European Portuguese. *Reading and Writing*. Epub ahead of print. DOI 10.1007/s11145-008-9156-4.

Valdois, S., Carbonnel, S., Juphard, A., Baciu, M., Ans, B., Peyrin, C., & Segebarth, C. (2006). Polysyllabic pseudo-word processing in reading and lexical decision: Converging evidence from behavioural data, connectionist simulations and functional MRI. *Brain Research*, 149–162.

Vigário, M., & Falé, I. (1994). A sílaba no português fundamental: Uma descrição e algumas considerações de ordem teórica [*The syllable in Portuguese: description and theoretical considerations*]. In Associação Portuguesa de Linguística (Ed.), *Actas do IX Encontro da Associação Portuguesa de Linguística* (pp. 465–478). Lisboa, Portugal: Colibri, Artes Gráficas.

Weekes, B. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology*, *50A*(2), 439–456.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29.

Ziegler, J., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, *9*(5), 429–436.

Ziegler, J., Perry, C., Jacobs, A., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, *12*(5), 379–384.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bi-directional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments and Computers*, *28*(4), 504–515.

APPENDIX 1

Stimuli used in Study 1

Words				*Nonwords*			
High frequency		*Low frequency*		*High baseword frequency*		*Low baseword frequency*	
5 letters	*6 letters*	*5 letters*	*6 letters*	*5 letters*	*6 letters*	*5 letters*	*6 letters*
berço	branco	banha	bastão	berpo	bralco	banfa	bascão
bispo	brinco	bilha	brecha	bismo	brilco	bilua	brelha
caixa	caixão	boina	bronco	caiva	caitão	boipa	broaco
circo	cancro	caule	calhau	cirto	canfro	caute	caliau
cobre	centro	ceifa	caução	coble	cengro	ceica	cautão
dança	ficção	dogma	charro	danta	ficlão	dogna	chabro
doido	filtro	farsa	crente	doigo	filbro	farla	crelte
febre	franco	feudo	crosta	febne	fraico	feuto	cronta
filme	grelha	fisga	desvão	filve	grenha	fisna	desfão
golfe	leilão	ganso	fausto	golte	leicão	ganvo	faucto
golpe	língua	genro	fresta	golme	líntua	genco	frenta
gordo	mancha	jarro	gralha	gorno	manlha	jarbo	granha
justo	montra	lacre	grémio	jusbo	monfra	lacte	grélio
lábio	padrão	leigo	guelra	lábuo	padião	leimo	gueira
lesão	parque	manco	lastro	lesio	pargue	manfo	lascro
marco	pátria	melga	mansão	marno	pátnia	melta	mantão
metro	planta	naipe	neutro	metso	plasta	naire	neucro
ninho	quinta	pauta	pinhão	ninto	quiata	pauda	pinião
pátio	rancho	plebe	plasma	pátuo	ranlho	plete	plaima
pausa	sombra	quina	prenda	pauna	sompra	quiba	presda
regra	tanque	rasgo	quelha	regma	tangue	rasdo	quenha
relva	tensão	repto	quisto	relta	tenlão	repno	quilto
salto	treino	sacro	rasgão	salco	treuno	sacno	rasbão
texto	triste	tango	térreo	texpo	trinte	tanco	térmeo
verde	versão	vulgo	tralha	verfe	verbão	vulfo	tranha

APPENDIX 2
Stimuli used in Study 2

Words

High baseword frequency			*Low baseword frequency*		
4 letters	*5 letters*	*6 letters*	*4 letters*	*5 letters*	*6 letters*
base	balão	bosque	bege	banha	bastão
bebé	berço	branco	bibe	bilha	brando
bife	bispo	brinco	bode	boina	brecha
café	caixa	caixão	caju	caule	calhau
cego	circo	centro	cepo	ceifa	cifrão
cume	cobre	contra	cone	cloro	caução
doce	dança	dragão	demo	dogma	dicção
face	febre	ficção	feno	farsa	franja
fome	filme	filtro	figo	feudo	fraque
fumo	fruto	franco	fole	fisga	fresta
gota	golfe	grelha	galã	ganso	gralha
gozo	golpe	grosso	gare	gorro	grémio
guia	gordo	guarda	grua	grilo	guelra
leão	lábio	leilão	laje	lacre	lastro
lixo	lesão	língua	lupa	leigo	lustro
maçã	marco	mancha	muco	manco	mansão
maré	metro	montra	musa	melga	mescla
nulo	ninho	núcleo	nexo	naipe	neutro
pêra	pátio	padrão	peru	pauta	plasma
puré	pausa	pátria	poça	plebe	prensa
rede	relva	rancho	rubi	rasgo	rasgão
sete	sócio	sombra	sapo	sacro	salmão
táxi	texto	tanque	tabu	tampo	térreo
tese	tigre	triste	tule	tango	tralha
vale	verde	versão	veto	vulgo	vítreo

Nonwords

High baseword frequency			*Low baseword frequency*		
4 letters	*5 letters*	*6 letters*	*4 letters*	*5 letters*	*6 letters*
bafe	baluo	bosgue	beze	balha	basdão
beté	berpo	branto	bine	binha	branfo
bifo	bismo	brinfo	bope	boipa	brenha
cané	caiva	caitão	caji	caute	calhiu
cefo	cirto	cendro	ceto	ceiga	ciftão
cuma	coble	confra	cove	cloto	cautão
dole	danta	dranão	dema	dogna	dicfão
fabe	febne	fictão	fene	farna	franta
foma	filve	filbro	fizo	feuco	frague
funo	frupo	frando	fope	fista	frelta
gote	golte	grenha	ganã	ganfo	granha
gobo	golme	grotso	garo	gorre	grébio

APPENDIX 2 (*Continued*)

Nonwords					
High baseword frequency			*Low baseword frequency*		
4 letters	*5 letters*	*6 letters*	*4 letters*	*5 letters*	*6 letters*
guea	gorno	guarma	grue	gripo	guelpa
leuo	lábuo	leifão	laja	lacde	lasbro
libo	lesio	líntua	lupe	leibo	luspro
matã	marno	manlha	muca	manzo	mantão
mafé	metso	mondra	musi	melfa	mesfla
nule	ninhe	núcteo	nefo	naile	neucro
pêma	pátuo	padlão	penu	pauda	plasta
pufé	pauna	pátnia	pola	plefe	prenfa
redo	relta	ranlho	ruli	rasno	rasfão
seto	sólio	sompra	salo	saclo	salgão
tápi	texco	tangue	tadu	tambo	térleo
tepe	tigle	trisve	tuve	tanfo	tranha
valu	verpe	verfão	vepo	vuldo	vítneo