COGNIT 981

COGNITION

Cognition 74 (2000) 1–12

www.elsevier.com/locate/cognit

Brief article

Graphemes are perceptual reading units

Arnaud Rey[a, b,*], Johannes C. Ziegler[c, d], Arthur M. Jacobs[e]

[a]*Harvard University, Cambridge, MA, USA*
[b]*LEAD-CNRS, Dijon, France*
[c]*Macquarie University, Sydney, Australia*
[d]*CREPCO-CNRS, Aix-en-Provence, France*
[e]*Katholische Universitat Eichstatt, Eichstatt, Germany*

Received 5 January 1999; received in revised form 2 April 1999; accepted 27 November 1999

Abstract

Graphemes are commonly defined as the written representation of phonemes. For example, the word 'BREAD' is composed of the four phonemes /b/, /r/, /e/ and /d/, and consequently, of the four graphemes 'B', 'R', 'EA', and 'D'. Graphemes can thus be considered the minimal 'functional bridges' in the mapping between orthography and phonology. In the present study, we investigated the hypothesis that graphemes are processed as perceptual units by the reading system. If the reading system processes graphemes as units, then detecting a letter in a word should be harder when this letter is embedded in a multi-letter grapheme than when it corresponds to a single-letter grapheme. In Experiment 1A, done in English, participants were slower to detect a target letter in a word when the target letter was embedded in multi-letter grapheme (i.e. 'A' in 'BEACH') than when it corresponded to a single-letter grapheme (i.e. 'A' in 'PLACE'). In Experiment 1B, this effect was replicated in French. In Experiment 2, done in English, this grapheme effect remained when phonemic similarity between the target letter alone and the target letter inside the word was controlled. Together, the results are consistent with the assumption that graphemes are processed as perceptual reading units in alphabetic writing systems such as English or French. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Visual word recognition; Reading units; Graphemes

* Corresponding author. Université de Bourgogne, LEAD-CNRS, 6, Boulevard Gabriel, 21000 Dijon, France. Tel.: + 33-3-80-39-57-22; fax: + 33-3-80-39-57-67.

E-mail address: arnaud.rey@u-bourgogne.fr (A. Rey)

PII: S0010-0277(99)00078-5

1. Introduction

A central question in the domain of reading research is whether readers process individual letters or higher-order letter clusters when reading printed words (e.g. Adams, 1979; Gibson, 1965; Hansen & Rodgers, 1965; McClelland & Johnston, 1977). Particularly in recent years a large number of studies investigated whether this is the case (e.g. Peereman & Content, 1997; Rey, Jacobs, Schmidt-Weigand & Ziegler, 1998; Treiman, Mullennix, Bijeljac-Babic & Richmond-Welty, 1995; Ziegler & Perry, 1998).

One reading unit that has received a lot of attention is the grapheme. Graphemes are commonly defined as the written representation of phonemes (e.g. Berndt, Reggia & Mitchum, 1987; Berndt, Lynne D'Autrechy & Reggia, 1994; Coltheart, 1978) (see also Henderson, 1985 for a discussion). Given this definition, the importance of graphemes for reading is clear: graphemes, as opposed to letters, provide a more direct mapping from orthography to phonology. For instance, in the word BREAD there is no one-to-one mapping between the individual letters B, R, E, A, and D and the phonemes /bred/. In contrast, the graphemes B, R, EA, and D map directly onto their respective phonemes. Accordingly, graphemes can be considered as functional phonographic units (Peereman & Content, 1997).

Two recent studies provided empirical evidence that graphemes influence visual word processing. Rastle and Coltheart (1998) observed that naming latencies were longer for non-words composed of few graphemes (i.e. FOOPH → /fuf/) than for non-words composed of many graphemes (i.e. FROLP → /frolp/), both groups having the same number of letters. In a similar experiment using words and a perceptual identification task, Rey et al. (1998) displayed this effect with English and French low frequency words. Thus, both studies conjointly showed that graphemes affect reading performance (see also Pring, 1981).

However, Rastle and Coltheart (1998) did not conclude that graphemes are reading units or that the reading system processes letter strings grapheme by grapheme. Indeed, using simulations of the dual-route cascaded model (Coltheart, Curtis, Atkins & Haller, 1993), they showed that only a non-lexical procedure that operates letter by letter would predict the grapheme effect. Therefore, Rastle and Coltheart (1998) made a strong theoretical claim: letters, not graphemes, are the functional units of the reading system.

The purpose of the present study is to test the opposite assumption that graphemes, not letters, are the functional units of the reading system. To this end, we used the letter detection procedure. This procedure can be considered as an empirical tool for the investigation of unitization processes (for a review see Healy, 1994). The primary idea of this procedure is the following: if the reading system processes particular letter groups as a whole, i.e. a reading unit, then it should be harder to detect the presence of a target letter when it is embedded in a reading unit than when it is not part of a reading unit. This is because subjects have to split the unit into its constituent letters in order to perform the task. For example, previous letter detection studies reported that subjects tend to miss a target letter (e.g. E) more often in high frequency short words (e.g. THE) than in other words containing the same target letter (e.g. Drewnowski & Healy, 1977). These authors argued that because words like THE are processed as a whole, it is harder to process these stimuli at the letter level.

In the present study, we followed the same logic with graphemes as reading units. For instance, searching for the letter A in the word BEACH (A being embedded in the multi-letter grapheme EA) should be harder than detecting the same letter A in the word PLACE. On the contrary, if letters, not graphemes, are the units of the reading system, then letter detections should not be affected by the grapheme manipulation. In addition, we included a word frequency manipulation to investigate the locus of the unitization process. If the unitization process is sublexical, as suggested by the reading unit assumption, no interaction with frequency should be obtained (e.g. Perfetti & Bell, 1991). Finally, to test the generality of the idea that graphemes are reading units, identical experiments were carried out in both English (Experiment 1A) and French (Experiment 1B).

2. Experiments 1A and 1B

2.1. Method

2.1.1. Participants

Twenty-one Australian-English speakers at Macquarie University participated in Experiment 1A. Ninety French speakers at the University of Provence participated in Experiment 1B.

2.1.2. Procedure

Participants were seated 50 cm in front of the computer screen. They started with a training session consisting of 10 trials. Then, the experiment started with 120 experimental trials presented in a randomized order for each participant. Each trial began with a 700 ms presentation of a target letter in the middle of the screen. Then a fixation mark (':') was presented for 1000 ms and was replaced by a stimulus word, which remained on the screen for 33 ms. The stimulus word was followed by a blank interval of 70 ms. Then a mask consisting of superimposed Xs and Os appeared until participants responded. Participants had to press a 'yes' button if the target letter was in the stimulus word and a 'no' button otherwise. The interval between trials was 1000 ms. The target letter was presented in uppercase and the stimulus word in lowercase. The same procedure was used for all experiments in the present report.

2.1.3. Stimuli and apparatus

The stimulus set was composed of 60 target-present trials (i.e. the target letter was in the word) and 60 target-absent trials. The 60 target-present trials were divided into four lists of 15 stimuli. The four lists corresponded to an orthogonal manipulation of word frequency (i.e. high versus low) and grapheme category (i.e. the target letter was embedded in a multi-letter grapheme, A in BEACH, versus in a single-letter grapheme, A in PLACE). Note that each list was constructed so that the same target letter would occur at the same position for any stimulus quadruple (e.g. Ziegler & Jacobs, 1995). All of the word stimuli were five-letter monosyllabic words. Frequency was estimated using the CELEX and the BRULEX frequency counts (Baayen, Piepenbrock & van Rijn, 1993; Content, Mousty & Radeau, 1990). For example, let A be the target letter. The four matched words in each experimental condition were: (1) HOARD in the low frequency/multi-letter grapheme condition; (2) BRASH in the low frequency/single-letter grapheme condition; (3) BOARD in the high frequency/multi-letter grapheme condition; (4) STAFF in the high frequency/single-letter grapheme condition. In Experiment 1A, the mean frequency values of each word list were 5.1, 5.1, 138.7 and 139.8 occurrences per million, respectively, and in Experiment 1B the mean frequency values were 4, 4, 95.5 and 88.5 occurrences per million, respectively. The stimuli were typed with font Courier, size 24. The complete list of stimuli can be found in Appendix A.

2.2. Results

Mean correct response times and error rates are reported in Table 1. The same trimming procedure and analyses of variance (ANOVAs) were used for all experiments in the present report. That is, the trimming procedure excluded scores smaller than 200 ms and greater than 3 SDs above the participant's overall response time. Analyses of variance (ANOVAs) were conducted using both participants (F_1) and items (F_2) as random factors, treating manipulated variables as within-participant factors.

Table 1
Mean correct response times (RT in ms), percentage of errors (% Err), and the corresponding standard errors (SE) for the four conditions in Experiment 1A (English) and in Experiment 1B (French)

	LF words		HF words	
	Multi-letter grapheme (A in HOARD)	Single-letter grapheme (A in BRASH)	Multi-letter grapheme (A in BOARD)	Single-letter grapheme (A in STAFF)
English (Exp. 1A)				
RT (ms)	562	531	557	542
SE	28	26	30	28
% Err	8.7	10.9	9.9	7.1
SE	1.7	2.7	2.4	2.3
French (Exp. 1B)				
RT (ms)	568	527	566	517
SE	28	23	27	22
% Err	10.3	6.1	9.3	7.5
SE	1.8	1.5	2.1	1.3

2.2.1. Experiment 1A: English

The ANOVAs exhibited a main effect of grapheme category ($F_1(1, 20) = 7.38$, $P < 0.05$; $F_2(1, 56) = 4.21$, $P < 0.05$). That is, letter detection times were longer when letters were embedded in a multi-letter grapheme than when letters were not part of a multi-letter grapheme. There was no significant effect of frequency (all $F < 1$) and the interaction between grapheme category and frequency was not significant ($F_1(1, 20) = 1.91$, $P > 0.1$; $F_2(1, 56) < 1$). No significant effects were obtained in the error data.

2.2.2. Experiment 1B: French

A main effect of grapheme category was also obtained, with letter detection times being longer when letters were embedded in a multi-letter grapheme than when letters were not part of a multi-letter grapheme ($F_1(1, 18) = 27.75$, $P < 0.0001$; $F_2(1, 56) = 4.53$, $P < 0.05$). There was no main effect of word frequency (all $F < 1$) and the interaction between grapheme category and word frequency was not significant (all $F < 1$). The error data showed a marginal effect of grapheme category. That is, subjects made more detection errors when the target letter was embedded in a multi-letter grapheme ($F_1(1, 18) = 3.34$, $0.05 < P < 0.1$; $F_2(1, 56) = 2.27$, $P > 0.1$). Error scores were not affected by word frequency (all $F < 1$) and the interaction between grapheme category and word frequency was not significant (all $F < 1$).

2.3. Discussion

The results of Experiments 1A and 1B can be summarized as follows. First, in two letter search experiments, we obtained longer response times when the target letter was embedded in a multi-letter grapheme (A in BEACH) than when it was not part of a multi-letter grapheme (A in PLACE). Second, this effect was present in both the English and French data. Third, letter detection latencies were not affected by the frequency of the target word in English or in French. Taken together the present data are consistent with the assumption that graphemes are processed as perceptual units by the reading system.

Before accepting such a strong conclusion, however, one alternative interpretation should be considered. Previous letter detection studies have shown that phonemic factors influence a subject's response in the letter detection task (for a review see Healy, 1994). For example, Read (1983) reported that subjects tend to miss the target letter F more often in OF than in control words. Read (1983) attributed this effect to the particular pronunciation of F in OF (i.e. F is pronounced as /v/ in OF). Similarly, in the present study, subjects may take a longer time to detect A in BEACH compared to A in PLACE, not because A is embedded in the multi-letter grapheme EA, but because the sound of the letter A alone is closer to A in PLACE than to A in BEACH. Therefore, one might argue that the present grapheme effect is in fact a phonemic similarity effect.

In French, phonemic similarity and grapheme category are confounded. When a letter is embedded in a multi-letter grapheme then its pronunciation is automatically different from the letter's name (e.g. A in TAUPE belongs to the multi-letter grapheme AU that is pronounced /o/, whereas A in SALLE is pronounced as A alone). In English, however, it is possible to manipulate phonemic similarity and grapheme category in an orthogonal fashion. One can find, for example, single-letter graphemes that are pronounced either in the same way as a letter's name (O in SLOPE) or in a different way (O in PROVE), and also letters embedded in multi-letter graphemes that are pronounced either in the same way as a letter's name (O in FLOAT) or in a different way (O in CLOUD). Therefore, if the grapheme effect obtained in Experiments 1A and 1B was due to phonemic similarity, this effect should not occur when the target letter's name is the same in both single- and multi-letter graphemes. This hypothesis was tested in Experiment 2.

3. Experiment 2

3.1. Method

3.1.1. Participants

Twenty-seven English speakers at Harvard University participated in Experiment 2.

3.1.2. Stimuli and apparatus

The stimulus set was composed of 64 target-present trials and 64 target-absent trials. Grapheme category (multi-letter versus single-letter grapheme) and phonemic similarity (same versus different) were manipulated orthogonally, leading to four lists of 16 stimuli. For example, let O be the target letter. The four matched words in each experimental condition were: (1) SLOPE in the single-letter grapheme/same phoneme condition; (2) FLOAT in the multi-letter grapheme/same phoneme condition; (3) PROVE in the single-letter grapheme/different phoneme condition; (4) CLOUD in the multi-letter grapheme/different phoneme condition. Word frequency and letter position in the word were matched as closely as possible for any stimulus quadruple. The mean frequency values of each word list were 139, 137, 139 and 140 occurrences per million, respectively. The mean letter position in the word was 2.88 for all lists. Word stimuli were five-letter monosyllabic English words. The complete list of stimuli can be found in Appendix A.

3.2. Results

Mean correct response times and error rates are reported in Table 2. As in Experiments 1A and 1B, we obtained a main effect of grapheme category ($F_1(1, 26) = 7.14$, $P < 0.05$; $F_2(1, 60) = 10.43$, $P < 0.01$). That is, letter detection times were longer when letters were embedded in a multi-letter grapheme than when letters were not part of a multi-letter grapheme. There was also a main effect of phonemic similarity that was marginally significant in the item analysis ($F_1(1, 26) = 4.87$, $P < 0.05$; $F_2(1, 60) = 3.21$, $0.05 < P < 0.1$). This effect indicates that letter detections were faster when the pronunciation of the target letter

Table 2
Mean correct response times (RT in ms), percentage of errors (% Err), and the corresponding standard errors (SE) for the four conditions in Experiment 3 (English)

	Same phoneme		Different phoneme	
	Multi-letter grapheme (O in FLOAT)	Single-letter grapheme (O in SLOPE)	Multi-letter grapheme (O in CLOUD)	Single-letter grapheme (O in PROVE)
English (Exp. 2)				
RT (ms)	565	532	567	557
SE	20	15	19	19
% Err	5.2	4.2	5.0	5.4
SE	1.4	0.9	1.3	1.2

in the word was identical to the letter's name. The interaction between grapheme category and phonemic similarity was marginally significant for subjects ($F_1(1,26) = 3.58,\ 0.05 < P < 0.1$) and failed to reach significance for items ($F_2(1,60) = 2.66,\ P > 0.1$). Most importantly, there was still a significant grapheme effect when the name of the target letter was identical across the grapheme category (i.e. O in SLOPE versus O in FLOAT; $F_1(1,26) = 13.35,\ P < 0.01$; $F_2(1,30) = 11.81, P < 0.01$). No significant effects were obtained in the error data.

Therefore, Experiment 2 demonstrates that, while controlling for phonemic similarity, we still obtained a letter detection disadvantage for letters embedded in multi-letter graphemes. This result rules out the interpretation of the grapheme effect strictly in terms of phonemic similarity and is consistent with the assumption that graphemes are processed as perceptual units by the reading system.

The present data also suggest that phonemic similarity between a letter's name and the pronunciation of the letter in the word might partially explain the effect obtained in Experiments 1A and 1B. This result is congruent with previous letter search studies. For example, Schneider, Healy and Gesi (1991) extended the result obtained by Read (1983) and reported that subjects also tend to miss the target letter O more often in OF than in control words. Although the effect obtained for O was weaker than for F, their results clearly showed that both the unitization and the phonemic interpretations have to be combined to account for this effect.

4. General discussion

The main finding of the present study is that detecting a letter embedded in a multi-letter grapheme (i.e. A in BEACH) is harder than detecting the same letter in a single-letter grapheme (i.e. A in PLACE). This result is consistent with the assumption that graphemes are automatically processed by the reading system as perceptual units. Conversely, this result casts doubt on the assumption that individual letters are the functional units of the adult reading system. Furthermore, the finding that word frequency did not influence detection latencies in Experiments 1A and 1B suggests that the grapheme effect is due to an automatic sublexical grouping of letters into multi-letter graphemes. Finally, Experiment 2 shows that the grapheme effect cannot be reduced to a phonemic similarity effect although both unitization and phonemic encoding processes are needed to account for the complete pattern (Healy, 1994; Schneider et al., 1991).

By assuming that graphemes are processed as perceptual units by the reading system, we claim that during the acquisition of reading readers develop internal representations of graphemes. Such internal representations can be seen as the result of recurrent associations between orthographic and phonological patterns. These representations are functional in reading because, contrary to letters, they provide a more direct mapping from orthography to phonology. Once these representations have been established within the reading system, skilled readers can activate them in the same way they activate word representations in current spreading of activation models of reading (e.g. Coltheart et al., 1993; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981).

Given the present empirical evidence in support of graphemes as functional reading units, how can we account for the finding of Rastle and Coltheart (1998) and Rey et al. (1998) that words and non-words with multi-letter graphemes take longer to process than those with single-letter graphemes? If graphemes are functional reading units, why do multi-letter graphemes slow down word recognition? If we assume that graphemes have internal representations, then this apparent contradiction can be solved in a straightforward manner. For letter strings that include a multi-letter grapheme, such as the word BREAD, activation will spread to the graphemes B, R, E, EA, A, and D at a grapheme level of representation. Therefore, the multi-letter grapheme unit EA will compete with the single-letter grapheme units E and A. This competition is time consuming as demonstrated by Rey (1998) and Richter (1999) in simulations with an interactive activation model that contained graphemes as intermediate functional units. In contrast, for words like CRISP, such competition is absent since the word is composed only of single-letter graphemes. Therefore, the assumption that graphemes are reading units can account for the present letter detection data as well as for the grapheme effects found by Rastle and Coltheart (1998) and Rey et al. (1998).

Although we provide evidence in the present study indicating that graphemes are processed as minimal perceptual units, we have not addressed, however, the question of the precise definition of these phonographic units. For example, in English, it has been argued that letter clusters like A plus final E (noted A_E) in words like STAGE should be considered as graphemes (e.g. Coltheart et al., 1993). Indeed, the presence of a final E seems to influence the pronunciation of the middle vowel (A in STAG is pronounced /æ/ whereas A in STAGE is pronounced /eɪ/, e.g. Berndt et al., 1994). One may therefore ask whether the reading system codes these letter clusters in the same way as graphemes with contiguous letters (e.g. EA in BEACH). Alternatively, given the non-contiguity between these letters, it is unclear how the reading system could code them as perceptual units. Although some models assume that they are coded just like graphemes with contiguous letters (e.g. see Rastle & Coltheart, 1999), more computational and empirical work is needed to clearly solve this question.

More generally, a number of studies in different languages (English, Spanish or French) indicate that orthographic units larger than graphemes exist and play a role in reading (e.g. Bowey, 1990; Carreiras, Alvarez & de Vega, 1993; Ferrand, 2000; Ferrand, Segui & Grainger, 1996; Nüerk, Rey, Graf & Jacobs, 2000; Peereman & Content, 1997; Treiman et al., 1995; Ziegler & Perry, 1998). Such units include onsets, nuclei, rimes, or syllables. These different types of functional orthographic units are likely to emerge during reading acquisition as suggested by developmental studies (e.g. Coltheart & Leahy, 1992; Goswami & Bryant, 1990). Their role during reading acquisition could be to increase parallel processing (Laberge & Samuel, 1974) and to enhance phonological recoding of letter strings (Share, 1995).

Acknowledgements

This research was supported by a grant from the Fyssen Fondation and a grant from the French Ministry of Education and Research (#95124, programme Sciences Cognitives) to A.R., together with a grant from the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) to A.M.J. We are grateful to Pierre Barrouillet, Alfonso Caramazza, Albert Costa, Deborah Edelman, Michele Miozzo and Pierre Peruchet for their helpful critical comments.

Appendix A

A.1. Stimuli used in Experiments 1A and 1B

For target-present trials, target letters are followed by four target words belonging to the four experimental conditions: (1) low frequency/multi-letter grapheme; (2) low frequency/single-letter grapheme; (3) high frequency/multi-letter grapheme; (4) high frequency/single-letter grapheme. For target-absent trials, target letters are followed by target-absent words.

A.1.1. Experiment 1A: English

Target-present trials: A, peach, stark, reach, class; A, yearn, tract, learn, black; A, lease, slack, peace, grass; A, leach, graft, teach, glass; A, poach, trash, coast, track; A, hoard, brash, board, staff; E, niece, speck, piece, dress; E, wield, theft, field, press; I, waive, skimp, paint, quick; I, maize, whisk, raise, stick; I, poise, twine, joint, shift; I, hoist, slick, noise, skill; U, rouse, pluck, count, trust; U, mourn, crumb, mouth, truth; U, louse, snuff, youth, stuff.

Target-absent trials: A, bench, brick, brush, chest, cloth, crude, ghost, grind, kneel, lodge, sheer, shore, slice, slope, split, spoil, steel, stiff, tribe, trout, truck, twist, weigh, witch; E, blind, bunch, crush, drain, drift, pitch, proof, smart; I, beard, beast, cease, false, flame, found, gleam, phase, shade, skull, snake, spare, spray, straw, wheat, worst; U, brave, chill, depth, faint, fence, frown, pride, shine, skirt, spell, sport, waist.

A.1.2. Experiment 1B: French

Target-present trials: A, taupe, casse, faute, vague; A, faune, bagne, jaune, basse; A, lange, latte, danse, balle; A, saule, panne, cause, salle; A, stand, stage, franc, grave; I, beige, tripe, neige, crime; I, moite, prime, poids, crise; I, poire, bribe, voile, prise; N, gland, gaine, blanc, peine; N, plant, liane, chant, haine; N, coing, frêne, poing, reine; O, sonde, colle, honte, poche; U, guêpe, pulpe, quart, russe; U, loupe, prune, sourd, plume; U, moule, rhume, douze, chute.

Target-absent trials: A, figue, gueux, louve, meule, bride, brune, douve, clore, flore, frise, globe, gnome, grive, noeud, pouls, poupe, prude, queue, store, vogue; I, baume, chaux, heurt, soude, brave, brume, brute, crabe, gnose, maure, sceau, slave, stock, taule; N, chaud, daube, joute, brise, clerc, glace, grade, proue, sbire, soute, trame, veule; O, sauce, frire, gaule, meute, plage, trace; U, craie, drain, frite, maint, prose, sport, stade, teint.

A.2. Stimuli used in Experiment 2: English

Target letters are followed by four target words belonging to the four experimental conditions: (1) single-letter grapheme/same phoneme; (2) single-letter grapheme/different phoneme; (3) multi-letter grapheme/same phoneme; (4) multi-letter grapheme/different phoneme. For target-absent trials, target letters are followed by target-absent words.

Target-present trials: A, crave, trash, braid, fraud; A, slate, slang, snail, plaid; A, baste, cache, gauge, vault; A, place, small, great, leave; A, trade, class, break, clear; A, shave, flare, steak, bleak; I, spice, crisp, aisle, saint; I, slice, trick, thigh, faith; I, write, bring, light, piece; I, price, stick, fight, raise; O, stone, wrong, throw, shout; O, slope, prove, float, cloud; O, vogue, dodge, poach, pouch; O, globe, frost, groan, stout; O, grope, shove, dough, couch; O, whole, whose, known, group.

Target-absent trials: A, bench, brick, brush, chest, cloth, crude, ghost, grind, kneel, lodge, sheer, shore, split, spoil, steel, stiff, tribe, trout, truck; E, blind, bunch, crush, drain, pitch, proof, smart; I, beard, beast, cease, false, flame, found, gleam, phase, shade, skull, snake, spare, spray, straw; O, share, sheet, truth, greet, trust, taste, fruit, plant; U, brave, chill, depth, faint, fence, frown, pride, shine, skirt, spell, sport, waist.

References

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology*, *11*, 133–176.

Baayen, R. H., Piepenbrock, R. & van Rijn, H. (1993). The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium. Philadelphia, PA: University of Pennsylvania.

Berndt, R. S., Reggia, J. A., & Mitchum, C. C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior Research Methods, Instruments, & Computers*, *19*, 1–9.

Berndt, R. S., Lynne D'Autrechy, C., & Reggia, J. A. (1994). Functional pronunciation units in English words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 977–991.

Bowey, J. A. (1990). Orthographic onsets and rimes as functional units of reading. *Memory and Cognition*, *18*, 419–427.

Carreiras, M., Alvarez, C. J., & De Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*, 766–780.

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

Coltheart, V., & Leahy, J. (1992). Children's and adult's reading of nonwords: effects of regularity and consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 718–729.

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

Content, A., Mousty, P., & Radeau, M. (1990). BRULEX: une base de données lexicales informatisée pour le Français écrit et parlé. *L'Année Psychologique*, *90*, 551–566.

Drewnowski, A., & Healy, A. F. (1977). Detection errors on *the* and *and*: evidence for reading units larger than the word. *Memory and Cognition*, *5*, 154–168.

Ferrand, L. (2000). Reading aloud polysyllabic words and nonwords: the syllabic-length effect re-examined. *Psychonomic Bulletin and Review*, in press.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of words and picture naming: the role of syllabic units. *Journal of Memory and Language*, *35*, 708–723.

Gibson, E. J. (1965). Learning to read. *Science*, *148*, 1066–1072.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*, Hillsdale, NJ: Lawrence Erlbaum.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, *103*, 518–565.

Hansen, D., & Rodgers, T. S. (1965). An exploration of psycholinguistic units in initial reading. *Proceedings of the symposium on the psycholinguistic nature of the reading process*, Detroit, MI: Wayne State University.

Healy, A. F. (1994). Letter detection: a window to unitization and other cognitive processes. *Psychonomic Bulletin and Review*, *1*, 333–344.

Henderson, L. (1985). On the use of the term 'grapheme'. *Language and Cognitive Processes*, *2*, 135–148.

Laberge, D., & Samuel, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

McClelland, J. L., & Johnston, J. L. (1977). The role of familiar units in the perception of words and nonwords. *Perception and Psychophysics*, *22*, 249–261.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, *88*, 375–407.

Nüerk, H. -C., Rey, A., Graf, R., & Jacobs, A. M. (2000). Phonographic sublexical units in visual word recognition. *Current Psychology Letters: Behaviour, Brain & Cognition*, submitted.

Peereman, R., & Content, A. (1997). Orthographic and phonological neighborhood in naming: not all neighbors are equally influential in orthographic space. *Journal of Memory and Language*, *37*, 382–410.

Perfetti, C. A., & Bell, L. C. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming. *Journal of Memory and Language*, *30*, 473–485.

Pring, L. (1981). Phonological codes and functional spelling units: reality and implications. *Perception and Psychophysics*, *30*, 573–578.

Rastle, K., & Coltheart, M. (1998). Whammy and double whammy: length effects in nonword naming. *Psychonomic Bulletin and Review*, *5*, 277–282.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 482–503.

Read, J. D. (1983). Detection of Fs in a single statement: the role of phonetic recoding. *Memory and Cognition*, *11*, 390–399.

Rey, A. (1998). Orthographe et phonologie dans la perception des mots écrits. Doctoral dissertation. Marseille: Université de Provence.

Rey, A., Jacobs, A. M., Schmidt-Weigand, F., & Ziegler, J. C. (1998). A phoneme effect in visual word recognition. *Cognition*, *68*, B71–B80.

Richter, K. (1999). A functional units model of visual word recognition. Diploma thesis. Marburg: Philipps-University.

Schneider, V. I., Healy, A. F., & Gesi, A. T. (1991). The role of phonetic processes in letter detection: a reevaluation. *Journal of Memory and Language*, *30*, 294–318.

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

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology*, *124*, 107–136.

Ziegler, J. C., & Jacobs, A. M. (1995). Phonological information provides early sources of constraint in the processing of letter strings. *Journal of Memory and Language*, *34*, 567–593.

Ziegler, J. C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: resolving neighborhood conflicts in the lexical decision task. *Cognition*, *68*, B53–B62.