Perception & Psychophysics
2002, 64 (5), 785-794

Does "whole-word shape" play a role in visual word recognition?

MANUEL PEREA and EVA ROSA
Universitat de València, València, Spain

To analyze the impact of outline shape on visual word recognition, the visual pattern of the stimuli can be distorted by size alternation. Contrary to the predictions of models that rely on outline shape (Allen, Wallace, & Weber, 1995), the effect of size alternation was greater for low-frequency words than for high-frequency words in a lexical decision task (Experiment 1). In Experiment 2, the effect of case type (lowercase vs. UPPERCASE) occurred for low-frequency words, but not for high-frequency words. The effect of neighborhood size was remarkably similar in the two experiments. The results can be readily explained in the framework of a resonance model (Grossberg & Stone, 1986), in which a mismatch between the original sensory pattern and the abstract orthographic code slows down the formation of a stable percept.

An important (and recurrent) issue in visual word recognition in alphabetic languages is whether words can be formed uniquely on the basis of abstract letter units or, instead, can also be formed on the basis of other sources (e.g., via word global shape). Undoubtedly, this is an issue that has implications for the teaching of both reading and spelling (Besner, 1983). Although early research suggested that words could be identified by the use of word shape (see Cattell, 1886), most theorists currently support the idea that words are initially formed from component letters (*analytical* models; e.g., the search model, Forster, 1976; the multiple read-out model, Grainger & Jacobs, 1996; the interactive-activation model, McClelland & Rumelhart, 1981; the activation-verification model, Paap, Newsome, McDonald, & Schvaneveldt, 1982). In these models, information about visual form is probably lost early in the process of word recognition, so that the particular visual form that a letter takes is *irrelevant* to this process.

Nonetheless, other investigators still argue that supraletter features, such as word shape, play a role in visual word recognition (e.g., Allen, Wallace, & Weber, 1995; Healy & Cunningham, 1992; Healy, Oliver, & McNamara, 1987). For instance, Healy and Cunningham found that the number of proofreading errors was affected by word shape in the lowercase passages, but not in the all-uppercase passages. However, Healy and Cunningham remained neutral regarding what stage this processing of word shape occurs in (an initial stage, a later "verification" stage, or a postaccess stage). The most detailed holistic model is probably Allen et al.'s *holistic biased hybrid* model. In this model, words can be formed either via letter-level codes (as in analytical models) or via word-level codes—in which "the spatial frequency pattern of an entire word" is the basic unit of analysis. Specifically, Allen et al. indicated, on the basis of neuropsychological studies (e.g., Schiller, Logothetis, & Charles, 1991), that the word-level channel uses low-frequency information that relies on the (fast) magnocellular visual pathway, unlike the letter-level channel, which would rely on the (slower) parvocellular channel. This "horserace" model predicts that highly familiar patterns (i.e., high-frequency words) can be identified by the fast and frequency-sensitive word-level channel, whereas low-frequency words will be identified, on many occasions, by the frequency-insensitive letter-level channel (see Figure 1).

Alternatively, Besner and Johnston (1989) proposed a multiple-route model in which word identification occurs only via letter-level codes. Nonetheless, Besner and Johnston suggested that a lexical decision response can be achieved by three routes (see Figure 2): (1) using a visual familiarity assessment (i.e., a "fast guess" via global word shape), (2) using an orthographic familiarity assessment (a "fast guess" based on overall lexical activation in the orthographic lexicon; see also Balota & Chumbley, 1984), and (3) via word identification (on the basis of letter-level codes). We should note that when making a *word* response on the basis of the familiarity mechanisms (Routes 1 and 2), participants are not assumed to be identifying the words. Instead, the visual/orthographic *familiarity* of letter sequences may have a *goodness* dimension, which contaminates word–nonword judgments (Besner & Johnston, 1989).

The main goal of the present series of experiments is to examine whether the effects of word frequency and number of neighbors (*neighborhood size*, or Coltheart's N[1]) are affected by the visual familiarity of the letter-string.

Parts of this research were reported at the 40th Annual Meeting of the Psychonomic Society, Los Angeles, 1999. E. R. was the recipient of a predoctoral grant from the Spanish Ministry of Science and Culture. We thank Sandy Pollatsek, Ken Paap, and Derek Besner for constructive and insightful criticism of earlier versions of the manuscript. Correspondence concerning this paper should be sent to M. Perea, Departament de Metodologia, Facultat de Psicologia, Av. Blasco Ibáñez, 21. 46010-València, Spain (e-mail: mperea@uv.es).

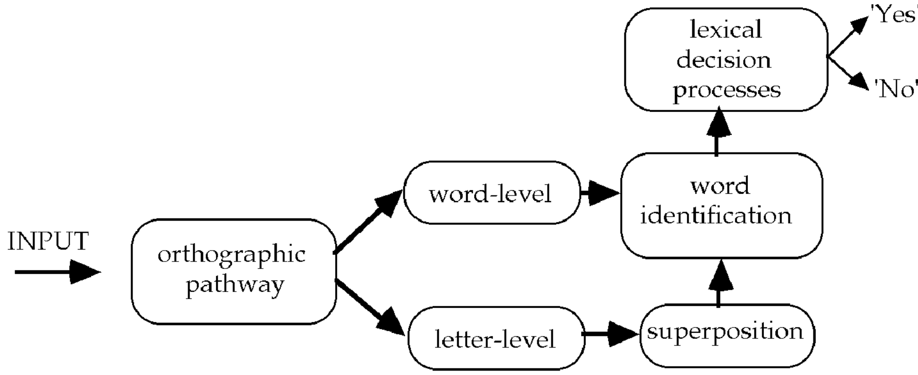

Figure 1. Processes involved in the lexical decision task according to the holistically biased hybrid model (based on Allen, Wallace, & Weber, 1995). The word-level channel encodes the spatial frequency pattern of entire words, and the letter-channel encodes the spatial frequency of individual letters of a given word. Only the word-level channel is frequency sensitive.

The rationale behind these experiments is the following: If the items are presented in a format that minimizes differences in word shape information (or distorts the visual pattern), participants will be less able to rely on whole-word information (or visual familiarity) as a basis for responding (relative to the standard lowercase print), which may affect high- and low-frequency words differentially. In Experiment 1, we chose to manipulate size alternation (instead of a manipulation such as cAsE aLtErNaTiOn), because it is a more direct—and less intrusive—manipulation of visual familiarity. Bear in mind that it is still unclear what steps in the word recognition process are hindered by alternating-case presentations. For instance, the number of possible patterns from which each letter must be recognized is doubled when a pure case condition is switched to mIxEd cAsE (Paap, Newsome, & Noel, 1984). More important, case alternation may also alter the normal mode for lexical access (Forster & Guess, 1996; Mayall & Humphreys, 1996; Mayall, Humphreys, & Olson, 1997; Paap et al., 1984). In Experiment 2, we manipulate the case of the letters: Items are presented either in lowercase or in UPPERCASE. The motivation behind this manipulation is that the outline shape with UPPERCASE LETTERS is less informative than that with the more distinctive, lowercase letters (e.g., in terms of ascending, descending, or neutral letters). Furthermore, if the lexical system takes letter shape into account, it is reasonable to expect that the effect of neighborhood size would be reduced for lowercase presentations: Lowercase letters have more distinct features, which could affect the degree of lexical similarity among word neighbors.

Alternating Size and Word Recognition

In a purely feedforward analytical model, lexical access follows the computation of a set of abstract letter identities, so that the particular visual form that a letter takes should be irrelevant to the process of word identification. By hypothesis, this computation would lead us to expect consistent stimuli to have no advantage over alternating-size stimuli. We must bear in mind that the quality of the

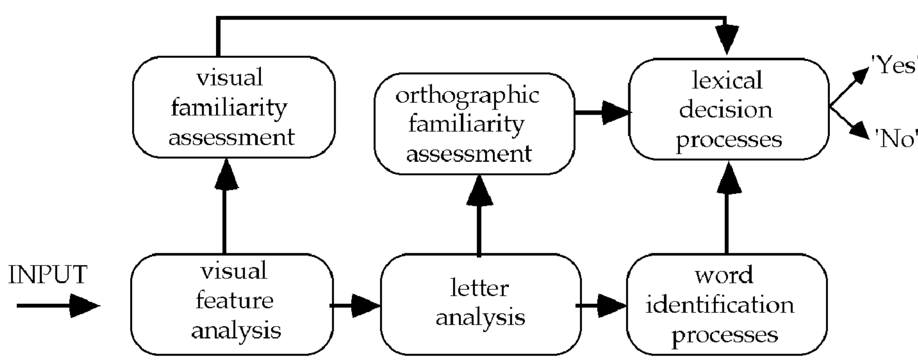

Figure 2. Processes involved in the lexical decision task according to the multiple-route model (based on Besner & Johnston, 1989; Besner & McCann, 1987). Lexical decision responses may be based on unique word identification (via an abstract letter level), on orthographic familiarity (via an abstract *wordlikeness* dimension), or on visual familiarity (via global word shape).

letters is not degraded in any way, and it could well be the case that alternating-size stimuli do not significantly disrupt letter identification. However, there is evidence indicating that naming and reading times are slowed by size alternation (Mayall et al., 1997; Rudnicky & Kolers, 1984; Smith, 1969), which implies that lexical access is hindered by this manipulation. What, then, is the locus of the size alternation effect?

One possibility is that letter identification and word identification overlap in time (i.e., the letter level provides continuous output to the word level, as in a cascade process; see McClelland, 1979) and that, because of pattern distortion, activation in the orthographic lexicon might rise more slowly when stimuli are presented in alternating size (Besner & McCann, 1987). If this is so, there could be less of an effect of letter-size variation on stimuli for which it is easier to contact stored knowledge (i.e., high-frequency words). In this light, it has been suggested that participants in a lexical decision experiment could use the strategy of selecting a rapid *word* response whenever lexical activity in the lexicon is high at the early stages of word processing. This "fast guess" mechanism has been posited as being responsible for the facilitative effects of *N* on low-frequency words (Grainger & Jacobs, 1996; Paap & Johansen, 1994; Pollatsek, Perea, & Binder, 1999; but see Andrews, 1997; Sears, Lupker, & Hino, 1999). (This mechanism is probably not operative with high-frequency words, since these items can be rapidly identified via unique word identification; see Grainger & Jacobs, 1996.) If we assume that activation in the orthographic lexicon rises more slowly when stimuli are presented in alternating size, this "fast guess" mechanism would encounter relatively low levels of early activation in the lexicon, and on such trials, participants might not set the "fast guess" threshold low enough to cause a *word* decision earlier than would occur with the other way of deciding *word* (when single entry reaches critical activation). If this reasoning is correct, the facilitative effect of *N* with low-frequency words should be weaker for alternating-size words than for consistent-size words.

Another possibility would be to argue that the effect of size alternation arises late, rather than early, in word processing. Specifically, the effect might occur when an attempt is made to map abstract information (from the word units in the lexicon) back onto the visual representation of the stimulus pattern that gave rise to it (e.g., via recurrent feedback from the word level to the letter and/or the feature level). For instance, in the framework of an adaptive resonance model (Grossberg & Stone, 1986; Stone & Van Orden, 1994; Van Orden & Goldinger, 1994), a mismatch between the original sensory pattern and the abstract orthographic code will slow down the formation of a stable percept. In this model, high-frequency word components tend to make a faster and larger contribution in the early stages of word processing (because of faster access to stored knowledge), and they will be less affected by the effects of format distortion. As a result, high-frequency words should be less affected by size alternation than are low-frequency words. The interaction between format distortion and word frequency should also involve a greater word frequency effect: The word frequency effect should be amplified when the lexical decision task involves extra cycles of feedback as a result of a slower resonance process (see Gibbs & Van Orden, 1998; Stone & Van Orden, 1993). The model is not explicit about the role of such factors as number of neighbors. At first glance, the adaptive resonance model (or the original version of the interactive-activation model; McClelland & Rumelhart, 1981) seems to predict an inhibitory effect of the number of (higher frequency) neighbors in visual word recognition (see Van Orden & Goldinger, 1994). However, it may be important to stress that lexical decision is a word/nonword discrimination task. A lexical decision response is made when the processing system can distinguish a word from a nonword, and hence, words for which there are many similarly spelled words can be classified as members of the *word* category more rapidly/accurately than can those words with few similarly spelled words, whereas increasing the number of neighbors of a nonword in the lexical decision task should slow correct responses to these nonwords and cause more errors (see Andrews, 1997). In this light, if the influence of the orthographic neighbors occurs at the level of an abstract orthographic code, its effect should be independent of visual familiarity.

Alternatively, in the holistic biased hybrid model, Allen and colleagues (Allen & Emerson, 1991; Allen et al., 1995) proposed that high-frequency words could be identified by the (fast) word-level route when the format is familiar. However, when the stimuli are presented in an unfamiliar format, the familiarity level of the *spatial frequency pattern* will be very low, and then the word-level channel will not be able to accept (or reject) words or nonwords in a lexical decision task (Allen & Emerson, 1991). In this case, high-frequency words will be identified by the frequency-insensitive letter-level channel. As a result, this model predicts a larger alternating-size disadvantage for higher than for lower frequency words. The model does not make any assumptions about the effects of neighborhood size, although it seems reasonable to assume that since high-*N* words tend to be more visually familiar than low-*N* words, the model would predict (if anything) a larger alternating-size disadvantage for high-*N* words.

Finally, the multiple-process model of Besner and Johnston (1989) predicts that when the visual format is familiar, lexical decision responses could be made on the basis of "word-specific visual pattern information" (i.e., it would reflect a form of holistic *recognition*, as opposed to holistic *word identification*, as in the Allen et al., 1995, model). However, when the format is unfamiliar (e.g., alternating-size stimuli), participants will ordinarily use the letter-level routes (i.e., the orthographic familiarity assessment process or the unique identification of the word). High-frequency words are the ones that have the most visual familiarity, so they are the ones for which alternating size should produce the biggest drop in assessed visual familiarity. If this reasoning is correct, one would expect an interaction between alternating size and word frequency (high-frequency words being the most affected by type size). What is the role of neighborhood size in the model? Increasing *N* provides increasing amounts of evidence (via the orthographic familiarity assessment process) favoring a *word* response. Orthographic familiarity will have more impact for low-frequency words than for high-frequency words, because the latter are capable of driving the corresponding word unit over the identification threshold rather quickly. Since the orthographic familiarity assessment can be operative even with alternating-size words (keeping in mind that this assessment is based on abstract units), the model does not predict an interaction between *N* and alternating size. (Nonetheless, as we said earlier, activation in the orthographic lexicon might rise more slowly when stimuli are presented in alternating size, and then the predictions would be more complicated; e.g., see Besner & McCann, 1987.)

In short, in Experiment 1 we factorially manipulated word frequency, neighborhood size, and size alternation in a lexical decision task. When the stimuli are presented in an unfamiliar format (alternating-size presentations), holistic models predict a larger alternating-size disadvantage for higher frequency words than for lower frequency words. Unlike holistic models, analytic models predict a larger alternating size disadvantage (if anything) for lower frequency words than for higher frequency words.

EXPERIMENT 1
Size Alternation

Method

Participants. Twenty-four psychology students from the University of València took part in the experiment for course credit. All of them had either normal vision or vision that was corrected to normal and were native speakers of Spanish.

Materials. A set of 128 disyllabic Spanish words five letters in length was selected from the Spanish word pool (Alameda & Cuetos, 1995) as a function of word frequency (low-frequency words vs. high-frequency words) and neighborhood size (low-*N* words vs. high-*N* words). A set of 128 orthographically legal nonwords were constructed for the purposes of the lexical decision task. In order to maximize our chances of obtaining a facilitative effect of *N* on low-frequency words (at least with consistent-size words), high-*N* words tended to have more neighbors than did the nonwords. (Keep in mind that the effects of *N* with Spanish words seem to be less robust than the effects of *N* with English words; e.g., see Carreiras, Perea, & Grainger, 1997b.) Half of the nonwords had been created by changing an interior letter from disyllabic Spanish words with five or six orthographic neighbors. The other half of the nonwords were constructed by combining two Spanish syllables and then checking that the constructed nonwords did not have any orthographic neighbors, while being orthographically legal in Spanish. The characteristics of the items used in the experiment are presented in Table 1. Word and nonword stimuli were counterbalanced across two experimental lists, so that if a letter string was presented in alternating size in the first list, it would be presented in consistent size in the second list.

Design. For words, type size (consistent, alternating), word frequency (high frequency, low frequency), and neighborhood size (high *N*, low *N*) were varied within subjects. For nonwords, type size (consistent, alternating) and neighborhood size (high *N*, low *N*) were varied within subjects. Each participant was given a total of 256 experimental trials: 128 word trials and 128 nonword trials.

Procedure. The participants were tested in groups of 4–8 in a quiet room. Presentation of the stimuli and recording of reaction times were controlled by Apple Macintosh Classic II microcomputers. The routines for controlling stimulus presentation and reaction time collection were obtained from Lane and Ashby (1987) and from Westall, Perkey, and Chute (1986), respectively. The stimuli were presented on the computer screen in 12-point Courier (consistent-size stimuli; e.g., favor) or in 10- and 14-point Courier (alternating-size stimuli; e.g., favor). At the beginning of each trial, the sequence "> <" was presented for 200 msec on the center of the screen. After a 50-msec blank, the target stimulus was presented. The target stimulus remained on the screen until the participant's response. The participants were instructed to press one of two buttons on the keyboard to indicate whether the letter string was a Spanish word or not. The participants used their dominant hands to make the *word* responses. This decision was to be made as rapidly and as accurately as possible. Reaction times were measured from the onset of the letter string until the participant's response. The intertrial interval was set to 400 msec. Each participant received a different random order of stimuli. Each participant received a total of 24 practice trials prior to the experimental phase. The session lasted approximately 15 min.

Results

Lexical decision latencies less than 250 msec or greater than 1,500 msec were excluded from the latency analyses (less than 0.85% for words and less than 1.6% for nonwords). Mean lexical latencies for correct responses and mean error rates were calculated across individuals and across items, and these means were submitted to separate analyses of variance (ANOVAs) for participants and items, respectively. For the word data, participant and item ANOVAs based on the participants' and items' response latencies and error rates were conducted on the basis of a 2 (word frequency, high or low) × 2 (neighborhood size, high or low) × 2 (type size, consistent size or alternating size) × 2 (list, List 1 or List 2) design. In this and subsequent analyses, the list factor was included as a dummy variable to extract the variance that was due to the error associated with the lists (see Pollatsek & Well, 1995). For

Table 1
Characteristics of the Words and Nonwords Tested in Experiments 1–2

	Word Frequency		Number of Neighbors	
Condition	*M*	Range	*M*	Range
Words				
High-frequency–Hi-*N*	102	26–515	9.9	8–16
High-frequency–Lo-*N*	115	27–667	0.5	0–1
Low-frequency–Hi-*N*	3.6	1–8	9.3	8–15
Low-frequency–Lo-*N*	3.7	1–9	0.5	0–1
Nonwords				
Lo-*N*			0.0	0–0
Hi-*N*			4.3	1–10

Note—Frequency is expressed as number of occurrences in a one-million corpus.

Table 2
Mean Lexical Decision Times (LDTs, in Milliseconds) and Error Rates (ERs, in Percentages) for the Word and Pseudoword Targets in Experiment 1

	Type size					
	Same Size		Alternating Size		Alternating − Same	
Condition	LDT	ER	LDT	ER	LDT	ER
			Words			
High-frequency–Hi-*N*	591	2.3	613	3.4	22	1.1
High-frequency–Lo-*N*	575	2.1	608	3.9	33	1.8
Low-frequency–Hi-*N*	629	6.3	672	12.2	43	5.9
Low-frequency–Lo-*N*	666	3.4	718	7.0	52	3.6
			Nonwords			
Low-*N*	719	2.6	723	3.4	4	1.8
High-*N*	755	9.8	766	6.3	11	−3.5

the nonword data, participant and item ANOVAs based on the participants' and items' response latencies and error rates were conducted on the basis of a 2 (neighborhood size, high or low) × 2 (type size, consistent size or alternating size) × 2 (list, List 1 or List 2) design. The .05 level of significance was adopted throughout. The mean lexical decision latencies and error rates from the participant analysis are presented in Table 2.

Word data. The ANOVAs on the latency data showed a significant effect of word frequency [$F_1(1,22) = 128.89$, $MS_e = 2{,}052.1$; $F_2(1,120) = 95.98$, $MS_e = 4{,}708.9$]. The main effect of neighborhood size was also significant [$F_1(1,22) = 5.94$, $MS_e = 1{,}925.8$; $F_2(1,120) = 5.10$, $MS_e = 4{,}708.9$]. The interaction between these two factors was significant [$F_1(1,22) = 36.97$, $MS_e = 887.5$; $F_2(1,120) = 11.03$, $MS_e = 4{,}708.9$], which reflected a facilitative effect of *N* for low-frequency words [$F_1(1,22) = 24.45$, $MS_e = 1{,}697.5$; $F_2(1,120) = 15.57$, $MS_e = 4{,}708.9$] and a small, nonsignificant inhibitory effect of *N* for high-frequency words [$F_1(1,22) = 2.47$, $MS_e = 1{,}115.8$, $p = .13$; $F_2 < 1$]. The effect of type size was significant [$F_1(1,22) = 51.48$, $MS_e = 1{,}312.2$; $F_2(1,120) = 57.24$, $MS_e = 1{,}992.1$]. The word-frequency × type size interaction was significant [$F_1(1,22) = 6.92$, $MS_e = 774.0$; $F_2(1,120) = 5.99$, $MS_e = 1{,}992.1$], which reflected a stronger effect of type size for low-frequency words [47.5 msec; $F_1(1,22) = 47.74$, $MS_e = 1{,}162.1$; $F_2(1,120) = 50.12$, $MS_e = 1{,}992.1$] than for high-frequency words [27.5 msec; $F_1(1,22) = 18.86$, $MS_e = 924.7$; $F_2(1,120) = 13.11$, $MS_e = 1{,}992.1$]. The other interactions were not significant (all ps >.15).

The ANOVAs on the error data showed a significant effect of word frequency [$F_1(1,22) = 52.73$, $MS_e = 51.03$; $F_2(1,120) = 23.49$, $MS_e = 152.8$]: The participants made more errors on low-frequency words than on high-frequency words. The main effect of neighborhood size was not significant [$F_1(1,22) = 2.09$; $F_2 < 1$]. The effect of type size was significant [$F_1(1,22) = 25.14$, $MS_e = 38.5$; $F_2(1,120) = 26.15$, $MS_e = 49.39$]. The word frequency × type size interaction was significant [$F_1(1,22) = 28.16$, $MS_e = 15.96$; $F_2(1,120) = 12.13$, $MS_e = 49.39$], which again reflected a reliable effect of type size for low-frequency words [$F_1(1,22) = 33.82$, $MS_e = 40.47$; $F_2(1,120) = 36.95$, $MS_e = 49.39$], but not for high-frequency words [$F_1(1,22) = 3.51$, $MS_e = 14.02$, $p = .074$; $F_2(1,120) = 1.33$]. The other interactions did not approach significance (all ps > .15).

Nonword data. The ANOVAs on the latency data showed a significant effect of neighborhood size [$F_1(1,22) = 83.96$, $MS_e = 448.4$; $F_2(1,124) = 30.80$, $MS_e = 4{,}015.0$]. However, the 7.5-msec effect of type size was not significant [$F_1(1,22) = 2.41$, $MS_e = 642.0$, $p = .135$; $F_2(1,124) = 1.73$, $MS_e = 2{,}273.3$, $p = .191$]. The interaction between these two factors was not significant (both Fs < 1).

The ANOVAs on the error data showed a significant effect of neighborhood size [$F_1(1,22) = 23.23$, $MS_e = 25.97$; $F_2(1,124) = 20.16$, $MS_e = 79.76$]. The main effect of type size was not significant [$F_1(1,22) = 2.27$, $MS_e = 19.79$, $p = .146$; $F_2(1,124) = 2.37$, $MS_e = 50.46$, $p = .126$]. Finally, the interaction between these two factors was significant [$F_1(1,22) = 4.30$, $MS_e = 25.78$; $F_2(1,124) = 5.85$, $MS_e = 50.46$], which reflected an effect of alternating size for wordlike words [$F_1(1,22) = 4.84$, $MS_e = 30.69$; $F_2(1,120) = 7.84$, $MS_e = 50.46$], but not for the hermit nonwords (both Fs < 1).

Discussion

The present results replicate several well-known findings, such as the interaction between neighborhood size and word frequency for word stimuli and the inhibitory effect of *N* for nonword stimuli (e.g., Andrews, 1997; see also Carreiras et al., 1997a, 1997b, for evidence of this effect in Spanish). But the most important finding is that, contrary to the predictions of holistic models of visual word recognition, the effect of type size was *substantially* greater for low-frequency words (47.5 msec) than for high-frequency words (27.5 msec). Interestingly, alternating-size presentations did not disrupt performance for nonwords (a nonsignificant 7.5-msec effect).[2]

Together, these findings suggest that alternating-size presentations have not slowed down letter identification, because if they had, an effect of type size should have been found for nonword stimuli (as occurs with aLtErNaTiNg-cAsE presentations; see Allen et al., 1995; Besner & Johnston, 1989; Kinoshita, 1987; Mayall & Humphreys, 1996). Keep in mind that the quality of the letters is not degraded in any way, and it could well be the case that alternating-size stimuli do not disrupt letter identification. Most theorists currently assume that correct responses to nonword stimuli are generated by a decision criterion set on the time dimension (i.e., a deadline mechanism; Coltheart, Davelaar, Jonasson, & Besner, 1977; Grainger & Jacobs, 1996). The value of this negative response criterion varies as a function of overall lexical activity, higher values being adopted in the presence of stimuli that generate high levels of global lexical activity. This mechanism captures the inhibitory effect of *N* on responses to nonword stimuli in the lexical decision task. Consistent with this interpretation, the time deadline for responding *no* was not modified as a function of size alternation. If this reasoning is correct, lexical activation at the early stages of word processing will not differ much from that of a sequence of lowercase letters in consistent size. As a result, the effects of neighborhood size should be similar with the two formats, as was actually observed.

It is interesting to consider whether the word frequency × size type interaction could be due to the fact that size alternation may have altered the normal reading processes. For instance, it could be argued that the a and the o in the word favor may be grouped together and form an inappropriate unit for visual lexical access, thus slowing word identification (Mayall & Humphreys, 1996). To obtain further evidence for the present pattern of results in a less disruptive situation, Experiment 2 uses a different manipulation of visual familiarity: Items will be presented in the standard lowercase print (e.g., favor) or in UPPERCASE (FAVOR).

EXPERIMENT 2
(Lowercase vs. Uppercase)

Reading times of sentences in uppercase words are slower than reading times of sentences in lowercase words (Smith, 1969; Tinker, 1963). This is not surprising, since we are much more familiar with words printed in lowercase than with words printed in uppercase. Furthermore, words printed in lowercase are more psychophysically distinct than words printed in uppercase (Paap et al., 1984). But is this *lowercase advantage* similar for all words? Recently, Mayall and Humphreys (1996, Experiment 1) found that the effect of case occurred for low-frequency words (38 msec), but not for high-frequency words (4 msec). However, Mayall and Humphreys did not discuss the implications of this finding. Furthermore, the fact that lowercase, uppercase, and mixed-case items were presented within the same block might have provoked the use by participants of some strategies with the uppercase items. (In fairness to Mayall and Humphreys, we should note that the main goal of their experiments was to test the impact of *mixed-case* items on visual word recognition.) In fact, an earlier experiment by Paap et al. (1984, Experiment 3) failed to find a stronger effect of case for high-frequency words, relative to low-frequency words (14 vs. 16 msec).

In addition, it may be of great interest to examine whether the effect of neighborhood size differs as a function of case (lowercase vs. UPPERCASE). Words printed in lowercase have characteristic shapes in terms of patterns of ascending, descending, and neutral letters. In contrast, in uppercase no such distinctions are made, because the letters are all of the same height. If the lexical system takes word shape (or letter shape) into account, it seems reasonable to expect that the effect of neighborhood size will be reduced for lowercase words. For instance, Havens and Foote (1963) suggested that similarly spelled words that are visually similar, such as *list* and *lint*, are closer "neighbors" than words with a different outline shape (e.g., *list* and *lift*). Nevertheless, in an analytical model, the format of the word is lost early in word processing and the effect of *N* should be similar for lowercase and uppercase words.

Method

Participants. Twenty-four psychology students from the University of València took part in the experiment for course credit. All of them had either normal vision or vision that was corrected to normal and were native speakers of Spanish. None of them had participated in the previous experiment.

Materials. The materials were the same as those in Experiment 1.

Design. For words, case type (lowercase, uppercase), word frequency (high frequency, low frequency), and neighborhood size (high *N*, low *N*) were varied within subjects. For nonwords, case

Table 3
Mean Lexical Decision Times (LDTs, in Milliseconds) and Error Rates (ERs, in Percentages) for the Word and Nonword Targets in Experiment 2

	Case Type					
	Lowercase		Uppercase		Upper-Lower	
Condition	LDT	ER	LDT	ER	LDT	ER
		Words				
High-frequency–Hi-*N*	612	1.6	620	1.6	8	0.0
High-frequency–Lo-*N*	602	2.1	601	1.8	−1	−0.3
Low-frequency–Hi-*N*	666	8.1	693	8.1	29	0.0
Low-frequency–Lo-*N*	692	10.9	721	9.4	27	−1.5
		Nonwords				
Low-*N*	707	2.2	723	2.1	16	−0.1
High-*N*	769	9.2	765	7.0	−4	−2.2

type (lowercase, uppercase) and neighborhood size (low N, high N) were varied within subjects. Each participant was given a total of 256 experimental trials: 128 word trials and 128 nonword trials.

Procedure. The procedure was the same as that in Experiment 1, except that the stimuli were presented on the computer screen in 12-point Courier (e.g., favor vs. FAVOR).

Results

Lexical decision latencies less than 250 msec or greater than 1,500 msec were excluded from the latency analyses (less than 0.62% for words and less than 1.9% for nonwords). Mean lexical latencies for correct responses and mean error rates were calculated across individuals and across items, and these means were submitted to separate ANOVAs for participants and items, respectively. The mean lexical decision latencies and error rates from the participant analysis are presented in Table 3.

Word data. Not surprisingly, the ANOVAs on the latency data showed a significant effect of word-frequency [$F_1(1,22) = 113.40$, $MS_e = 3{,}039.4$; $F_2(1,120) = 99.89$, $MS_e = 5{,}192.3$]. The main effect of neighborhood size was not significant [$F_1(1,22) = 1.62$, $MS_e = 1{,}195.3$, $p > .15$; $F_2(1,120) = 1.27$, $MS_e = 5{,}192.3$, $p > .15$]. The interaction between these two factors was significant [$F_1(1,22) = 30.33$, $MS_e = 667.12$; $F_2(1,120) = 7.41$, $MS_e = 5{,}192.3$] which reflected a facilitative effect of N on low-frequency words [$F_1(1,22) = 14.64$, $MS_e = 1{,}183.94$; $F_2(1,120) = 7.42$, $MS_e = 5{,}192.3$] and an inhibitory effect of N on high-frequency words in the analysis by participants [$F_1(1,22) = 7.12$, $MS_e = 678.4$; $F_2(1,120) = 1.27$, $MS_e = 5{,}192.3$, $p > .15$]. The effect of case type was significant [$F_1(1,22) = 6.80$, $MS_e = 1{,}785.0$; $F_2(1,120) = 11.81$, $MS_e = 1793.6$]. More important, the word frequency × case type interaction was significant [$F_1(1,22) = 6.46$, $MS_e = 1{,}086.7$; $F_2(1,120) = 7.33$, $MS_e = 1{,}793.6$], which reflected a reliable effect of case type for low-frequency words [28 msec; $F_1(1,22) = 10.16$, $MS_e = 1{,}851.46$; $F_2(1,120) = 18.88$, $MS_e = 1{,}793.6$], but not for high-frequency words (−4.5 msec, both Fs < 1). The other interactions were not significant (all ps > .15).

The ANOVAs on the error data showed a significant effect of word frequency [$F_1(1,22) = 58.84$, $MS_e = 44.15$; $F_2(1,120) = 25.30$, $MS_e = 136.9$]. The other effects were not significant (all ps > .15).

Nonword data. Not surprisingly, the ANOVAs on the latency data showed a significant effect of neighborhood size [$F_1(1,22) = 69.74$, $MS_e = 923.8$; $F_2(1,124) = 59.90$, $MS_e = 3{,}138.1$]. The other effects were not significant (all ps > .10).

As in Experiment 1, the ANOVAs on the error data showed a significant effect of neighborhood size [$F_1(1,22) = 58.85$, $MS_e = 14.63$; $F_2(1,124) = 37.04$, $MS_e = 61.97$]. The other effects were not significant (all ps > .10).

Discussion

The pattern of results in Experiment 2 resembles closely that of Experiment 1, in the sense that low-frequency words are the most affected by visual familiarity. Specifically, high frequency words do not seem to be affected by case type, whereas there is a robust 28-msec effect of case for low-frequency words, replicating the results of Mayall and Humphreys (1996). In addition, the results also show an interaction between neighborhood size and word frequency for word stimuli and an inhibitory effect of N for nonword stimuli.[3] Finally, as in Experiment 1, latencies on nonwords are not affected by format presentation.

GENERAL DISCUSSION

The main findings of the present experiments can be summarized as follows. First, there is a reliable effect of format presentation (both size alternation and case type), which is greater for low-frequency words than for high-frequency words. Second, the facilitative effect of neighborhood size on low-frequency words is not modified by using less visually familiar items (i.e., words printed in alternating-size or in UPPERCASE). Taken together, these results have clear implications with respect to the locus of visual familiarity in visual word recognition.

The results of Experiments 1 and 2 clearly weaken the view that high-frequency words can be identified via a fast word-level route as a function of visual familiarity. If global word shape had played a role in the process of lexical access via a fast word-level channel (as was proposed by Allen et al., 1995), the effect of size alternation (or case type) should have been greater (rather than smaller) for high-frequency words than for low-frequency words. Furthermore, the lack of an effect of case on high-frequency words (Experiment 2; see also Mayall & Humphreys, 1996) can also be considered evidence against the *word shape* hypothesis: Keep in mind that, ordinarily, both high- and low-frequency words are much more often seen in lowercase than in UPPERCASE. Thus, the present findings suggest that recognition of familiar words does not seem to rely on word shape information.

Nonetheless, the main effect of size alternation for words—and the effect of case type for low-frequency words—implies that, at some stage, visual familiarity *plays a role* in the process of lexical decisions. Interestingly, the fact that the effect of size alternation (or case type) occurs for word rather than for nonword stimuli suggests that this effect occurs *late* in word processing (see e.g., Besner, 1983, 1989, Forster & Guess, 1996, and Kinoshita, 1987, for a similar account). In other words, alternating-size presentations have not slowed down the initial stage of letter processing, because, if they had, an alternating-size effect (or an effect of case type) should have been found for nonword stimuli, as occurs with aLtErNaTiNg-cAsE presentations (see, e.g., Besner & Johnston, 1989; Kinoshita, 1987). Instead, if the confluence of abstract orthographic information and visual information is sensitive to visual familiarity (e.g., via recurrent feedback from the word level to the letter or the feature level in an interactive-activation model), the decrement on performance that accompanies the effect of pattern distortion for words follows naturally (Besner, 1989).

In this light, the present results can be interpreted in the framework of a *resonance* model (see, e.g., Grossberg & Stone, 1986; Stone & Van Orden, 1993, 1994; Van Orden & Goldinger, 1994), in which the network "senses" the degree of match between subpatterns of the total bottom-up (the original sensory pattern) and top-down (the abstract orthographic code) input. In this framework, high-frequency word components tend to make a faster and larger contribution in the early stages of word processing (because familiar words accelerate toward resonance very quickly and have straighter trajectories; see Gibbs & Van Orden, 1998; Stone & Van Orden, 1993; Van Orden & Goldinger, 1994), and they will be less affected by the format of bottom-up information. A mismatch between the original sensory pattern and the abstract orthographic code will slow down the formation of a stable percept. In other words, the processing of a word printed in alternating size (or in uppercase) may require some extra feedback cycles to reach stability, especially for the lower frequency words.[4] Thus, the word frequency effect should be greater for words printed in alternating size (or in uppercase) than for those words printed in plain lowercase letters, just as our results showed (the size of the word frequency effect was 84.5 vs. 64.5 msec in Experiment 1 and 96.5 vs. 72 msec in Experiment 2, respectively). Consistent with this proposal, the participants made substantially more errors on low-frequency words when they were presented in an unfamiliar format than when they were presented in plain lowercase format (i.e., the deadline for responding *no* might have been set too early in a number of trials; see Grainger & Jacobs, 1996), whereas that was not the case for high-frequency words. In addition, the participants made more false positive errors on wordlike nonwords when they were presented in a familiar format (plain lowercase letters) than when they were presented in a less familiar format.

If the previous account is correct, lexical activation in the early stages of word processing for an alternating-size word (or for a word printed in uppercase) will not differ much from that of a sequence of lowercase letters with the same size. As a result, a facilitative effect of *N* for low-frequency words and an inhibitory effect of *N* for nonwords would be expected via a "fast guess" on the basis of global lexical activation (Grainger & Jacobs, 1996), as actually occurred. We must bear in mind that this computation is made on the basis of abstract orthographic codes that should be independent of visual form. In this light, it may be of interest to indicate that the magnitude of the neighborhood size effect for low-frequency words was not modified by case type (lowercase vs. UPPERCASE; Experiment 2), despite the fact that lowercase letters are psychophysically more distinct than uppercase letters. This finding suggests that information about visual form is probably lost early in the process of word recognition, as is predicted by analytical models.

Is it necessary to postulate the presence of recurrent feedback in order to interpret the present results? To analyze this possibility, it may be of interest to examine whether or not a feedforward model, such as the multiple-route model (Besner & Johnson, 1989), can accommodate the present results.[5] As was stated in the introduction, a positive lexical decision response is initiated when the *yes* unit exceeds the criterion from any of three sources (visual familiarity, orthographic familiarity, and unique word identification). Increasing degrees of visual familiarity (via the visual familiarity assessment process; see Figure 2) provide increasing amounts of evidence favoring a *yes* response. This boost is more likely to affect low-frequency words that have weaker connections and, hence, slower activation times, as occurred in our experiments. The overall effect of visual familiarity is also greater when the manipulation produces a bigger distortion (i.e., alternating size rather than just the less common uppercase format). In addition, increasing *N* provides increasingly greater amounts of evidence (via the orthographic familiarity assessment process) favoring a *yes* response. Similar to the effect of visual familiarity, orthographic familiarity will have more impact for low-frequency words than for high-frequency words, because the latter are capable of driving the *yes* unit over the word identification criterion early and by themselves (see, e.g., Grainger & Jacobs, 1996). Finally, with respect to the inhibitory effect of *N* on nonword responses, it could be argued that the deadline is adjusted on line, depending on total lexical activity (i.e., the output of the orthographic familiarity mechanism). Thus, it seems that a feedforward model such as the multiple-route model (Besner & Johnston, 1989) gives a reasonable account of the overall pattern of results obtained in the present experiments. (Of course, it might be argued that computer simulations are needed to obtain specific predictions from the model.) By Occam's razor, a feedforward model should be preferred to a top-down interactive model when the two models predict a similar pattern of results (see Norris, McQueen, & Cutler, 2000, for a defense of feedforward models in word recognition). Although we believe that models of word recognition might tolerate local complexity (i.e., recurrent feedback) in the interest of global simplicity and biological plausibility (see Grossberg, 2000; Luce, Goldinger, & Vitevitch, 2000), this is currently a controversial issue.

To conclude, the present experiments provide evidence against models of visual word recognition that assume that word-level codes (via a word's overall shape) are computed by the reading system. In this light, eye movement research has failed to obtain empirical evidence to suggest that visual information is combined across saccades during reading in an integrative visual buffer (see Rayner, 1998). As Paap et al. (1984) concluded, word shape does not seem to be in good shape for the race to the lexicon.

REFERENCES

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencia de las unidades lingüísticas del castellano* [Dictionary of word frequency in

Spanish]. Oviedo: Servicio de publicaciones de la Universidad de Oviedo.

Allen, P. A., & Emerson, P. L. (1991). Holism revisited: Evidence for independent word-level and letter-level processors during word and letter processing. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 489-511.

Allen, P. A., Wallace, B., & Weber, T. A. (1995). Influence of case type, word frequency, and exposure duration on visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 914-934.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 234-254.

Andrews, S. (1997). The effects of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, **4**, 439-461.

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

Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology*, **37**, 429-438.

Besner, D. (1989). On the role of outline shape and word-specific visual patterns in the identification of function words: NONE. *Quarterly Journal of Experimental Psychology*, **41A**, 91-105.

Besner, D., & Johnston, J. C. (1989). Reading and the mental lexicon: On the uptake of visual information. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 291-316). Cambridge, MA: MIT Press.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII*: *The psychology of reading* (pp. 201-219). Hillsdale, NJ: Erlbaum.

Carreiras, M., Perea, M., & Grainger, J. (1997a, November). *Blocking by frequency and neighborhood density*. Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia.

Carreiras, M., Perea, M., & Grainger, J. (1997b). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **23**, 857-871.

Cattell, J. M. (1886). The time taken by cerebral operations. *Mind*, **11**, 220-242.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). Hillsdale, NJ: Erlbaum.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. W. Walker (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North-Holland.

Forster, K. I., & Guess, K. (1996, October). *Effects of prime duration and visual degradation in masked priming*. Paper presented at the 37th Annual Meeting of the Psychonomic Society, Chicago.

Gibbs, P., & Van Orden, G. C. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1162-1187.

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

Grossberg, S. (2000). Brain feedback and adaptive resonance in speech perception. *Behavioral & Brain Sciences*, **23**, 332-333.

Grossberg, S., & Stone, G. O. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review*, **93**, 46-74.

Havens, L. L., & Foote, W. E. (1963). The effect of competition on visual duration thresholds and its independence of stimulus frequency. *Journal of Experimental Psychology*, **65**, 6-11.

Healy, A. F., & Cunningham, T. F. (1992). A developmental evaluation of the role of word shape in word recognition. *Memory & Cognition*, **20**, 141-150.

Healy, A. F., Oliver, W. L., & McNamara, T. P. (1987). Detecting letters in continuous text: Effects of display size. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 279-290.

Kinoshita, S. (1987). Case alternation effect: Two types of word recognition? *Quarterly Journal of Experimental Psychology*, **39A**, 701-720.

Lane, D. M., & Ashby, B. (1987). PsychLib: A library of machine language routines for controlling psychology experiments on the Apple Macintosh computer. *Behavior Research Methods, Instruments, & Computers*, **19**, 246-248.

Luce, P. A., Goldinger, S. D., & Vitevitch, M. S. (2000). It's good... but is it ART? *Behavioral & Brain Sciences*, **23**, 336.

Mayall, K., & Humphreys, G. W. (1996). Case mixing and the task-sensitive disruption of lexical processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 278-294.

Mayall, K., Humphreys, G. W., & Olson, A. (1997). Disruption to word or letter processing: The origins of case-mixing effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1275-1286.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* **86**, 287-330.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375-407.

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

Paap, K. R., & Johansen, L. S. (1994). The case of the vanishing frequency effect: A retest of the verification model. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1129-1157.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, **89**, 573-594.

Paap, K. R., Newsome, S. L., & Noel, R. W. (1984). Word shape's in poor shape for the race to the lexicon. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 413-428.

Pollatsek, A., Perea, M., & Binder, K. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1142-1158.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **123**, 372-422.

Rudnicky, A. I., & Kolers, P. A. (1984). Size and case of type as stimuli in reading. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 231-249.

Schiller, P., Logothetis, N., & Charles, E. (1991). Parallel pathways in the visual system: Their role in perception at isoluminance. *Neuropsychologia*, **29**, 433-441.

Sears, C. R., Lupker, S. J., & Hino, Y. (1999). Orthographic neighborhood effects in perceptual identification and semantic categorization tasks: A test of the multiple read-out model. *Perception & Psychophysics*, **61**, 1537-1554.

Smith, F. (1969). Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science*, **14**, 261-263.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 744-774.

Stone, G. O., & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1248-1268.

Tinker, M. A. (1963). *Legibility of print.* Ames: Iowa State University.

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

Westall, R., Perkey, M. N., & Chute, D. L. (1986). Accurate millisecond timing on Apple's Macintosh using Drexel's MilliTimer. *Behavior Research Methods, Instruments, & Computers*, **18**, 307-311.

NOTES

1. Coltheart's *N* is defined as the number of words differing by a single letter from the stimulus, preserving letter positions; for example, *worse*, and *house* are orthographic neighbors of *horse* (Coltheart, Davelaar, Jonasson, & Besner, 1977; for recent reviews, see Andrews, 1997; Grainger & Jacobs, 1996; Pollatsek, Perea, & Binder, 1999).

2. We should like to note that the lack of an effect of size alternation for nonwords is in sharp contrast with the findings of Mayall et al. (1997) with the naming task. Specifically, Mayall et al. found that alternating the size of letters had a much greater disruptive effect on naming nonwords than on naming words (96 vs. 25 msec, respectively). These results seem to suggest that the nonlexical route in naming is particularly disrupted by size alternation.

3. We should note that the small inhibitory effect of *N* for high-frequency words in the analysis by subjects, which also occurred in Experiment 1 (albeit the *F* ratio was not significant), is not new (see, e.g., Carreiras, Perea, & Grainger, 1997a; see also Andrews, 1992, for this same pattern of data in English), and it can be captured by the multiple read-out model (see Grainger & Jacobs, 1996).

4. Nonetheless, we acknowledge that simulations on an implemented *resonance* model would be necessary to examine whether or not it can capture the basic pattern of observed effects. As Ken Paap and Sandy Pollatsek indicated, there is the possibility that pattern distortion might disrupt low-frequency words more than high-frequency words or even the reverse, depending on the choice of parameter settings.

5. We thank Ken Paap for providing this reasoning for the multiple-route model.

(Manuscript received August 9, 2000;
revision accepted for publication August 23, 2001.)